**Gil Daniel Bregieiro Carraco**

# Regulating the Pace of the Embryo Clock: RNA Molecules in Temporal Control of Vertebrate Development


Faculdade de Medicina e Ciências Biomédicas

2021

**Gil Daniel Bregieiro Carraco**

# Regulating the Pace of the Embryo Clock: RNA Molecules in Temporal Control of Vertebrate Development

**PhD Program in Mechanisms of Disease**

**and Regenerative Medicine**

**Work developed under the supervision of:**

Prof. Doutora Raquel P. Andrade


Faculdade de Medicina e Ciências Biomédicas

2021

*This page is intentionally left blank*

# Regulating the Pace of the Embryo Clock:

# RNA Molecules in Temporal Control of Vertebrate Development

*Declaração de autoria de trabalho*

Declaro ser o autor deste trabalho, que é inédito e original. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

.

*This work was performed at:*

**Temporal Control of Cell Differentiation Lab**

Center for Biomedical Research (CBMR)

Algarve Biomedical Center Research Institute (ABC-RI)

Faculdade de Medicina e Ciências Biomédicas da Universidade do Algarve

Campus de Gambelas, 8005-139 Faro, Portugal

**Schier Lab**

Department of Molecular and Cellular Biology

Harvard University, 52 Oxford Street, Cambridge, MA 02138, United States of America

This work was supported by National Portuguese funding through FCT – Fundação para a Ciência e a Tecnologia, scholarship SFRH/BD/101609/2014 and Fulbright Commission Portugal.

*This page is intentionally left blank*

## Acknowledgements

Em primeiro lugar gostaria de agradecer à minha orientadora e amiga Professora Raquel Andrade por todos os ensinamentos que me passou ao longo destes anos, por me ter ajudado neste nesta longa caminhada que começou ainda antes doutoramento e muito importante pela paciência que teve. Agradeço por me ter dado a oportunidade de fazer investigação no seu grupo, mesmo sabendo que não teria muitas bases nesta incrível área do conhecimento que é a Biologia de Desenvolvimento. Agradeço também pela força que me deu, mesmo nas alturas mais complicadas, tanto no decorrer dos trabalhos como nos momentos pessoais mais complicados.

I would like to thank Professor Alex Schier, for receiving me to work the research group under his leadership. It was a stressful and demanding time, but above all, it was an awesome experience and an important growth moment as a researcher. I am thankful for all the advice and brainstorming moments.

Ao Programa Doutoral ProRegeM e a toda a direção por me terem permitido conduzir o meu trabalho dentro deste fantástico programa. Agradeço em especial aos Professores José Belo, Rui Martinho e José Bragança por toda a ajuda e apoio que me deram ao longo deste anos.

Agradeço também a todo o grupo de laboratório desde o memento em comecei esta aventura na Biologia do Desenvolvimento. In the University of Minho, I would like to thank Sheeba, Hugo, Rute Moura and Mariana for taking me on board and teach me the first steps in this field. À Ana Gonçalves, a minha primeira companheira de armas. A todos os membros do laboratório em especial à Cristina, ao Tomás e à Isabel Duarte por toda a ajuda e amizade que tornou ainda mais fácil cada dia no lab.

To all Scheir lab for all the support during my time in the United States. A special thanks to Tessa, Adam, Nate, Jeff and Katheryn for teaching me all that I need to work with zebrafish. To Vassillis for the long conversations and support.

A toda à minha família em especial aos meus pais e aos meus avós pelo suporte que me deram durante todos estes anos. E, sem dúvida, que foi graças a vocês e há educação que me deram que me permitiu chegar este ponto!

A todos os meus amigos que me deram suporte ao longo destes anos e que me aturam e que especialmente me acompanharam nas loucuras. Especialmente à Rafaela, ao Gonçalo, ao Dino, ao Daniel, à Susana e ao Ricardo, que sempre me apoiaram e me ajudaram nos tempos difíceis e festejaram comigo nas vitórias.

A todos vós muito obrigado!!!

*This page is intentionally left blank*

# Resumo

Durante as últimas décadas têm-se estudado com profundidade os mecanismos genéticos que levam ao correto desenvolvimento do embrião vertebrado. Estes mecanismos têm sido estudados especialmente na caraterização da expressão genética nos tecidos e de que modo a sua presença ou ausência se reflete na correta diferenciação de uma célula, no desenvolvimento de um tecido ou órgão e na correta formação do organismo. Contudo para além da localização e intensidade da expressão, uma outra dimensão muito menos explorada é o Tempo. A desregulação temporal da expressão génica pode levar a que os tecidos se diferenciem num momento ou ritmo impróprio, podendo ter consequências graves e levar mesmo à morte do organismo.

Um dos processos que evidencia uma periodicidade durante o desenvolvimento, é a somitogénese. À medida que o embrião alonga ao longo do eixo antero-posterior graças à ingressão de células no botão caudal, as células da mesoderme pré-somítica (PSM) anterior formam periodicamente pares de estruturas esféricas transitórias que flanqueiam a notocorda, conhecidas como sómitos. Os sómitos diferenciam-se sendo responsáveis pela formação do esqueleto axial, bem como da musculatura esquelética e a derme das costas. A desregulação temporal da somitogénese, provoca doenças congénitas graves que afetam especialmente a formação das vertebras, como a disostose espondilocostal.

Existe um relógio embrionário (EC) subjacente à periodicidade da somitogénese. O EC foi primeiramente descrito por Palmeirim e colegas (1997) no desenvolvimento do embrião de galinha. Este grupo descobriu que o gene *hairy1* tem ciclos de expressão na PSM com o mesmo período que a somitogénese, 90 minutos. Posteriormente, este mecanismo foi descoberto noutros organismos e hoje em dia, sabe-se que é conservado em todos os vertebrados. O período EC é específico de cada espécie, sendo, por exemplo, 30 minutos no peixe-zebra, 120 no ratinho e entre 5 e 6 horas no Humano. Sabe-se também que vários genes pertencentes a várias vias de sinalização possuem expressão cíclica. Mais concretamente, pertencentes às vias de Notch, FGF e WNT.

Dez anos depois da descoberta do EC, foi descrito que este mecanismo não era exclusivo da PSM. Pascoal e colegas (2007) descobriram que o relógio embrionário estava presente no crescimento do membro da galinha. Mais concretamente, o grupo descreveu que o gene *hairy2* tem uma expressão cíclica, apresentando um período de seis horas em vez dos 90 minutos na PSM. O EC é caracterizado por mecanismos de regulação por *feedback*

negativo em que um gene é expresso e traduzido e a proteína vai reprimir a sua própria expressão. Para além disso, o mecanismo é refinado por atrasos em diversas etapas, desde a transcrição, passando pela exportação nuclear do mRNA, pela tradução e terminando na degradação do mRNA e da proteína. Contudo, a regulação do EC, mais propriamente, como é capaz de apresentar diferentes períodos no mesmo organismo, ainda permanece por esclarecer, constituindo a questão central do presente trabalho.

Neste trabalho foi explorada a hipótese da regulação da periodicidade do EC em diferentes tecidos ser exercida ao nível das moléculas de RNA. Para abordar esta hipótese, foram aplicadas diversas metodologias experimentais. Em primeiro lugar foram identificados os transcritos produzidos por *hairy1* e analisado o seu impacto na segmentação do tronco embrionário (Capítulo 2). Para tal, foi feita uma sobre-expressão de ambos os transcritos identificados na PSM do embrião de galinha. Os nossos resultados indicam que a sobre-expressão do *s-hairy1,* tal como de *hairy1*, leva a um atraso no alongamento dos embriões, e a um atraso no início da somitogénese. Neste capítulo, adicionalmente, procurámos aprofundar o conhecimento acerca do relógio embrionário no membro. Estudámos a expressão de *hairy1*, descobrindo que também cicla com um período de 6 horas no mesênquima distal do membro.

A região 3'UTR do mRNA é um dos maiores responsáveis pela regulação do tempo de meia-vida do próprio mRNA. Tendo isso em mente, estudámos os 3'UTRs de três genes centrais do relógio embrionário do peixe-zebra, *her1, her7* e *deltaC* (*dlc*) (Capítulo 3). No caso de *her1 e her7* deletou-se 3'UTR alternativos, contudo os resultados obtidos não revelaram um papel para os 3'UTRs alternativos no funcionamento do EC. No caso do gene *dlc,* dividiu-se o 3'UTR em três regiões de regulação (RR), e procedeu-se à deleção dessas regiões isoladamente, ou em combinação. Quando deletadas RR1, RR2 ou a combinação das duas (RR1&2), não foram detetadas alterações significativas à expressão de *dlc.* Contudo, quando foram removidas RR3 ou RR2&3 obteve-se uma estabilização da expressão dos genes do EC. Uma avaliação mais aprofundada das deleções RR2 e RR3 revelou que os embriões produziam menos sómitos e mais pequenos, resultado mais notório na deleção RR3. Em ambos os casos também se verificou uma estabilização da expressão de *her7.*

Por último, decidiu-se estudar se os miRNAs podem estar envolvidos na regulação dos períodos do EC (Capítulo 4). Para tal, analisou-se o *miRNoma* por sequenciação de RNA a partir de três tecidos que apresentam oscilações do EC com diferentes ritmos: PSM determinada (D-PSM), PSM indeterminada (U-PSM) e o domínio distal cíclico do membro superior (DCD). Foram identificados 926 miRNAs conhecidos. Para além disso descobrimos mais 1141 possíveis novos miRNAs em galinha (quase duplicando os que se encontram anotados na base de dados miRBase). Procedemos, posteriormente, a uma análise de

expressão diferencial e à validação dos dados por RT-qPCR. Adicionalmente, os miRNAs diferencialmente expressos entre os tecidos foram comparados com os genes negativamente regulados nos mesmos tecidos e procedemos a uma análise de enriquecimento funcional GO e REACTOME.

Em suma, utilizando uma nossa estratégia plural para testar a nossa hipótese, levou-nos a um aprofundar do conhecimentos dos mecanismos que controlam a expressão genética periódica nos tecidos onde o EC opera.

**Palavras-chave:** Relógio Embrionário, Somitogénese, Crescimento do Membro, Transcritos Alternativos, 3'UTR, miRNA

*This page is intentionally left blank*

# Abstract

Vertebrate embryo body segmentation occurs progressively over time. An embryo clock (EC) characterized by gene expression oscillations underlies somitogenesis periodicity. The EC periodicity is species-specific, *i.e.,* 30 minutes and 90 minutes in zebrafish and chicken, respectively. The EC was also found in chicken forelimb outgrowth, however, with a 6-hour pace. Robust EC gene expression oscillations are built by negative feedback loops, which require tight regulation of mRNA stability. This thesis will present work addressing the hypothesis that alternative transcripts and/or miRNAs may account for different paces of the EC.

Different approaches were followed to address this hypothesis: 1) we analyzed the alternative transcripts of the chicken EC *hairy1* gene and found that both *s-hairy1* and *hairy1* overexpression promoted trunk developmental delay. We identified *hairy1* gene expression oscillations in the limb bud with a periodicity of 6 hours. 2) Using zebrafish mutants with truncated 3'UTRs in EC mRNAs we found that deletion of the distal portion of the *dlc* 3'UTR stabilizes EC expression and promotes segmentation alterations. 3) We performed miRNA sequencing (miRNA-Seq) of three chicken embryo tissues: determined PSM (D-PSM), undetermined PSM (U-PSM), and forelimb distal cyclic domain (DCD). We identified 926 known miRNAs and 1141 candidate novel miRNAs not previously described in chicken. Furthermore, a differential expression analysis was conducted, and specific miRNAs were selected for validation and functional analysis.

Altogether, by tackling our hypothesis in multiple mechanisms involving RNA regulation, we gave a step further in understanding the temporal control of the EC.



**Keywords:** Embryo Clock, Somitogenesis, Limb Outgrowth, Alternative Transcripts, 3'UTR, miRNA

*This page is intentionally left blank*

# Table of Contents

*This page is intentionally left blank*

# List of Figures

*This page is intentionally left blank*

# List of Tables

*This page is intentionally left blank*

# Abbreviations

| | |
|---|---|
| **3‘UTR** | Three Prime Untranslated Region |
| **5‘UTR** | Five Prime Untranslated Region |
| **A-P** | Anterior-Posterior |
| **AER** | Apical Ectodermal Ridge |
| **AGO** | Argonaute |
| **APA** | Alternative Polyadenylation |
| **ARE** | AU-rich element |
| **bHLH** | basic Helix-Loop-Helix |
| **ChIP** | Chromatin immunoprecipitation |
| **CPD** | Central Mesenchymal Domain |
| **DCD** | Distal Cyclic Domain |
| **DF** | Differentiation Front |
| **dlC** | DeltaC |
| **dlD** | DeltaD |
| **Dll1** | Delta-like 1 |
| **dpf** | days post-fertilization |
| **Dusp6** | Dual specificity phosphatase 6 |
| **EC** | Embryo Clock |
| **eIF4F** | eukaryotic Initiation Factor 4F |
| **EMT** | Epithelial to Mesenchymal Transition |
| **GFP** | Green Fluorescent Protein |
| **FGF** | Fibroblast Growth Factor |
| **ESC** | Embryonic Stem Cells |
| **Her** | Hairy and enhancer of split-related |
| **HES** | Hairy/Enhancer-of-Split |
| **HH4** | Hamburger-Hamilton stage 4 |
| **hpf** | hours post fertilization |
| **Hox** | Homeobox protein |
| **Lfng** | Lunatic fringe |
| **Mfng** | Manic Fringe |
| **miRNA** | microRNA |
| **miR-Seq** | micro-RNA sequencing |
| **mRNA** | Messenger RNA |
| **MZ** | Maternal to Zygote Transition |
| **ncRNA** | non-coding RNA |
| **NICD** | Notch-Intracellular Domain |
| **NPC** | neural progenitor cells |
| **P-D** | Proximal-Distal |
| **PABP** | poly-A binding protein |
| **PAS** | polyadenylation sites |

| | |
|---|---|
| **Poly(A)** | Polyadenylation |
| **PPD** | Postirior Positive Domain |
| **PRE** | Pumilio ResponseEelement |
| **pre-miRNA** | Precursor miRNA |
| **pri-miRNA** | Primary miRNA |
| **PS** | Primitive Streak |
| **PSM** | Presomitic Mesoderm |
| **PSM_D / D-PSM** | Determined Presomitic Mesoderm |
| **PSM_U / U-PSM** | Undetermined Presomitic Mesoderm |
| **PZ** | Progress Zone |
| **RA** | Retinoic Acid |
| **RACE** | Rapid Amplification of cDNA Ends |
| **RBP** | RNA-biding proteins |
| **RBP-jk** | Recombination signal Binding Protein for immunoglobulin kappa J |
| **Rfng** | Radical fringe |
| **RISC** | RNA-induced silencing complex |
| **RNA** | Ribonucleic acids |
| **rRNA** | ribosomal RNA |
| **RR** | Regulatory regions |
| **Shh** | Sonic Hedgehog |
| **tRNA** | transfer RNA |
| **TSS** | Transcription Start Sites |
| **WNT** | Wingless-wint |
| **WT** | Wild-Type |
| **YFP** | Yellow Fluorescent Protein |
| **YSL** | Yolk Syncytial Layer |
| **ZPA** | Zone of Polarizing Activity |

# Chapter 1

## General introduction

*This page is intentionally left blank*

## 1.1 Embryo development over time

The study of embryo development has been targeted by human curiosity since thousands of years ago. Egyptians and the Babylonians were interested in embryo malformations, and in Ancient Greece, for example, Hippocrates and Aristoteles already studied the chicken egg and the chicken embryo (Horder, 2010). Modern Development Biology dates from early 1900 with the work of Hans Spemann and Hilde Mangold, when they first described the Spemann organizer in the newt embryo (Sander and Faessler, 2001). Since then, Developmental Biologists have discovered and characterized numerous embryonic processes, uncovering how an embryo develops in time and space.

Embryo development starts with fertilization, after which the zygote enters the cleavage phase, in which the embryo undergoes multiple and fast cell divisions until it forms a mass of cells called the morula. The morula cells continue to divide and enter the blastula phase. For the first time during embryo development, the embryo cells have an organization and commitment. The human blastocyst is divided into two cell populations, the trophoblast, that gives rise to the extraembryonic tissues, such as the placenta, and the inner cell mass that originates the embryo proper (Firmin and Maître, 2021; Gilbert and Barresi, 2019). It is during this phase, between the fifth and sixth day after fertilization, that Human embryo implantation occurs (Wilcox *et al*., 1999).

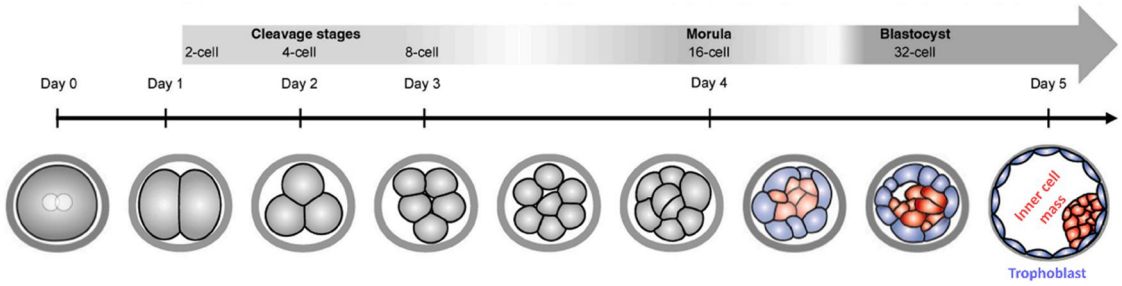



**Figure 1.1 - First stages of the human embryo development: from Zygote to Blastocyst.** During the first days of Human embryo development, there is no morphogenesis, and the cells only divide until they form the morula at day 4. After internalization and lumen formation, there are two distinct cell populations in the blastocyst, the trophoblast (cells in blue) and the inner cell mass (cells in red). Adapted from (Firmin and Maître, 2021).

Upon the end of the blastula stage, the embryo enters the gastrulation phase. As Lewis Wolpert once said, it is “not birth, marriage or death, but gastrulation which is truly the most important time in your life.” (Stern, 2021). Indeed, gastrulation is defined by the formation of the three embryonic germ layers: Ectoderm, Mesoderm, and Endoderm, that will give rise to the entire embryo (Gilbert and Barresi, 2019). For example, the ectoderm gives rise to the

epidermis, brain, and nerves; mesoderm differentiates into connective tissue, muscle, blood, heart, skeleton, gonads, and kidneys; and endoderm originates the gut and respiratory systems (Gilbert and Barresi, 2019).

The amniote chicken embryo has been used as a model to study Human gastrulation due to its striking similarities. Before gastrulation takes place, the chicken embryo is only composed of two cell layers, a dorsal layer, the epiblast, and a ventral layer, the hypoblast. From the epiblast, the three germ layers are formed: ectoderm, mesoderm, and endoderm (Gilbert and Barresi, 2019). Gastrulation starts at Hamburger-Hamilton stage 4 (HH4) (Hamburger and Hamilton, 1951) when the primitive streak is formed. From stage HH4 onwards, the epiblast cells start to ingress through the primitive streak to the blastocoel by undergoing an epithelial to mesenchymal transition (EMT) and form the mesoderm and endoderm (Figure 1.2 A) (Viebahn, 1995). The cell position along the primitive streak dictates the fate of ingressing cells (Figure 1.2 B) (Psychoyos and Stern, 1996). The cells in the Hensen's node ingress ventrally and then move anteriorly. These cells will originate the head and axial mesoderm structures. The cells that ingress through the primitive streak move laterally and originate the remaining mesoderm and endoderm. The antero-posterior (A-P) position of the prospective mesoderm cells in the primitive streak dictates their fate, and they can be divided into four regions from the embryo midline to its flanks (Figure 1.2 B). The axial mesoderm is accumulated in the midline and originates the prechordal mesoderm and notochord. Flanking the axial mesoderm is deposited the paraxial mesoderm, will give rise to the somites through segmentation of the presomitic mesoderm (PSM). Adjacent to the paraxial mesoderm, the intermediate mesoderm is formed, responsible for the development of the urogenital system, such as kidneys and gonads. At last, the fourth and most lateral mesoderm region is the lateral plate mesoderm. It will differentiate in multiple tissues, including the limbs and the circulatory system (Freitas *et al.*, 2001; Psychoyos and Stern, 1996).

An important feature of gastrulation is that the cells that gastrulate in early stages differentiate into the anterior structures, and those in later stages differentiate into more posterior tissues. On the anterior portion of the primitive streak, the Hensen's node is formed and once the primitive streak reaches its full extension, it starts to regress posteriorly along the embryo midline. As it regresses, it leaves behind the dorsal endoderm and axial mesoderm. As it continues to regress, the cells that are anterior to the Hensen's node start their differentiation program. At the same time, the cells continue to ingress in the Hensen's node and primitive streak. Consequently, this results in a "gradient" of cell maturity, where anterior cells are progressively more differentiated than the posterior cells (Sawada and Aoyama, 1999).

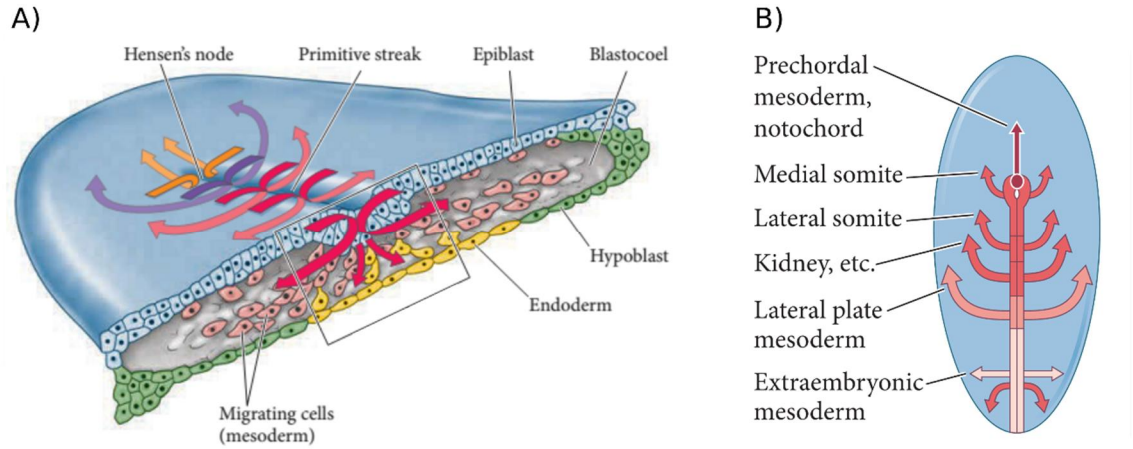



**Figure 1.2 - Schematics of chicken gastrulation and mesoderm fate map.**
**A)** Artwork showing the avian gastrulation cell movements to generate the mesoderm and endoderm layers. **B)** Fate map of the mesodermal cells after gastrulation at stage HH4. Adapted from (Gilbert and Barresi, 2019)

Non-amniote organisms have their gastrulation slightly different. For instance, the zebrafish gastrulation stage is characterized by four cell movements: epiboly, convergence, internalization and extension. The first of the movements is the epiboly. During epiboly, the cells from the blastoderm start to engulf the yolk cell. Meantime, convergence occurs, the blastoderm cells migrate to one end of the embryo close to the YSL to form the shield (at 50% epiboly). It is at this moment that internalization and extension start (Figure 1.3). The cells from the epiblast start to internalize and form the hypoblast. This new layer of cells will originate the mesoderm and the endoderm (Solnica-Krezel, 2005). Gastrulation terminates as the epiboly also finishes. At this time, the animal cells completely engulfed the yolk cell forming the tail-bud.

During the initial stages of development, the embryo relies solely on maternally deposited RNA and proteins. However, since fertilization, the newly formed zygote starts to gradually degrade the maternal material and slowly activates the zygote transcriptional program. During gastrulation, the vertebrate embryo completely shifted from the maternal influence to solely relying on its own transcriptional program. This program is known Maternal to Zygote Transition (MZT) (Giraldez *et al*., 2006; Hwang *et al*., 2018; Langley *et al*., 2014; Raffaelli and Stern, 2020; Tadros and Lipshitz, 2009). Nowadays, it is known that the embryo relies on mechanisms to specifically clear the maternal RNA. For instance, zebrafish embryos rely on miR-430. If this mechanism fails, the embryo accumulates the maternal RNA, which impairs its development (Giraldez *et al.*, 2006; Hwang *et al.*, 2018; Langley *et al.*, 2014; Raffaelli and Stern, 2020; Tadros and Lipshitz, 2009).

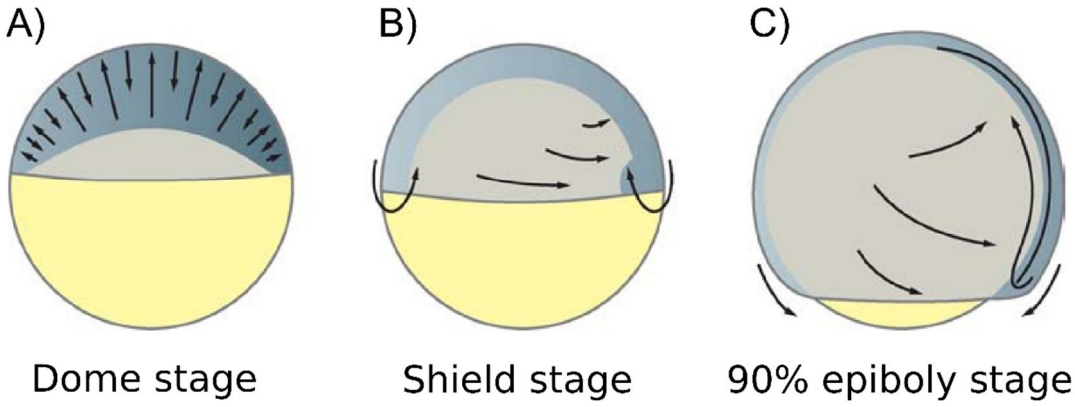



**Figure 1.3 - Zebrafish cell movements prior to and during gastrulation.**
**A)** Dome stage embryo. Radial cell movement that contributes to the epiboly **B)** The gastrulation starts at the shield stage as the internalization movements start. Epiboly, compaction, and extension continue through this process. **C)** As epiboly progresses, the movements continue until the end of gastrulation. Adapted from (Schier and Talbot, 2005)

### 1.1.1 Presomitic mesoderm (PSM) and somite differentiation

The unsegmented paraxial mesoderm, also known as presomitic mesoderm (PSM), differentiates into round epithelial cell structures called somites. Somites are the precursor structures that give rise to the vertebrae, ribs, occipital bone, skeletal muscle, and dermis from the back (Reviewed in (Christ *et al*., 2007)). The first PSM tissue comes directly from the primitive streak. However, once it halts (at chicken HH12 stage), the embryo continues to elongate, and new cells are constantly being added to the PSM by the tailbud, a functional remnant of Hensen's node, maintaining its size relatively stable (Figure 1.4) (Cambray and Wilson, 2002; Catala *et al*., 1995). Once a cell enters the PSM, it takes around eighteen hours until they incorporate into a new somite (Palmeirim *et al*., 1997). As the anterior cells pass the determination front, they trigger an irreversible genetic cascade that ultimately leads to them being incorporated into a new somite (Figure 1.4 B).

Somitogenesis takes place at a species-specific pace. For instance, a new pair of somites is formed every 30 minutes in the zebrafish embryo, 90 minutes in chicken, 120 minutes in mice, and around 4 to 5 hours in humans. Also, the number of segments varies among the vertebrates. Zebrafish form around 31 pairs of somites, chicken 55 pairs, and humans between 42 to 44 pairs. Also, certain snake species can have more than 300 pairs (Gomez *et al*., 2008).

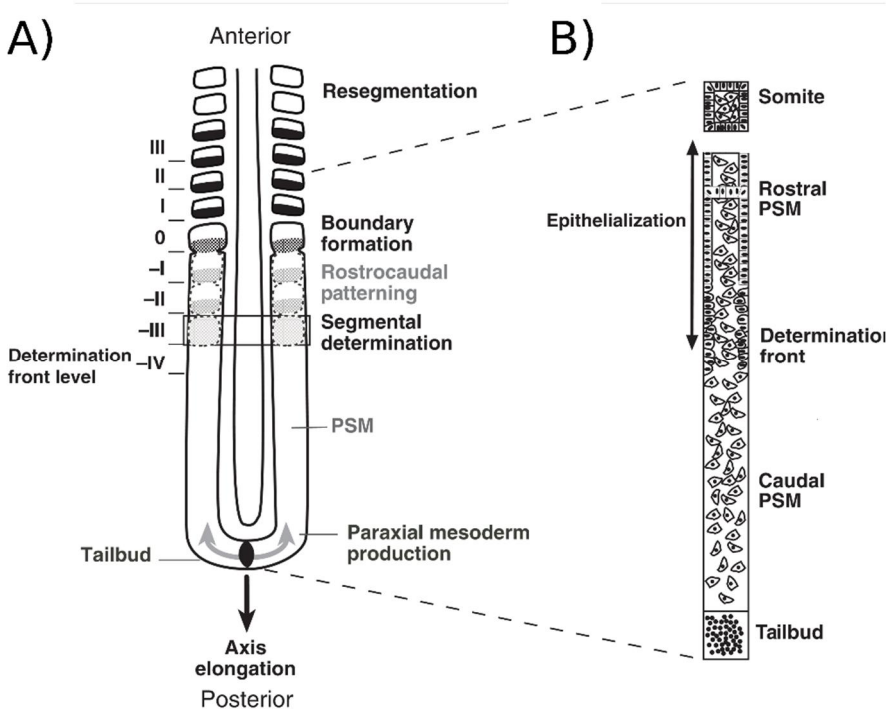



**Figure 1.4 - Schematics of the somitogenesis process.**
**A)** Representation of a caudal portion of a two-day-old chicken embryo. Cells are dividing in the tailbud and ingressing the PSM. At the presumptive somite -III, the cells get committed to forming a somite as it passes the determination front. After a couple of cycles of somite segmentation, the cells epithelialize and form a new somite. **B)** Magnification of one side of the PSM, showing the epithelialization process until a new somite is formed. Adapted from (Dubrulle and Pourquie, 2004)

As the embryo elongates, the most anterior somites start to differentiate. By undergoing EMT, the ventral somite cells originate distinct cell populations that ultimately differentiate into different adult tissues. At first, the somite is sub-divided into the sclerotome and the dermomyotome. The sclerotome cells, resultant from the EMT process, are the group of cells that generate the vertebrae and ribs. The cells adjacent to the notochord form the vertebral body (Figure 1.5 pink cells), the central sclerotome cells give rise to the proximal ribs (Figure 1.5 blue cells) and flanking the central sclerotome there two small cell populations, one adjacent to the neural tube that originates the spinal process (Figure 1.5 orange cells) and in the opposite the cells that originate distal ribs (Figure 1.5 green cells). The dermomyotome originates the dermis of the back (Figure 1.5 light pink), and the myotome (Figure 1.5 coral pink), from which all skeletal muscles of trunk and limbs will be formed. There is also a small cell group between the sclerotome and dermatomyotome that forms the tendons, the syndetome (Figure 1.5 white cells) (Reviewed in (Kalcheim, 2016; Nobrega *et al*., 2021)).

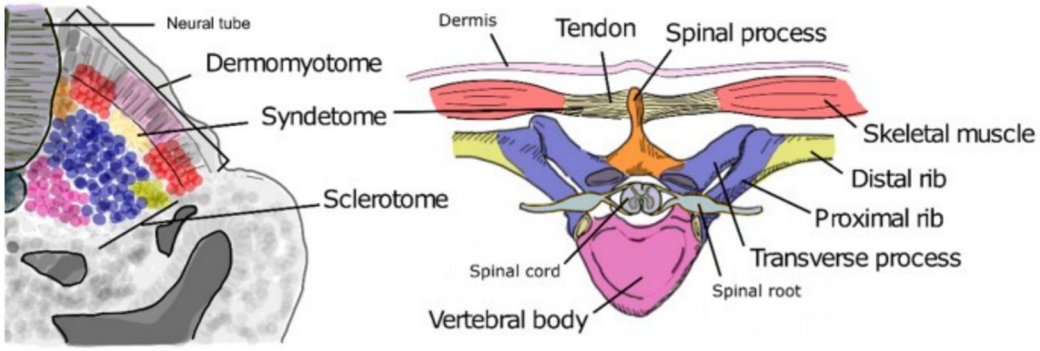



**Figure 1.5 - Representation of somite cell specification and differentiation.**
At the left, it shows different somite cells populations that will differentiate into the tissues represented on the right. The dermomyotome is divided into the dermatome and myotome. The dermatome originates the dermis from the back, and the myotome differentiates into the skeletal muscle. A small cell population, the syndetome differentiates into the tendons. Lastly, the sclerotome originates vertebrae and the ribs, and part of the occipital bone. Adapted from (Nobrega *et al.*, 2021)

During the somite differentiation process, the sclerotome undergoes a second segmentation, the resegmentation. In fact, the vertebral body is not composed of cells of a single sclerotome. Rather, it is composed of two halves of sclerotomes. In amniotes, such as chickens and mice, upon the sclerotome resegmentation, the posterior half joins together with the anterior half of the adjacent somite (Figure 1.6) (Aoyama and Asamoto, 2000; Stockdale *et al*., 2000). This process is crucial for the vertebral column to bend. This way, since the dermomyotome does not resegment, each skeletal muscles segment is inserted into two successive vertebrae, allowing the vertebral colon to bend laterally (Saga and Takeda, 2001; Scaal, 2016). This is different in fish, such as zebrafish. Due to a “leaky” resegmentation, one-half of the sclerotome cells can be incorporated in more than one somite (Figure 1.6) (Morin-Kensicki *et al*., 2002).

**The Hox code**

A critical aspect of the body axis formation is the identity of each segment. Depending on the axial position of a specific somite, it will give rise to different types of bone segments. The most anterior somites originate the occipital, and the subsequent somites originate the cervical, thoracic, lumbar, sacral, and caudal vertebrae (when applicable), by this order, along the A-P axis. Each of these bone segments has particular morphological traits (Mallo *et al*., 2010). Although the morphological differences are only spotted further in development, once the bone is being formed, the identity of each segment is determined during early development. By transplanting the PSM tissue that originates the thoracic vertebrae in a different embryo, in

another developmental time, the tissue maintains the donor identity and differentiates into rib-containing vertebrae (Kieny *et al*., 1972).

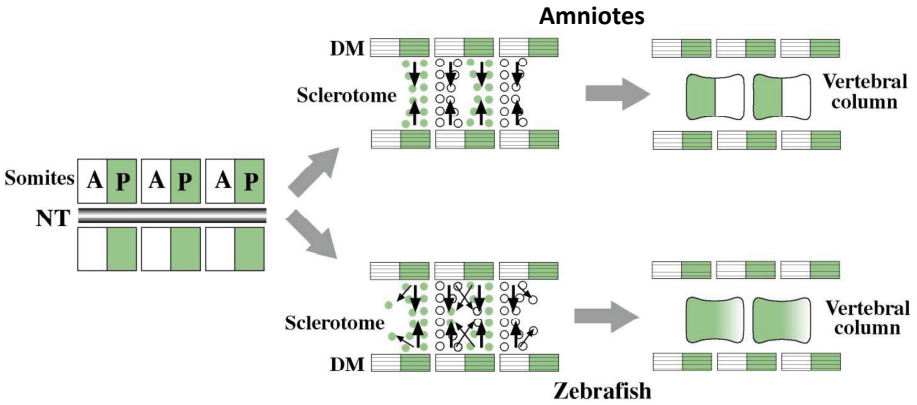



**Figure 1.6 - Sclerotome resegmentation.**
In the amniotes like chicken and mice (upper figure) the sclerotome resegments and the posterior part of the one somite fuses with the anterior part of the subsequent somite to form a vertebra. In zebrafish (lower figure), the vertebrae have contributions from both anterior and posterior parts of the sclerotome, due to a process called leaky resegmentation. Adapted from (Takahashi, 2005).

The *Hox* genes were first described in Drosophila. Nowadays, it is known that the colinear expression activation of the homeobox-containing Hox transcription factors is responsible for A-P axial vertebral identity. Moreover, it was found that *Drosophila Hox* genes were responsible for giving the identity to each one of the insect segments (Lewis, 1978; Nusslein-Volhard and Wieschaus, 1980). Humans and mammals, in general, have 39 *Hox* genes divided into thirteen paralogs divided into four clusters, *HoxA, HoxB, HoxC,* and *HoxD* (Figure 1.7). The temporal colinear activation of the *Hox* cluster is directly correlated with its chromosomic position, occurring from 3' to 5'. The first group of *Hox* genes (*HoxA1, HoxB1,* and *HoxD1*) starts to be expressed as early as the beginning of gastrulation, defining the occipital identity, and the last *Hox* paralog group (*Hox13*) is only activated much later, during somitogenesis stages. This phenomenon results in different *Hox* expression combinations along the A-P axis (the so-called *Hox* code), which determines the specific identity of each vertebra (Figure 1.7) (Mallo *et al.*, 2010).

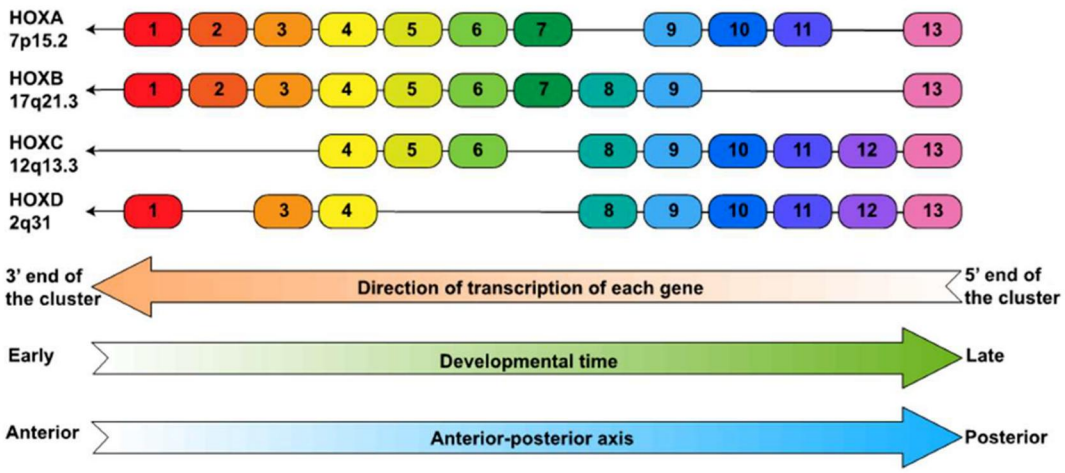



**Figure 1.7 - The Hox code.**
In amniotes, such as humans, the *Hox* genes are divided into four clusters in different chromosomes, *HoxA, HoxB, HoxC* and *HoxD*. The expression of the *Hox* genes is collinearly activated. The first group of paralogs, *HoxA1, HoxB1,* and *HoxD1*, is expressed as early as the beginning of gastrulation and defines more anterior structures, and the subsequent paralogs are activated later and more posteriorly. The different combinations of *Hox* gene expression result in specific vertebrae identity. Adapted from (Luo *et al.*, 2019).

## 1.2 The Embryo Molecular Clock

Concomitant with the segmentation of the PSM into the somites, there is the operation of the Embryo Clock (EC) in the PSM. The EC is characterized by a wave of gene expression that starts in the posterior PSM and propagates to the anterior PSM (Figure 1.8 left panel). The end of each cycle of expression results in the formation of a new pair of somites. As the segmentation process, the EC periodicity is species-specific. For instance, the EC pace has 30 minutes, 90 minutes, 120 minutes, and four to five hours in zebrafish, chicken, mice, and humans, respectively (Aulehla and Johnson, 1999; Forsberg *et al.*, 1998; Holley *et al.*, 2000; Jiang *et al.*, 2000; Matsuda *et al.*, 2020a; Palmeirim *et al.*, 1997; William *et al.*, 2007).

The wave effect of expression in the PSM is due to an intrinsic genetic oscillator that is coupled with the neighbor cells (Figure 1.8 middle panel) by Notch signaling (Figure 1.8 right panel). One cell undergoes several cycles of gene expression until it is incorporated into a somite. For instance, in chicken, when a cell enters the PSM from the tailbud, it undergoes twelve gene expression cycles until it integrates a somite (Palmeirim *et al.*, 1997).

In 1976, Cook and Zeeman proposed a model that underlying the somitogenesis there should exist an intrinsic molecular oscillator responsible for generating a rhythmic signal through the PSM and a wavefront responsible for arresting the oscillation and committing the

cells towards a somite fate (Cooke and Zeeman, 1976). However, it was only in 1997 that the first gene that fits this model was discovered. Palmeirim and colleagues found that chicken *hairy1* has a cyclic expression behavior in the PSM with the exact pace of somite formation, 90 minutes (Palmeirim *et al.*, 1997). This discovery was considered a XX century milestone in Developmental Biology by the Nature publishing group (Skipper, 2004). In each cycle, *hairy1* expression starts in the posterior PSM and propagates anteriorly through the PSM. Once the signal reaches the anterior PSM, the gene expression stabilizes in a single stripe, and a new pair of somites are formed (Figure 1.9) (Palmeirim *et al.*, 1997).

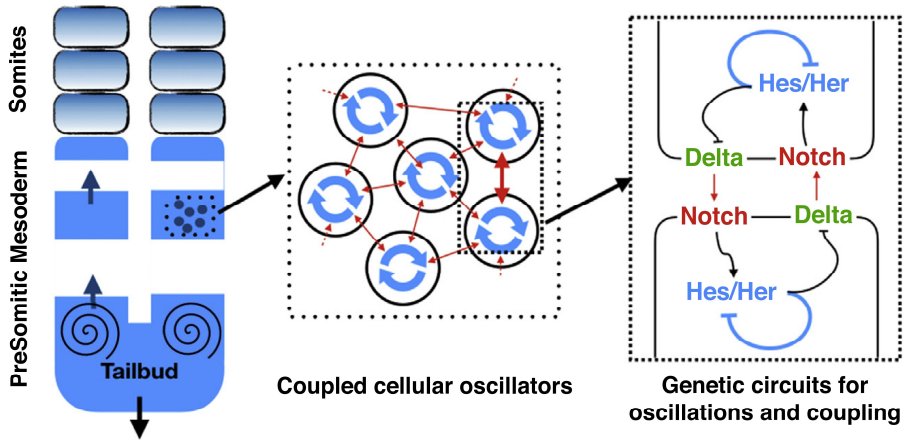



**Figure 1.8 - Representation of the vertebrate Segmentation Clock.**
Left panel: Representation of one gene expression oscillation in the PSM. In blue, the expression of a cyclic gene is shown, moving anteriorly (blue arrows), while the embryo is still elongating. (Black arrow) Middle panel: Cell-autonomous gene expression oscillations (Blue circular arrows), are mutually coupled to their neighbors (Red arrows). Right panel: Representation of the conserved intracellular feedback loop and intercellular coupling through Delta-Notch signaling. Adapted from (Oates, 2020)

Since then, several genes have been detected to oscillate in the chicken PSM, such as *hairy2* and *lunatic fringe* (*lfng*) (Jouve *et al.*, 2000; McGrew *et al.*, 1998). Furthermore, it was discovered cyclic expressing genes in several other vertebrates, such as *hes1* (Hirata *et al.*, 2002), *hes7* (Bessho *et al.*, 2001), and *lfng* (Aulehla and Johnson, 1999; Forsberg *et al.*, 1998) in mouse; *her1* (Elmasri *et al.*, 2004) and *her7* (Gajewski *et al.*, 2006) in Medaka and *her1* (Holley *et al.*, 2000), *her7* (Oates and Ho, 2002) and *deltaC* (Jiang *et al.*, 2000) in zebrafish; *hes5.5* and *hes9.1* in *Xenopus* (Li *et al.*, 2003); among many others in these and other species especially genes belonging to Notch, FGF and WNT signaling pathways (Dale *et al.*, 2006; Dequéant *et al.*, 2006; Krol *et al.*, 2011). Nowadays, it is believed that this is a conserved mechanism among all vertebrates (Bailey and Dale, 2015), including in humans. Thanks to recent technology development, it was found that *hes1* and *hes7* have a cyclic expression in

cells cultures and organoids, with a period of 4 to 5 hours (Diaz-Cuadros *et al.*, 2020; Matsuda *et al.*, 2020a; Matsuda *et al.*, 2020b; William *et al.*, 2007).

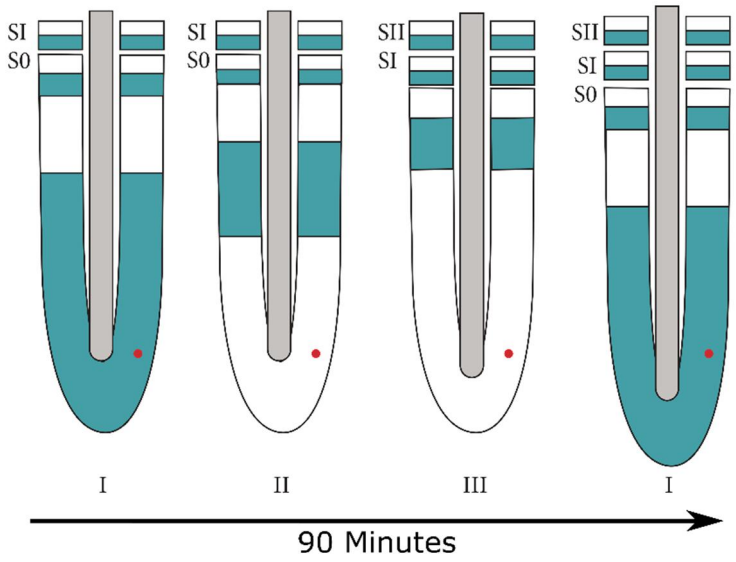



**Figure 1.9 - Cyclic expression of chicken *hairy1* in the Segmentation Clock.**
Representation of the three phases of the EC. Gene expression propagates from the posterior to the anterior PSM. A new pair of somites is formed at the end of each cycle (end of phase III). Cells in the PSM pass through periods of active and inactive expression (red dot). Adapted from (Resende *et al.*, 2014).

Although EC oscillating genes are a common feature of the segmentation of all vertebrates, the specific set of genes with cyclic expression is unique in each species (Krol *et al.*, 2011). However, there is a family of genes that have an oscillatory behavior in all vertebrates, the Hes (Hairy and enhancer of split)/Her (Hairy and enhancer of split related) gene family. Hes/Her genes encode transcription factors, more precisely, transcriptional repressors. The Hes proteins have four characteristic conserved domains, a basic domain, an *helix-loop-helix* domain, an orange domain, and a final WRPW domain. The basic domain allows binding to N-boxes in the target gene promoters. The *helix-loop-helix* is used for Hes proteins to homo or heterodimerize. This motif is highly important since Hes proteins are known to operate as dimers. The Orange domain is also relevant for the Hes dimerization, and WRPW is crucial for its repression activity (Kageyama *et al.*, 2007; Kang *et al.*, 2005; Taelman *et al.*, 2004). A particular aspect of this gene family is that they are able to repress their own transcription (Chen *et al.*, 2005; Hirata *et al.*, 2002). This particular feature creates a feedback loop, which is the core mechanism of the embryo clock (Figure 1.10) (Chen *et al.*, 2005; Hirata *et al.*, 2002). In teleost fish such as the zebrafish, *her* genes, almost alone, integrate the feedback loop mechanism. Notch or FGF signaling activate *her1* and *her7* transcription, upon

the translation, the Her1 homodimerizes and Her7 Heterodimerizes with Hes6, and the dimers are translocated to the nucleus, where they bind with their own promoter, which represses their transcription (Further detailed in section 1.4)

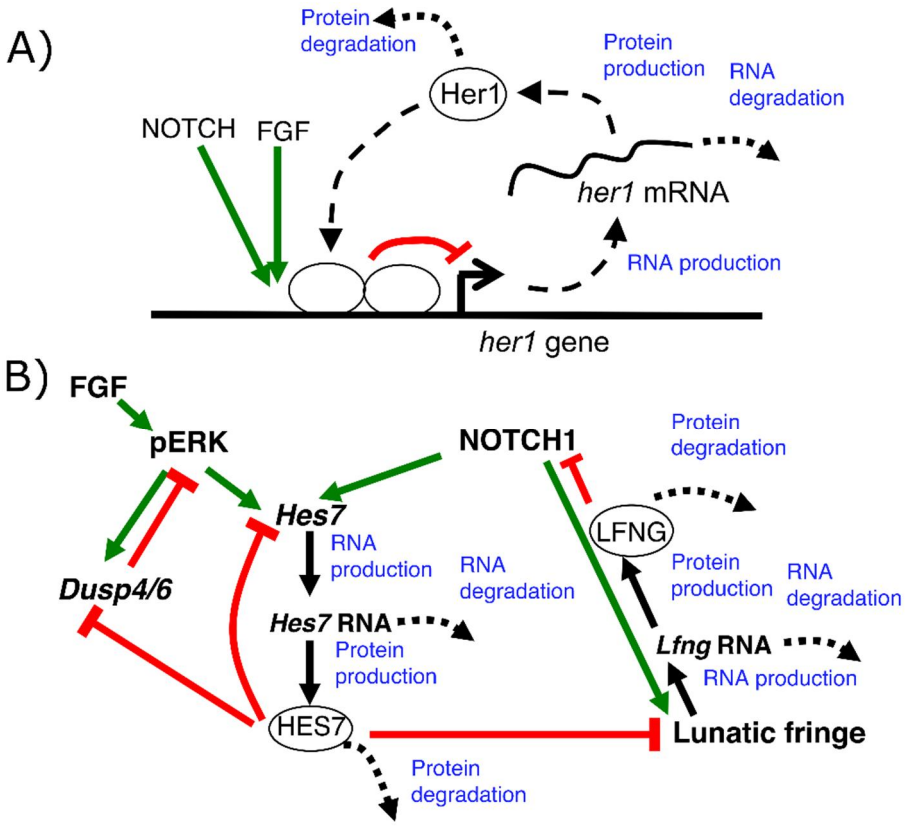



**Figure 1.10 - Zebrafish and Mouse EC negative feedback loop regulation.**
**A)** Zebrafish feedback loop network. *Her1* or *her7* are activated by FGF or Notch signaling. After transcription and subsequent translation, Her1 and or Her7 proteins homodimerize or heterodimerize and inhibit their own transcription, restarting the cycle. **B)** Mouse negative feedback loop. Hes7 protein is activated by either FGF or Notch signaling. Once the HES7 protein is activated, it represses its own transcription. Besides inhibiting its own transcription, HES7 suppresses Dusp4 and Duspt6, which inhibit ERK phosphorylation and decrease Hes7 transcription. Furthermore, HES7 represses Lunatic fringe, which inhibits Notch activity, resulting in a decreased Hes7 transcription. Delays in transcription, translation, and mRNA and protein decays account for the EC regulation. Adapted from (Wahi *et al.*, 2016).

EC oscillations are driven by cell-autonomous negative feedback loops and are synchronized with neighboring cells through the Notch/Delta signaling pathway (Figure 1.11). Binding of ligand transmembrane proteins, such as Delta-Like1 (Dll1), to Notch receptors in adjacent cells, triggers the cleavage and release of the Notch intercellular domain (NICD) that translocates to the nucleus and activates target genes, such as Hes and Lfng (Wahi *et al.*, 2016) (Further detailed in section 1.4).

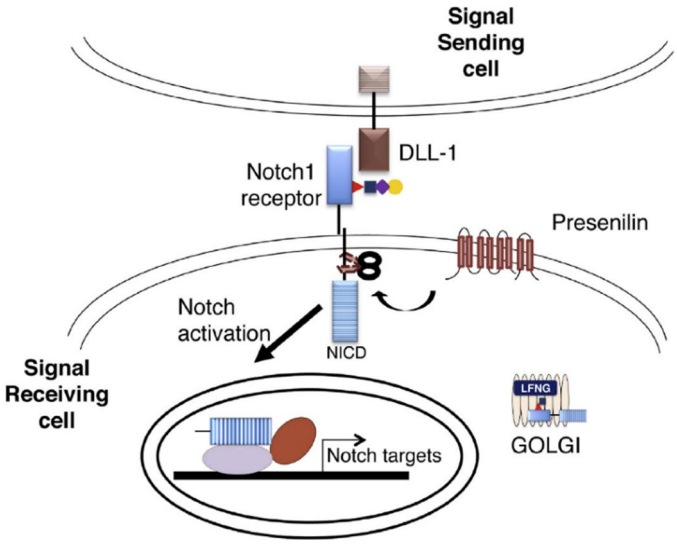



**Figure 1.11 - Notch signaling pathway and intercellular synchronization of the EC.**
Notch proteins are cellular membrane receptors. When a Notch receptor recognizes a Notch ligand such as Delta-like ligand 1 (Dll1) from a neighboring cell, it triggers a series of protein cleavages. The last cleavage performed by Presenilin1 or presenilin2 (catalytic subunit of the γ-secretase intramembrane protease complex) releases the Notch Intracellular Domain (NICD). NICD is then translocated to the nucleus and activates gene expression of, for example, Hes genes, initiating the feedback loop. LFNG modifies the Notch receptor by adding sugar moieties to the EGF repeats. Adapted from (Wahi *et al.*, 2016).

Synchronized EC oscillations take place along the A-P PSM axis, but are decelerated towards the rostral end, upon entering the determined PSM tissue (Figure 1.12) (Dubrulle *et al.*, 2001). Upon the formation of a somite, EC gene expression stabilized.

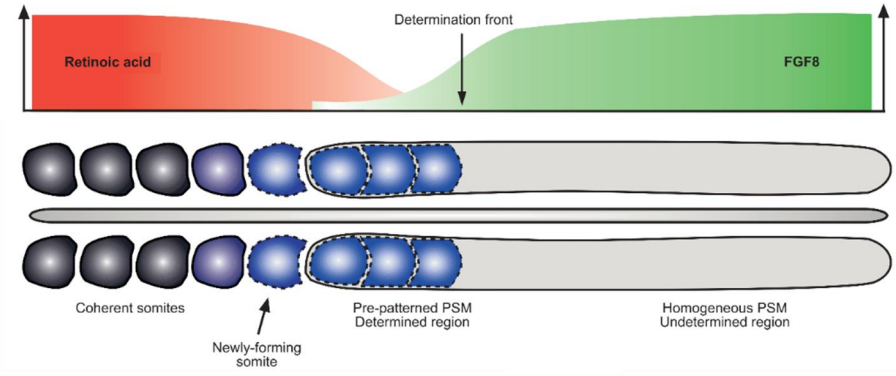



**Figure 1.12 - Schematics of the determination front.**
The determination front is specified by counter gradients of FGF8/WNT (green) and Retinoic Acid (red). When the PSM cells reach the FGF8/WNT threshold (Determination front), they get committed towards a somite fate. The determination front marks the separation between the Undetermined PSM (non-committed cells) and Determined PSM (committed cells). Adapted from (Baker *et al.*, 2009).

## 1.3 Limb bud outgrowth in time and space

Besides somitogenesis, several other processes require a correct positioning of cells in time and space during embryo development. Limb outgrowth is one of these processes. The limb is divided into three distinct bone segments: stylopod, zeugopod, and autopod. To explain how limb bone segments are specified, several models were proposed (Figure 1.13). The first model, proposed in 1973, was the Progress Zone model (Figure 1.13 A) (Summerbell *et al.*, 1973). In this model, Summerbell and colleagues proposed that cells in the progress zone (PZ) (cells 300µm beneath the Apical Ectodermal Ridge – AER) were maintained in an undifferentiated state under the influence of the AER. This model proposed an intrinsic timer that would measure the time cells were under the influence of the AER before exiting and initiating their differentiation program. The cells that escaped first would give rise to the more proximal segments, and the ones that escaped later would give rise to the more distal ones (Summerbell *et al.*, 1973).

The progress zone model prevailed for approximately 30 years until the early specification model was proposed (Figure 1.13 B). In this model, the authors proposed that all three bone segments are already specified in the early limb bud. The more proximal region gives rise to the stylopod, the distal region the autopod, and the intermediate gives rise to the zeugopod (Dudley *et al.*, 2002). Later, each early cell population would only expand in number.

In 2007 a third model was proposed, and is currently the most accepted model, the two signal model (Figure 1.13 C) (Tabin and Wolpert, 2007). This model is based on two counter gradients, the Retinoic Acid (RA) gradient (proximal-distal gradient) and an FGF gradient that comes from the AER (distal-proximal gradient). In early limb development, all cells are under the influence of both signals. However, once the limb bud grows, three different populations arise, one more proximal under the influence of RA alone (expressing *meis1* and *meis2*), a more distal population under the influence of FGF alone (expressing *hoxa13*), and an intermediate one that is under the joint influence of lower levels of each signal (expressing *hoxa11*). The three regions give rise to the stylopod, autopod, and zeugopod, respectively (Tabin and Wolpert, 2007).

### 1.3.1 The Limb Molecular Clock

Driven by the question of how cells count time during limb bud formation, Pascoal *et al*. analyzed the expression of the embryo clock gene *hairy2* during limb bud outgrowth, from the HH14 until the end of the digit formation, HH32 (Pascoal *et al.*, 2007). *Hairy2* was found to be expressed as early as stage HH14 in the presumptive forelimb region. In early limb bud development, at the stages HH16 and HH17, the expression of *hairy2* was uniform throughout

the tissue. From stage HH18 onwards, the expression was differentially localized in several domains: *hairy2* was constantly expressed in both the AER and in the distal posterior limb bud area, that includes the zone of polarizing activity (ZPA), while it presented different expression patterns in the distal limb mesenchyme in limb buds within the same developmental stage. Also, they found that *hairy2* was also present in the proximal dorsal and ventral mesenchymal domains. At stage HH32, *hairy2* was found in the joints of the forming digits.

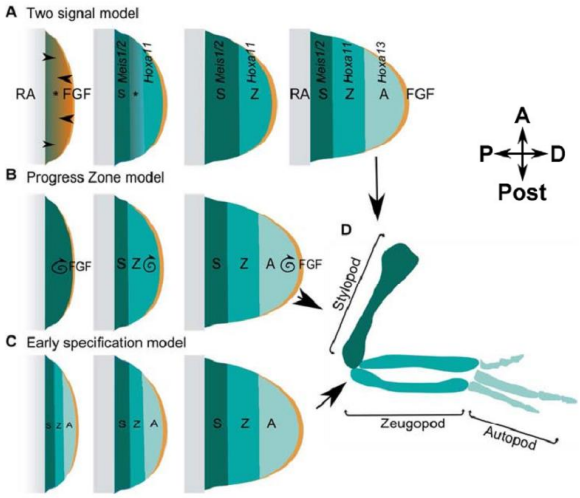



**Figure 1.13 - Models for limb bud proximal-distal patterning.**
**A)** Two signal model; **B)** Progress Zone model; **C)** Early specification model; **D)** Representation of the chicken forelimb bone segments, the stylopod, the zeugopod, and the autopod. Adapted from (Sheeba, 2011)

The authors hypothesized that the expression of *hairy2* could be dynamic over time in the distal limb mesenchyme. To evaluate that, they thoroughly analyzed the expression pattern of *hairy2* in limb buds between the stages HH20 to HH28. The authors found a region of variable expression in the distal limb. They noticed that between the stages HH20 to HH22, the variable expression was contained between the limb central mesenchyme and the ZPA; after that, it expands anteriorly, becoming localized between the limb central mesenchyme and the AER (between stage HH23 and HH26). Between the stages HH27 and 28, the variable expression domain was detected only in the posterior part of the distal limb mesenchyme. Curiously, the domain of variable *hairy2* expression coincides with where the chondrogenic precursor cells (Summerbell, 1974) are located (Pascoal *et al.*, 2007).

At this point, Pascoal and colleagues decided to determine if the expression of *hairy2* in the distal limb bud was cyclic. For that, the authors, microsurgically, removed one of the limb buds from the embryo and fixed it, then they let the embryo grow for more 1.5, 3, 6, 9 or 12 hours. After that time, they collected the second limb bud and fixed it. After analyzing the

expression of *hairy2* in both limbs, the one fixed at time 0 and the incubated one, the authors found that ~ 90% of the limb pairs had the same expression pattern after 6 or 12 hours of incubation. In the other time points, the limbs pairs with the same expression were never more than 30%. The authors concluded that *hairy2* had a cyclic expression in the chondrogenic precursor cells, with a 6-hour periodicity (Pascoal *et al.*, 2007). Based on the frequency of the different patterns, they established three phases of the "limb molecular clock": a negative phase, which takes around 1 hour and 20 minutes (Phase I), an intermediate phase, that takes around 1 hour and 10 minutes (Phase II) and a positive phase, that takes 3 hours and thirty minutes (Phase III) (Figure 1.14).

At last, Pascoal et al. found that it takes twelve hours to form one forelimb bone segment, that is, the precise time of two cycles of *hairy2* expression. As in somitogenesis, where one vertebra is formed from cells belonging to two contiguous somites (each somite takes 90 minutes to be formed, which also corresponds to a cycle of expression of the EC), a limb bone segment also could be the result of two cycles of limb molecular clock expression (Pascoal *et al.*, 2007).

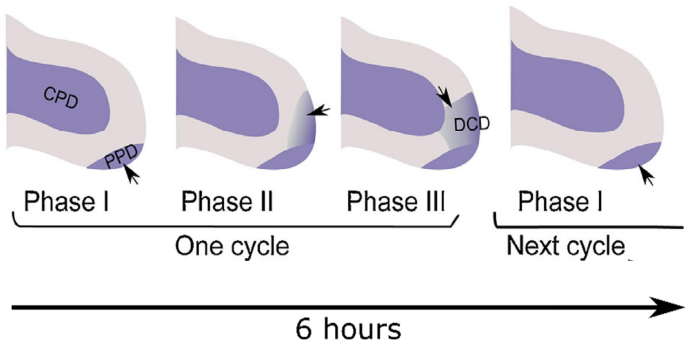



**Figure 1.14 - Schematic representation of the distinct gene expression phases of the Limb Molecular Clock.**
Between stages HH23 and HH26, chicken *hairy2* presents a non-variable positive expression in the central mesenchymal domain (CPD) and in the posterior distal limb, including the ZPA (PPD). Moreover, it presents a cyclic expression in the distal limb mesenchyme (DCD), with a period of 6 hours, divided into three phases: Phase I, a negative phase, with no expression of *hairy2* detected; Phase II, intermediate phase, with an expression close to the AER, but not in all the domain; Phase III, positive phase, with *hairy2* being expressed in all the domain. Adapted from (Sheeba *et al.*, 2016b)

By considering the existence of a molecular clock, Sheeba and collaborators proposed a new model for the proximal-distal (P-D) limb patterning, the Integrated Space-Time model (Figure 1.15). According to this model, the stylopod is specified in the early limb outgrowth, when *hairy2* presents a non-oscillatory behavior due to the simultaneous influence of the proximal RA signal and distal FGF signal. Once the mesenchymal cells start to get distanced

from the proximal RA signal and getting influenced only by FGF (AER), *hairy2* starts to oscillate. A posterior-anterior gradient of Gli-activator to Gli-repressor (Gli-A/Gli-R) is established by Sonic Hedgehog (Shh) signaling from the ZPA, which allows the cycles of activation and repression of *hairy2.* The expression oscillations of *hairy2* constitute a time-counting mechanism that gives the positional information for the cells to form the zeugopod and the autopod (Sheeba *et al.*, 2014).

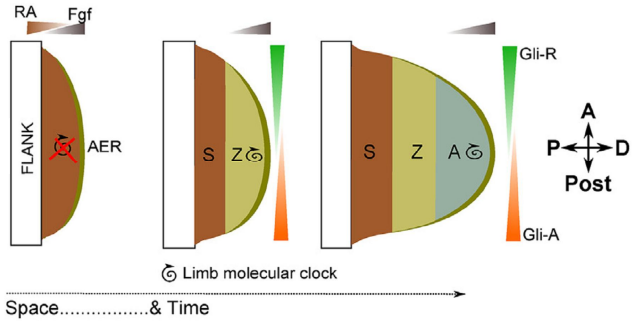



**Figure 1.15 - Integrated Space-Time model.**
This model integrates the Progress Zone model and the Two Signal model. It considers the counter gradients of RA and FGF that give spatial information for the limb mesenchymal cells, and the existence of a timer gives the right time for differentiation. In the Integrated Space-Time model, the timer is proposed to be the *hairy2* cyclic expression. Adapted from (Sheeba *et al.*, 2016a)

### 1.3.2 Molecular parallelisms between PSM segmentation and limb patterning

The Limb bud and PSM are morphologically very different. However, the signaling networks that rule their patterning is strikingly similar (Figure 1.16). In the PSM, several genes that participate in the EC have been identified in multiple organisms (Dequéant *et al.*, 2006; Krol *et al.*, 2011). These include chicken *hairy2*, that is also cyclically expressed in the distal limb mesenchyme (Pascoal *et al.*, 2007). *Hairy1* expression was also reported in the developing limb (Vasiliauskas *et al.*, 2003), but there is no information regarding its dynamics. It is well established that Notch signaling regulates *HES* cyclic expression, as well cyclic production of the Notch intracellular domain (NICD) in the PSM (Bessho *et al.*, 2003). In the limb bud, *notch2* and *serrate1* are expressed and may also regulate the expression of *hairy2* (Sheeba, 2011). The opposing gradients of FGF and RA that regulate the somitogenesis are also important for limb patterning. In the PSM, *fgf8* is transcribed in the tailbud, where it is responsible for initiating the expression of *Hes7* (Niwa *et al.*, 2011). FGF signaling forms a distal-proximal gradient in the forelimb, being transcribed in the AER, and activates the

expression of *hairy2* (Sheeba *et al.*, 2012). Moreover, the *wnt3a* gradient is also present in both tissues. In the PSM, Wnt3a activates *HES* genes (Gibb *et al.*, 2009) and engages in a positive feedback loop with *fgf8* (Naiche *et al.*, 2011). However, in the limb, these interactions are still to be proved. Another important regulator that is present in both tissues is *shh*. In the PSM, *shh* is produced in the notochord and is a positive regulator of *fgf8* and *hairy2* in the adjacent PSM (Resende *et al.*, 2010). However, high levels of FGF downregulate *shh* (Ribes *et al.*, 2009) (Diez del Corral *et al.*, 2003). During limb outgrowth, *shh* is produced in the ZPA, and it was shown that *FGF* and *shh* are in a positive feedback loop (Fallon *et al.*, 1994). ZPA-derived Shh also exerts a positive regulation on *hairy2* (Sheeba *et al.*, 2012). Retinoic Acid (RA) gradient is also present in both tissues, and in both cases, it counters FGF and WNT3a gradients, defining the differentiation front.

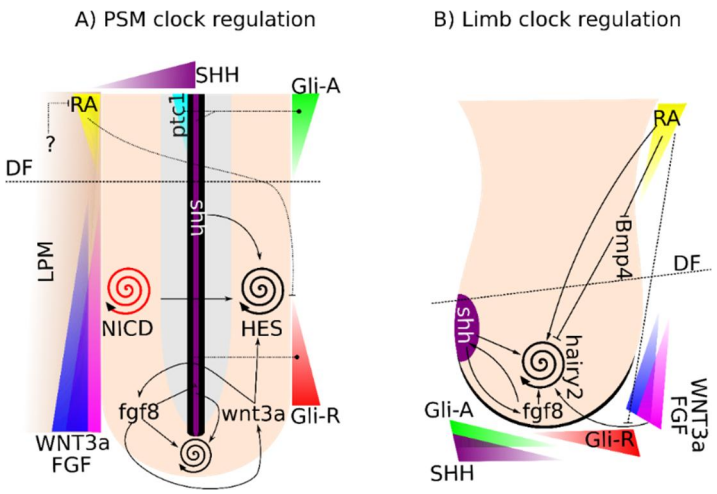



**Figure 1.16 - Parallelism between the PSM molecular clock and the Limb molecular clock.**
**A)** Schematic of the major EC regulators of the embryo clock operating in the PSM. **B)** Schematic of the major regulators of the hairy2 oscillations in the distal forelimb mesenchyme. Spirals indicate gene expression oscillations. Triangles indicate graded expression. DF: differentiation front. LPM: lateral plate mesoderm. NICD: Notch Intercellular Domain.

## 1.4 "The Vertebrate Embryo Clock: Common Players Dancing to a Different Beat" *(Review manuscript in preparation)*

Gil Carraco[1,2,3,¥], Ana P. Martins-Jesus[1,2,¥] and Raquel P. Andrade[1,2,4*]

1. ABC-RI, Algarve Biomedical Center Research Institute, Faro, Portugal
2. Faculdade de Medicina e Ciências Biomédicas (FMCB), Universidade do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal
3. ProRegeM-PhD Program in Mechanisms of Disease and Regenerative Medicine, Faro, Portugal
4. Champalimaud Research Program, Champalimaud Center for the Unknown, Lisbon, Portugal.
¥ Equal contribution

*Corresponding author: rgandrade@ualg.pt

### Highlights

- The vertebrate EC oscillates with species-specific periodicities
- The EC oscillates with distinct paces in different tissues and developmental stages within the same organism
- Experimental changes to the pace of the EC provide information on its regulation

### Abstract

Great attention has been dedicated to studying the molecular mechanisms involved in correct spatial positioning of cells/tissues/organs during embryo development, while the nature and dynamics of gene expression over time has been an under-represented concern. The discovery of a molecular embryonic clock (EC) underlying the temporal dynamics of somite formation gave way to a dramatic shift in this trend. The EC is now known to be present in all vertebrate organisms studied and this mechanism was also described in limb development and neuronal cell differentiation. An outstanding question, however, remains unanswered: what sets the different EC *paces* observed among organisms and tissues? This review aims to summarize the available knowledge regarding the pace of the clock and its experimental manipulation and to expose new questions that might help shed light on what is still to unveil.

## Introduction

Vertebrate embryo development comprises several processes that are tightly regulated in time, for example, somitogenesis. This process is characterized by the formation of metameric structures, the somites, in a periodic fashion along the anterior-posterior axis of the presomitic mesoderm (PSM).

Nearly half a century ago, Cooke and Zeeman (1976) proposed a theoretical model that aimed to explain the formation of periodic structures during vertebrate development. In their Clock and Wavefront model, the authors were the first to propose the existence of two players: a molecular oscillator (clock), responsible for the generation of rhythmic signal at the PSM-level, and the wavefront, a gradient along the PSM that arrests the oscillator and, as a consequence, creates a periodic pattern. Together, these two players translate temporal information into a spatial pattern. Moreover, according to this model, somite size and number were determined by the period of the clock's oscillations (Cooke and Zeeman, 1976; Oates *et al.*, 2012). However, breakthroughs regarding the identity of the Clock and the Wavefront were only made 20 years later.

The term *somitogenesis clock* arose from the discovery that the mRNA of chick *hairy1* (now termed *hes4*), a factor of the Hairy Enhancer of Split (HES) transcription repressor family, oscillated in the chicken embryo PSM with a 90 min periodicity, concomitant with the formation of a new pair of somites (Palmeirim *et al.*, 1997). In their study, the authors first observed that same stage chicken embryos displayed different patterns of *hairy1* expression, leading them to hypothesize that its expression could be cyclic. Indeed, by bisecting an embryo, and culturing one half for a given time while the other was immediately fixed, *hairy1* expression was shown to recapitulate after 90 min. Moreover, *hairy1* oscillations in the PSM were found to be an intrinsic property of the system. Since then, other genes that display an oscillatory behaviour during somitogenesis have been identified in multiple organisms, such as *her1* and *her7* in zebrafish and medaka (Elmasri *et al.*, 2004; Gajewski *et al.*, 2006; Holley *et al.*, 2000; Oates and Ho, 2002), *Hes1* and *Hes7* in mouse (Bessho *et al.*, 2001; Hirata *et al.*, 2002), and *hes9.1* in *Xenopus* (Li *et al.*, 2003), suggesting that this is a well conserved mechanism among vertebrates (Krol *et al.*, 2011). The somitogenesis clock is characterized for being an intricate oscillatory genetic network, that comprises genes belonging to Notch, Wnt and FGF pathways (Dequéant *et al.*, 2006). Studies done about the subject have shown that the onset of mouse *Hes7* expression in the tailbud is FGF-dependent, although its maintenance and propagation throughout the PSM requires Notch signalling (Niwa *et al.*, 2007, 2011). Moreover, a nuclear β-catenin gradient was shown to control key features of PSM maturation and somite formation (Aulehla *et al.*, 2008) and, for proper segmentation to occur, it was reported that Notch and

Wnt signalling oscillations are required to be in-phase (Sonnen *et al.*, 2018), further evidencing the importance of the interplay between these signalling pathways during somitogenesis.

Several studies have linked the existence of negative feedback regulation to the core mechanism of the clock's oscillations (Hirata *et al.*, 2002; Bessho *et al.*, 2003; Lewis, 2003; Chen *et al.,* 2005). Yet, regulation by negative feedback is influenced by other processes and, altogether, the feedback loop delay influences the periodicity associated to the expression of a given gene.

Using mathematical modelling, Lewis (2003), Monk (2003) and Jensen *et al.* (2003) showed, in independent studies, that oscillations in gene expression are influenced by transcriptional delays in the segmentation clock negative feedback loop. Giudicelli *et al.* (2007) experimentally observed a delay from nuclear mRNA production to mature mRNA localization in the cytoplasm and estimated the transcriptional delays of *her1*, *her7* and *deltaC* – zebrafish's core clock components. There are several factors that influence the overall duration of transcription, namely transcriptional elongation, mRNA splicing, and nuclear export (Hanisch *et al.*, 2013).

Lewis (2003) proposed that the time it takes to mature a transcript was one of the major accountants for transcriptional delay, so one would assume lengthier genes would have larger transcriptional delays. Elongation was, however, not found to have a major contribution to these delays – RNA polymerase II elongation rate measured in intact zebrafish embryos showed that the time needed to transcribe *her1* and *her7* is negligible and elongation kinetics of *Hes7* and *Lfng* determined using mouse cells also occurred at a fast rate (Hanisch *et al.*, 2013; Hoyle and Ish-Horowicz, 2013). Takashima *et al.*, (2011) addressed the contribution of **mRNA splicing** to gene expression oscillations. They used transgenic mice carrying all or none of *Hes7* introns, together with a luciferase reporter, and assessed the time of *Hes7* transcription and protein production in both situations. Mice carrying all of *Hes7* introns showed a delay of approximately 19 minutes in *Hes7* expression, when compared to the intron-null mice. When employing this delay to a mathematical model, *Hes7* oscillations were abolished. Hoyle and Ish-Horowicz (2013) corroborated that splicing, but also nuclear export phenomena in its majority, account for most of the transcriptional delays. Additionally, they compared the splicing and export delays – both *ex* and *in vivo* – for mouse, chicken and zebrafish, and concluded that organisms that have longer delays in these processes also present longer clock periods.

Another aspect to take into consideration is the **half-life times** of messenger RNA (mRNA) and proteins. 3'-UTR and microRNAs – have been found to influence the

presence/absence of mRNA and, consequently, the maintenance of the negative feedback loop responsible for gene expression oscillations.

While studying the mechanisms that control segmental gene expression in *Xenopus*, *hes4* (formerly known as *hairy2a*) mRNA stability was found to be influenced by its 3'-UTR sequence. Moreover, when the 3'-UTR of *hes4* was substituted by the 3'-UTR of other *hes* related genes (either from *Xenopus* or other vertebrate species), *hes4* expression pattern remained similar, suggesting that this is a conserved motif amongst species (Davis *et al.*, 2001). However, a more detailed study regarding this matter was performed by Hilgers and colleagues. In their work, two plasmids were generated – one containing a Flag-*Fgf8* sequence followed by the endogenous *Fgf8* 3'-UTR sequence, a gene expressed in a gradient fashion in the PSM, and another with the 3'-UTR of *lunatic fringe* (*Lfng*), a known cyclic gene –, and electroporated into the neural tube of chicken embryos, using a tetracyclin-dependent inducible Tet-Off system. After doxycycline exposure to different time periods to stop transcription, embryos electroporated with the plasmid containing the *Lfng* 3'UTR did not have *Fgf8* positive cells, suggesting that the transcripts where rapidly degraded, unlike that was observed for the other plasmid containing the *Fgf8* 3'-UTR. To access the influence of each 3'-UTR in mRNA decay, a GFP reporter was used instead of the Flag sequence. It was shown that a decrease in GFP intensity was directly correlated to mRNA decay, and that the GFP intensity decrease occurred faster in the L*fng* 3'-UTR condition (Hilgers *et al.*, 2005).

Evidence that oscillatory gene expression was 3'-UTR-dependent was further given by the work of Nitanda *et al.* (2014). By bisecting the mouse PSM and culturing one half in actinomyocin D to inhibit transcription, while the other was immediately fixed, quantitative PCR analysis showed that *Lfng* mRNA is less stable that *Hes7* mRNA in the mouse PSM. They later linked this instability to be due to the lengthier 3'-UTR of *Lfng*, since cells transfected with a reporter vector followed by either *Hes7* or *Lfng* 3'-UTR, showed bigger mRNA instability in the latter condition. Furthermore, they generated two transgenic mice lines both containing the *Hes7* promoter followed by a reporter, but with different 3'-UTR – one from *Hes7,* and another from *Lfng*. The transgenic line with the *Lfng* 3'-UTR showed a severed reduction in the reporter mRNA, further emphasizing the role of the 3'-UTR in mRNA stability. Moreover, the reporter mRNA presented the same expression pattern of its endogenous counterpart, i.e., the reporter containing the *Lfng* 3'-UTR presents the same pattern as endogenous *Lfng*, being the same true for the *Hes7* transgenic, showing that the 3'-UTR is crucial for the cyclic expression pattern displayed by a given gene. However, contradictory results have been reported by Davis *et al.* (2001). They generated two constructs with the same *hes4* promoter, albeit with different 3'-UTR, one from *hes4* and other from *hes5.7* (former *esr9*). They observed that both transgenes presented an expression that overlap the one from endogenous *hes4* expression.

Furthermore, Fujimuro *et al.* (2014) developed a knock-in mice, where a large intron derived from a human gene was inserted into the 3'UTR sequence of the mouse *Hes7* gene and, due to this insertion, the transcribed gene lost its 3'UTR region. Since this insertion led to the loss of oscillatory expression of *Hes7*, the authors hypothesized that it could be due to insufficient levels of mRNA or protein in the knock-in mice. They found out that transcription levels of *Hes7* mRNA were reduced, and that Hes7 protein was hardly detectable in the mouse PSM, compared to wildtype embryos. As expected, since the protein was not being correctly produced, *Hes7* transcription inhibition was impaired, which compromised the maintenance of the oscillations.

Similar findings have also been described in zebrafish. Gajewski and colleagues' data suggested the existence of a signal in the *her1* 5'-UTR that destabilized *her1* mRNA. However, in their study, the reporter construct they used lacked the 3'-UTR of this gene (Gajewski *et al.*, 2003). While studying the mechanism behind *hes6* expression in an anterior-posterior gradient in the zebrafish PSM, Kawamura *et al.* (2016) reported the existence of a regulatory region downstream of the *hes6* sequence that they thought was responsible for its expression pattern. However, since zebrafish mutants that had this regulatory region coupled to a GFP reporter did not recapitulate *hes6* gradient expression in the PSM, they conjectured that the *hes6* 3'-UTR could be at play. To test this hypothesis, the authors generated a mutant that had the *hes6* 3'-UTR coupled to the GFP-regulatory region. In these mutants, *hes6* was expressed in an anterior-posterior gradient, and restricted to the PSM and tailbud. Moreover, it was also shown that these mutants recapitulated endogenous *hes6* expression, and that *hes6* 3'-UTR was responsible for the mRNA decay rate necessary for the replication of the endogenous *hes6* expression. Following this work, Fujino *et al.* (2018) examined how the half-lives of *her1*, *her7* and *hes6* are regulated. To do so, transgenic zebrafish lines containing either *her1* 3'-UTR, *her7* 3'-UTR or *hes6* 3'-UTR coupled to the EGFP reporter, under the control of the *hes6* enhancer were generated. Analysis of the mRNA decay by quantitative PCR, preceded by a heat-shock treatment to induce the 3'-UTR mRNA expression, revealed that *her7*'s 3'-UTR mRNA was degraded more rapidly than *her1*'s 3'-UTR mRNA, and *hes6*'s 3'-UTR mRNA was the one with the slowest degradation rate. These results were further confirmed to visualize the distribution in the expression of the transgenes in the PSM. In situ hybridization revealed that the expression domain of *her1* 3'-UTR and *her7* 3'-UTR mRNAs was restricted to the more posterior PSM, compared with the *hes6* 3'-UTR mRNA. These results further corroborated the fact that 3'-UTRs are a key component in mRNA stability. Furthermore, the differences in 3'-UTR-mediated mRNA turnover could be due to distinct poly(A) lengths. In the work mentioned before, Fujino and colleagues measured the lengths of the poly(A) tails of *her1*, *her7* and *hes6*, and observed that the first two genes, that display a cyclic expression, have shortened poly(A)

tails, while the statically expressed *hes6* has a longer one. Through the inhibition of the deadenylase complex CCR4-NOT, the authors were able to lengthen the poly(A) tails of *her1* and *her7*. The influence of poly(A) lengthening in mRNA stability, assessed by quantitative PCR, showed that mRNA levels increased about 2-3-fold, when the deadenylation was inhibited, indicating an increase in mRNA stability. These data suggest that poly(A) lengths also contribute to mRNA instability (Fujino *et al.*, 2018).

Lastly, delays in the negative feedback loop can also be achieved through **microRNAs** (miRNAs). These RNA silencing molecules were described to be highly expressed in a tissue-specific fashion during zebrafish segmentation (Wienholds *et al.*, 2005). Using a previously proposed model (Monk, 2003), Xie *et al.* (2007) theorized that miRNAs could either stabilize or destabilize mRNA expression, depending on its effect in mRNA degradation rate. Experimental evidence of oscillatory gene modulation by a miRNA was brought by Bonev *et al.* (2012), who reported that *Hes1* is a direct target of microRNA-9 (mir-9) in the central nervous system. This study reported that mir-9 regulates *Hes1* expression, since mir-9 inhibition led to the stabilization of *Hes1* mRNA. Additionally, *Hes1* oscillations were also found to be modulated by this miRNA; interestingly, *Hes1* oscillations were dampened either if mir-9 was overexpressed or knocked-down, suggesting that *Hes1* oscillations are only maintained within a certain threshold of mir-9. Similar findings were also reported in zebrafish hindbrain development. Her6 protein, the zebrafish Hes1 homologue, was found to be expressed in a fluctuating non-periodic fashion in NPC. However, the onset of Her6 oscillations at the protein levels coincides with the expression of mir-9 in the hindbrain. Furthermore, the authors showed that mutating the mir-9 binding site to *her6* lead to an increase in the noise of protein levels, and cells are unable to transit to a differentiated state, evidencing the role of Her6 in the zebrafish neurogenesis (Soto *et al.*, 2020). During somitogenesis, mir-125a-5p was found to be upregulated in the chicken PSM, as well as to directly target the *Lfng* 3'-UTR. Inhibition of chicken mir-125a-5p activity resembles a phenotype previously described, when *Lfng* is ubiquitously expressed in the chicken PSM (Dale *et al.*, 2003), arguing that miRNA activity plays a role in the *Lfng* mRNA turnover. Moreover, *Lfng* and *hairy1* expression domain was assessed in chicken embryos, 8h after they were electroporated to disrupt the binding of mir-125a-5p to *Lfng* transcripts. Both genes lost their oscillatory expression pattern is the chicken PSM, further evidencing that miRNA regulation is necessary for the destabilization of *Lfng* transcripts (Riley *et al.*, 2013). Taking the results obtained by Riley and colleagues, mathematical modelling performed by Jing *et al.* (2015) shed some insights regarding miRNA role in the segmentation clock. Their work suggests that the interaction between *Lfng* and mir-125a-5p affects both the amplitude and period of the oscillations, thus acting as a fine-tuning mechanism to Notch activity during somitogenesis. Since *Lfng* targeting by mir-125a-5p was

also reported in mouse, suggesting this regulatory mechanism is conserved amongst organisms where *Lfng* plays a role in the segmentation clock (Riley *et al.*, 2013), a study done by Wahi *et al.* (2017) aimed to shed insight about a possible mir-125a-5p-mediated *Lfng* regulation in mouse. Mice expressing a *Venus* reporter construct, under the control of *Lfng* promoter, together with the wildtype *Lfng* 3'-UTR or with mutations in the putative mir-125a-5p binding sites were generated. Double in situ hybridization for *Venus* and endogenous *Lfng* showed that these two expression domains did not always overlap, suggesting that mir-125a-5p regulates the *Lfng* transcript turnover. *Lfng* expression pattern in mice lacking mir-125a-5p was not affected and displayed all the described phases of *Lfng* expression. Interestingly, and unlike what was reported in chicken embryos (Riley *et al.*, 2013), mir-125a-5p does not affect mouse somitogenesis, and therefore has no influence in the maintenance of *Lfng* oscillations (Wahi *et al.*, 2017).

Even though translational delays are not accounted to influence oscillations (Hoyle and Ish-Horowicz, 2013), **protein stability** plays a crucial role. Hirata *et al.*, (2004) addressed what would happen if Hes7 protein half-life time increased from 20 min (wildtype conditions) to 30 min, and found that this provokes a dampening in both Hes7 mRNA and protein oscillations over time. Interestingly, lysine residues were found to play a key role in Hes7 protein stability. The authors generated Hes7 protein mutants, by introducing lysine-to-arginine mutations for each of the seven lysine residues in Hes7 sequence, and found that different mutations gave rise to proteins with a half-life that differed from the wildtype. Moreover, Ishii *et al.*, (2008) reported that some of the lysine mutants lost transcriptional repressor activity, although they were found to be more stable than its wildtype counterpart, thus evidencing the role of these lysine residues to Hes7 protein stability. Studies done by Lewis (2003) and Giudicelli *et al.* (2007) also stated that *her* proteins half-life time had to be short, compared to the zebrafish's segmentation clock pace. Mathematical modelling performed by Ay *et al.* (2013) reiterated the finding that proteins with a short half-life time are an essential requirement for the maintenance of the period of oscillations in the wildtype zebrafish segmentation clock. They further confirmed this finding by determining that Her7 protein has a half-life time about 10 times inferior to the zebrafish segmentation clock period.

Oscillations are a cell autonomous property. In dissected PSM, it has been observed that Notch pathway genes *hes4*, *lunatic fringe* (chicken) and *Hes1* (mouse) expression remains cyclic (Maroto *et al.*, 2005; Masamizu *et al.*, 2006; Palmeirim *et al.*, 1997). Cyclic gene expression was also evaluated in dissociated chicken PSM cells, where it occurred in asynchronous among cells (Maroto *et al.*, 2005) and, in individual mouse PSM cells, *Hes1* oscillations were occurred in a random fashion (Masamizu *et al.*, 2006). Webb *et al.* (2016) also reported that *her1* gene retains an oscillatory expression in cells cultured from zebrafish

tailbud. However, oscillations in individual cells presented a longer period and were less robust, compared with the intact tissue. Altogether, these results suggest that cell-cell communication is a key requirement for oscillations to be in phase within a tissue. This was corroborated by the work of Tsiairis and Aulehla (2016), that showed that a re-aggregation of mouse PSM cells is able to synchronize oscillations. Moreover, oscillatory *Hes1* expression was also shown in mouse ES cells and individual fibroblasts (Kobayashi *et al.*, 2009; Masamizu *et al.*, 2006). Strikingly, synchronization is not mandatory for all cells. For instance, *Hes1* oscillations are asynchronous between individual neural stem cells, and their period varies (Shimojo *et al.,* 2008). This will be further discussed in subsection *Different tissues*.

To form a pair of somites, undifferentiated PSM cells have to be in the same oscillatory phase, i.e., synchronization between them is required. Local synchrony within a tissue is achieved through Delta-Notch signalling (Jiang *et al.*, 2000). Horikawa *et al.* (2006) observed that zebrafish embryos treated with Notch inhibitor DAPT presented a more disperse cellular localization of *her1* mRNA, indicating loss of synchrony and, as a consequence, a defective somitogenesis (Özbudak and Lewis, 2008). The same was observed by Delaune *et al.* (2012) using single cell imaging, who also reported that neighbour cells were in different oscillation phases. Interestingly, two players are crucial for the synchrony of the zebrafish segmentation clock – *deltaC* and *deltaD* (Mara *et al.*, 2007). While *deltaD* is responsible for the priming and onset of the oscillations at the tailbud level, *deltaC* plays in role in maintaining and amplifying the oscillations in adjacent cells along the posterior PSM. Soza-Ried *et al.* (2014) further evidenced the role of Notch-Delta signalling in maintaining the oscillations synchronized, by using a *deltaC* zebrafish mutant (*bea* mutant) line, where the onset of *deltaC* expression was driven by a heat-shock. Furthermore, they employed this system to address how changes in *deltaC* expression would affect the zebrafish segmentation, and reported that it is disrupted after the formation of five somites. However, when these *bea* mutants are exposed to heat-shock treatment, creating an artificial pulse of *deltaC*, cells are able to synchronize for a while, and the next somites to form have normal morphology and well-defined boundaries. Moreover, the longer the intervals between heat-shocks, the larger the somite width, evidencing that is possible to alter the oscillations of the segmentation clock by modulating the timing of successive pulses of Notch activation.

The periodicity of the clock oscillations varies between species, tissues within an organism, and over developmental time. However, what determines this period is poorly understood. Are the same mechanisms employed? Answering these and other pressing questions would allow us to understand how TIME is set and perceived during development. This review aims to summarize the available knowledge regarding the pace of the clock in

different organisms and under different experimental conditions, hoping it will raise new key questions and novel experimental approaches.

## The Embryo Clock (EC) periodicity varies among different organisms

### *Time of somite formation*

The time each pair of somites takes to form is species-specific. Studies done using several animal models showed it displays great variability between different organisms – it ranges from 30 min in zebrafish to about 4-6 hours in humans (Table 1).

**Table 1. Reported time of somite formation in different vertebrate organisms.**

| Organism | Period | Reference |
|---|---|---|
| **Human** | 4 – 5 h | Müller & O'Rahilly, 1986 |
| **Mouse** | 2 h | Tam, 1981 |
| **Chicken** | 115 min<br>90 min (37ºC) | Menkes *et al.*, 1961<br>Pourquié, 2004 |
| **Quail** | ~90 min | Inferred from Packard 1980 |
| **Emu** | 100 – 110 min | Nagai *et al.*, 2011 |
| **Zebrafish** | 30 min | Stickney, Barresi & Devoto, 2000 |
| **Medaka** | 60 min | Elmasri *et al.*, 2004 |
| ***Xenopus*** | 40 min | Cooke and Zeeman 1976; Li et al. 2003 |
| **House snake** | 60 min | Gomez *et al.*, 2008 |
| **Corn snake** | 100 min | |
| **Whiptail lizard** | 4 h* | |

* tail

One curious aspect is that the size of the organism does not influence the time a pair of somites take to form. For instance, somites have a relatively similar time of formation in the chicken and the emu, although dimension-wise these two birds are very distinct. Likewise, the time of somite formation does not depend on the phylogenetics, since vertebrates belonging to different phyla have the same somitogenesis period: 60 min for medaka and the house snake. Overall, the time somites take to form is shorter in teleost fish and amphibians and longer in higher organisms, such mammals (Table 1, Figure 1).

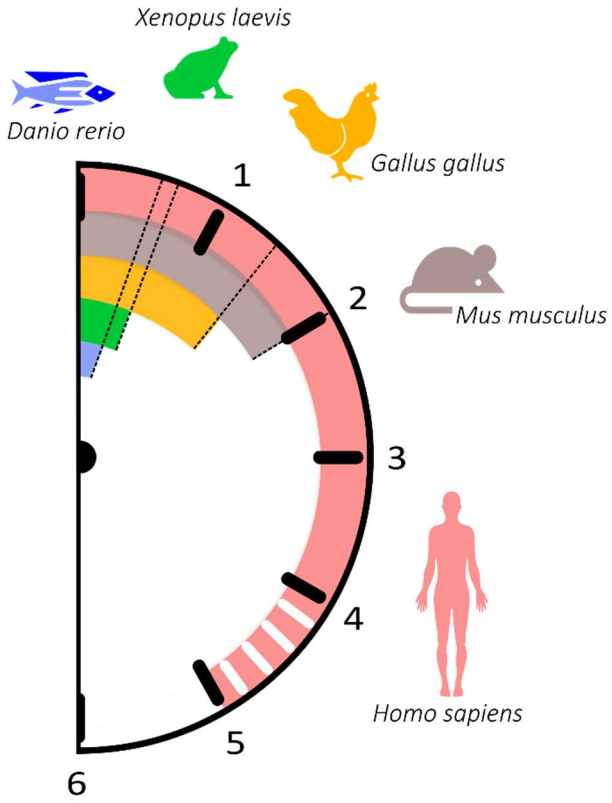



**Figure 1. The somitogenesis clock ticks with different paces amongst vertebrates.** *Danio rerio*: 30 min (blue); *Xenopus laevis:* 40 min (green); *Gallus gallus*: 90 min (orange); *Mus musculus*: 120 min (brown); *Homo sapiens*: 4 to 5 hours (pink).

***EC gene expression dynamics in the presomitic mesoderm***

Since the existence of a molecular segmentation clock was proposed by Palmeirim *et al.* (1997), other genes have been described has having a role during somitogenesis (Dequéant *et al.*, 2006; Krol *et al.*, 2011). However, the vast majority is just described has having a cyclic or dynamic expression in the PSM (Dale *et al.*, 2006; Elmasri *et al.*, 2004; Feller *et al.*, 2008; Gajewski *et al.*, 2006; Leimeister *et al.*, 2000b; Mastromina *et al.*, 2018). Thus, we will only focus on the cyclic genes with a defined period of oscillation (Table 2).

The genes involved in the segmentation clock oscillate with a periodicity characteristic of each species, which correlates with the time of somite formation (see Table 1). The signalling pathways that comprises these genes are well conserved; however, data suggests that cyclic genes display an evolutionary plasticity, since they differ in the studied organisms (Krol *et al.*, 2011). Interestingly, both in zebrafish and medaka, only genes belonging to the Notch signalling pathway have been described so far, as having an oscillatory behaviour during somitogenesis.

**Table 2: Oscillation period of known genes with cyclic expression.**

| Organism | Gene | Tissue/cell line | Period | Technique | Reference |
|---|---|---|---|---|---|
| Human | *HES1* | Mesenchymal stem cells | 5 h | qPCR | (William *et al.*, 2007) |
| | *HES7* | iPSC | 5 h | Live imaging | (Diaz-Cuadros *et al.*, 2018) |
| | | PSM-like cells derived from iPSC | 5 – 6 h | Luciferase reporter assay | (Matsuda *et al.*, 2019b) |
| | | | 5.36 h | Luciferase reporter assay | (Matsuda *et al.*, 2019a) |
| | | ESC | ~5 h | Luciferase reporter assay | (Chu *et al.*, 2019) |
| Mouse | *Ascl* | NPC | 2.92 h | Live imaging | (Imayoshi *et al.*, 2013) |
| | *Axin2* | PSM | 1.7 h | Microarray | (Dequéant *et al.*, 2006) |
| | | | 2 h | In situ hybridization (ISH) | (Aulehla *et al.*, 2003) |
| | *Dact1* | PSM | 2 h | ISH | (Suriben *et al.*, 2006) |
| | *Dll1* | PSM | 2 h | ISH | (Bone *et al.*, 2014; Maruhashi *et al.*, 2005) |
| | | NPC | 2 h | Live imaging | (Shimojo *et al.*, 2008) |
| | *Dusp4* | PSM | 2 h | ISH | (Niwa *et al.*, 2007) |
| | *Hes1* | Myoblasts, fibroblasts, neuroblastoma and teratocarcinoma cells | 2 h | qPCR | (Hirata *et al.*, 2002) |
| | | C2C12 myoblasts | | | (William *et al.*, 2007) |
| | | Fibroblasts (C3H 10T1/2 cells) | 2.03 h | Bioluminescence imaging | (Masamizu *et al.*, 2006) |
| | | PSM | 2.67 h | | |
| | | Dissociated PSM cells | 2.58 h | | |
| | | Neural stem cells | 2 – 3 h | Live imaging | (Shimojo *et al.*, 2008) |
| | | ESC (MG1.19 cell line) | 3 – 5 h | Live imaging | (Kobayashi *et al.*, 2009) |
| | *Hes5* | NPC | 2 h | Live imaging | (Imayoshi *et al.*, 2013) |
| | | Spinal cord | 3.3 h | Live imaging | (Manning *et al.*, 2019) |
| | *Hes7* | PSM | 2 h | ISH | (Bessho *et al.*, 2001) |
| | | Induced PSM from ESC | 2.5 – 3 h | Live imaging | (Matsumiya *et al.*, 2018) |
| | | PSM-derived cells form iPSC | 2.03 h | Luciferase reporter assay | (Matsuda *et al.*, 2019a) |
| | *Lfng* | PSM | 2 h | ISH | (Aulehla and Johnson, 1999; Forsberg *et al.*, 1998) |
| | *Mesp2* | PSM | 2 h | ISH | (Cole *et al.*, 2002) |
| | *Nkd1* | PSM | 1.87 h | Microarray | (Dequéant *et al.*, 2006) |
| | | | 2 h | ISH | (Ishikawa *et al.*, 2004) |
| | *Notch1* | PSM | 2 h | ISH | (Bone *et al.*, 2014; Takahashi *et al.*, 2000) |
| | *Nrarp* | PSM | 2 h | ISH | (Sewell *et al.*, 2009) |
| | *Smad6* | Fibroblasts (C3H 10T1/2 cells) | 2 h | qPCR | (Yoshiura *et al.*, 2007) |
| | *Snail1* | PSM | 2 h | ISH | (Dale *et al.*, 2006) |
| | *Sprouty4* | PSM | 2 h | ISH | (Hayashi *et al.*, 2009) |

**Table 2: Oscillation period of known genes with cyclic expression.** (Continuation)

| Organism | Gene | Tissue/cell line | Period | Technique | Reference |
|---|---|---|---|---|---|
| Chicken | *hairy2* | PSM | 1.5 h | ISH | (Jouve *et al.*, 2000) |
| | | Limb bud | 6 h | | (Pascoal *et al.*, 2007) |
| | *HES4* | PSM | 1.5 h | ISH | (Palmeirim *et al.*, 1997) |
| | *HEY2* | PSM | 1.5 h | ISH | (Leimeister *et al.*, 2000a) |
| | *LFNG* | PSM | 1.5 h | ISH | (McGrew *et al.*, 1998) |
| | *NRARP* | PSM | 1.5 h | ISH | (Wright *et al.*, 2009) |
| Medaka | *her1/11* | PSM | 1 h | ISH | (Elmasri *et al.*, 2004) |
| | *her5* | PSM | 1 h | ISH | (Gajewski *et al.*, 2006) |
| | *her7* | PSM | 1 h | ISH | (Gajewski *et al.*, 2006) |
| Xenopus | *hes5.5* | PSM | 0.67 h | ISH | (Li *et al.*, 2003) |
| | *hes9.1* | PSM | 0.67 h | ISH | (Li *et al.*, 2003) |
| Zebrafish | *deltaC* | PSM | 0.5 h | ISH | (Jiang *et al.*, 2000) |
| | *her1* | PSM | 0.5 h | ISH | (Holley *et al.*, 2000) |
| | *her2* | PSM | 0.5 h | Microarray | (Krol *et al.*, 2011) |
| | *her4* | PSM | 0.5 h | Microarray | (Krol *et al.*, 2011) |
| | *her7* | PSM | 0.5 h | ISH | (Oates and Ho, 2002) |
| | *her11* | PSM | 0.5 h | ISH | (Sieger *et al.*, 2004) |
| | | PSM | 0.5 h | ISH | (Gajewski *et al.*, 2006) |
| | *her12* | | | | (Shankaran *et al.*, 2007) |
| | *her15* | PSM | 0.5 h | ISH | (Shankaran *et al.*, 2007) |
| | *hey1* | anterior PSM | 0.5 h | ISH | (Winkler *et al.*, 2003) |
| | *nrarp-a* | PSM | 0.5 h | ISH | (Wright *et al.*, 2009) |
| | *tbx16* | PSM | 0.5 h | qPCR | (Krol *et al.*, 2011) |

Legend: iPSC – induced Pluripotent Stem Cells; ESC – Embryonic Stem Cells; ISH – *In situ* hybridization; NPC – Neural Progenitor Cells; qPCR – quantitative PCR.

Work done using mathematical modelling, showed that *Hes7* and *her1* oscillations slowdown in the anterior PSM, in mouse (about 1.5-fold) and in zebrafish, respectively (Giudicelli *et al.*, 2007; Niwa *et al.*, 2011). Shih *et al.* (2015) corroborated these findings using live imaging in a transgenic zebrafish line with a reporter protein. They stated that the periodicity of the segmentation clock increases by 1.5-fold in the anterior PSM, comparatively with the posterior PSM. To better understand the dynamics of the segmentation clock slowing in the anterior PSM, the reporter expression of the cells that will form either side of somite boundaries was assessed. Interestingly, it was found that within the same presumptive somites, the clock oscillations are arrested from a posterior-to-anterior direction, since cell-tracking experiments showed that cells that are incorporated in a posterior somite boundary cease oscillations prior to cells that are incorporated in the anterior boundary of the same somite. Moreover, the authors also reported that cells at a one-somite distance are initially synchronized in the posterior PSM, while in the anterior PSM – as the clock slows down –, they shift into anti-phase, giving rise to non-synchronized waves of cyclic expression. However, opposite findings have also been described using the *Looping* zebrafish reporter line, which expresses a fusion of the Her1 protein with YFP. Using a multidimensional time-lapse setup

that allow to compare the periods of morphological segmentation and genetic oscillations, Soroldoni and colleagues reported that the period of oscillations in the zebrafish anterior PSM matches the segmentation period, while the posterior PSM displays a longer period. The authors proposed a model where this is due to a Doppler effect together with a dynamic wavelength. As the embryo elongates, PSM length gets shorter over time, resulting in a relative motion of the anterior end of the PSM towards the posterior end, i.e., the anterior end of the PSM is moving towards the kinematic waves – which also decrease in number as segments are formed –, thus resulting in a faster period of oscillation in the anterior PSM (Soroldoni *et al.*, 2014).

**Table 3: Oscillation period of known cyclic proteins.**

| Organism | Protein | Tissue/cell line | Period | Technique | Reference |
|---|---|---|---|---|---|
| Mouse | Delta1 | PSM | 2 h | Immunochemistry | (Bone *et al.*, 2014) |
| | | | 2.45 h | Live imaging | (Shimojo *et al.*, 2016) |
| | | NPC | 2.38 h | Live imaging | (Shimojo *et al.*, 2016) |
| | | Pancreas | ~1.5 h | Live imaging | (Seymour *et al.*, 2019) |
| | Dusp4 | PSM | 2 h | Immunochemistry | (Niwa *et al.*, 2007, 2011) |
| | Hes1 | Myoblasts | 2 h | | (Hirata *et al.*, 2002) |
| | | NPC | 2.5 h | Live imaging | (Imayoshi *et al.*, 2013) |
| | | Pancreas | ~1.5 h | Live imaging | (Seymour *et al.*, 2019) |
| | Hes7 | PSM | 2 h | Immunochemistry | (Bessho *et al.*, 2003) |
| | NICD | PSM | 2 h | Immunochemistry | (Huppert *et al.*, 2005) |
| | | | | | (Niwa *et al.*, 2011) |
| | Notch1 | PSM | 2 h | Immunochemistry | (Bone *et al.*, 2014) |
| | | | | | (Morimoto *et al.*, 2005) |
| | p-ERK | Fibroblasts (CH3 10T1/2 cells) | 2h | Western Blot | (Nakayama *et al.*, 2008) |
| | | PSM | | Immunochemistry | (Niwa *et al.*, 2011) |
| | p-Smad1/5/8 | Fibroblasts (CH3 10T1/2 cells) | 2 h | Western Blot | (Yoshiura *et al.*, 2007) |
| | Olig2 | NPC | 2 h | Live Imaging | (Imayoshi *et al.*, 2013) |
| | Smad6 | Fibroblasts (CH3 10T1/2 cells) | 2 h | Western Blot | (Yoshiura *et al.*, 2007) |
| Chiken | LFNG | PSM | 1.5 h | Western Blot | (Dale *et al.*, 2003) |
| Zebrafish | deltaC | anterior PSM | 0.5 h | Immunochemistry | (Giudicelli *et al.*, 2007) |
| | ERK | PSM | 0.5 h | Live imaging | (Wong *et al.*, 2019) |
| | Her6 | Neural progenitors | 1.2-1.4h | Live imaging | (Soto *et al.*, 2019) |
| | Her13.2 | PSM | 0.5 h | Immunochemistry | (Schröter *et al.*, 2012) |
| | Tbx6 | PSM | 0.5 h | Immunochemistry | (Wanglar *et al.*, 2014) |

The differences regarding the segmentation clock *modus operandi* in anterior and posterior PSM were elucidated by other studies. Choorapoikayil *et al.* (2012) developed zebrafish mutants for *her1* and *her7*, where the respective proteins are unable to form dimers. *Her1* mutants showed defects in the pre-patterning of the first 3 somites and the cyclicity of *deltaC* expression was only lost in the anterior PSM. However, at 10-12 somite stage, pre-patterning defects were no longer observable, and the anterior PSM displayed a single stripe

of *deltaC*, while different phases of *deltaC* oscillatory expression were detected in the posterior PSM. On the other hand, *her7* mutants displayed somite border defects between somites 8 and 17, although appearing unaffected after. Similar findings were described by Schröter *et al.* (2012) and Hanisch *et al.* (2013), using the zebrafish mutants described before. Segmentation defects in *her1* mutants were observed only in the anterior trunk, while segments in *her7* mutants became defective by somites 10 or 11. Overall, these results suggest the existence of two independent genetic modules operating in the zebrafish anterior and posterior PSM.

There is now also evidence that spatiotemporal metabolic changes are connected with PSM development and differentiation. The first link between metabolism and segmentation came to light with the work of Özbudak and colleagues. Microarray analysis performed in bissected PSM and tailbud samples from zebrafish embryos revealed that anterior and posterior PSM have different sets of genes being upregulated: cell cycle and DNA metabolic functions are enriched in the posterior PSM, while oxidative metabolism is enriched both in anterior PSM and somites. Moreover, the authors also reported an ATP 2-fold increase between the posterior PSM and the formed somites, as well as 2.5-fold increase in Cytochrome C oxidase activity in the last five-formed somites compared with the tailbud (Özbudak *et al.* 2010). Recently, a gradient of glycolytic activity was linked to the mesoderm development in mouse and chicken embryos. In mice, Bulusu *et al.* (2017) described that glycolytic enzymes expression, as well as the glycolytic activity, were higher in the posterior PSM – in comparison with its anterior counterpart –, thus being the first to provide evidence for the metabolic differences across the mouse PSM during embryonic development. These findings were corroborated by Oginuma *et al.* (2017) who also reported that metabolites involved in glycolysis were detected at significantly higher levels in the posterior PSM of chicken embryos. Wondering if these metabolic differences were relevant in a physiological context, mouse PSM explants were cultured in different conditions (glucose, and pyruvate supplemented culture medium) in order to assess the consequences of bypassing glycolytic activity. While explants cultured in normal conditions (glucose supplemented medium) developed normally, pyruvate-cultured explants displayed several defects, mainly in the posterior PSM, while the anterior PSM developed normally under the same conditions. Moreover, this phenotype coincided with the loss of the Notch signalling pathway gene *lunatic fringe* in the posterior PSM, although its expression was maintained in the anterior PSM. The authors also developed a genetically encoded sensor for pyruvate, in order to monitor metabolic transitions during PSM differentiation in real-time. They report that pyruvate levels, i.e. glycolytic activity, decrease as cells transit towards an anterior PSM-like, more differentiated, state. Akin to what was described for mouse (Bulusu *et al.* 2017), microarray analysis of chicken PSM fragments also showed that glycolytic enzymes were expressed in a posterior-like gradient. Moreover, cytochrome c oxidase activity was higher in the anterior PSM, both in mouse and in chicken

(Oginuma *et al.*, 2017), similarly to the findings reported in zebrafish by Özbudak *et al.* (2010). Strikingly, the domain where the glucose uptake takes place in the posterior PSM of the chicken embryo overlaps with the region where *Fgf8* is expressed. Components of the FGF pathway have been shown to control the oscillations during somitogenesis (Naiche *et al.*, 2011; Wahl *et al.*, 2007), as well as to influence embryo elongation (Boulet and Capecchi, 2012), and thus the authors went on to assess the role of glycolysis in the elongation of the posterior embryonic axis. To do so, chicken embryos were treated with 2-deoxy-D-glucose (2DG), a competitive inhibitor of the glycolytic enzyme hexokinase, and with sodium azide ($NaN_3$) – a respiration inhibitor. Embryos treated with 2DG displayed severe elongation defects, even though somite formation occurred normally, while $NaN_3$-treated embryos had somite segmentation blocked.

***Different developmental stages***

Palmeirim *et al.* (1997) showed that, in chicken embryos with 15-20 somites (HH12-13$^+$), oscillations occurred with a 90-min periodicity. However, does the *pace* of the clock change throughout development?

While studying what drives the end of somitogenesis, Tenin and colleagues (2010) observed that the last 5 to 8 somites are formed slowly in the chicken embryo. Moreover, they also showed that *lunatic fringe* oscillates with a period of 150 min (HH23), and that this time is concomitant with somite formation in the late stages of development. By stage HH24, *lunatic fringe* oscillation becomes undetectable, and the molecular oscillator is arrested.

Knowledge about the EC is still scarce when it comes to the early stages of the chicken development. Rodrigues *et al.* (2006) characterized the onset of several Notch-related genes in the anterior-most somites (somites 1-10), and reported that their expression emerges at different somite levels, getting more complex by somite 10. During gastrulation, the EC was proposed to already be active. Jouve et al. (2002) reported the existence of pulses of gene expression of *hairy2* and *lunatic fringe* in the prospective PSM of chicken embryos.
What triggers the onset of the Embryo Clock, and the existence of a periodicity in these early developmental stages, remains elusive.

***Different tissues***

Throughout development, the same gene regulatory networks are employed differently, giving rise to different outcomes. Besides the segmentation of the axial vertebrate body plan, EC's genes have been reported to be involved in other processes, where their behaviour differs from the one presented in the PSM during somitogenesis. In this subchapter, we will focus on neural cell differentiation and on the chicken limb growth.

Embryonic stem cells (ESC) are able differentiate into different cell types belonging to all three germ layers – mesoderm, endoderm and ectoderm. However, during development, the balance between their self-renewal capacity and pluripotency needs to be maintained and this is achieved through the heterogeneity of the population. Evidence that oscillatory gene expression played an important role in the differentiation of ESC was described by Kobayashi and colleagues (2009). Using a mouse ESC cell line transformed with a luciferase reporter coupled to the *Hes1* promoter, the authors reported that EC gene *Hes1* expression levels were variable between individual cells, and later confirmed that *Hes1* expression was indeed cyclic in mouse ESC. Moreover, they described that *Hes1* expression oscillations occurred with a periodicity of 3 to 5 h – longer than the *Hes1* period described for other cell types isolated from mice (see Table 2, (Hirata *et al.*, 2002; Masamizu *et al.*, 2006) –, and that they were asynchronous. To assess whether different expression levels of *Hes1* would influence the differentiation outcome of ESC, the authors sorted, what they termed, Hes1-high and Hes1-low cells, and differentiated them into a neural fate. Hes1-high cells differentiated more efficiently into mesodermal cells, while Hes1-low cells into neurons. To further confirm this finding, when *Hes1*-null cells and ESC were cultured under neural differentiation conditions, all Hes1-null cells differentiated into neural cells, while only a small subset of ESC did so (Kobayashi *et al.*, 2009). Altogether, these results emphasize the contribution that different *Hes1* expression levels have when it comes to cell fate decisions.

Expression of some clock genes was also found to be oscillatory in neural progenitor cells (NPC) during mouse development. Using real-time imaging, *Hes5* and *Dll1* mRNA were determined to oscillate with a 2 hour-periodicity (Imayoshi *et al.*, 2013; Shimojo *et al.*, 2008, 2016), while for *Hes1* this period has been reported to be comprised between 2 h to 3 h (Shimojo *et al.*, 2008; Imayoshi *et al.*, 2013) – similar to the period described for EC genes in the mouse PSM (see Table 2). Likewise, Hes1 and Delta1 proteins display an identical period of oscillation (see Table 3; Imayoshi *et al.*, 2013). Studies performed in mouse NPC reiterated the importance of gene expression oscillations for cell fate determination. During neurogenesis, the NPC population is maintained through the repression of fate determination factors that promote neural differentiation, such as *Neurogenin2*, *Ascl1/Mash1* and *Olig2* by *Hes1* and *Hes5* oscillatory expression levels. As a consequence, *Neurogenin2* and *Ascl1* also display an oscillatory mRNA levels (Imayoshi *et al.*, 2013; Shimojo *et al.*, 2008). Imayoshi *et al.*, (2013) reported that the expression level of gene oscillations plays an important role in NPC differentiation. After segregating NPC according to their levels of expression – high and low – of *Hes1*, *Ascl1* and *Olig2*, and culturing them in a differentiation medium, the authors found that higher or lower levels of expression give rise to different differentiation outcomes. For instance, Hes1-high NPC differentiate into an astrocyte lineage, while Hes1-low NPC into neurons. In the same study, it was also reported that towards the differentiation into a given

lineage, the co-expression of the fate determination factors that exists in NPC is lost and, as a result, one of the factors accumulates, while the other two are lost. Unlike somitogenesis, oscillations during neural development are asynchronous. While undergoing differentiation, neural cells inhibit their neighbours to differentiate into the same cell type through lateral inhibition mediated by the Notch signalling pathway.

Studies have also linked components of the Notch signalling pathway to pancreatic development. Mice deficient for *Dll1* and *RBP-Jk* displayed a phenotype where the differentiation of pancreatic endocrine cells was accelerated, and severe hypoplasia was observed in mice lacking functional *Hes1* (Apelqvist *et al.*, 1999; Jensen *et al.*, 2000). *Hes1* protein is able to bind to the *Neurogenin3* promoter, inhibiting its expression. This will repress the activation of *Neurod*, an insulin gene transcriptional activator, and as a consequence, inhibit the differentiation into an endocrine fate (Jensen *et al.*, 2000; Lee *et al.*, 2001). Similar to what is observed during neural development, *Hes1* expression is responsible for the maintenance of multipotency, and undifferentiated neighbour cells also undergo a mechanism of lateral inhibition, to give rise to different cell types (de Lichtenberg *et al.*, 2018). Even though pancreas development shares common players with somitogenesis and neural development, it is not known or described if it these players also share an oscillatory behaviour. The only clue regarding this matter is the work submitted by Seymour and colleagues, where they report that both *Hes1* and *Dll1* proteins oscillate with a 90-minute periodicity in cultured pancreatic explants isolated from mice (Seymour *et al.*, 2019).

Oscillatory gene expression was also reported during the chicken limb development. *C-hairy2* was first described as a cyclic gene in the chicken PSM, with an expression pattern and periodicity similar to *hes4* (90 min) (Jouve *et al.*, 2000). The existence of a molecular clock operating during limb development was first proposed by Pascoal *et al.*, (2007). Using the proxy that if *c-hairy2* expression was cyclic, its cyclicity would be correlated with the formation of a new autopod limb skeletal element, it was determined that a new element takes about 12 hours to be formed. Furthermore, to assess the correlation between this finding and *c-hairy2* expression, one of the wings was surgically removed and immediately fixed, while the remaining embryo was incubated for certain time periods. After in situ hybridization, the authors reported that *c-hairy2* expression had a 6-hour periodicity, indicating that the autopod chondrogenic precursor cells need to undergo two cycles of *c-hairy2* expression until a new skeletal segmented is formed. This was the first evidence that a molecular clock is also operating during limb development, a process where temporal control is also fundamental. *C-hairy1* was also described to be expressed in the distal mesenchyme of the avian limb (Vasiliauskas *et al.*, 2003). However, its cyclic expression during limb development was only recently reported. Bhat et al. cultured cells from chicken pre-cartilage leg mesenchyme and assessed *hairy1* expression period, retrieving a period of oscillation of 6h (Bhat *et al.*, 2019).

The cases mentioned above are prime examples that the same mechanism – the molecular clock – can have distinct outputs. In both PSM and limb – during somitogenesis and the formation of new autopod limb elements, respectively –, cells need to be synchronized in order to be able to aggregate and give rise to new segmented structures. However, in neurogenesis, the asynchronous oscillatory expression is essential for the proliferation of NPC. Once the oscillation ceases, neural differentiation is promoted. In this case, since the output is cell fate, the need for synchrony between neighbouring cells is not a key requirement.

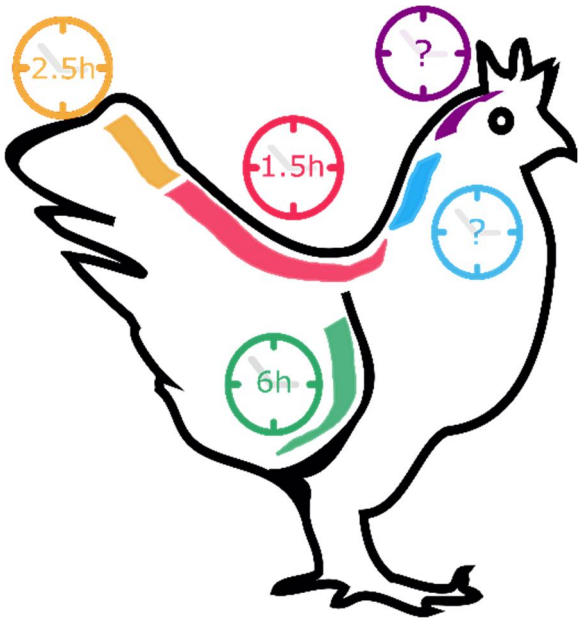



**Figure 2.** The pace of *Gallus gallus* embryonic clock in different tissues. During somitogenesis (red and orange), the EC pace ranges from 1.5 to 2.5 hours, while during forelimb development (green) a cycle lasts 6 hours. The existence – and if so, the pace – of the EC in the early stages of somitogenesis (blue) and during neural development (purple) is still unknown.

## Experimental manipulations of the EC

To fully comprehend the mechanisms behind ultradian rhythms, it is not enough to rely only on the description of the genes' cyclic behavior, such as transcription, translation, and degradation. Since its discovery, researchers have been manipulating EC genes to understand it better. Many experiments, such as deleting core cyclic genes or EC related pathways, give insight into how this intricate and mysterious mechanism works. Several groups also attempted to test the influence of specific embryonic structures on the pace of EC. These experiments turn out to be imperative to understand diseases related to EC.

Interestingly when the EC is experimentally perturbed, its periodicity is not redefined, even though the organism presents totally different EC paces in different tissues, such as the chicken 90 min EC in the PSM and 6h in the forelimb outgrowth (Palmeirim *et al.*, 1997; Pascoal *et al.*, 2007). The majority of the experiments either totally disrupt the EC oscillations, or slightly increase or decrease the pace of the EC oscillations; others have no effect at all. What is even more puzzling is that the same manipulation in different organisms rarely have the same output, which turns it even more difficult to understand this mechanism. To have a more comprehensive understanding of the EC, in this chapter, we will review the major experiments performed leading to both total disruption of the EC oscillations or the ones that could increase or decrease the oscillation period.

Some manipulations lead to disruption of EC oscillations. Since its discovery, a broad range of EC gene manipulations has led to a complete arrest of the embryo clock in different organisms and tissues (Table 4). Although the pace of the EC clock seems to be robust, small changes in the EC machinery seem to have a huge impact on the whole mechanism.

The first and more intuitive way to perturb the EC is to manipulate genes that are directly involved. Niwa and colleagues deleted mouse *Hes7*, a core mouse EC gene. The oscillations were completed disrupted by deleting it, leading to a steady expression of core clock components such as NCID and Mesp2 (Niwa *et al.*, 2011). In fact, to disrupt the oscillation of the EC it is enough to interfere with the delays of transcript production. By deleting all three introns from the *Hes7* gene Takashima and collaborators interfered with elongation and, more importantly, with splicing time. The team measured that splicing and transcription of the mouse *Hes7* introns account for a 19 min delay, which was proven to be crucial to maintain a robust oscillation since its ablation leads to a total loss of *hes7* oscillations (Takashima *et al.*, 2011). Similar results were found when mouse *dll1* introns were removed or increased (Shimojo *et al.*, 2016).

**Table 4: EC disruption**

| Organism | Tissue / Cell line | EC Pace_min | Manipulation | Somitogenesis phenotype | Expression levels | Reference |
|---|---|---|---|---|---|---|
| Mouse | Tailbud sections | 120 | uncoupling oscillations of notch and wnt (d-PSM) | Stops segmentation | Cells continue to oscillate and arrest later | (Sonnen *et al.,* 2018) |
| | PSM | 120 | Hypoxia in hes7+/- and Mesp2+/- | Somite defects and axial strutures defects | Notch pathway and FGF are downregulated | (Sparrow *et al.,* 2012) |
| | Embryo | 123 | Hes7 intron deletion | Fused somites and consequent vetebral problems | Steady and continuous Hes7 expression | (Takashima *et al.,* 2011) |
| | Embryo | 120 | Dll1 elongation manipulation | fusion of somites | Steady and continuous Dll1 protein expression and dampened oscillations of Hes1 and Hes7 | (Shimojo *et al.,* 2016) |
| | Embryo | 120 | Hes7 KO | segmentation and patterning defects | Steady and continuous NCID expression and steady expression of MESP2 | (Niwa *et al.,* 2011) |
| | Embryo | 120 | LFNG Dominant alele (RLFNG) resistant to Golgi degradation and non secreted | absent or disorganized intersomitic boundaries | Clock genes lose their cyclic behavior (Dll1, Notch and Hes7 affected) | (Williams *et al.,* 2016) |
| | Embryo | 120 | Hes7 3'UTR insertion of 5, 10 or 20kb | Somite defects and axial strutures defects | Clock genes have their oscillations dampened (LFNG and Hes7 affected) | (Fujimuro *et al.,* 2014) |
| Chicken | Embryo | 90 | Mir-125-5p manipulation | absent or disorganized intersomitic boundaries | Steady and continuous hairy1 expression and no expression of lfng | (Riley *et al.,* 2013) |
| | Forelimb | 6h | Abrogate FGF signaling via AER ablation or inhibiting drugs | N/A | Absence of hairy2 expression in the Distal Cyclic Domain | (Sheeba *et al.,* 2012) |
| | Forelimb | 6h | Abrogate Shh signaling via ZPA ablation or inhibiting drugs | N/A | Absence of hairy2 expression in the Distal Cyclic Domain | (Sheeba *et al.,* 2012) |
| Medaka | Embryo | 60 | Pharmacological modulation of ROS levels using NAC and DPI treatments | complete or partial loss of somite boundary formation and irregular somite sizes and shapes | her4 and hey1 downregulated | (Ventre *et al.,* 2015) |
| Zebrafish | Embryo | 30 | Double Mutants her1 and her7 | Somite shape seems to be compromized but not the segmentation of the body. | Complete disruption of the embryo clock. Steady gene expression throughout the PSM. | (Lleras-Forero *et al.,* 2018) |
| | Embryo | 30 | Double Morphant her1 and her7 | Somite shape and boundeies are compromized | DeltaD, Mesp and Notch have their expression disrupted. | (Henry *et al.,* 2002) |
| | Embryo | 30 | Mutation in her1 or her7 | her1 mutation disrupts the three anterior-most somite borders; her7 mutation display somite border defects from somite 8 to 17. | Steady and continuous deltaC, her1 and her7 expression and steady expression of mesp2 | (Choorapoikayil *et al.,* 2012) |
| | Embryo | 30 | Greb1 MO leads to a loss of her 7 expression | Somite shape and boundeies are compremized | Her1 expression still ok, but her7 is downregulated | (Prajapati *et al.,* 2020) |
| | Embyo/hindbrain | 1,2-1,4h | miR-9 manipulation that targets her6 | N/A | Overall levels of her6 stabilize, due to a noise increase | (Soto *et al.,* 2020) |

Also, *hes7*'s 3'UTR plays an important role in the oscillation. By deleting portions of 5, 10, and 20 kb and maintaining the coding region intact, was enough to stabilize not only *hes7* expression but also other core cyclic genes such as *lfng* (Fujimuro *et al.*, 2014). Experiments in other animal models show similar results. By ablating zebrafish *her* genes in different ways also lead to a complete disruption of the embryo clock. The first study with morpholinos against *her1* and *her7* compromised the expression of genes crucial to somitogenesis, such as *notch, mesp2,* and *deltaD,* which promoted somite shape defects (Henry *et al.*, 2002). These results were corroborated by Lleras-Forero and colleagues, by using TALEN, they deleted *her1* and *her7* they found the same morphological defects (Lleras-Forero *et al.*, 2018). However, these genes are expressed simultaneously throughout all somitogenesis, they seem to be important in different somitogenesis stages (Choorapoikayil *et al.*, 2012). By mutating these two genes and disabling their capacity to form dimers, it was found that *her1* is important for *deltaC* patterning in the anterior PSM during the formation of the first three somites, showing defects in these respective somites. After that, embryos regain expression of *deltaC* in this tissue, but with only a stripe, and a normal oscillation in the posterior PSM. On the other hand, *her7* mutation only showed impaired somite border formation between somite 8 and 17 (Choorapoikayil *et al.*, 2012). Thus, different core genes of the EC seem to have distinct roles in embryo segmentation, which shows how complex and far from being understood this system is. Moreover, these genes are also regulated independently, *greb1* morphant completely disrupted the oscillation of *her7* by downregulating its expression but does not have any implication in the expression levels of *her1* or in its oscillatory behavior (Prajapati *et al.*, 2020).

*Lunatic fringe (lfng)* is also being broadly used to study embryo clock oscillations and the molecular segmentation mechanism. Moreover, disruption of the expression of this gene leads to severe oscillatory problems. In 2013, Riley and colleagues, by inhibiting the repression action of miR-125a-5p on *lfng*, completely disrupted its oscillation and other chicken clock genes such as *hes4* (Riley *et al.*, 2013). The same phenotype was not found in mice (Wahi *et al.*, 2017). Avoiding the export of this enzyme, also leads to the disruption of the embryo clock. Since it is involved in NCID cleavage, it means that this gene could still have a different unknown role on the EC and/or in somitogenesis (Williams *et al.*, 2016).

Other factors, such as hypoxia or ROS manipulation, have been shown to be important for normal EC behavior. Pregnant mice carrying heterozygous pups for *Hes7* or *Mesp2* had their oscillatory behavior disrupted only when they were exposed to hypoxia. This could explain why not all heterozygous humans present developmental problems, even in the same family (Sparrow *et al.*, 2012). One way that hypoxia could modulate the EC, is by changing the levels of ROS levels. In Medaka, pharmacological modulation of ROS levels using NAC and DPI was

enough to downregulate *her4* and *hey1,* two EC core genes in this organism (Ventre *et al.*, 2015).

In all these works, small changes in the EC lead to an abrogation of the EC, which could simply be because it changes the EC's expression dynamics and changes the communication between the different EC pathways. Sonnen and colleagues found that uncoupling Notch and Wnt signaling only in the determined PSM was enough to stop somitogenesis (Sonnen *et al.*, 2018). Thus, mutation or changes in any clock gene, could simply lead to a shift in their expression and impair somitogenesis.

There is also an effort to understand the EC clock dynamics in other tissues. Sheeba and collaborators mechanically removed the FGF and Shh sources, Apical Ectodermal Ridge (AER), and Zone of Polarizing Activity (ZPA), respectively, during limb outgrowth. By removing these tissues, *hairy2* cyclic expression is abolished completed from the cyclic domain (Sheeba *et al.*, 2012). However, it is still unknown how losing the EC in the limb impacts the limb outgrowth and patterning. The clock has also been studied in neurogenesis, and in this case, the work in the modulation of EC is even scarcer. Soto and collaborators, by inhibiting mir-9, stabilize the expression noise of *her6.* This could impair proper neural differentiation (Soto *et al.*, 2020).

Table 5 presents a graphical overview of the consequences of EC gene manipulation.

Besides experiments that lead to a total disrupt EC oscillations, there are several successful experiments where it was possible to alter the periodicity of EC oscillations (Table 6). Hirama and colleagues formulated a model which was capable of correctly predicting that deleting all 3 introns of *hes7* would completely disrupt the EC, as previously described by (Takashima *et al.*, 2011). Using this model, they further predicted that the deletion of the first and second introns would lead to an EC pace acceleration of 11 minutes, shortening a cycle of expression from 123 to 112 minutes (Harima *et al.*, 2013).

## Table 5: Resume of the effect of EC mutation in mice and zebrafish

| MOUSE | | NOTCH | | | | | | | | | FGF | | | | | | WNT |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Hes1 | Hes7 | Hey2 | Hey1 | Hes5 | Lfng | Nrarp | NICD | Mesp2 | Sprouty2 | Sprouty4 | Snail1 | Dusp4 | Dusp6 | p-Erk | Axin2 |
| WT | | ~ | ~ | ~ | ~ | ~ | ~ | ~ | ~ | ~ | ~ | ~ | ~ | ~ | ~ | ~ | ~ |
| NOTCH | Hes7 KO | | | | | | | | | | ~ | | ~ | | ~ | | ~ |
| | Hes7 constitutive | | | | | | | | | | | | | | | | |
| | Dll1 KO | | | | | | | | | | | | | ~ | | | ~ |
| | Dll3 KO | | | | | | | | | | | | | | | | |
| | RBPJκ KO | | | | | | | | | | | | | ~ | | | |
| | Lfng KO | | | | | | | | | | | ~ | | ~ | | | |
| | Lfng constitutive | | | | | | | | | | | | | | | | |
| | Psen1-/-; Psen2-/- | | | | | | | | | | | | | | | | |
| FGF | FGFR1 cKO | | | | | | | | | | | | | | | | |
| WNT | Wnt3a KO | | | | | | | | | | | | | | | | |
| | Ctnnb1 KO | | | | | | | | | | | | | | | | |

| ZEBRAFISH | | NOTCH | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | her1 | her7 | dlC | her12 | her15 | her4 | her11 | Mesp2 |
| WT | | ~ | ~ | ~ | ~ | ~ | ~ (?) | ~ | |
| NOTCH | MO her1 | | | | ~ | ~ | | ~ | |
| | her1 mutant | | ~ | | | | | | |
| | MO her7 | | | | | | | | |
| | her7 mutant | | | | | | | | |
| | dlC KO (bae) | | | | | | | | |
| | MO dlC | | | | | | | | |
| | dlD KO (aei) | | | | | | | | |
| | MO hes6 | | | | | | | | |
| | notch1 KO (des) | | | | | | | | |
| | NICD Over activation | | | | | | | | |
| | OE her12 | | | | | | | | |
| | OE her15 | | | | | | | | |
| | MO her12 | | | | | | | | |
| | Su (H) MO | | | | | | | | |
| FGF | | | | | | | | | |
| WNT | tbx6 KO (fss) | | | | | | | | |
| | | | | | | | | | |
| | MO RPTPψ | | | | | | | | |
| | Greb1 MO | | | | | | | | |

| | |
|---|---|
| | No expression |
| | Altered/disrupted dynamics |
| | Down-regulated, altered/disrupted dynamics |
| ~ | Oscillatory expression |
| | data not available |

**Table 6: EC pace manipulation**

| Organism | Manipulation | Inicial_Pace_time | Final_Pace_time | Time_alteration | Reference |
|---|---|---|---|---|---|
| Mouse | Deletion of introns 1 and 2 | 123 | 112 | (-)8,94% | (Harima *et al.,* 2013) |
| | Hes7 K14R mutation (HES7 prot half-life increase from 22 to 30 min) | 121,4 | 131,6 | (+)8,4% | (Hirata *et al.,* 2004) |
| | KO of Nrarp | 106 | 111 | (+)4,5% | (Kim *el al.,* 2011) |
| | LiCl 20 mM treatment | 2,5h | 2,9h | (+)16% | (González *et al.,* 2013) |
| | LiCl 40 mM treatment | 2,5h | 3,6h | (+)44% | (González *et al.,* 2013) |
| | CKI-7 100 mM treatment | 2,5h | 3,3h | (+)32% | (González *et al.,* 2013) |
| | XAV939 1 mM treatment | 2,5h | 2,5h | 0% | (González *et al.,* 2013) |
| Zebrafish | Damascus mutant (~100 copies of DeltaD) | 24.7 ± 0.6 | 23.1 ± 0.8 | (-)6,4% | (Liao *et al.,* 2016) |
| | MO hes6 | N/A | N/A | (-)6.5% ± 1.2% | (Schröter *et al.,* 2010) |
| | mind bomb mutant | N/A | N/A | (+)19% | (Herrgen *et al.,* 2010) |
| | aei/deltaD mutant | N/A | N/A | (+)23% | (Herrgen *et al.,* 2010) |
| | des/notch1a mutant | N/A | N/A | (+)7% | (Herrgen *et al.,* 2010) |
| | Blocking Notch signaling with saturating DAPT concentrations (R 40 mM) | N/A | N/A | (+)18% | (Herrgen *et al.,* 2010) |
| | Her7 hetero:hes6 mutant | N/A | N/A | (+)6% | (Schröter *et al.,* 2012) |
| | Her7 Mutant:hes6 mutant | N/A | N/A | (+)5% | (Schröter *et al.,* 2012) |
| Chicken | Shh inhibition / notochord removal | 90 min | ~2h 45min | (+)~85% | (Resende *et al.,* 2010) |

Similar to a change the mRNA half-life, changing the protein half-life can have an impact on EC. Through replacement of a Lysine for an Arginine in position 14, Hirata and colleagues were able to increase the half-life of the protein HES7 by 10 minutes, which led to an increase of 10,2 minutes in the global pace of the embryo clock (Hirata *et al.*, 2004). Also, in mice, the knock-out of Nrarp, a downstream effector of Notch, led to an EC pace slowdown of 5 min (Kim *et al.*, 2011).

A great effort has also been made using zebrafish, Herrgen and collaborators mutated several genes, leading to the EC slowdown. In their work, the Knock-out of *notch1a*, *mib1*, and *deltaD* slowed down the EC pace in 7, 19, and 23%, respectively. To evidence the importance of the notch signaling pathway, the authors blocked it using DAPT and delayed the EC in 18% (Herrgen *et al.*, 2010). In the same year, by mutating hes6 (gradient gene), Schröter and colleagues were able to slow down the clock in 6.5%. Knowing that the KO of *deltaD* have quite a huge impact on EC, knock-in extra copies of *deltaD* performed the inverse experiment. Liao and colleagues created two fish lines, one with 7 extra copies of *deltaD* and the other with 100 extra copies, *Dover* and *Damascus* mutants, respectively. Curiously, only in the *Damascus* mutant was observed a slight EC pace alteration. The pace in this mutant was 1.6 minutes faster than the wildtype counterpart (Liao *et al.*, 2016).

## Pressing questions and future perspectives

Since it was first described in 1997, the EC has been characterized in several species of vertebrates, making it a widely conserved mechanism. However, there are two major aspects that differ depending on the organism: the pace of the EC and the genes involved in its mechanism, although they participate in the same signalling pathways. What sets the pace of the clock? Is there a core component transversal to all vertebrates?

The EC mechanism seems to be a robust process, although small changes in its machinery dramatically impact the whole mechanism. However, problems with the clock's machinery have more dramatic effects in higher organisms, such as mice and humans. The EC biological function is tightly correlated with the segmentation of paraxial mesoderm, and mutations in EC-related genes give rise to phenotypes associated with spondylocostal dysostosis (Sparrow *et al.*, 2012; Nobrega and Maia-Fernandes *et al.*, 2021). In zebrafish, however, segmentation seems to be less dependent on this mechanism.

In chicken, the time of somite formation and the EC pace during PSM segmentation is conventionally described as 90 minutes. Nevertheless, it is known that the EC pace in the PSM varies during development. During the formation of the last 5 to 8 somites, the pace of the clock increases to 150 min. What makes the pace increase at the end of somitogenesis? Is it

metabolism related? Is it linked to PSM shortage? Another explanation could be the existence of different regulators. Most of the EC genes have more than one polyadenylation site (PAS), and these structures are known to influence the half-life time of the transcripts. Therefore, could different PAS underlie different EC paces?

The EC is not only required for proper PSM segmentation, but it is also important for the patterning and/or differentiation of other tissues. Among these are mouse neural progenitor cell differentiation and the chicken forelimb outgrowth. In the mouse, *Hes1* asynchronous oscillations are essential for the timely differentiation of the NPC into other cell types. However, it is still unknown if there are other genes that display cyclic expression and, if so, do they have the same pace as *Hes1*? Cyclic gene expression in NPC was only reported in mice, so it remains to be determined if it is also observed in other vertebrates during neural development.

Until now, only *hairy2* was described as having a cyclic expression during the chicken limb development. This discovery yields several questions still to be answered. Besides *hairy2*, are there more genes that are cyclically expressed? What is the relevance of the EC in limb patterning? And finally, could it be a conserved mechanism among vertebrates? For now, the information about the EC role in the limb is very scarce.

**Acknowledgments**

This study was supported by the Portuguese Fundação para a Ciência e Tecnologia (FCT) grants PTDC/BEX-BID/5410/2014 to RPA. GC was supported by the FCT scholarship SFRH/BD/101609/2014.

**Author contributions**

G.C., A.P.M-J. and R.P.A. drafted the manuscript.

**Competing interests**

The authors declare no competing interests.

## References

**Apelqvist, A., Li, H., Beatus, P., Anderson, D. J., Honjo, T., Hrabe de Angelis, M., et al.** (1999). Notch signalling controls pancreatic cell differentiation. Nature 400, 877–881.

**Aulehla, A., and Johnson, R. L.** (1999). Dynamic expression of lunatic fringe suggests a link between notch signaling and an autonomous cellular oscillator driving somite segmentation. Dev. Biol. 207, 49–61. Doi:10.1006/dbio.1998.9164.

**Aulehla, A., and Pourquié, O.** (2008). Oscillating signaling pathways during embryonic development. Curr. Opin. Cell Biol. 20, 1–6. Doi:10.1016/j.ceb.2008.09.002.

**Aulehla, A., Wehrle, C., Brand-Saberi, B., Kemler, R., Gossler, A., Kanzler, B., et al.** (2003). Wnt3a Plays a Major Role in the Segmentation Clock Controlling Somitogenesis. Dev. Cell 4, 395–406. Doi:10.1016/s1534-5807(03)00055-8.

**Aulehla, A., Wiegraebe, W., Baubet, V., Wahl, M. B., Deng, C., Taketo, M., et al.** (2008). A β-catenin gradient links the clock and wavefront systems in mouse embryo segmentation. Nat. Cell Biol. 10, 186–193. Doi:10.1038/ncb1679.

**Ay, A., Knierer, S., Sperlea, A., Holland, J., and Özbudak, E. M.** (2013). Short-lived her proteins drive robust synchronized oscillations in the Zebrafish segmentation clock. Development 140, 3244–3253. Doi:10.1242/dev.093278.

**Bessho, Y., Hirata, H., Masamizu, Y., and Kageyama, R**. (2003). Periodic repression by the bHLH factor Hes7 is an essential mechanism for the somite segmentation clock. Genes Dev. 17, 1451–1456. Doi:10.1101/gad.1092303.GENES.

**Bessho, Y., Sakata, R., Komatsu, S., Shiota, K., Yamada, S., and Kageyama, R**. (2001). Dynamic expression and essential functions of Hes7 in somite segmentation. Genes Dev. 15, 2642–2647. Doi:10.1101/gad.930601.2642.

**Bhat, R., Glimm, T., Linde-medina, M., Cui, C., and Newman, S.** A. (2019). Synchronization of Hes1 oscillations coordinates and refines condensation formation and patterning of the avian limb skeleton. Mech. Dev. 156, 41–54. Doi:10.1016/j.mod.2019.03.001.

**Bone, R. A., Bailey, C. S. L., Wiedermann, G., Ferjentsik, Z., Appleton, P. L., Murray, P. J., et al. (**2014). Spatiotemporal oscillations of Notch1, Dll1 and NICD are coordinated across the mouse PSM. Development 141, 4806–4816. Doi:10.1242/dev.115535.

**Bonev, B., Stanley, P., and Papalopulu,** N. (2012). MicroRNA-9 modulates hes1 ultradian oscillations by forming a double-negative feedback loop. Cell Rep. 2, 10–18. Doi:10.1016/j.celrep.2012.05.017.

**Boulet, A. M., and Capecchi, M. R.** (2012). Signaling by FGF4 and FGF8 is required for axial elongation of the mouse embryo. Dev. Biol. 371, 235–245. Doi:10.1016/j.ydbio.2012.08.017.

**Bulusu, V., Prior, N., Snaebjornsson, M. T., Schultz, C., Sauer, U., Aulehla, A., et al.** (2017). Spatiotemporal Analysis of a Glycolytic Activity Gradient Linked to Mouse Embryo Mesoderm Article Spatiotemporal Analysis of a Glycolytic Activity Gradient Linked to Mouse Embryo Mesoderm Development. Dev. Cell 40, 331-341.e1-e4. Doi:10.1016/j.devcel.2017.01.015.

**Chen, J., Kang, L., and Zhang, N**. (2005). Negative Feedback Loop Formed by Lunatic Fringe and Hes7 Controls Their Oscillatory Expression During Somitogenesis. Genesis 43, 196–204. Doi:10.1002/gene.20171.

**Choorapoikayil, S., Willems, B., Ströhle, P., and Gajewski, M**. (2012). Analysis of her1 and her7 Mutants Reveals a Spatio Temporal Separation of the Somite Clock Module. PloS Comput. Biol. 7, 1–7. Doi:10.1371/journal.pone.0039073.

**Chu, L.-F., Mamott, D., Ni, Z., Bacher, R., Liu, C., Swanson, S., et al.** (2019). An In Vitro Human Segmentation Clock Model Derived from Embryonic Stem Cells. Cell Rep. 28, 2247-2255.e5. doi:10.1016/j.celrep.2019.07.090.

**Cole, S. E., Levorse, J. M., Tilghman, S. M., and Vogt, T. F**. (2002). Clock Regulatory Elements Control Cyclic Expression of Lunatic fringe during Somitogenesis. Dev. Cell 3, 75–84.

**Cooke, J., and Zeeman, E. C**. (1976). A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. J. Theor. Biol. 58, 455–476. Doi:10.1016/S0022-5193(76)80131-2.

**Dale, J. K., Malapert, P., Chal, J., Vilhais-Neto, G., Johnson, T., Maroto, M., et al.** (2006). Oscillations of the Snail Genes in the Presomitic Mesoderm Coordinate Segmental Patterning and Morphogenesis in Vertebrate Somitogenesis. Dev. Cell 10, 355–366. Doi:10.1016/j.devcel.2006.02.011.

**Dale, J. K., Maroto, M., Dequeant, M. L., Malapert, P., McGrew, M., and Pourquie, O**. (2003). Periodic Notch inhibition by lunatic fringe underlies the chick segmentation clock. Nature 421, 275–278. Doi:10.1038/nature01244.

**Davis, R. L., Turner, D. L., Evans, L. M., and Kirschner, M. W.** (2001). Molecular Targets of Vertebrate Segmentation: Two Mechanisms Control Segmental Expression of Xenopus hairy2 during Somite Formation. Dev. Cell 1, 553–565. Doi:10.1016/S1534-5807(01)00054-5.

**De Lichtenberg, K. H., Seymour, P. A., Jørgensen, M. C., Kim, Y.-H., Grapin-Botton, A., Magnuson, M. A., et al.** (2018). Notch Controls Multiple Pancreatic Cell Fate Regulators Through Direct Hes1-mediated Repression. bioRxiv, 336305. Doi:10.1101/336305.

**Delaune, E. A., François, P., Shih, N. P., and Amacher, S. L.** (2012). Single-Cell-Resolution Imaging of the Impact of Notch Signaling and Mitosis on Segmentation Clock Dynamics. Dev. Cell 23, 995–1005. Doi:10.1016/j.devcel.2012.09.009.

**Dequéant, M.-L., Glynn, E., Gaudenz, K., Wahl, M., Chen, J., Mushegian, A., et al.** (2006). A Complex Oscillating Network of Signaling Genes Underlies the Mouse Segmentation Clock. Science (80-. ). 314, 1595–1598. Doi:10.1126/science.1133141.

**Diaz-Cuadros, M., Wagner, D., Budjan, C., Hubaud, A., Touboul, J., Michaut, A., et al.** (2018). In vitro characterization of the human segmentation clock. bioRxiv, 461822. Doi:10.1101/461822.

**Elmasri, H., Liedtke, D., Lücking, G., Volff, J. N., Gessler, M., and Winkler, C.** (2004). Her7 and hey1, but not lunatic fringe show dynamic expression during somitogenesis in medaka (Oryzias latipes). Gene Expr. Patterns 4, 553–559. Doi:10.1016/j.modgep.2004.02.003.

**Feller, J., Schneider, A., Schuster-Gossler, K., and Gossler, A**. (2008). Noncyclic Notch activity in the presomitic mesoderm demonstrates uncoupling of somite compartmentalization and boundary formation. Genes Dev. 22, 2166–2171. Doi:10.1101/gad.480408.

**Forsberg, H., Crozet, F., and Brown, N. A.** (1998). Waves of mouse Lunatic fringe expression, in four-hour cycles at two-hour intervals, precede somite boundary formation. Curr. Biol. 8, 1027–1030. Doi:10.1016/S0960-9822(07)00424-1.

**Fujimuro, T., Matsui, T., Nitanda, Y., Matta, T., Sakumura, Y., Saito, M., et al.** (2014). Hes7 3'UTR is required for somite segmentation function. Sci. Rep. 4, 1–9. Doi:10.1038/srep06462.

**Fujino, Y., Yamada, K., Sugaya, C., Ooka, Y., Ovara, H., Ban, H., et al.** (2018). Deadenylation by the CCR4-NOT complex contributes to the turnover of hairy-related mRNAs in the zebrafish segmentation clock. FEBS Lett. 592, 3388–3398. Doi:10.1002/1873-3468.13261.

**Gajewski, M., Elmasri, H., Girschick, M., Sieger, D., and Winkler, C.** (2006). Comparative analysis of her genes during fish somitogenesis suggests a mouse/chick-like mode of oscillation in medaka. Dev. Genes Evol. 216, 315–332. Doi:10.1007/s00427-006-0059-6.

**Gajewski, M., Sieger, D., Alt, B., Leve, C., Hans, S., Wolff, C., et al.** (2003). Anterior and posterior waves of cyclic her1 gene expression are differentially regulated in the presomitic mesoderm of zebrafish. Development 130, 4269–4278. Doi:10.1242/dev.00627.

**Giudicelli, F., Özbudak, E. M., Wright, G. J., and Lewis, J.** (2007). Setting the Tempo in Development: An Investigation of the Zebrafish Somite Clock Mechanism. PloS Biol. 5, 1309–1323. Doi:10.1371/journal.pbio.0050150.

**Gomez, C., Özbudak, E. M., Wunderlich, J., Baumann, D., Lewis, J., and Pourquié, O.** (2008). Control of segment number in vertebrate embryos. Nature 454, 335–339. Doi:10.1038/nature07020.

**González, A., Manosalva, I., Liu, T., and Kageyama, R.** (2013). Control of Hes7 Expression by Tbx6, the Wnt Pathway and the Chemical Gsk3 Inhibitor LiCl in the Mouse Segmentation Clock. PLOS ONE 8, e53323. 10.1371/journal.pone.0053323.

**Hanisch, A., Holder, M. V, Choorapoikayil, S., Gajewski, M., Özbudak, E. M., and Lewis, J.** (2013). The elongation rate of RNA polymerase II in zebrafish and its significance in the somite segmentation clock. Development 453, 444–453. Doi:10.1242/dev.077230.

**Harima, Y., Takashima, Y., Ueda, Y., Ohtsuka, T., and Kageyama, R.** (2013). Accelerating the tempo of the segmentation clock by reducing the number of introns in the Hes7 gene. Cell Rep 3, 1-7. 10.1016/j.celrep.2012.11.012.

**Hayashi, S., Shimoda, T., Nakajima, M., Tsukada, Y., Sakumura, Y., Dale, J. K., et al.** (2009). Sprouty4, an FGF inhibitor, Displays Cyclic Gene Expression under the Control of the Notch Segmentation Clock in the Mouse PSM. PloS One 4, e5603. Doi:10.1371/journal.pone.0005603.

**Henry, C.A., Urban, M.K., Dill, K.K., Merlie, J.P., Page, M.F., Kimmel, C.B., and Amacher, S.L.** (2002). Two linked hairy/Enhancer of split-related zebrafish genes, her1 and her7, function together to refine alternating somite boundaries. Development 129, 3693-3704.

**Herrgen, L., Ares, S., Morelli, L.G., Schroter, C., Julicher, F., and Oates, A.C.** (2010). Intercellular coupling regulates the period of the segmentation clock. Curr Biol 20, 1244-1253. 10.1016/j.cub.2010.06.034.

**Hilgers, V., Pourquié, O., and Dubrulle, J.** (2005). In vivo analysis of mRNA stability using the Tet-Off system in the chicken embryo. Dev. Biol. 284, 292–300.

**Hirata, H., Bessho, Y., Kokubu, H., Masamizu, Y., Yamada, S., Lewis, J., et al.** (2004). Instability of Hes7 protein is crucial for the somite segmentation clock. Nat. Genet. 36, 750–754. Doi:10.1038/ng1372.

**Hirata, H., Yoshiura, S., Ohtsuka, T., Bessho, Y., Harada, T., Yoshikawa, K., et al.** (2002). Oscillatory expression of the BHLH factor Hes1 regulated by a negative feedback loop. Science (80-. ). 298, 840–843. Doi:10.1126/science.1074560.

**Holley, S. A., Geisler, R., and Nüsslein-Volhard, C.** (2000). Control of her1 expression during zebrafish somitogenesis by a Delta-dependent oscillator and an independent wave-front activity. Genes Dev. 14, 1678–1690.

**Horikawa, K., Ishimatsu, K., Yoshimoto, E., Kondo, S., and Takeda, H.** (2006). Noise-resistant and synchronized oscillation of the segmentation clock. Nature 441, 719–723. Doi:10.1038/nature04861.

**Hoyle, N. P., and Ish-Horowicz, D.** (2013). Transcript processing and export kinetics are rate-limiting steps in expressing vertebrate segmentation clock genes. Proc. Natl. Acad. Sci. 110, E4316 LP-E4324. Doi:10.1073/pnas.1308811110.

**Huppert, S. S., Ilagan, M. X. G., De Strooper, B., and Kopan, R.** (2005). Analysis of Notch Function in Presomitic Mesoderm Suggests a γ-Secretase-Independent Role for Presenilins in Somite Differentiation. Dev. Cell 8, 677–688. Doi:10.1016/j.devcel.2005.02.019.

**Imayoshi, I., Isomura, A., Harima, Y., Kawaguchi, K., Kori, H., Miyachi, H., et al.** (2013). Oscillatory control of factors determining multipotency and fate in mouse neural progenitors. Science (80-. ). 342, 1203–1208. Doi:10.1126/science.1242366.

**Ishii, A., Kobayashi, T., and Kageyama, R.** (2008). Requirement of multiple lysine residues for the transcriptional activity and the instability of Hes7. Biochem. Biophys. Res. Commun. 372, 142–146. Doi:10.1016/j.bbrc.2008.05.015.

**Ishikawa, A., Kitajima, S., Takahashi, Y., Kokubo, H., Kanno, J., Inoue, T., et al**. (2004). Mouse Nkd1, a Wnt antagonist, exhibits oscillatory gene expression in the PSM under the control of Notch signaling. Mech. Dev. 121, 1443–1453.

**Jensen, J., Pedersen, E. E., Galante, P., Hald, J., Heller, R. S., Ishibashi, M., et al.** (2000). Control of endodermal endocrine development by Hes-1. Nat. Genet. 24, 36–44.

**Jensen, M. H., Sneppen, K., and Tiana, G.** (2003). Sustained oscillations and time delays in gene expression of protein Hes1. FEBS Lett. 541, 176–177. Doi:10.1016/S0014-5793(03)00279-5.

**Jiang, Y.-J., Aerne, B. L., Smithers, L., Haddon, C., Ish-Horowicz, D., and Lewis, J.** (2000). Notch signalling and the synchronization of the somite segmentation clock. Nature 408, 475–479. Doi:10.1038/35044091.

**Jing, B., Yuan, J., Yin, Z., Lv, C., Lu, S., Xiong, H., et al.** (2015). Dynamic properties of the segmentation clock mediated by microRNA. Int. J. Clin. Exp. Pathol. 8, 196–206.

**Jouve, C., Iimura, T., and Pourquié, O**. (2002). Onset of the segmentation clock in the chick embryo: evidence for oscillations in the somite precursors in the primitive streak. Development 129, 1107–17. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11874907.

**Jouve, C., Palmeirim, I., Henrique, D., Beckers, J., Gossler, A., Ish-Horowicz, D., et al.** (2000). Notch signalling is required for cyclic expression of the hairy-like gene HES1 in the presomitic mesoderm. Development 127, 1421–1429.

**Kawamura, A., Koshida, S., Hijikata, H., Sakaguchi, T., Kondoh, H., and Takada, S.** (2005). Zebrafish Hairy/Enhancer of split protein links FGF signaling to cyclic gene expression in the periodic segmentation of somites. Genes Dev. 19, 1156–1161. Doi:10.1101/gad.1291205.mesoderm.

**Kawamura, A., Ovara, H., Ooka, Y., Kinoshita, H., Hoshikawa, M., Nakajo, K., et al.** (2016). Posterior-anterior gradient of zebrafish hes6 expression in the presomitic mesoderm is established by the combinatorial functions of the downstream enhancer and 3'UTR. Dev. Biol. 409, 543–554. Doi:10.1016/j.ydbio.2015.11.010.

**Kim, W., Matsui, T., Yamao, M., Ishibashi, M., Tamada, K., Takumi, T., Kohno, K., Oba, S., Ishii, S., Sakumura, Y., and Bessho, Y.** (2011). The period of the somite segmentation clock is sensitive to Notch activity. Mol Biol Cell 22, 3541-3549. 10.1091/mbc.E11-02-0139.

**Kobayashi, T., Mizuno, H., Imayoshi, I., Furusawa, C., Shirahige, K., and Kageyama, R.** (2009). The cyclic gene Hes1 contributes to diverse differentiation responses of mouse embryonic stem cells. Gene Dev. 23, 1870–1875. Doi:10.1016/j.neures.2010.07.2164.

**Krol, A. J., Roelling, D., Dequéant, M.-L., Tassy, O., Glynn, E., Hattem, G., et al.** (2011). Evolutionary plasticity of segmentation clock networks. Development 138, 2783–2792.

**Lee, J. C., Smith, S. B., Watada, H., Lin, J., Scheel, D., Wang, J., et al.** (2001). Regulation of the Pancreatic Pro-Endocrine Gene. Diabetes 50, 3–11.

**Leimeister, C., Dale, K., Fischer, A., Klamt, B., Hrabe de Angelis, M., Radtke, F., et al.** (2000a). Oscillating Expression of c-Hey2 in the Presomitic Mesoderm Suggests That the Segmentation Clock Mey Use Combinatorial Signalling through Multiple Interacting bHLH Factors. Dev. Biol. 227, 91–103.

**Leimeister, C., Schumacher, N., Steidl, C., and Gessler, M.** (2000b). Analysis of HeyL expression in wild-type and Notch pathway mutant mouse embryos. Mech. Dev. 98, 175–178. Doi:10.1016/S0925-4773(00)00459-7.

**Lewis, J.** (2003). Autoinhibition with Transcriptional Delay: A Simple Mechanism for the Zebrafish Somitogenesis Oscillator. Curr. Biol. 13, 1398–1408. Doi:10.1016/S0960-9822(03)00534-7.

**Li, Y., Fenger, U., Niehrs, C., and Pollet, N.** (2003). Cyclic expression of esr9 gene in Xenopus presomitic mesoderm. Differentiation 71, 83–89. Doi:10.1046/j.1432-0436.2003.700608.x.

**Liao, B.K., Jorg, D.J., and Oates, A.C.** (2016). Faster embryonic segmentation through elevated Delta-Notch signalling. 7, 11861. 10.1038/ncomms11861.

**Lleras Forero, L., Narayanan, R., Huitema, L.F., VanBergen, M., Apschner, A., Peterson-Maduro, J., Logister, I., Valentin, G., and Morelli, L.G.** (2018). Segmentation of the zebrafish axial skeleton relies on notochord sheath cells and not on the segmentation clock. 7. 10.7554/eLife.33843.

**Manning, C. S., Biga, V., Boyd, J., Kursawe, J., Ymisson, B., Spiller, D. G., et al.** (2019). Quantitative single-cell live imaging links HES5 dynamics with cell-state and fate in murine neurogenesis. Nat. Commun. 10. Doi:10.1038/s41467-019-10734-8.

**Mara, A., Schroeder, J., Chalouni, C., and Holley, S. A.** (2007). Priming, initiation and synchronization of the segmentation clock by deltaD and deltaC. Nat. Cell Biol. 9, 523–530. Doi:10.1038/ncb1578.

**Maroto, M., Dale, J. K., Dequéant, M. L., Petit, A. C., and Pourquié, O.** (2005). Synchronised cycling gene oscillations in presomitic mesoderm cells require cell-cell contact. Int. J. Dev. Biol. 49, 309–315. Doi:10.1387/ijdb.041958mm.

**Maruhashi, M., Van De Putte, T., Huylebroeck, D., Kondoh, H., and Higashi, Y.** (2005). Involvement of SIP1 in positioning of somite boundaries in the house embryo. Dev. Dyn. 234, 332–338. Doi:10.1002/dvdy.20546.

**Masamizu, Y., Ohtsuka, T., Takashima, Y., Nagahara, H., Takenaka, Y., Yoshikawa, K., et al.** (2006). Real-time imaging of the somite segmentation clock: Revelation of unstable oscillators in the individual presomitic mesoderm cells. Proc. Natl. Acad. Sci. 103, 1313–1318. Doi:10.1073/pnas.0508658103.

**Mastromina, I., Verrier, L., Silva, J. C., Storey, K. G., and Dale, J. K.** (2018). Myc activity is required for maintenance of the neuromesodermal progenitor signalling network and for segmentation clock gene oscillations in mouse. Development 145. Doi:10.1242/dev.161091.

**Matsuda, M., Hayashi, H., Garcia-Ojalvo, J., Yoshioka-Kobayashi, K., Kageyama, R., Yamanaka, Y., et al**. (2019a). Species-specific oscillation periods of human and mouse segmentation clocks are due to cell autonomous differences in biochemical reaction parameters. bioRxiv, 650648. Doi:10.1101/650648.

**Matsuda, M., Yamanaka, Y., Uemura, M., Osawa, M., Saito, M. K., Nagahashi, A., et al.** (2019b). Modeling the Human Segmentation Clock with Pluripotent Stem Cells. bioRxiv, 562447. Doi:10.1101/562447.

**Matsumiya, M., Tomita, T., Yoshioka-Kobayashi, K., Isomura, A., and Kageyama, R.** (2018). ES cell-derived presomitic mesoderm-like tissues for analysis of synchronized oscillations in the segmentation clock. Development 145. Doi:10.1242/dev.156836.

**McGrew, M. J., Dale, J. K., Fraboulet, S., and Pourquié, O**. (1998). The lunatic Fringe gene is a target of the molecular clock linked to somite segmentation in avian embryos. Curr. Biol. 8, 979–982. Doi:10.1016/S0960-9822(98)70401-4.

**Menkes, B., Miclea, C., Elias, S., and Deleanu, M.** (1961). Researches on the formation of axial organs. I. Studies on the differentiation of somites. Stud. Cercet. Stiint. Med. 8.

**Monk, N. A. M.** (2003). Oscillatory Expression of Hes1, p53, and NF-κB Driven by Transcriptional Time Delays. Curr. Biol. 13, 1409–1413. Doi:10.1016/S0960-9822(03)00494-9.

**Morimoto, M., Takahashi, Y., Endo, M., and Saga, Y**. (2005). The Mesp2 transcription factor establishes segmental borders by suppressing Notch activity. Nature 435, 354–359. Doi:10.1038/nature03591.

**Müller, F., and O'Rahilly, R.** (1986). Somitic-vertebral correlation and vertebral levels in the human embryo. Am. J. Anat. 177, 3–19. Doi:10.1002/aja.1001770103.

**Nagai, H., Mak, S. S., Weng, W., Nakaya, Y., Ladher, R., and Sheng, G**. (2011). Embryonic development of the emu, Dromaius novaehollandiae. Dev. Dyn. 240, 162–175. Doi:10.1002/dvdy.22520.

**Naiche, L. A., Holder, N., and Lewandoski, M. (**2011). FGF4 and FGF8 comprise the wavefront activity that controls somitogenesis. Proc. Natl. Acad. Sci. U. S. A. 108, 4018–4023. Doi:10.1073/pnas.1007417108.

**Nakayama, K., Satoh, T., Igari, A., Kageyama, R., and Nishida, E**. (2008). FGF induces oscillations of Hes1 expression and Ras/ERK activation. Curr. Biol. 18, 332–334. Doi:10.1016/j.cub.2008.03.013.

**Nitanda, Y., Matsui, T., Matta, T., Higami, A., Kohno, K., Nakahata, Y., et al**. (2014). 3′-UTR-dependent regulation of mRNA turnover is critical for differential distribution patterns of cyclic gene mRNAs. FEBS J. 281, 146–156. Doi:10.1111/febs.12582.

**Niwa, Y., Masamizu, Y., Liu, T., Nakayama, R., Deng, C. X., and Kageyama, R.** (2007). The Initiation and Propagation of Hes7 Oscillation Are Cooperatively Regulated by Fgf and Notch Signaling in the Somite Segmentation Clock. Dev. Cell 13, 298–304. Doi:10.1016/j.devcel.2007.07.013.

**Niwa, Y., Shimojo, H., Isomura, A., González, A., Miyachi, H., and Kageyama, R**. (2011). Different types of oscillations in notch and Fgf signaling regulate the spatiotemporal periodicity of somitogenesis. Genes Dev. 25, 1115–1120. Doi:10.1101/gad.2035311.

**Nobrega, A., Maia-Fernandes, A.C., and Andrade, R.P.** (2021). Altered Cogs of the Clock: Insights into the Embryonic Etiology of Spondylocostal Dysostosis. J Dev Biol 9. 10.3390/jdb9010005.

**Oates, A. C., and Ho, R. K.** (2002). Hairy/E(spl)-related (Her) genes are central components of the segmentation oscillator and display redundancy with the Delta/Notch signaling pathway in the formation of anterior segmental boundaries in the zebrafish. Development 129, 2929–2946. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12050140.

**Oates, A. C., Morelli, L. G., and Ares, S.** (2012). Patterning embryos with oscillations: structure, function and dynamics of the vertebrate segmentation clock. Development 139, 625–639. Doi:10.1242/dev.063735.

**Oginuma, M., Moncuquet, P., Xiong, F., Karoly, E., Guevorkian, K., and Pourquie, O.** (2017). A Gradient of Glycolytic Activity Coordinates FGF and Wnt Signaling during Elongation of the Body Axis

in Amniote Embryos Article A Gradient of Glycolytic Activity Coordinates FGF and Wnt Signaling during Elongation of the Body Axis in Amniote Emb. Dev. Cell 40, 342–353. Doi:10.1016/j.devcel.2017.02.001.

**Özbudak, E. M., and Lewis, J.** (2008). Notch signalling synchronizes the zebrafish segmentation clock but is not needed to create somite boundaries. PloS Genet. 4. Doi:10.1371/journal.pgen.0040015.

**Özbudak, E. M., Tassy, O., and Pourquié, O.** (2010). Spatiotemporal compartmentalization of key physiological processes during muscle precursor differentiation. PNAS 107, 4224–4229. Doi:10.1073/pnas.0909375107.

**Packard, D. S. J.** (1980). Somitogenesis in cultured embryos of the Japanese quail, Coturnix coturnix japonica. Am. J. Anat. 158, 83–91.

**Palmeirim, I., Henrique, D., Ish-Horowicz, D., and Pourquié, O.** (1997). Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. Cell 91, 639–648. Doi:10.1016/S0092-8674(00)80451-1.

**Pascoal, S., Carvalho, C. R., Rodriguez-León, J., Delfini, M. C., Duprez, D., Thorsteinsdóttir, S., et al.** (2007). A Molecular Clock Operates During Chick Autopod Proximal-distal Outgrowth. J. Mol. Biol. 368, 303–309. Doi:10.1016/j.jmb.2007.01.089.

**Pourquié, O.** (2004). The chick embryo: A leading model in somitogenesis studies. Mech. Dev. 121, 1069–1079. Doi:10.1016/j.mod.2004.05.002.

**Prajapati, R.S., Mitter, R., Vezzaro, A., and Ish-Horowicz, D.** (2020). Greb1 is required for axial elongation and segmentation in vertebrate embryos. Biology Open 9. 10.1242/bio.047290.

**Resende, T.P., Ferreira, M., Teillet, M.-A., Tavares, A.T., Andrade, R.P., and Palmeirim, I.** (2010). Sonic hedgehog in temporal control of somite formation. Proceedings of the National Academy of Sciences 107, 12907-12912. 10.1073/pnas.1000979107.

**Riley, M. F., Bochter, M. S., Wahi, K., Nuovo, G. J., and Cole, S. E**. (2013). Mir-125a-5p-Mediated Regulation of Lfng Is Essential for the Avian Segmentation Clock. Dev. Cell 24, 554–561. Doi:10.1016/j.devcel.2013.01.024.

**Rodrigues, S., Santos, J., and Palmeirim, I.** (2006). Molecular characterization of the rostral-most somites in early 51omatic stages of the chick embryo. Gene Expr. Patterns 6, 673–677. Doi:10.1016/j.modgep.2006.01.004.

**Schröter, C., Ares, S., Morelli, L. G., Isakova, A., Hens, K., Soroldoni, D., et al.** (2012). Topology and dynamics of the zebrafish segmentation clock core circuit. PloS Biol. 10, 11. Doi:10.1371/journal.pbio.1001364.

**Schroter, C., and Oates, A.C**. (2010). Segment number and axial identity in a segmentation clock period mutant. Curr Biol 20, 1254-1258. 10.1016/j.cub.2010.05.071.

**Sewell, W., Sparrow, D. B., Smith, A. J., Gonzalez, D. M., Rappaport, F., Dunwoodie, S. L., et al.** (2009). Cyclical expression of the Notch/Wnt regulator Nrarp requires modulation by Dll3 in somitogenesis. Dev. Biol. 329, 400–409. Doi:10.1016/j.ydbio.2009.02.023.Cyclical.

**Seymour, P. A., Collin, C. A., Egeskov-Madsen, A. I. R., Jørgensen, M. C., Shimojo, H., Imayoshi, I., et al.** (2019). Jag1 modulates an oscillatory Dll1-Notch-Hes1 signaling module to coordinate growth and fate of pancreatic progenitors. bioRxiv, 336529. Doi:10.1101/336529.

**Shankaran, S. S., Sieger, D., Schröter, C., Czepe, C., Pauly, M. C., Laplante, M. A., et al.** (2007). Completing the set of h/E(spl) cyclic genes in zebrafish: her12 and her15 reveal novel modes of expression and contribute to the segmentation clock. Dev. Biol. 304, 615–632. Doi:10.1016/j.ydbio.2007.01.004.

**Sheeba, C.J., Andrade, R.P., and Palmeirim, I.** (2012). Joint interpretation of AER/FGF and ZPA/SHH over time and space underlies hairy2 expression in the chick limb. Biol Open 1, 1102-1110. 10.1242/bio.20122386.

**Shih, N. P., Francois, P., Delaune, E. A., and Amacher, S. L.** (2015). Dynamics of the slowing segmentation clock reveal alternating two-segment periodicity. Development 142, 1785–1793. Doi:10.1242/dev.119057.

**Shimojo, H., Isomura, A., Ohtsuka, T., Kori, H., Miyachi, H., and Kageyama, R.** (2016). Oscillatory control of Delta-like1 in cell interactions regulates dynamic gene expression and tissue morphogenesis. Genes Dev. 30, 102–116. Doi:10.1101/gad.270785.115.

**Shimojo, H., Ohtsuka, T., and Kageyama, R**. (2008). Oscillations in Notch signalling regulate the maintenance of neural progenitors. Neuron 58, 52–64.

**Sieger, D., Tautz, D., and Gajewski, M**. (2004). Her11 is involved in the somitogenesis clock in zebrafish. Dev. Genes Evol. 214, 393–406. Doi:10.1007/s00427-004-0427-z.

**Sonnen, K. F., Lauschke, V. M., Uraji, J., Falk, H. J., Petersen, Y., Funk, M. C., et al.** (2018). Modulation of Phase Shift between Wnt and Notch Signaling Oscillations Controls Mesoderm Segmentation. Cell 172, 1079-1090.e12. doi:10.1016/j.cell.2018.01.026.

**Soroldoni, D., Jörg, D. J., Morelli, L. G., Richmond, D. L., Schindelin, J., Jülicher, F., et al.** (2014). A Doppler effect in embryonic pattern formation. Science (80-. ). 345, 222–225.

**Soto, X., Biga, V., Kursawe, J., Lea, R., Doostdar, P., Thomas, R., and Papalopulu, N.** (2020). Dynamic properties of noise and Her6 levels are optimized by miR-9, allowing the decoding of the Her6 oscillator. The EMBO Journal 39, e103558. https://doi.org/10.15252/embj.2019103558.

**Soza-Ried, C., Ozturk, E., Ish-Horowicz, D., and Lewis, J.** (2014). Pulses of Notch activation synchronise oscillating somite cells and entrain the zebrafish segmentation clock. Development 141, 1780–1788. Doi:10.1242/dev.102111.

**Sparrow, D.B., Chapman, G., Smith, A.J., Mattar, M.Z., Major, J.A., O'Reilly, V.C., Saga, Y., Zackai, E.H., Dormans, J.P., Alman, B.A., et al.** (2012). A mechanism for gene-environment interaction in the etiology of congenital scoliosis. Cell 149, 295-306.

**Stickney, H. L., Barresi, M. J. F., and Devoto, S. H.** (2000). A PEER REVIEWED FORUM Somite Development in Zebrafish. Dev. Dyn. 303, 287–303. Doi:10.1002/1097-0177(2000)9999:9999<::AID-DVDY1065>3.0.CO;2-A.

**Suriben, R., Fisher, D. A., and Cheyette, B. N. R.** (2006). Dact1 Presomitic Mesoderm Expression Oscillates in Phase with Axin2 in the Somitogenesis Clock of Mice. Dev. Dyn. 235, 3177–3183. Doi:10.1002/dvdy.20968.

**Takahashi, Y., Koizumi, K. I., Takagi, A., Kitajima, S., Inoue, T., Koseki, H., et al.** (2000). Mesp2 initiates somite segmentation through the Notch signalling pathway. Nat. Genet. 25, 390–396. Doi:10.1038/78062.

**Takashima, Y., Ohtsuka, T., González, A., Miyachi, H., and Kageyama, R.** (2011). Intronic delay is essential for oscillatory expression in the segmentation clock. Proc. Natl. Acad. Sci. 108, 3300 LP – 3305. Doi:10.1073/pnas.1014418108.

**Tam, P. P. L.** (1981). The control of somitogenesis in mouse embryos. J. Embryol. Exp. Morphol. 65 Suppl, 103–28. Available at: http://www.ncbi.nlm.nih.gov/pubmed/6801176.

**Tenin, G., Right, D., Ferjentsik, Z., Bone, R., McGrew, M. J., and Maroto, M.** (2010). The chick somitogenesis oscillator is arrested before all paraxial mesoderm is segmented into somites. BMC Dev. Biol. 10. Doi:10.1186/1471-213x-10-24.

**Tsiairis, C. D., and Aulehla, A.** (2016). Self-Organization of Embryonic Genetic Oscillators into Spatiotemporal Wave Patterns. Cell 164, 656–667. Doi:10.1016/j.cell.2016.01.028.

**Vasiliauskas, D., Laufer, E., and Stern, C. D.** (2003). A role for hairy1 in regulating chick limb bud growth. Dev. Biol. 262, 94–106. Doi:10.1016/S0012-1606(03)00360-9.

**Ventre, S., Indrieri, A., Fracassi, C., Franco, B., Conte, I., Cardone, L., and di Bernardo, D.** (2015). Metabolic regulation of the ultradian oscillator Hes1 by reactive oxygen species. J Mol Biol 427, 1887-1902. 10.1016/j.jmb.2015.03.007.

**Wahi, K., Friesen, S., Coppola, V., and Cole, S. E.** (2017). Putative binding sites for mir-125 family miRNAs in the mouse Lfng 3′UTR affect transcript expression in the segmentation clock, but mir-125a-5p is dispensable for normal somitogenesis. Dev. Dyn. 246, 740–748. Doi:10.1002/dvdy.24552.

**Wahl, M. B., Deng, C., Lewandowski, M., and Pourquié, O.** (2007). FGF signaling acts upstream of the NOTCH and WNT signaling pathways to control segmentation clock oscillations in mouse somitogenesis. Development 134, 4033–4041. Doi:10.1242/dev.009167.

**Wanglar, C., Takahashi, J., Yabe, T., and Takada, S.** (2014). Tbx Protein Level Critical for Clock-Mediated Somite Positioning Is Regulated through Interaction Between Tbx and Ripply. PloS One 9, 1–12. Doi:10.1371/journal.pone.0107928.

**Webb, A. B., Lengyel, I. M., Jörg, D. J., Valentin, G., Jülicher, F., Morelli, L. G., et al.** (2016). Persistence, period and precision of autonomous cellular oscillators from the zebrafish segmentation clock. Elife 5, 1–17. Doi:10.7554/eLife.08438.

**Wienholds, E., Kloosterman, W. P., Miska, E., Alvarez-Saavedra, E., Berezikov, E., de Bruijn, E., et al.** (2005). MicroRNA expression in zebrafish embryonic development. Science 309, 310–311. Doi:10.1126/science.1114519.

**William, D. A., Saitta, B., Gibson, J. D., Traas, J., Markov, V., Gonzalez, D. M., et al.** (2007). Identification of Oscillatory Genes in Somitogenesis from Functional Genomic Analysis of a Human Mesenchymal Stem Cell Model. Dev. Biol. 305, 172–186. Doi:10.3899/jrheum.121180.Response.

**Williams, D.R., Shifley, E.T., Braunreiter, K.M., and Cole, S.E.** (2016). Disruption of somitogenesis by a novel dominant allele of Lfng suggests important roles for protein processing and secretion. Development 143, 822-830. 10.1242/dev.128538.

**Winkler, C., Elmasri, H., Klamt, B., Volff, J.-N., and Gessler, M**. (2003). Characterization of hey bHLH genes in teleost fish. Dev. Genes Evol. 213, 541–553. Doi:10.1007/s00427-003-0360-6.

**Wong, K.-L., Akiyama, R., Bessho, Y., and Matsui, T.** (2019). ERK Activity Dynamics during Zebrafish Embryonic Development. Int. J. Mol. Sci. 20. Doi:10.3390/ijms20010109.

**Wright, D., Ferjentsik, Z., Chong, S.-W., Qiu, X., Yun-Jin, J., Malapert, P., et al.** (2009). Cyclic Nrarp mRNA Expression Is Regulated by the Somitic Oscillator but Nrarp Protein Levels Do Not Oscillate. Dev. Dyn. 238, 3043–3055. Doi:10.1002/dvdy.22139.

**Xie, Z. R., Yang, H. Te, Liu, W. C., and Hwang, M. J.** (2007). The role of microRNA in the delayed negative feedback regulation of gene expression. Biochem. Biophys. Res. Commun. 358, 722–726. Doi:10.1016/j.bbrc.2007.04.207.

**Yoshiura, S., Ohtsuka, T., Takenaka, Y., Nagahara, H., Yoshikawa, K., and Kageyama, R**. (2007). Ultradian oscillations of Stat, Smad, and Hes1 expression in response to serum. PNAS 104, 11292–11297.

## 1.5 RNA-dependent regulation of the Embryo Clock

Ribonucleic acids (RNA) are polymeric molecules composed of four nitrogenous monomers: adenosine, cytidine, guanosine, and uridine. RNA has critical roles in cell and organism homeostasis. One of the best-known functions is to be the intermediate molecule between DNA and protein, through messenger RNAs (mRNA). The mRNA is a transient molecule with an average length of 2000 to 2500 nucleotides that is transcribed from the DNA and is translated into a protein in the cytoplasm (Ota *et al.*, 2004). Because of its transient nature, mRNA half-life is critical for tuning the amount of protein produced by a given gene in the cell.

Besides the protein-coding sequence, the mRNA is composed of four more regulatory regions, the 5'Cap, 5'Utranslated Region (5'UTR), the 3'Untranslated Region (3'UTR), and the Poly(A) tail (Figure 1.17) (Hajj and Whitehead, 2017). The 5'Cap, also known as 7-methylguanosine Cap ($m^7G$), corresponds to a methylated guanosine located at the 5'-most end of the mRNA molecule (Shatkin, 1976). The 5'Cap protects the mRNA against degradation by 5'endonucleases (Evdokimova *et al.*, 2001). Moreover, it binds to the eukaryotic Initiation Factor 4F (eIF4F) complex and helps to initiate translation (Haghighat and Sonenberg, 1997). The 5'UTR has several regulatory functions, especially related to the translation of the coding sequence (Leppek *et al.*, 2018). It harbors specific sequences where the eIF4F complex binds and consequently unwinds the mRNA and recruits ribosomes to initiate translation (Mignone *et al.*, 2002). Another untranslated region, the 3'UTR, is an AU-rich sequence that is especially important for mRNA stability. It harbors several sequences that are targeted by RNA-biding proteins (RBP) that contribute to promote RNA degradation or stabilization. Moreover, it can also have microRNA (miRNA) target sites, which are related to mRNA decay (Bartel, 2004; Behm-Ansmant *et al.*, 2006; Chen and Shyu, 1995; Moor *et al.*, 2005). The last mRNA regulatory region is the Poly(A) tail. For instance, in humans an average of 250 to 300 adenosines are added (Elkon *et al.*, 2013). The Poly(A) tail promotes mRNA stability; also, it is implicated in the mRNA nuclear export and regulates translation. The Poly(A) tail is degraded over time, and when it gets to a critical short length, it leads to mRNA degradation (Guhaniyogi and Brewer, 2001).

The "central dogma" of Biology states that one gene codifies for a mRNA, which is translated into a protein. However, the reality is much more complex: each DNA locus can be transcribed into several different mature RNAs (due to the use of alternative transcription start sites (TSS), alternative 3'UTRs, and alternative splicing, for example), which can result in different proteins and even different functional RNAs, with potentially different functions (Shabalina *et al.*, 2010).

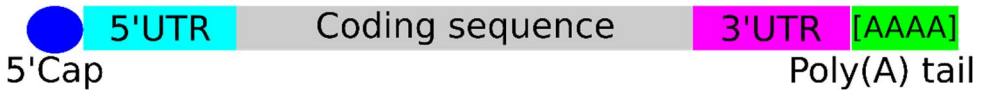



**Figure 1.17 - Schematics of the mRNA structure.**
A mRNA molecule is divided into five regions, the 5'Cap, 5'UTR, the coding sequence, the 3'UTR and a poly-A tail.

### 1.5.1 Alternative 3' untranslated regions (3'UTRs) in Embryo Clock regulation

The 3'UTR has a variety of crucial roles for mRNA regulation. It can control mRNA stability, mRNA localization, function, and translation efficiency (Reviewed in (Elkon *et al.*, 2013)). It is estimated that 70% to 75% of all human genes present more than one polyadenylation site, originating multiple RNA species with different 3'UTRs (Shi, 2012).

There are four different types of Alternative Polyadenylation (APA) sites, with varying impact on mRNA usage and/or protein production (Figure 1.18). The first and simplest APA is the one present inside the 3'UTR region (tandem 3'UTR APA). In this case, only the untranslated region is altered, not affecting the potential protein translation. This type of alternative 3'UTR is tightly correlated with mRNA stability. The lengthening of the 3'UTR was correlated with senescent cells, which leads to gene expression downregulation. In contrast, younger dividing cells preferentially employ more proximal APA, leading to an upregulation of these genes (Chen *et al.*, 2018). Longer 3'UTRs potentially harbor more cis-regulatory elements, and, consequently, mRNAs with longer 3'UTRs are more tightly regulated than mRNAs with shorter 3'UTR (Reviewed in (Elkon *et al.*, 2013)). RNA with shorter 3'UTRs are correlated with diseases such as cancer due to the lack of regulation (Mayr and Bartel, 2009; Xia *et al.*, 2014). Other types of APA change the translation potential of a given mRNA. By alternative splicing, the alternative exon could have a poly(A) signal, or it could be inside of a non-spliced intron. In both cases, it could give rise to smaller proteins lacking the more distal domains and with different functions and/or cellular localization. The last type of APA is inside of an exon. As the intron APA, this one also generates smaller mRNAs that could be translated into smaller proteins, with different functions and cellular localization (Reviewed in (Elkon *et al.*, 2013)).

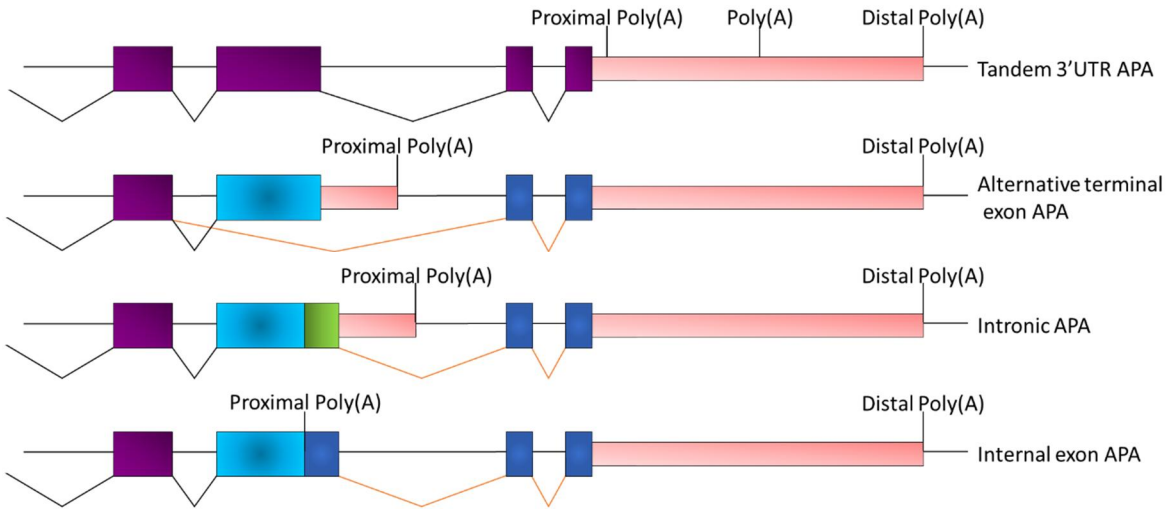



**Figure 1.18 - Schematics of different Alternative Polyadenylation (APA).**
There are four types of APA sites. Tandem 3'UTR APA, Alternative terminal exon APA, Intronic APA and Internal exon APA. In purple are represented common exons; in light and dark blue are represented alternative exons; in green is represented an extended CDS; and in red the 3'UTRs.

Alternative APA is essential for multiple functions in different tissues, and it is related to several pathologies. As stated above, shorter 3'UTRs potentially harbor fewer *cis*-regulatory elements, such as miRNA and RBP target sites, and are prone to be less regulated than the longer isoforms. The shorter isoforms are commonly associated with proliferative tissues and cancer cells (Mayr and Bartel, 2009; Xia *et al.*, 2014). On the other hand, longer 3'UTR isoforms are more expressed in differentiated tissues and polarized cells (which generally requires more regulation) (Guvenek and Tian, 2018; Miura *et al.*, 2013; Zhang *et al.*, 2005).

Since the miRNAs link mostly to the 3'UTR, this structure also assumes an important role in regulating the mRNA dynamics. It was found that different lengths of 3'UTRs are responsible for different mRNA turnover rates. In mouse, it was demonstrated that the 3'UTR of *hes7* and *lfng* have different lengths and have different impact in the turnover of an mRNA. The smaller 3'UTR (from *hes7*) gives more stability to the mRNA than a longer one (from *lfng*), and that influences the pattern of expression of each one of the genes (Nitanda *et al.*, 2014).

### 1.5.2 Gene expression regulation by miRNAs

Most of the work conducted in gene networks and regulation only considers a small percentage of the total RNA produced in a cell. In higher organisms, mRNA represents less than 3% of the total RNA (Li and Liu, 2019). The rest of the RNA is considered non-coding RNA, RNA that is not translated into a protein. However, they have an immense number of critical roles for cell homeostasis. As an example, rRNAs and tRNAs are crucial for protein

production. Alone, these two RNA categories account for more than 90% of all of the RNA produced by a cell.

Most of the other non-coding RNAs are involved in gene regulation, such as long non-coding RNAs (lncRNAs) and microRNAs (miRNAs). miRNAs are small single-strand RNA molecules with ~22 nucleotides (Bartel, 2018). These molecules are mainly present in the cytoplasm, where they are active. They are present in most living organisms, including animals, plants, and even viruses. The primary and more studied mode of action is by blocking translation or by driving mRNA degradation. In any of the cases, it negatively regulates the production of proteins (Reviewed in (Bartel, 2018)).

There is a standard nomenclature for miRNAs, a code that identifies each miRNA (Figure 1.19) (Bernardo *et al.*, 2012). The "name" of a given miRNA provides a lot of information. Namely, the species of origin (in red), if it refers to the mature sequence (capital R) or if it is related to its precursor miRNA, pre-mir (lowercase r). In green, is the identification number of the given miRNA. This number is common among species. Light blue identifies a miRNA slightly different but with the same seed sequence. In violet, the different genomic locations from which the miRNA is produced, and in orange, information on the arm of the pre-miRNA it comes from. It is common to use an asterisk to refer to the "Passenger" strand, that is less frequently used miRNA and the one without an asterisk refers to the more common mature miRNA produced by a given hairpin (Bernardo *et al.*, 2012).

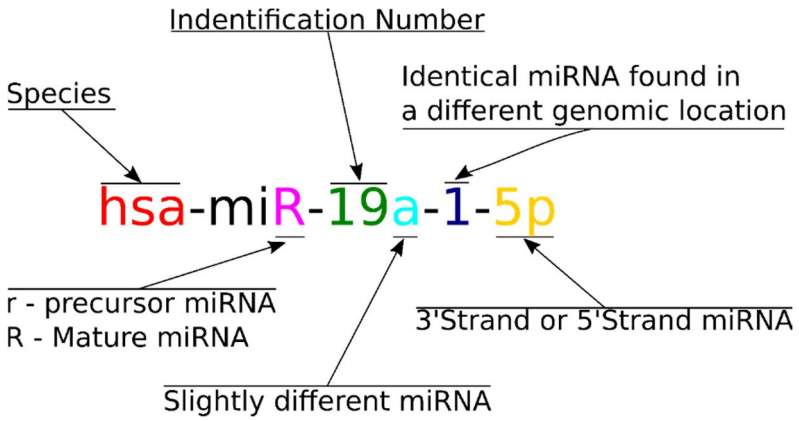



**Figure 1.19 - Nomenclature of a miRNA.**
Each miRNA has a specific code name, and here is the example of the human miRNA 19a-1-5p. In red, it identifies the species to which the miRNA belongs. Pink represents the maturity of the miRNA, "r" for the precursor, and "R" for the mature miRNA. In green, the identification number, specific for a given miRNA gene. The letter in light blue represents slightly different miRNAs but with the same targets. The number in dark blue refers to the number of identical miRNAs that come from different genomic locations. "3p" or "5p" represents the chosen strand for the mature miRNA.

The first miRNA was described in *Caenorhabditis elegans* in 1993 by two groups that joined forces to characterized it (Lee et al., 1993; Wightman *et al.*, 1993). They found that the heterochronic gene *lin-4* gave rise to two small RNAs, one with 61 nucleotides (pre-lin-4) and one with 22 nucleotides (mature lin-4). They also found that these RNAs had a small sequence that is the complement of a sequence in the 3'UTR of another heterochronic gene, *lin*-14. Thus, they postulated that this small RNA had the capability of targeting *lin-14* by RNA complementation, generating a double strand of RNA, leading to degradation. This process is extremely important for *C. elegans* development. The transcription of *lin-4,* and consequent increasing levels of the mature miRNA, targets *lin-14* mRNA, decreasing its expression. This expression shift culminates in the L1 stage larvae to advance to the L2 stage (Lee *et al.*, 1993; Wightman *et al.*, 1993).

At the time, the scientific community considered it a sporadic event and did not pursue the hypothesis of miRNAs being a major gene expression regulator. The second miRNA was only described seven years later, in 2000, by Reinhart and collaborators (Reinhart *et al.*, 2000). The team described a small RNA of 21 nucleotides that they called *let-7.* This second miRNA is also implicated in the development of *C. elegans.* In this case, it is involved in the L4 larval stage to adult transition. The loss of function of *let-7* caused the resurgence of "larval" cells in the adult, and the overexpression caused the early appearance of "adult" cells during larval stages. Loss of function of *lin-41* could compensate for the lack of *let-7* activity. In fact, *lin-41* is a target of *let-7*. The deletion of *let-7* target sites in the *lin-41* 3'UTR or *let-7* loss of function, both led to sustained expression of *lin-41* and an improper transition from larvae to adult (Grishok *et al.*, 2001; Pasquinelli *et al.*, 2000; Reinhart *et al.*, 2000).

The discovery of *let-7* was a turning point in the miRNA research. Unlike *lin-4*, *let-7* is phylogenetically conserved from *C. elegans* to humans. This triggered the scientific community to search for the function of *let-7* in other organisms and also the search for new miRNAs. Nowadays, thousands of miRNAs are known to be present among animals, plants, and even viruses (Bartel, 2004; Kozomara and Griffiths-Jones, 2011).

#### 1.5.2.1 miRNA biogenesis

There are several pathways that produce miRNAs. They can be transcribed or result as a "by-product" of the transcription of other genes, namely from an intron after splicing. In the canonical biogenesis pathway (Figure 1.20), the miRNA gene is transcribed by RNA polymerase II, the RNA is polyadenylated and caped, as would an mRNA (Lee and Kim, 2004). The resulting RNA molecule is called primary-miRNA (pri-miRNA). In the alternative pathway,

the pri-miRNA can be generated by the release of an intron upon splicing of a coding mRNA (pri-mirtron) (Tang and Maxwell, 2008). Several miRNAs can be located at the same genetic locus, and they are transcribed all at once. The pri-miRNA does not have an established size. It can range from hundreds to thousands of nucleotides (Tanzer and Stadler, 2004). Due to its sequence, the pri-miRNA folds and forms a hairpin with a specific conformation recognized by an enzymatic microprocessor in the nucleus. DiGeorge syndrome critical region 8 (DGCR8) and RNase II endonuclease III Drosha join with some other factors to form the microprocessor. The microprocessor subunit DGCR8 recognizes the stem-loop on the top of the hairpin, and the other subunit Drosha, cleaves and detaches the hairpin to form the precursor miRNA (pre-miRNA) (Denli *et al.*, 2004; Lee *et al.*, 2003; Yeom *et al.*, 2006). This molecule, of around 60 to 110 nucleotides, is then exported to the cytoplasm by Exportin-5 (XPO-5) (Yi *et al.*, 2003). Once in the cytoplasm, the hairpin is now recognized by another RNase II endonuclease III, Dicer. With the help of TRBP (transactivation response element RNA-binding protein) and PACT (protein activator of PKR), Dicer cleaves the stem-loop from the pre-miRNA and leaves a small double-strand RNA (Daniels *et al.*, 2009; Koscianska *et al.*, 2011). The small double-strand RNA is recognized by Argonaute (AGO 1-4), and with it, all the RISC (RNA-induced silencing complex) complex is recruited upon the choice of one strand (MacRae *et al.*, 2008; Miyoshi *et al.*, 2009). At this point, the mature miRNA is active.

There are also other ways to produce miRNAs, known as non-canonical pathways. In some cases, a given intron (mirtron), upon being processed by the spliceosome and the lariat is resolved, folds and forms the hairpin (pre-miRNA), which is immediately exported to the cytoplasm, and is processed by Dicer. There are more particular cases of non-canonical miRNA processing dependent on Drosha but not on Dicer. In these cases, the AGO2 trims the miRNA and loads it into RISC (Reviewed in (García-López *et al.*, 2013)).

There is still some controversy about how one of the strands of the miRNA duplex is selected. That is because, in many cases, both strands of the miRNA duplex generate a mature miRNA. However, the preference is for the strand that starts by an Uracil at the 5’ end and is rich in A/G nucleotides. That occurs because this end of the double-strand is richer in week interactions than the other end of the duplex. AGO engulfs the 5’ loose end and protects this strand (guide strand). With the help of a helicase, the duplex is unfolded, and the other strand (passenger strand) is released and degraded (Okamura *et al.*, 2009; Shin, 2008).

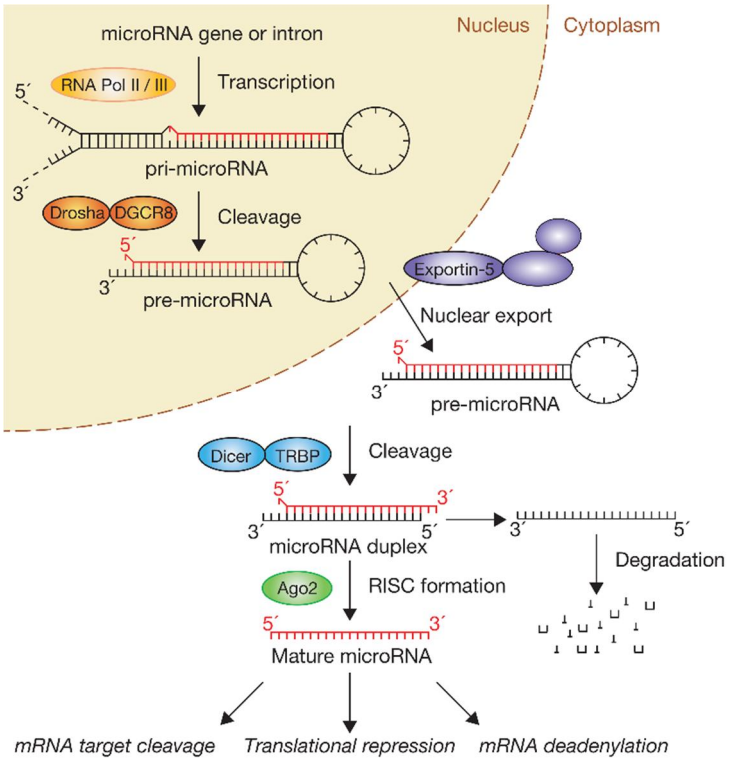



**Figure 1.20 - Canonical miRNA Biogenesis.**
A microRNA gene is transcribed by RNA Polymerase II or results from a hairpin in an intron (pri-microRNA). The hairpin is recognized by the microprocessor (Drosha and DGCR8), which cleaves the hairpin from the pri-miRNA forming the pre-miRNA. The pre-miRNA is exported to the cytoplasm by Exportin-5. In the cytoplasm, the pre-miRNA is recognized by Dicer, which cleaves the stem-loop, leaving a small RNA duplex. AGO proteins recognize the duplex and select one of the strands. The active single strand miRNA targets mRNAs and leads to mRNA cleavage, translation repression, and mRNA deadenylation. Adapted from (Winter *et al.*, 2009).

### 1.5.2.2 miRNA mechanisms of action

The mature miRNA generally recognizes its target by sequence paring of a small region in the 5' end with the target site (Figure 1.21). This small sequence is called seed sequence. The seed sequence has between 6 to 7 nucleotides, ranging from the second position of the miRNA on the 5' end (Lewis *et al.*, 2003). In addition to this seed, the first uracil may contribute to recognizing the target site, increasing the strength of the recognition (Bartel, 2009).

Once the miRNA recognizes the 3'UTR of an mRNA, one of two things may happen: if the base paring is partial, there is translation repression; if the pairing is perfect, there is a cut of the double strand, promoting RNA degradation. It is important to refer that a mature miRNA could be integrated into the RISC complex with AGO 1 to 4. However, only AGO2 (Liu *et al.*, 2004; Meister *et al.*, 2004) and AGO3 (found in recent studies) (Park *et al.*, 2017) have catalytic activity. For that reason, only when RISC has AGO2 or AGO3, the miRNA is able to degrade the mRNA directly.

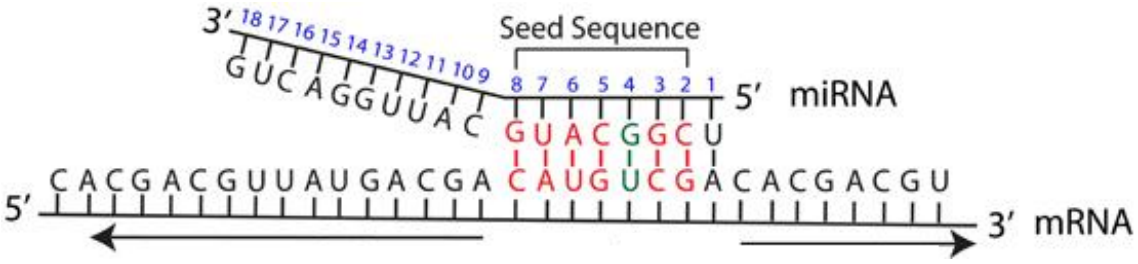



**Figure 1.21 - Schematics of miRNA targeting an mRNA.**
A miRNA recognizes its targets mainly by the 5' nucleotides 2-8 (Seed Sequence). The uracil in position 5' 1 can increase the strength of the interaction. Generally, the paring is made by GC interactions (Red), but in sporadic cases, it can be recognized with some GU pair (Wobble paring) (Green). Adapted from (Peterson et al., 2014)

In most cases, the miRNA stops translation and promotes mRNA degradation by deadenylating the mRNA and consequently destabilizing it. For that, the miRNA recruits GW182 adaptor, which interacts with poly-A binding protein (PABP) and at the same time recruits CCR4-NOT deadenylase complex. The CCR4-NOT complex starts to deadenylate the mRNA, which shortens the half-life of the mRNA. Once the Poly-A is too short, the mRNA is decapped and degraded from the 5'end (Couttet *et al.*, 1997; Coutts and Brawerman, 1993). The deadenylase complex has a negative effect on protein production by inhibiting translation initiation (Collart, 2016). In some cases, the miRNA is stored in miRNA reservoirs. These reservoirs have mRNAs, lncRNAs, and circRNAs that are rich in miRNA target sites that sequester a given miRNA until the cell needs it (Reviewed in (García-López *et al.*, 2013)).

miRNAs typically bind the mRNA in the 3'UTR, causing translation repression and/or mRNA degradation. However, there some reports show that miRNAs are able to bind to different mRNA regions. miRNAs are able to attach to 5'UTR and lead to either translation repression or activation. It may also bind to the coding region blocking the translation. In some particular cases, it was also reported that the miRNA is able to promote translation when it attaches to the 3'UTR (Reviewed in (Ling *et al.*, 2013)).

There are thousands of miRNAs in a cell, and since it only needs six or seven nucleotides to have action, they can regulate, in some form, the expression of virtually all genes. Consequently, they are involved in many diseases such as cancer, immune-related diseases, neurodegenerative diseases, and even viral infections. Besides the huge effort to mechanistically find miRNAs' role in these diseases, the scientific community is also studying miRNAs as potential biomarkers and therapeutic targets.

#### 1.5.2.3 miRNAs in development

Most of the miRNAs are generated by the canonical machinery. Thus, one way to explore miRNA roles in development during the past years has been by deleting core proteins of the canonical biogenesis pathway. For example, several groups deleted Drosha or Dicer and evaluated the impact of the absence of miRNAs during embryo development. Deletion of Dicer in mice causes embryonic lethality around the day E7.5 (Bernstein *et al.*, 2003). The same happens when DGCR8 is deleted, around the day E6.5 (Chong *et al.*, 2008; Wang *et al.*, 2007). Ago2 null mutant also arrests development between E5.5 and E6.5 (Morita *et al.*, 2007). In *C. elegans,* the deletion of Argonaute causes embryonic arrest during morphogenesis (Vasquez-Rifo *et al.*, 2012). Giraldez and colleagues showed that when embryos are depleted of the maternal contribution of *Dicer1,* zebrafish embryos undergo abnormal morphogenesis and die around day five post-fertilization (dpf) (Giraldez *et al.*, 2006). However, zygotic *dicer1* depleted embryos only have a developmental arrest around day ten dpf (Wienholds *et al.*, 2003). Moreover, in *Drosophila,* when mother and zygotic AGO1 are deleted, the larvae die with severe developmental problems of the nervous system (Kataoka *et al.*, 2001). However, when only the zygotic miRNA machinery is depleted in *Drosophila* and *C. elegans*, the embryo develops without significant problems, besides being sterile. Thus, showing the importance of the mother's inherited miRNA machinery to maintain the germline in these embryos (Alberti and Cochella, 2017; Denli *et al.*, 2004; Knight and Bass, 2001; Martin *et al.*, 2009).

Despite the dramatic consequences for embryo development of the overall depletion of miRNAs, the same is not observed when specific miRNAs are deleted. There are three major explanations for this to happen. First, the redundancy, as most miRNAs do not have an exclusive seed sequence. That makes it possible for other miRNAs to compensate for the lack of one particular miRNA. Moreover, one mRNA could have several miRNA target sites for different miRNA species with different seed sequences. This may help these other miRNA species to compensate for the lack of each other. For that reason, the entire miRNA family is often studied instead of one particular miRNA. Second, the genetic regulatory networks where miRNAs participate are resistant to perturbations, and the miRNAs contribute to that resistance. In fact, when specific miRNAs are deleted, the organism is more susceptible to environmental changes (Burke *et al.*, 2015; Kasper *et al.*, 2017; Li *et al.*, 2009). However, the embryos do not have an apparent phenotype in all the miRNA deficient embryos. Third, some miRNAs are expressed in a specific place and time of development. That means that the deleted miRNA could be irrelevant in another specific tissue or time of development.

### 1.5.3 miR-9 and miR-125a-5p in Embryo Clock regulation

Somitogenesis requires a very tightly regulated genetic network in order to produce each pair of somites at a precise pace. For instance, in chicken, the embryonic clock has a pace of 90 minutes, which means that the genes are expressed and downregulated with this period (Palmeirim *et al.*, 1997). miRNAs are perfect candidates as negative regulators for the EC and may explain the fast mRNA turnover of these genes. The EC in the forelimb outgrowth has a pace of six hours, four times the period of the EC in the PSM (Pascoal *et al.*, 2007). Based on this, we hypothesized that miRNAs could be responsible for finetuning the pace of the embryo clock and be responsible for the difference of the EC pace in the PSM and the forelimb outgrowth.

Previously in our lab, we showed that the machinery to produce the miRNAs is present in the tissues where the embryonic clock is active, more precisely in the PSM during the somitogenesis stages and in the forelimb outgrowth at stages in which the EC is operating (HH20 to HH25) (Carraco *et al.*, 2014). Moreover, in somitogenesis, it has been shown that these small RNAs have an essential role in the turnover of the mRNA. Riley and collaborators in 2013 showed that miR-125a-5p is essential for *lfng* mRNA turnover. When this miRNA is absent, or its activity is absent, *lfng* loses cyclic expression, and instead, it has a static expression pattern with increased mRNA and protein levels (Riley *et al.*, 2013).

EC gene *hes1* also has a critical role in neuronal differentiation. Also, here, it was found an essential role for miRNAs. miR-9 targets the mouse *hes1* mRNA. In this case, pri-mir-9 transcription is in anti-phase with *hes1* transcription, as Hes1 inhibits its expression. However, the amount of mature miR-9, which is stable, increases at each cycle. Once the mature miR-9 reaches a critical concentration in the cell, it permanently downregulates *hes1,* and pri-mir-9 is permanently transcribed. Thus, it results in neuronal differentiation (Bonev *et al.*, 2012). More recently, it was shown that miR-9 is also important for zebrafish neurogenesis by targeting *her6*, the orthologue of *hes1* in zebrafish (Soto *et al.*, 2020).

## 1.6 Objectives

During embryo development, processes such as somitogenesis and limb outgrowth need a tight regulation in space and time. Both processes rely on oscillatory gene expression, the embryo clock (EC), to ensure correct segmentation and a correct identity to each segment. During somitogenesis, chicken EC genes, such as *hairy1* and *hairy2*, have a cyclic expression with a periodicity of 90 minutes. The cyclic expression of these genes slows down in the anterior PSM until it remains static upon somite formation. Moreover, in the same organism, *hairy2* has a cyclic expression of six hours in the distal forelimb mesenchyme. This leads to the question: How is EC periodicity regulated in the different tissues?

Since the EC timing is dependent on gene expression and mRNA stability, we hypothesized that different RNA molecules may regulate the Embryo Clock in different tissues, promoting different EC periods.

To investigate this hypothesis, the following topics were explored in the different Chapters of this thesis:

1. Alternative transcription fine-tunes the Embryo Clock pace.
    a. Alternative Transcription Start Sites (TSS) generate different 5'transcripts that may have a role in Embryo Clock regulation (Chapter 2).
    b. Alternative 3'UTRs influence mRNA stability and consequently regulate the Embryonic Clock (Chapter 3).
2. Tissue-specific miRNA species regulate the EC period, accounting for different EC paces in different tissues (Chapter 4).

## 1.7 References


**Alberti, C., and Cochella, L.** (2017). A framework for understanding the roles of miRNAs in animal development. Development *144*, 2548-2559. https://doi.org/10.1242/dev.146613.

**Aoyama, H., and Asamoto, K.** (2000). The developmental fate of the rostral/caudal half of a somite for vertebra and rib formation: experimental confirmation of the resegmentation theory using chick-quail chimeras. Mech Dev *99*, 71-82. https://doi.org/10.1016/s0925-4773(00)00481-0.

**Aulehla, A., and Johnson, R.L.** (1999). Dynamic expression of lunatic fringe suggests a link between notch signaling and an autonomous cellular oscillator driving somite segmentation. Dev Biol *207*, 49-61. https://doi.org/10.1006/dbio.1998.9164.

**Bailey, C., and Dale, K.** (2015). Somitogenesis in Vertebrate Development. In eLS, (John Wiley & Sons, Ltd). https://doi.org/10.1002/9780470015902.a0003820.pub2.

**Baker, R.E., Schnell, S., and Maini, P.K.** (2009). Waves and patterning in developmental biology: vertebrate segmentation and feather bud formation as case studies. Int J Dev Biol *53*, 783-794. https://doi.org/10.1387/ijdb.072493rb.

**Bartel, D.P.** (2004). MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell *116*, 281-297. https://doi.org/10.1016/S0092-8674(04)00045-5.

**Bartel, D.P.** (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233. https://doi.org/10.1016/j.cell.2009.01.002.

**Bartel, D.P.** (2018). Metazoan MicroRNAs. Cell *173*, 20-51. https://doi.org/10.1016/j.cell.2018.03.006.

**Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E.** (2006). MRNA degradation by miRNAs and GW182 requires both CCR4 : NOT deadenylase and DCP1 : DCP2 decapping complexes. Genes & Development *20*, 1885-1898. https://doi.org/10.1101/gad.1424106.

**Bernardo, B.C., Charchar, F.J., Lin, R.C., and McMullen, J.R.** (2012). A microRNA guide for clinicians and basic scientists: background and experimental techniques. Heart Lung Circ *21*, 131-142. https://doi.org/10.1016/j.hlc.2011.11.002.

**Bernstein, E., Kim, S.Y., Carmell, M.A., Murchison, E.P., Alcorn, H., Li, M.Z., Mills, A.A., Elledge, S.J., Anderson, K.V., and Hannon, G.J.** (2003). Dicer is essential for mouse development. Nat Genet *35*, 215-217. https://doi.org/10.1038/ng1253.

**Bessho, Y., Hirata, H., Masamizu, Y., and Kageyama, R.** (2003). Periodic repression by the bHLH factor Hes7 is an essential mechanism for the somite segmentation clock. Genes Dev *17*, 1451-1456. https://doi.org/10.1101/gad.1092303.

**Bessho, Y., Sakata, R., Komatsu, S., Shiota, K., Yamada, S., and Kageyama, R.** (2001). Dynamic expression and essential functions of Hes7 in somite segmentation. Genes Dev *15*, 2642-2647. https://doi.org/10.1101/gad.930601.

**Bonev, B., Stanley, P., and Papalopulu, N.** (2012). MicroRNA-9 Modulates Hes1 Ultradian Oscillations by Forming a Double-Negative Feedback Loop. Cell Reports *2*, 10-18. https://doi.org/10.1016/j.celrep.2012.05.017.

**Burke, S.L., Hammell, M., and Ambros, V.** (2015). Robust Distal Tip Cell Pathfinding in the Face of Temperature Stress Is Ensured by Two Conserved microRNAS in Caenorhabditis elegans. Genetics *200*, 1201-1218. https://doi.org/10.1534/genetics.115.179184.

**Cambray, N., and Wilson, V.** (2002). Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. Development *129*, 4855-4866. https://doi.org/ 10.1242/dev.129.20.4855

**Carraco, G., Gonçalves, A.N., Serra, C., and Andrade, R.P.** (2014). MicroRNA processing machinery in the developing chick embryo. Gene Expression Patterns *16*, 114-121. https://doi.org/10.1016/j.gep.2014.09.002.

**Catala, M., Teillet, M.A., and Le Douarin, N.M**. (1995). Organization and development of the tail bud analyzed with the quail-chick chimaera system. Mech Dev *51*, 51-65. https://doi.org/10.1016/0925-4773(95)00350-a.

**Chen, C.Y., and Shyu, A.B.** (1995). AU-rich elements: characterization and importance in mRNA degradation. Trends Biochem Sci *20*, 465-470. https://doi.org/10.1016/s0968-0004(00)89102-1.

**Chen, J., Kang, L., and Zhang, N.** (2005). Negative feedback loop formed by Lunatic fringe and Hes7 controls their oscillatory expression during somitogenesis. Genesis *43*, 196-204. https://doi.org/10.1002/gene.20171

**Chen, M., Lyu, G., Han, M., Nie, H., Shen, T., Chen, W., Niu, Y., Song, Y., Li, X., Li, H., et al.** (2018). 3' UTR lengthening as a novel mechanism in regulating cellular senescence. Genome Res. https://doi.org/10.1101/gr.224451.117.

**Chong, M.M.W., Rasmussen, J.P., Rudensky, A.Y., and Littman, D.R.** (2008). The RNAseIII enzyme Drosha is critical in T cells for preventing lethal inflammatory disease. J Exp Med *205*, 2005-2017. https://doi.org/10.1084/jem.20081219.

**Christ, B., Huang, R., and Scaal, M.** (2007). Amniote somite derivatives. Dev Dyn *236*, 2382-2396. https://doi.org/10.1002/dvdy.21189.

**Collart, M.A.** (2016). The Ccr4-Not complex is a key regulator of eukaryotic gene expression. Wiley Interdiscip Rev RNA *7*, 438-454. https://doi.org/10.1002/wrna.1332.

**Cooke, J., and Zeeman, E.C.** (1976). A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. Journal of Theoretical Biology *58*, 455-476. https://doi.org/10.1016/S0022-5193(76)80131-2.

**Couttet, P., Fromont-Racine, M., Steel, D., Pictet, R., and Grange, T.** (1997). Messenger RNA deadenylylation precedes decapping in mammalian cells. Proc Natl Acad Sci U S A *94*, 5628-5633. https://doi.org/10.1073/pnas.94.11.5628.

**Coutts, M., and Brawerman, G.** (1993). A 5′ exoribonuclease from cytoplasmic extracts of mouse sarcoma 180 ascites cells. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression *1173*, 57-62. https://doi.org/10.1016/0167-4781(93)90242-6.

**Dale, J.K., Malapert, P., Chal, J., Vilhais-Neto, G., Maroto, M., Johnson, T., Jayasinghe, S., Trainor, P., Herrmann, B., and Pourquie, O.** (2006). Oscillations of the snail genes in the presomitic mesoderm coordinate segmental patterning and morphogenesis in vertebrate somitogenesis. Dev Cell *10*, 355-366. https://doi.org/10.1016/j.devcel.2006.02.011.

**Daniels, S.M., Melendez-Pena, C.E., Scarborough, R.J., Daher, A., Christensen, H.S., El Far, M., Purcell, D.F.J., Laine, S., and Gatignol, A.** (2009). Characterization of the TRBP domain required for Dicer interaction and function in RNA interference. Bmc Molecular Biology *10*. 38. https://doi.org/10.1186/1471-2199-10-38.

**Denli, A.M., Tops, B.B.J., Plasterk, R.H.A., Ketting, R.F., and Hannon, G.J.** (2004). Processing of primary microRNAs by the Microprocessor complex. Nature *432*, 231-235. https://doi.org/10.1038/nature03049

**Dequéant, M.-L., Glynn, E., Gaudenz, K., Wahl, M., Chen, J., Mushegian, A., and Pourquié, O.** (2006). A Complex Oscillating Network of Signaling Genes Underlies the Mouse Segmentation Clock. Science *314*, 1595-1598. https://doi.org/10.1126/science.1133141.

**Diaz-Cuadros, M., Wagner, D.E., Budjan, C., Hubaud, A., Tarazona, O.A., Donelly, S., Michaut, A., Al Tanoury, Z., Yoshioka-Kobayashi, K., Niino, Y., et al.** (2020). In vitro

characterization of the human segmentation clock. Nature *580*, 113-118. https://doi.org/10.1038/s41586-019-1885-9.

**Diez del Corral, R., Olivera-Martinez, I., Goriely, A., Gale, E., Maden, M., and Storey, K.** (2003). Opposing FGF and retinoid pathways control ventral neural pattern, neuronal differentiation, and segmentation during body axis extension. Neuron *40*, 65-79. https://doi.org/10.1016/s0896-6273(03)00565-8.

**Dubrulle, J., McGrew, M.J., and Pourquié, O.** (2001). FGF Signaling Controls Somite Boundary Position and Regulates Segmentation Clock Control of Spatiotemporal Hox Gene Activation. Cell *106*, 219-232. https://doi.org/10.1016/S0092-8674(01)00437-8.

**Dubrulle, J., and Pourquie, O.** (2004). Coupling segmentation to axis formation. Development *131*, 5783-5793. https://doi.org/10.1242/dev.01519.

**Dudley, A.T., Ros, M.A., and Tabin, C.J.** (2002). A re-examination of proximodistal patterning during vertebrate limb development. Nature *418*, 539-544. https://doi.org/10.1038/nature00945.

**Elkon, R., Ugalde, A.P., and Agami, R.** (2013). Alternative cleavage and polyadenylation: extent, regulation and function. Nat Rev Genet *14*, 496-506. https://doi.org/10.1038/nrg3482.

**Elmasri, H., Liedtke, D., Lücking, G., Volff, J.N., Gessler, M., and Winkler, C.** (2004). her7 and hey1, but not lunatic fringe show dynamic expression during somitogenesis in medaka (Oryzias latipes). Gene Expr Patterns *4*, 553-559. https://doi.org/10.1016/j.modgep.2004.02.003.

**Evdokimova, V., Ruzanov, P., Imataka, H., Raught, B., Svitkin, Y., Ovchinnikov, L.P., and Sonenberg, N.** (2001). The major mRNA-associated protein YB-1 is a potent 5' cap-dependent mRNA stabilizer. Embo j *20*, 5491-5502. https://doi.org/10.1093/emboj/20.19.5491.

**Fallon, J.F., Lopez, A., Ros, M.A., Savage, M.P., Olwin, B.B., and Simandl, B.K.** (1994). FGF-2: apical ectodermal ridge growth signal for chick limb development. Science *264*, 104-107. https://doi.org/10.1126/science.7908145.

**Firmin, J., and Maître, J.-L.** (2021). Morphogenesis of the human preimplantation embryo: bringing mechanics to the clinics. Seminars in Cell & Developmental Biology. https://doi.org/10.1016/j.semcdb.2021.07.005.

**Forsberg, H., Crozet, F., and Brown, N.A.** (1998). Waves of mouse Lunatic fringe expression, in four-hour cycles at two-hour intervals, precede somite boundary formation. Curr Biol *8*, 1027-1030. https://doi.org/10.1016/s0960-9822(07)00424-1.

**Freitas, C., Rodrigues, S., Charrier, J.B., Teillet, M.A., and Palmeirim, I.** (2001). Evidence for medial/lateral specification and positional information within the presomitic mesoderm. Development *128*, 5139-5147. https://doi.org/10.1242/dev.128.24.5139

**Gajewski, M., Elmasri, H., Girschick, M., Sieger, D., and Winkler, C.** (2006). Comparative analysis of her genes during fish somitogenesis suggests a mouse/chick-like mode of oscillation in medaka. Dev Genes Evol *216*, 315-332. https://doi.org/10.1007/s00427-006-0059-6.

**García-López, J., Brieño-Enríquez Miguel, A., and del Mazo, J.** (2013). MicroRNA biogenesis and variability. BioMolecular Concepts. *4*, 367-380. https://doi.org/10.1515/bmc-2013-0015

**Gibb, S., Zagorska, A., Melton, K., Tenin, G., Vacca, I., Trainor, P., Maroto, M., and Dale, J.K.** (2009). Interfering with Wnt signalling alters the periodicity of the segmentation clock. Dev Biol *330*, 21-31. https://doi.org/10.1016/j.ydbio.2009.02.035.

**Gilbert, F.S., and Barresi, M.J.F.** (2019). Developmental biology, 12th Edition (Sinauer Associates, Inc.).

**Giraldez, A.J., Mishima, Y., Rihel, J., Grocock, R.J., Van Dongen, S., Inoue, K., Enright, A.J., and Schier, A.F.** (2006). Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science *312*, 75-79. https://doi.org/10.1126/science.1122689

**Gomez, C., Ozbudak, E.M., Wunderlich, J., Baumann, D., Lewis, J., and Pourquie, O.** (2008). Control of segment number in vertebrate embryos. Nature *454*, 335-339. https://doi.org/10.1038/nature07020.

**Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C.** (2001). Genes and Mechanisms Related to RNA Interference Regulate Expression of the Small Temporal RNAs that Control C. elegans Developmental Timing. Cell *106*, 23-34. http://dx.doi.org/10.1016/S0092-8674(01)00431-7.

**Guhaniyogi, J., and Brewer, G.** (2001). Regulation of mRNA stability in mammalian cells. Gene *265*, 11-23. https://doi.org/10.1016/S0378-1119(01)00350-X.

**Guvenek, A., and Tian, B.** (2018). Analysis of alternative cleavage and polyadenylation in mature and differentiating neurons using RNA-seq data. Quant Biol *6*, 253-266. https://doi.org/10.1007/s40484-018-0148-3.

**Haghighat, A., and Sonenberg, N.** (1997). eIF4G dramatically enhances the binding of eIF4E to the mRNA 5'-cap structure. J Biol Chem *272*, 21677-21680. https://doi.org/10.1074/jbc.272.35.21677.

**Hajj, K.A., and Whitehead, K.A.** (2017). Tools for translation: non-viral materials for therapeutic mRNA delivery. Nature Reviews Materials *2*, 17056. 10.1038/natrevmats.2017.56.

**Hamburger, V., and Hamilton, H.L.** (1951). A SERIES OF NORMAL STAGES IN THE DEVELOPMENT OF THE CHICK EMBRYO. Journal of Morphology *88*, 49-&. https://doi.org/10.1002/jmor.1050880104.

**Hirata, H., Yoshiura, S., Ohtsuka, T., Bessho, Y., Harada, T., Yoshikawa, K., and Kageyama, R.** (2002). Oscillatory Expression of the bHLH Factor Hes1 Regulated by a Negative Feedback Loop. Science *298*, 840-843. https://doi.org/10.1126/science.1074560.

**Holley, S.A., Geisler, R., and Nüsslein-Volhard, C.** (2000). Control of her1 expression during zebrafish somitogenesis by a delta-dependent oscillator and an independent wave-front activity. Genes Dev *14*, 1678-1690. https://doi.org/10.1101/gad.14.13.1678

**Horder, T.** (2010). History of Developmental Biology. In eLS. https://doi.org/10.1002/9780470015902.a0003080.pub2.

**Hwang, Y.S., Seo, M., Kim, S.K., Bang, S., Kim, H., and Han, J.Y.** (2018). Zygotic gene activation in the chicken occurs in two waves, the first involving only maternally derived genes. eLife *7*, e39381. https://doi.org/10.7554/eLife.39381.

**Jiang, Y.J., Aerne, B.L., Smithers, L., Haddon, C., Ish-Horowicz, D., and Lewis, J.** (2000). Notch signalling and the synchronization of the somite segmentation clock. Nature *408*, 475-479. https://doi.org/10.1038/35044091.

**Jouve, C., Palmeirim, I., Henrique, D., Beckers, J., Gossler, A., Ish-Horowicz, D., and Pourquie, O.** (2000). Notch signalling is required for cyclic expression of the hairy-like gene HES1 in the presomitic mesoderm. Development *127*, 1421-1429. https://doi.org/10.1242/dev.127.7.1421

**Kageyama, R., Ohtsuka, T., and Kobayashi, T.** (2007). The Hes gene family: repressors and oscillators that orchestrate embryogenesis. Development *134*, 1243-1251. https://doi.org/10.1242/dev.000786.

**Kalcheim, C.** (2016). Epithelial–Mesenchymal Transitions during Neural Crest and Somite Development. Journal of Clinical Medicine *5*, 1. https://doi.org/10.3390/jcm5010001

**Kang, S.A., Seol, J.H., and Kim, J.** (2005). The conserved WRPW motif of Hes6 mediates proteasomal degradation. Biochem Biophys Res Commun *332*, 33-36. https://doi.org/10.1016/j.bbrc.2005.04.089.

**Kasper, D.M., Moro, A., Ristori, E., Narayanan, A., Hill-Teran, G., Fleming, E., Moreno-Mateos, M., Vejnar, C.E., Zhang, J., Lee, D., et al.** (2017). MicroRNAs Establish Uniform Traits during the Architecture of Vertebrate Embryos. Dev Cell *40*, 552-565.e555. https://doi.org/10.1016/j.devcel.2017.02.021.

**Kataoka, Y., Takeichi, M., and Uemura, T.** (2001). Developmental roles and molecular characterization of a Drosophila homologue of Arabidopsis Argonaute1, the founder of a novel gene superfamily. Genes Cells *6*, 313-325. https://doi.org/10.1046/j.1365-2443.2001.00427.x.

**Kieny, M., Mauger, A., and Sengel, P. (**1972). Early regionalization of somitic mesoderm as studied by the development of axial skeleton of the chick embryo. Dev Biol *28*, 142-161. https://doi.org/10.1016/0012-1606(72)90133-9.

**Knight, S.W., and Bass, B.L.** (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. Science *293*, 2269-2271. https://doi.org/10.1126/science.1062039.

**Koscianska, E., Starega-Roslan, J., and Krzyzosiak, W.J.** (2011). The Role of Dicer Protein Partners in the Processing of MicroRNA Precursors. Plos One *6*. e28548

https://doi.org/10.1371/journal.pone.0028548.

**Kozomara, A., and Griffiths-Jones, S.** (2011). miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Research *39*, D152-D157. https://doi.org/10.1093/nar/gkq1027.

**Krol, A.J., Roellig, D., Dequeant, M.L., Tassy, O., Glynn, E., Hattem, G., Mushegian, A., Oates, A.C., and Pourquie, O.** (2011). Evolutionary plasticity of segmentation clock networks. Development *138*, 2783-2792. https://doi.org/10.1242/dev.063834.

**Langley, A.R., Smith, J.C., Stemple, D.L., and Harvey, S.A.** (2014). New insights into the maternal to zygotic transition. Development *141*, 3834-3841. https://doi.org/10.1242/dev.102368.

**Lee, O., and Kim, V.N.** (2004). Evidence that microRNA genes are transcribed by RNA polymerase II. Cell Structure and Function *29*, 68-68. https://doi.org/10.1038/sj.emboj.7600385

**Lee, R.C., Feinbaum, R.L., and Ambros, V.** (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854. http://dx.doi.org/10.1016/0092-8674(93)90529-Y.

**Lee, Y., Ahn, C., Han, J.J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., and Kim, V.N.** (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415-419. https://doi.org/10.1038/nature01957.

**Leppek, K., Das, R., and Barna, M.** (2018). Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. Nat Rev Mol Cell Biol *19*, 158-174. https://doi.org/10.1038/nrm.2017.103.

**Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P., and Burge, C.B.** (2003). Prediction of mammalian microRNA targets. Cell *115*, 787-798. https://doi.org/10.1016/s0092-8674(03)01018-3.

**Lewis, E.B.** (1978). A gene complex controlling segmentation in Drosophila. Nature *276*, 565-570. https://doi.org/10.1038/276565a0.

**Li, J., and Liu, C.** (2019). Coding or Noncoding, the Converging Concepts of RNAs. Frontiers in Genetics *10*. https://doi.org/10.3389/fgene.2019.00496.

**Li, X., Cassidy, J.J., Reinke, C.A., Fischboeck, S., and Carthew, R.W.** (2009). A microRNA imparts robustness against environmental fluctuation during development. Cell *137*, 273-282. https://doi.org/10.1016/j.cell.2009.01.058.

**Li, Y., Fenger, U., Niehrs, C., and Pollet, N.** (2003). Cyclic expression of esr9 gene in Xenopus presomitic mesoderm. Differentiation *71*, 83-89. https://doi.org/10.1046/j.1432-0436.2003.700608.x.

**Ling, H., Fabbri, M., and Calin, G.A.** (2013). MicroRNAs and other non-coding RNAs as targets for anticancer drug development. Nat Rev Drug Discov *12*, 847-865. https://doi.org/10.1038/nrd4140.

**Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J.** (2004). Argonaute2 is the catalytic engine of mammalian RNAi. Science *305*, 1437-1441. https://doi.org/10.1126/science.1102513.

**Luo, Z., Rhie, S.K., and Farnham, P.J.** (2019). The Enigmatic HOX Genes: Can We Crack Their Code? Cancers (Basel) *11*, 323. https://doi.org/10.3390/cancers11030323.

**MacRae, I.J., Ma, E., Zhou, M., Robinson, C.V., and Doudna, J.A.** (2008). In vitro reconstitution of the human RISC-loading complex. Proceedings of the National Academy of Sciences of the United States of America *105*, 512-517. https://doi.org/10.1073/pnas.0710869105.

**Mallo, M., Wellik, D.M., and Deschamps, J.** (2010). Hox genes and regional patterning of the vertebrate body plan. Dev Biol *344*, 7-15. https://doi.org/10.1016/j.ydbio.2010.04.024.

**Martin, R., Smibert, P., Yalcin, A., Tyler, D.M., Schäfer, U., Tuschl, T., and Lai, E.C.** (2009). A Drosophila pasha mutant distinguishes the canonical microRNA and mirtron pathways. Mol Cell Biol *29*, 861-870. https://doi.org/10.1128/MCB.01524-08.

**Matsuda, M., Hayashi, H., Garcia-Ojalvo, J., Yoshioka-Kobayashi, K., Kageyama, R., Yamanaka, Y., Ikeya, M., Toguchida, J., Alev, C., and Ebisuya, M.** (2020a). Species-specific segmentation clock periods are due to differential biochemical reaction speeds. Science *369*, 1450-1455. https://doi.org/10.1126/science.aba7668.

**Matsuda, M., Yamanaka, Y., Uemura, M., Osawa, M., Saito, M.K., Nagahashi, A., Nishio, M., Guo, L., Ikegawa, S., Sakurai, S., et al.** (2020b). Recapitulating the human segmentation clock with pluripotent stem cells. Nature *580*, 124-129. 10.1038/s41586-020-2144-9.

**Mayr, C., and Bartel, D.P.** (2009). Widespread Shortening of 3′UTRs by Alternative Cleavage and Polyadenylation Activates Oncogenes in Cancer Cells. Cell *138*, 673-684. https://doi.org/10.1016/j.cell.2009.06.016.

**McGrew, M.J., Dale, J.K., Fraboulet, S., and Pourquié, O.** (1998). The lunatic Fringe gene is a target of the molecular clock linked to somite segmentation in avian embryos. Current Biology *8*, 979-982. https://doi.org/10.1016/S0960-9822(98)70401-4.

**Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T**. (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Mol Cell *15*, 185-197. https://doi.org/10.1016/j.molcel.2004.07.007.

**Mignone, F., Gissi, C., Liuni, S., and Pesole, G.** (2002). Untranslated regions of mRNAs. Genome Biol *3*, Reviews0004. https://doi.org/10.1186/gb-2002-3-3-reviews0004.

**Miura, P., Shenker, S., Andreu-Agullo, C., Westholm, J.O., and Lai, E.C.** (2013). Widespread and extensive lengthening of 3' UTRs in the mammalian brain. Genome Res *23*, 812-825. https://doi.org/10.1101/gr.146886.112.

**Miyoshi, K., Okada, T.N., Siomi, H., and Siomi, M.C.** (2009). Characterization of the miRNA-RISC loading complex and miRNA-RISC formed in the Drosophila miRNA pathway. Rna-a Publication of the Rna Society *15*, 1282-1291. https://doi.org/10.1261/rna.1541209.

**Moor, C.H.d., Meijer, H., and Lissenden, S.** (2005). Mechanisms of translational control by the 3′ UTR in development and differentiation. Seminars in Cell & Developmental Biology *16*, 49-58. https://doi.org/10.1016/j.semcdb.2004.11.007.

**Morin-Kensicki, E.M., Melancon, E., and Eisen, J.S.** (2002). Segmental relationship between somites and vertebral column in zebrafish. Development *129*, 3851-3860. https://doi.org/10.1242/dev.129.16.3851

**Morita, S., Horii, T., Kimura, M., Goto, Y., Ochiya, T., and Hatada, I.** (2007). One Argonaute family member, Eif2c2 (Ago2), is essential for development and appears not to be involved in DNA methylation. Genomics *89*, 687-696. https://doi.org/10.1016/j.ygeno.2007.01.004.

**Naiche, L.A., Holder, N., and Lewandoski, M.** (2011). FGF4 and FGF8 comprise the wavefront activity that controls somitogenesis. Proc Natl Acad Sci U S A *108*, 4018-4023. https://doi.org/10.1073/pnas.1007417108.

**Nitanda, Y., Matsui, T., Matta, T., Higami, A., Kohno, K., Nakahata, Y., and Bessho, Y.** (2014). 3'-UTR-dependent regulation of mRNA turnover is critical for differential distribution patterns of cyclic gene mRNAs. Febs J *281*, 146-156. https://doi.org/10.1111/febs.12582

**Niwa, Y., Shimojo, H., Isomura, A., Gonzalez, A., Miyachi, H., and Kageyama, R.** (2011). Different types of oscillations in Notch and Fgf signaling regulate the spatiotemporal periodicity of somitogenesis. Genes Dev *25*, 1115-1120. https://doi.org/10.1101/gad.2035311.

**Nobrega, A., Maia-Fernandes, A.C., and Andrade, R.P.** (2021). Altered Cogs of the Clock: Insights into the Embryonic Etiology of Spondylocostal Dysostosis. J Dev Biol *9*. https://doi.org/10.3390/jdb9010005.

**Nusslein-Volhard, C., and Wieschaus, E.** (1980). Mutations affecting segment number and polarity in Drosophila. Nature *287*, 795-801. https://doi.org/10.1038/287795a0.

**Oates, A.C.** (2020). Waiting on the Fringe: cell autonomy and signaling delays in segmentation clocks. Current Opinion in Genetics & Development *63*, 61-70. https://doi.org/10.1016/j.gde.2020.04.008.

**Oates, A.C., and Ho, R.K.** (2002). Hairy/E(spl)-related (Her) genes are central components of the segmentation oscillator and display redundancy with the Delta/Notch signaling pathway in the formation of anterior segmental boundaries in the zebrafish. Development *129*, 2929-2946. https://doi.org/10.1242/dev.129.12.2929

**Okamura, K., Liu, N., and Lai, E.C.** (2009). Distinct mechanisms for microRNA strand selection by Drosophila Argonautes. Molecular cell *36*, 431-444. https://doi.org/10.1016/j.molcel.2009.09.027.

**Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., et al.** (2004). Complete sequencing and characterization of 21,243 full-length human cDNAs. Nature Genetics *36*, 40-45. https://doi.org/10.1038/ng1285.

**Palmeirim, I., Henrique, D., Ish-Horowicz, D., and Pourquie, O.** (1997). Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. Cell *91*, 639-648. https://doi.org/ 10.1016/S0092-8674(00)80451-1

**Park, M.S., Phan, H.-D., Busch, F., Hinckley, S.H., Brackbill, J.A., Wysocki, V.H., and Nakanishi, K.** (2017). Human Argonaute3 has slicer activity. Nucleic Acids Research *45*, 11867-11877. https://doi.org/10.1093/nar/gkx916.

**Pascoal, S., Carvalho, C.R., Rodriguez-León, J., Delfini, M.-C., Duprez, D., Thorsteinsdóttir, S., and Palmeirim, I.** (2007). A Molecular Clock Operates During Chick Autopod Proximal-distal Outgrowth. Journal of Molecular Biology *368*, 303-309. http://dx.doi.org/10.1016/j.jmb.2007.01.089.

**Pasquinelli, A.E., Reinhart, B.J., Slack, F., Martindale, M.Q., Kuroda, M.I., Maller, B., Hayward, D.C., Ball, E.E., Degnan, B., Muller, P., et al**. (2000). Conservation of the sequence

and temporal expression of let-7 heterochronic regulatory RNA. Nature *408*, 86-89. https://doi.org/10.1038/35040556

**Peterson, S.M., Thompson, J.A., Ufkin, M.L., Sathyanarayana, P., Liaw, L., and Congdon, C.B.** (2014). Common features of microRNA target prediction tools. Front Genet *5*, 23. https://doi.org/10.3389/fgene.2014.00023.

**Psychoyos, D., and Stern, C.D.** (1996). Fates and migratory routes of primitive streak cells in the chick embryo. Development *122*, 1523-1534. https://doi.org/ 10.1242/dev.122.5.1523

**Raffaelli, A., and Stern, C.D.** (2020). Chapter Three - Signaling events regulating embryonic polarity and formation of the primitive streak in the chick embryo. In Current Topics in Developmental Biology, L. Solnica-Krezel, ed. (Academic Press), pp. 85-111. https://doi.org/10.1016/bs.ctdb.2019.10.001.

**Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., and Ruvkun, G.** (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature *403*, 901-906. https://doi.org/10.1038/35002607.

**Resende, T.P., Andrade, R.P., and Palmeirim, I.** (2014). Timing Embryo Segmentation: Dynamics and Regulatory Mechanisms of the Vertebrate Segmentation Clock. BioMed Research International *2014*, 718683. https://doi.org/10.1155/2014/718683.

**Resende, T.P., Ferreira, M., Teillet, M.-A., Tavares, A.T., Andrade, R.P., and Palmeirim, I.** (2010). Sonic hedgehog in temporal control of somite formation. Proceedings of the National Academy of Sciences *107*, 12907-12912. https://doi.org/10.1073/pnas.1000979107.

**Ribes, V., Le Roux, I., Rhinn, M., Schuhbaur, B., and Dolle, P.** (2009). Early mouse caudal development relies on crosstalk between retinoic acid, Shh and Fgf signalling pathways. Development *136*, 665-676. https://doi.org/10.1242/dev.016204.

**Riley, M.F., Bochter, M.S., Wahi, K., Nuovo, G.J., and Cole, S.E.** (2013). Mir-125a-5p-mediated regulation of Lfng is essential for the avian segmentation clock. Dev Cell *24*, 554-561. https://doi.org/10.1016/j.devcel.2013.01.024

**Saga, Y., and Takeda, H.** (2001). The making of the somite: molecular events in vertebrate segmentation. Nat Rev Genet *2*, 835-845. https://doi.org/10.1038/35098552.

**Sander, K., and Faessler, P.E.** (2001). Introducing the Spemann-Mangold organizer: experiments and insights that generated a key concept in developmental biology. Int J Dev Biol *45*, 1-11.

**Sawada, K., and Aoyama, H.** (1999). Fate maps of the primitive streak in chick and quail embryo: ingression timing of progenitor cells of each rostro-caudal axial level of somites. Int J Dev Biol *43*, 809-815.

**Scaal, M.** (2016). Early development of the vertebral column. Semin Cell Dev Biol *49*, 83-91. https://doi.org/10.1016/j.semcdb.2015.11.003.

**Schier, A.F., and Talbot, W.S.** (2005). Molecular genetics of axis formation in zebrafish. Annu Rev Genet *39*, 561-613. https://doi.org/10.1146/annurev.genet.37.110801.143752.

**Shabalina, S.A., Spiridonov, A.N., Spiridonov, N.A., and Koonin, E.V.** (2010). Connections between Alternative Transcription and Alternative Splicing in Mammals. Genome Biology and Evolution *2*, 791-799. https://doi.org/10.1093/gbe/evq058.

**Shatkin, A.J.** (1976). Capping of eucaryotic mRNAs. Cell *9*, 645-653. https://doi.org/10.1016/0092-8674(76)90128-8.

**Sheeba, C. (2011).** Molecular parallelisms between vertebrate limb development and somitogenesis. PhD Thesis. Universidade do Minho, Portugal. http://hdl.handle.net/1822/19672.

**Sheeba, C.J., Andrade, R.P., and Palmeirim, I.** (2012). Joint interpretation of AER/FGF and ZPA/SHH over time and space underlies hairy2 expression in the chick limb. Biol Open *1*, 1102-1110. https://doi.org/10.1242/bio.20122386.

**Sheeba, C.J., Andrade, R.P., and Palmeirim, I.** (2014). Limb Patterning: From Signaling Gradients to Molecular Oscillations. Journal of Molecular Biology *426*, 780-784. http://dx.doi.org/10.1016/j.jmb.2013.11.022.

**Sheeba, C.J., Andrade, R.P., and Palmeirim, I.** (2016a). Getting a handle on embryo limb development: Molecular interactions driving limb outgrowth and patterning. Semin Cell Dev Biol *49*, 92-101. https://doi.org/10.1016/j.semcdb.2015.01.007.

**Sheeba, C.J., Andrade, R.P., and Palmeirim, I.** (2016b). Mechanisms of vertebrate embryo segmentation: Common themes in trunk and limb development. Seminars in Cell & Developmental Biology. http://dx.doi.org/10.1016/j.semcdb.2016.01.010.

**Shi, Y.** (2012). Alternative polyadenylation: new insights from global analyses. RNA *18*, 2105-2117. https://doi.org/10.1261/rna.035899.112.

**Shin, C.** (2008). Cleavage of the star strand facilitates assembly of some microRNAs into Ago2-containing silencing complexes in mammals. Mol Cells *26*, 308-313.

**Skipper, M.** (2004). Time for segmentation. Nature Milestones Development. https://doi.org/

**Solnica-Krezel, L.** (2005). Conserved patterns of cell movements during vertebrate gastrulation. Curr Biol *15*, R213-228. https://doi.org/10.1016/j.cub.2005.03.016.

**Soto, X., Biga, V., Kursawe, J., Lea, R., Doostdar, P., Thomas, R., and Papalopulu, N.** (2020). Dynamic properties of noise and Her6 levels are optimized by miR-9, allowing the decoding of the Her6 oscillator. The EMBO Journal *39*, e103558. https://doi.org/10.15252/embj.2019103558.

**Stern, C.D.** (2021). Lewis Wolpert (1929-2021). Developmental Cell *56*, 1218-1220. https://doi.org/10.1016/j.devcel.2021.04.010.

**Stockdale, F.E., Nikovits, W., Jr., and Christ, B.** (2000). Molecular and cellular biology of avian somite development. Dev Dyn *219*, 304-321. https://doi.org/10.1002/1097-0177(2000)9999:9999<::AID-DVDY1057>3.0.CO;2-5.

**Summerbell, D.** (1974). A quantitative analysis of the effect of excision of the AER from the chick limb-bud. Development *32*, 651-660. https://doi.org/10.1242/dev.32.3.651.

**Summerbell, D., Lewis, J.H., and Wolpert, L.** (1973). Positional information in chick limb morphogenesis. Nature *244*, 492-496. https://doi.org/10.1038/244492a0.

**Tabin, C., and Wolpert, L.** (2007). Rethinking the proximodistal axis of the vertebrate limb in the molecular era. Genes Dev *21*, 1433-1442. https://doi.org/10.1101/gad.1547407.

**Tadros, W., and Lipshitz, H.D.** (2009). The maternal-to-zygotic transition: a play in two acts. Development *136*, 3033-3042. https://doi.org/10.1242/dev.033183.

**Taelman, V., Van Wayenbergh, R., Solter, M., Pichon, B., Pieler, T., Christophe, D., and Bellefroid, E.J.** (2004). Sequences downstream of the bHLH domain of the Xenopus hairy-related transcription factor-1 act as an extended dimerization domain that contributes to the selection of the partners. Dev Biol *276*, 47-63. https://doi.org/10.1016/j.ydbio.2004.08.019.

**Takahashi, Y.** (2005). Common mechanisms for boundary formation in somitogenesis and brain development: shaping the 'chic' chick. Int J Dev Biol *49*, 221-230. https://doi.org/10.1387/ijdb.041960yt.

**Tang, G.Q., and Maxwell, E.S**. (2008). Xenopus microRNA genes are predominantly located within introns and are diffferentially expressed in adult frog tissues via post-transcriptional regulation. Genome Research *18*, 104-112. https://doi.org/10.1101/gr.6539108.

**Tanzer, A., and Stadler, P.F.** (2004). Molecular Evolution of a MicroRNA Cluster. Journal of Molecular Biology *339*, 327-335. https://doi.org/10.1016/j.jmb.2004.03.065.

**Vasiliauskas, D., Laufer, E., and Stern, C.D**. (2003). A role for hairy1 in regulating chick limb bud growth. Developmental Biology *262*, 94-106. http://dx.doi.org/10.1016/S0012-1606(03)00360-9.

**Vasquez-Rifo, A., Jannot, G., Armisen, J., Labouesse, M., Bukhari, S.I.A., Rondeau, E.L., Miska, E.A., and Simard, M.J.** (2012). Developmental characterization of the microRNA-specific C. elegans Argonautes alg-1 and alg-2. PloS one *7*, e33750-e33750. https://doi.org/10.1371/journal.pone.0033750.

**Viebahn, C.** (1995). Epithelio-mesenchymal transformation during formation of the mesoderm in the mammalian embryo. Acta Anat (Basel) *154*, 79-97. https://doi.org/10.1159/000147753.

**Wahi, K., Bochter, M.S., and Cole, S.E.** (2016). The many roles of Notch signaling during vertebrate somitogenesis. Seminars in Cell & Developmental Biology *49*, 68-75. https://doi.org/10.1016/j.semcdb.2014.11.010.

**Wang, Y., Medvid, R., Melton, C., Jaenisch, R., and Blelloch, R.** (2007). DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nature Genetics *39*, 380-385. https://doi.org/10.1038/ng1969.

**Wienholds, E., Koudijs, M.J., van Eeden, F.J.M., Cuppen, E., and Plasterk, R.H.A.** (2003). The microRNA-producing enzyme Dicer1 is essential for zebrafish development. Nat Genet *35*, 217-218. http://www.nature.com/ng/journal/v35/n3/suppinfo/ng1251_S1.html.

**Wightman, B., Ha, I., and Ruvkun, G.** (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell *75*, 855-862. https://doi.org/10.1016/0092-8674(93)90530-4.

**Wilcox, A.J., Baird, D.D., and Weinberg, C.R.** (1999). Time of Implantation of the Conceptus and Loss of Pregnancy. New England Journal of Medicine *340*, 1796-1799. https://doi.org/10.1056/nejm199906103402304.

**William, D.A., Saitta, B., Gibson, J.D., Traas, J., Markov, V., Gonzalez, D.M., Sewell, W., Anderson, D.M., Pratt, S.C., Rappaport, E.F., and Kusumi, K.** (2007). Identification of oscillatory genes in somitogenesis from functional genomic analysis of a human mesenchymal stem cell model. Dev Biol *305*, 172-186. https://doi.org/10.1016/j.ydbio.2007.02.007.

**Winter, J., Jung, S., Keller, S., Gregory, R.I., and Diederichs, S.** (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. Nature Cell Biology *11*, 228-234. https://doi.org/10.1038/ncb0309-228.

**Xia, Z., Donehower, L.A., Cooper, T.A., Neilson, J.R., Wheeler, D.A., Wagner, E.J., and Li, W.** (2014). Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3′-UTR landscape across seven tumour types. Nature Communications *5*, 5274. https://doi.org/10.1038/ncomms6274.

**Yeom, K.H., Lee, Y., Han, J.J., Suh, M.R., and Kim, V.N.** (2006). Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing. Nucleic Acids Research *34*, 4622-4629. https://doi.org/10.1093/nar/gkl458.

**Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R.** (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes & Development *17*, 3011-3016. https://doi.org/10.1101/gad.1158803.

**Zhang, H., Lee, J.Y., and Tian, B.** (2005). Biased alternative polyadenylation in human tissues. Genome Biol *6*, R100. https://doi.org/10.1186/gb-2005-6-12-r100.

# Chapter 2

## Assessing the role of alternative *hairy1* transcripts in temporal control of tissue-specific embryo clock oscillations

*This page is intentionally left blank*

## 2.1 Chapter Introduction

Embryo segmentation is a conserved process among many animal organisms, from insects to vertebrates. To ensure proper segmentation, animals rely on intricate and complex genetic networks. In *Drosophila*, the genetic network that dictates the segmentation and subsequent identity of the body segments is initiated even before fertilization (Gonzalez-Reyes *et al.*, 1997). During oocyte maturation, the nurse cells produce mRNAs and proteins needed for the oocyte and the future embryo's first development steps. Before fertilization, the oocyte has different mRNAs in the anterior and posterior ends, such as *bicoid* and *nanos*, respectively. Upon fertilization, the translation of these transcripts is activated, generating antero-posterior morphogen gradients in the early syncytial embryo. This generates different combinations of proteins (which are transcription factors) along the A-P axis, which results in the activation of distinct Gap genes. The Gap genes are the first segmentation genes to be activated in the *Drosophila* zygote, and they are expressed in broad domains of the A-P axis. The different combination of Gap genes results in the activation of the Pair-rule genes, such as *hairy*. The Pair-rule genes are expressed in seven stripes, and each stripe is about the size of two segments. At last, the Pair-rule genes activate the expression of the Segment polarity genes, which are also expressed in stripes, but only one segment wide, totalizing 14 stripes. The different combination of Gap, Pair-rule, and Segment polarity genes in each segment activates a different Homeotic gene. The Homeotic genes are ultimately responsible for giving identity to each embryonic segment (Figure 2.1 A) (Gilbert and Barresi, 2019).

*Drosophila* pair-rule gene *hairy* is *a* basic Helix-Loop-Helix (bHLH) transcription repressor that is expressed in seven stripes, with an expression domain of around two segments wide (Figure 2.1 B). Although *hairy* gene can be transcribed into two different transcripts, one larger with 2335 nucleotides and a second with 2139 nucleotides, it only produces one protein of 337 amino acids (Rushlow *et al.*, 1989). Both mRNA forms differ in their 5'UTR, sharing the same coding sequence (Figure 2.1 C). Until these days, it is not known the functionality of both isoforms.

*Hes* genes in vertebrates can also present alternative 5'UTRs, generating alternative mRNA species. It is the case of the mouse Hes3 (Hirata *et al.*, 2000). Hirata and colleagues found that Hes3 has two transcription start sites (TSS) producing two distinct transcripts. Both can be translated into protein, however, with distinct roles. The longer transcript produces a smaller protein relevant to the cerebellum (Hes3a), and the shorter transcript is translated into a larger protein that is produced throughout embryo development (Hes3b). Hes3a lacks the basic domain, and it is unable to bind DNA. Because of that, it cannot bind promoter N-boxes and, consequently, cannot repress transcription directly (Hirata *et al.*, 2000).

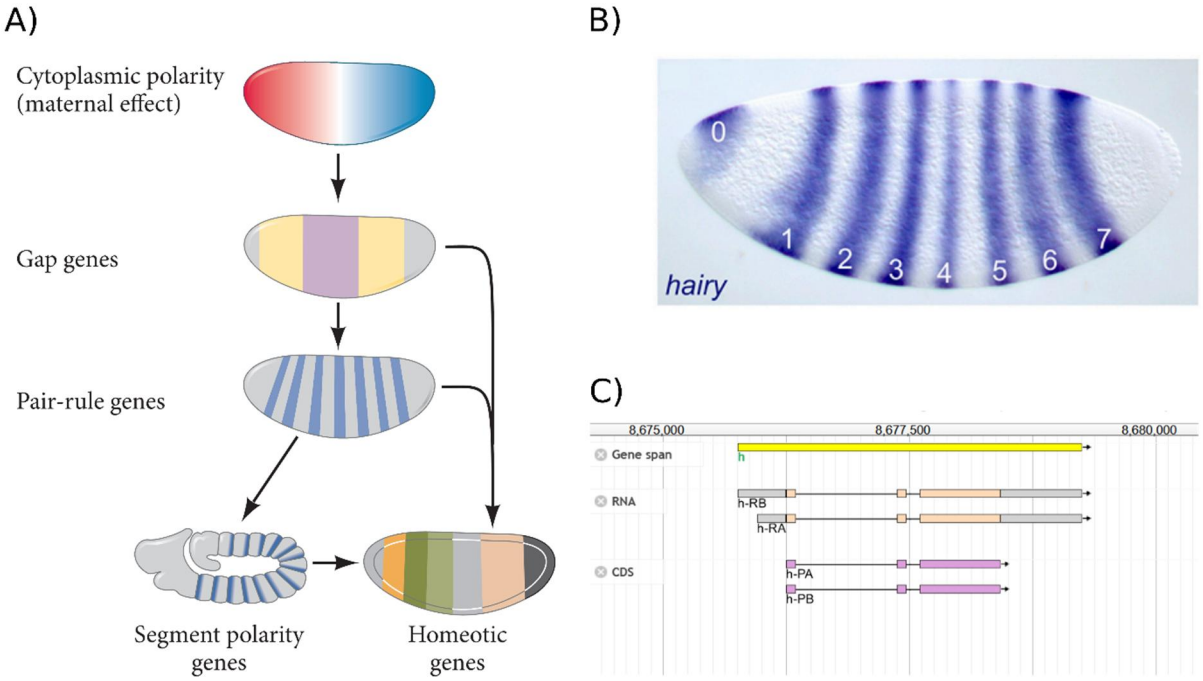



**Figure 2.1 - Drosophila antero-posterior patterning and the pair-rule *hairy* gene.**
**A)** Schematic representation of Drosophila antero-posterior patterning. mRNA deposited by the mother produces morphogenetic proteins that establish two counter gradients. The combination of these two gradients activates different Gap genes, which establish broad territories in the embryo. The Gap genes activate the Pair-rule genes, which in turn activate Segment-polarity genes, that divide the embryo into fourteen segments. The specific combination of Gap, Pair-rule, and Segment polarity genes, activate a specific homeotic gene. The homeotic genes give identity to each segment (Adapted from (Gilbert and Barresi, 2019)). **B)** Expression of the Pair-rule gene *hairy* by whole-mount *in situ* hybridization (Adapted from (Small and Arnosti, 2020)). **C)** *Drosophila hairy* has two transcription start sites, generating two different mRNAs (h-RA and h-RB). Since only the 5'UTR is altered, both transcripts have the potential to be translated into the same protein (Larkin et al., 2020).

In chicken, it was described already that *hairy1* has alternative protein isoforms (Vasiliauskas *et al.*, 2003). Vasiliauskas and collaborators found the *hairy* has two splicing variants (*hairy1A* and *hairy1B*). *hairy1B* encodes a protein with fourteen extra amino acids at the beginning of the basic domain. Although both are present in the limb mesenchyme, when they are over-expressed, *hairy1A* produces a stronger phenotype – shortening of the forelimb – than the misexpression of *hairy1B*. Limbs misexpressing *hairy1A* also have perturbation in cartilage mineralization and have delayed bone marrow development. The same was not evident upon misexpression of *hairy1B.* The authors postulated that the stronger phenotype of *hairy1A* misexpression could be due to one of four reasons: 1) Since the Hairy1B fourteen extra nucleotides are in the basic domain, it has less affinity for DNA than Hairy1A; 2) Hairy1B

totally lost the capacity to bind DNA, however, it has a function independent of DNA binding; 3) Hairy1B could have a dominant-negative function by heterodimerizing with Hairy1A or other proteins; 4) The targets of Hairy1A could be different from the targets of Hairy1B (Vasiliauskas *et al.*, 2003).

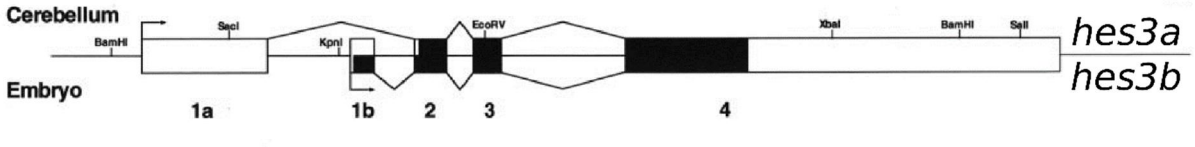



**Figure 2.2- Mouse *Hes3* gene architecture.**
The mouse *Hes3* locus produces two transcripts with different TSS. One longer, that gives rise to a shorter protein present in the cerebellum and another shorter transcript translated into a bigger protein. Adapted from (Hirata *et al.*, 2000).

At the time of Vasiliauskas and collaborator's work, the existence of an embryo clock working in the limb mesenchyme was still elusive. Nowadays, it is known that another Hes gene is expressed in the forelimb, *hairy2* (Pascoal *et al.*, 2007). Moreover, it was found that this gene has a cyclic behavior with a periodicity of six hours. Two cycles of *hairy2* expression correspond to the precise time for a bone segment to be formed (Pascoal *et al.*, 2007). Curiously, this gene is also involved in the segmentation clock that is working in the PSM, but with a pace of one hour and thirty minutes, four times faster than in the forelimb (Jouve *et al.*, 2000).

In this chapter, we explore the hypothesis that alternative transcripts of *hairy1* in the PSM and the forelimb mesenchyme are responsible for the different paces of the EC. We searched for new 5' and 3' alternative transcripts of *hairy1* and characterized the ones that could influence the pace of the EC in these two tissues.

## 2.2 Material and Methods

### Bioinformatic search 3'UTR

The bioinformatic search for alternative 3'UTRs was performed by using the Geneprox web tool APADB (Müller *et al.*, 2014). It was accessed via http://tools.genxpro.net/apadb/browse/, and in the options it was chosen chicken as organism,

to search in the whole organism, and at last, we chose HES1 to search for alternative 3'UTR of *hairy1.*

### Embryos

Fertilized *Gallus gallus* eggs were obtained from a commercial source (Pintobar Exploração Avícola, Lda, Portugal) and were incubated at 38ºC in a humidified atmosphere for up to four days until they reached the desired Hamburger and Hamilton (HH) stage (Hamburger and Hamilton, 1951).

### Tissue collection

Caudal portions were isolated from embryos at stages HH12-13. Embryos were collected from fertilized eggs incubated for 48h, placed dorsal side up in phosphate buffer saline (PBS) solution. Caudal portions (embryonic tissue compressed from the tailbud until the first somite border) were isolated using a dissecting knife. Sixty pairs of PSM portions were pooled together for each sample. The samples were snap-frozen in liquid nitrogen and stored at -80ºC.

Forelimbs were isolated from embryos at stages HH20-22. Embryos were collected from fertilized eggs incubated for four days. Embryos were placed in PBS, and the limb buds were manually dissected using forceps. Twenty forelimb pairs were pooled together for each sample, snap-frozen in liquid nitrogen, and stored at -80ºC.

### RNA extraction

Biological samples were defrosted on ice. Total RNA was extracted using RNAqueous™-Micro Total RNA Isolation Kit (Invitrogen™ #AM1931) according to the manufacturer's instructions. RNA was resuspended in 20 µL of Elution Solution, and it was immediately quantified using NanoDrop 2000 (Thermo Scientific) and stored at -80ºC.

### RNA quality control

RNA quality control was performed using Experion™ RNA StdSens Analysis Kit (BioRad #700-7103). Only samples with an RQI (RNA Quality Indicator) equal to or above nine were used for subsequent steps.

**Rapid Amplification of cDNA Ends (RACE)**

5' and 3' were performed using FirstChoice™ RLM-RACE Kit (Invitrogen™ AM1700M) according to the manufacturer's instructions for 10µg of total RNA as starting material. During the 5'RACE execution and after Calf-intestinal alkaline phosphatase (CIP) treatment, phenol:chloroform (Sigma-Aldrich #P2069) was used to stop the reaction and purify the Caped RNA.

PCR was performed using *hairy1* specific primers that were designed using primer BLAST (Ye et al., 2012) with attention to FirstChoice™ RLM-RACE Kit guidelines. Sequences of the primers used are in Table 2.1

**Table 2.1 - Specific PCR primers used for *hairy1* PCR after RACE**

| | |
|---|---|
| hairy1_5'_Outer | 5'-GGCGTTGGCGTAGAGTGGGA-3' |
| hairy1_5'_Inner | 5'-GTGGTGGTAGGAGGCAGTTGGAC-3' |
| hairy1_3'_Outer | 5'-CCATGAACTACCTGCCACCC-3' |
| hairy1_3'_Inner | 5'-GTCCAACTGCCTCCTACCAC-3' |
| hairy1_Control | 5'-TGACGGTGAAGCACCTGAGGAAC-3' |

Outer and Inner PCRs were performed using the FirstChoice™ RLM-RACE Kit guidelines and were performed using DreamTaq DNA Polymerase (Thermo Scientific™ #EP0701) and run in a C1000 Touch™ Thermal Cycler (Bio-Rad #1851196). 5µL of Outer and Inner PCR were run in 2% agarose gel with the DNA GreenSafe Premium (Nzytech #MB13201). A ChemiDoc XRS+ Imaging system (Bio-Rad #1708265) acquired gel images.

**Protein extraction and quantification**

Tissues previously collected and stored at -80ºC were defrosted on ice (entire HH22 embryo, pool of twelve HH12-13 embryos, pool of 20 pairs of forelimbs, and pool of 60 caudal portions). It was added 50µL of lysis buffer (Kling *et al.*, 2002) and homogenized using Kontes™ Pellet Pestle™ Cordless Motor (DWK Life Sciences Kimble™ #K749540-0000). Protein quantification was performed by using Bio-Rad Protein Assay (Bio-Rad #5000006). BSA (Sigma #9048-46-8) was used to produce the standard curve. Quantification was performed using NanoDrop 2000 (Thermo Scientific).

**Western Blot**

10µg of protein of each condition was mixed with 2x Laemmli Sample Buffer (Bio-Rad # 1610737) and β-Mercaptoethanol and denatured for 10 minutes at 100ºC. Denatured protein was loaded in a 12% SDS-PAGE minigel and subjected to electrophoresis at a constant voltage of 100V for four hours. Proteins were then transferred from the gel to Hybond-C extra Nitrocellulose membrane (Amersham # GERPN303D) for 1 hour at a constant electric current of 0.25A. Transfer efficiency was performed by staining the membrane using a Pauceau (Sigma-Aldrich # P3504) solution. Blot was probed using a custom-made monoclonal mouse antibody raised against recombinant chicken Hairy1 (8c9) diluted 1:200 in a BSA solution overnight at 4ºC. After blots were probed with a secondary rabbit anti-mouse IgG HCL (Abcam #ab6728) diluted 1:4000 in milk for 45 minutes. Blots were developed using Super Signal West Femto Substrate (Thermo Fisher # 34095). A ChemiDoc XRS+ Imaging system (Bio-Rad #1708265) acquired the blot signal.

**Electroporation**

Electroporation was performed as described (Voiculescu *et al.*, 2008). HH4 embryos were placed into filter paper and removed from the egg as described (Chapman *et al.*, 2001). The embryo was then placed in CUY700-P2E electrode (NEPA GENE) filled with Tyrod's saline solution. The electroporation solution containing the vector DNA at 0.5 to 1 µg/µL with Fast Green (Sigma-Aldrich #F7252) was injected between the vitelline membrane and the epiblast, under the prospective PSM tissue. Finally, anode electrode CUY700-P2L (NEPA GENE) was placed onto the Tyrod's saline solution surface, and five 50ms pulses of 9V with 350ms intervals were applied using BTX ECM-830 ElectroSquarePorator 22 (BTX Harvard Apparatus) (Figure 2.3 A). Electroporated embryos were cultured at 38ºC in a humidified atmosphere as described by (Chapman *et al.*, 2001). Limb electroporation was performed by optimizing the *in ovo* forelimb electroporation described by (Ueda *et al.*, 2017)

**Plasmids**

The empty vector (pCAT) and the *hairy1* plasmid (pCAT-hairy1) were previously constructed in the laboratory (Figure 2.3 B and C) (Andrade et al., unpublished). The plasmid used for overexpressing *s-hairy1* (pCAT-s-hairy1) (Figure 2.3 D) was synthesized by removing the exclusive region of the *hairy1* transcript from the plasmid pCAT-hairy1 (digestion with restriction enzymes BstSI and XcmI). The plasmid was then sent for sequencing to ensure that it only would express the smaller transcript. These plasmids were optimized for chicken codon usage, and they have a ubiquitous chicken promoter (chicken β-actin promoter). The transcript

of interest is transcribed with GFP. After the stop codon of our genes of interest, it has an IRES (Internal Ribosome Entry Site) followed by GFP. With GFP, it is possible to verify the success of the overexpression.

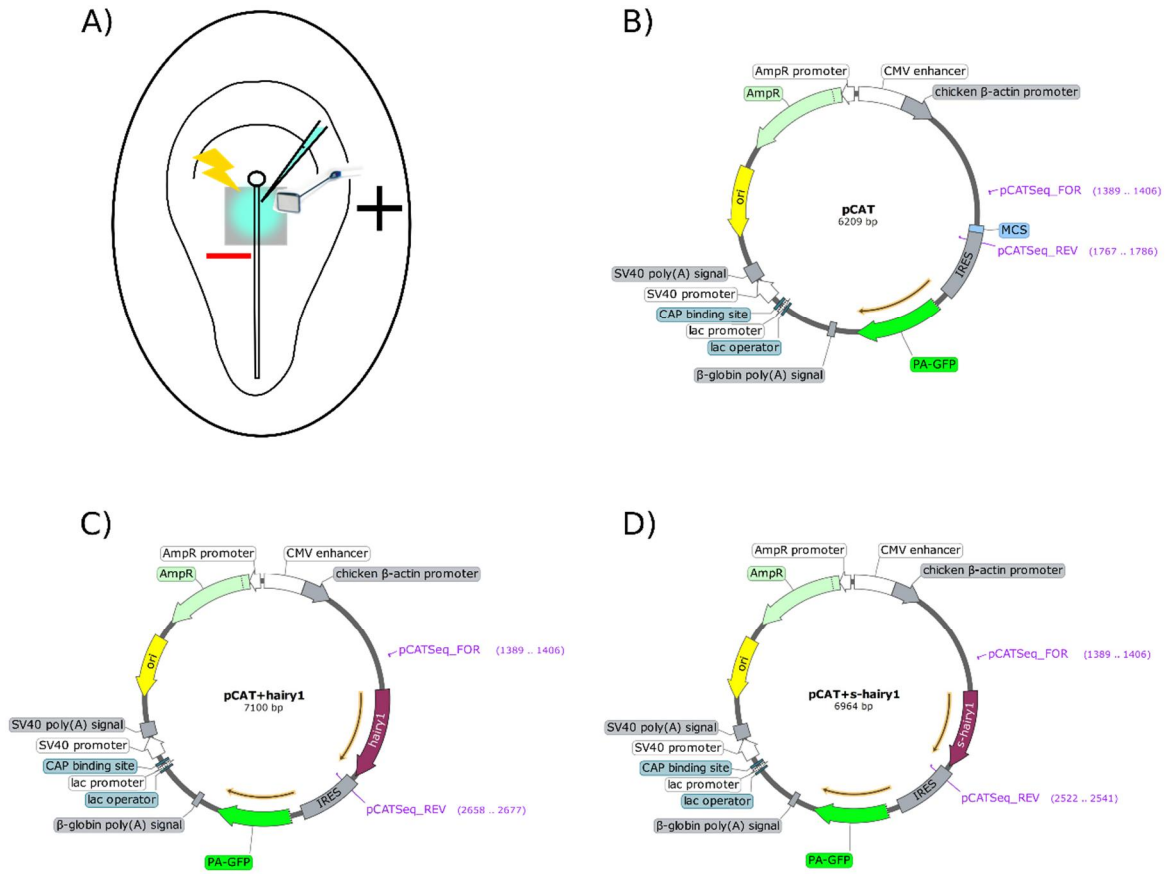



**Figure 2.3 - HH4 embryo electroporation and vectors used.**
**A)** Electroporation scheme of an HH4 embryo, evidencing the electroporation area. **B)** Empty pCAT-GFP vector, used as control. **C)** pCAT-hairy1 vector, used to overexpress *hairy1* in the PSM precursors. **D)** pCAT-s-hairy1 vector, used to overexpress *s-hairy1* in the PSM precursors.

### Whole-mount *in situ* hybridization

Whole-mount *in situ* hybridization was performed as described by (Henrique et al., 1995). Digoxigenin-labeled RNA probes were synthesized by *in vitro* transcription of linearized previously described plasmids: *dll1* (Henrique *et al.*, 1995); *serrate1* (Myat *et al.*, 1996); *hairy1* (Palmeirim *et al.*, 1997); *lfng* (Sakamoto *et al.*, 1997); *HoxB9* (Nelson *et al.*, 1996).

**Imaging**

All embryo imaging (Electroporated embryos, whole embryos, and forelimbs from *in situ* hybridization and Alcian blue-stained forelimbs) were imaged using Zeiss LUMAR V12 Stereoscope coupled to a Zeiss Axiocam 503 color 3MP camera. All measurements performed in the embryo images were performed in Zeiss AxioVision version 4.9.1 software.

For time-lapse imaging, the fluorescence of electroporated embryos was evaluated five hours post electroporation. The embryo with the highest GFP expression in the prospective PSM was selected for time-lapse imaging. The embryo was placed in a Stage Top Incubator (Oko-Lab) preheated at 38ºC and with controlled humidity. The camera was placed under Zeiss LUMAR V12 Stereoscope coupled to a Zeiss Axiocam 503 color 3MP camera that acquires one image every three minutes during six to sixteen hours.

**Chromatin immunoprecipitation**

HH8 embryos were bisected and half was immediately fixed and the other half was cultured for thirty or ninety minutes. Afterward, the second half was also collected. Haves were fixed using a fresh formaldehyde (Sigma-Aldrich # 47608) solution 1.27% for 5 min. Quenching of formaldehyde with glycine 2.5M (1/20) was performed for 5 minutes at room temperature. The collected halves were then stored at -80ºC. Chromatin was purified and immunoprecipitated using True MicroChIP kit (Diagnose), according to the manufacturer's instructions, with slight optimizations for using embryonic tissues. During the protocol chromatin was sheared to fragments of 100-600 bp using Bioruptor Pico (Diagenode). The optimal sonication time for chromatin of half embryo was one minute. It was used a custom monoclonal mouse Antibody against chicken Hairy1 (8c9).

**Chromatin qPCR**

Quantification of the DNA from *HoxB9* promotor precipitated with Hairy1 antibody was performed by qPCR using CFX96$^{TM}$ system (Bio-Rad). Nine different *HoxB9* primer pairs encompassing ~7kb were used in order to evaluate the dynamics of Hairy1 binding to the *HoxB9* promotor. qPCR was normalized to the input chromatin before immunoprecipitation. The reactions were performed using SsoFast_EvaGreen Supermix (BioRad) and CFX96 equipment (BioRad).

**Table 2.2 - qPCR primers for different locations of *HoxB9* promoter**

| Primer pair name | Forward (5' to 3') | Reverse (5' to 3') |
|---|---|---|
| Ch_HoxB9_1 | TGGCTTTTCCTTCAGAGGTT | GGTTTGTCTGGAGGTGTCCT |
| Ch_HoxB9_2 | AAGGGGAGGAGAGGGAAAGT | GCGCGTTGTTTCTAAGGAAG |
| Ch_HoxB9_3 | GAAAGAGCCCGTCTGAAAGA | AAAGGGCACAGGTACGAGAA |
| Ch_HoxB9_4 | TCTTCTCACCACTTACGGATTTT | AGCCTGAAAGAAGAGGGTTCT |
| Ch_HoxB9_5 | GCAGCTGTGTGTTTCAAAGG | AAACCCAGAGGCTTCCTAGC |
| Ch_HoxB9_6 | AACAAACGGAAACCTCCAGA | TGAGGCAAAGAGCAGGAGTT |
| Ch_HoxB9_7 | ATGGACAGCTTCTTCCCAGC | CCCATCAGCGGTGTTTTTCG |
| Ch_HoxB9_8 | CGCAATAAACGCATTTGTCA | CCTGGACGAGATGATTGA |
| Ch_HoxB9_9 | AACATTTCGGCAGCAAAATC | ATGCAATGTTGCAAAACGTC |

### Cartilage staining

Limbs in HH26 were collected in 1X Phosphate buffer solution (PBS) and were dehydrated in 95% ethanol solution overnight at 4ºC. Afterwards, forelimbs were incubated with a 0.02% Alcian blue (Sigma-Aldrich #33864-99-2) solution, overnight at room temperature. Forelimbs were transparentized with a Glycerol solution containing KOH until the cartilage surrounding tissue was completely transparent.

### Statistical analysis

Paired Student's t-test was performed to validate the difference of *HoxB9* expression length of the electroporated embryos. The differences were considered significant if the p-value < 0,05. GraphPad Prism 5 was used to perform the statistical analyses.

## 2.3 Results

### 2.3.1 Identification of tissue-specific alternative *hairy1* transcripts

*hairy1* is one of the best characterized chicken embryo clock genes (Palmeirim *et al.*, 1997). However, little is known about the transcripts it produces. Vasiliauskas and colleagues showed that *hairy1* has an alternative splice variant. However, it is present in both PSM and forelimb (Vasiliauskas *et al.*, 2003). So, the mere existence of this second variant cannot explain, *per se*, the different paces of the embryo clock. In order to deeper explore the

possibility of that alternative *hairy1* transcripts may underlie distinct clock oscillation periodicities, we decided to look to other types of variants, alternative TSS and alternative polyadenylation sites (PAS).

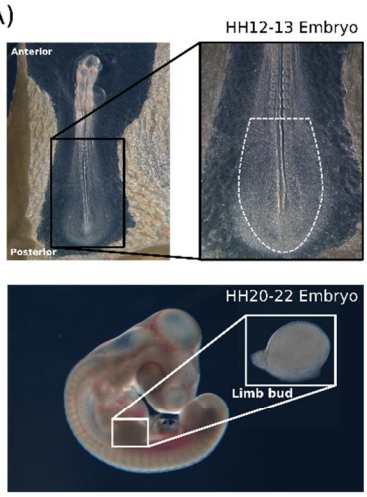
To find new 5' and 3' alternative transcripts and infer if any of them is specific to the tissues where the EC is operating, we designed a RACE (Rapid Amplification of cDNA Ends) experiment. For this experiment, caudal portions of embryos were collected from the last somite until the end of the tailbud, as well as entire forelimb buds. The tissues were collected from embryos at stage HH12 to HH13 and HH20 to HH22, respectively (Figure 2.4 A). We identified both alternative 5' and 3' transcripts in each tissue. Regarding the 5'RACE, we identified an alternative smaller transcript in the caudal portion of the somitogenesis stage embryos (Figure 2.4 B). It is 294 nucleotides shorter than the canonical transcript and is only present in the caudal portion, the tissue where the clock has a faster pace.

By 3'RACE, we found an alternative polyadenylation site. These two polyadenylation sites produce two alternative 3'UTR with a length of 417 and 282 nucleotides. Contrarily to the alternative 5'end transcript, in this case, both 3'ends were found in the analyzed tissues (Figure 2.4 B and C).

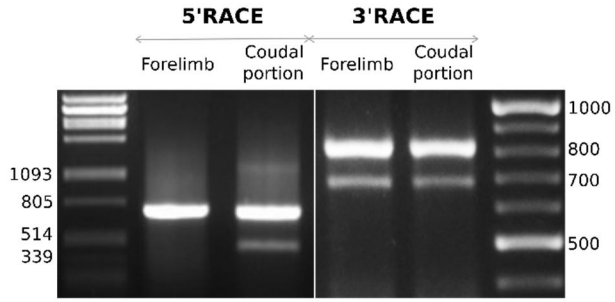
Using genxpro's tool, APADB, we looked for *hairy1* alternative 3'UTRs based on RNA sequencing data (http://tools.genxpro.net/apadb/browse/chicken/whole/HES1/)(Müller *et al.*, 2014). The reads deposited in the database match two possible 3'UTRs, a bigger one with 417 nucleotides and a smaller, less abundant, with 282 nucleotides (Figure 2.4 D). This data confirms the existence of the two 3'UTRs identified by 3'RACE.

The presence of alternative *hairy1* mRNAs suggests that alternative protein forms could be responsible for finetuning the EC periodicity in different tissues. In order to pursue this hypothesis, we performed Western blot analysis in the same tissues to find if there are any different protein forms. No commercial chick-specific antibodies were available, so we used a costume-made antibody available in the lab, but that has not been yet optimized. The assay resulted in a large amount of bands obtained and we could not identify alternative forms of Hairy1 protein between the tissues analyzed (Figure 2.5 A). Despite many optimization steps taken, we concluded that the antibody used was highly unspecific, which can mask alternative protein forms of Hairy1. New antibodies are now available in the lab and are presently being optimized. Once the antibody optimization is completed, this experiment should be repeated to find Hairy1 protein forms.

A)




B)




C)

```
ATTGCCGGCCCCGCCGCCGCCAGCGTCGCACGCGTCCGTTGCCCTTTCGGAGCGCCTTCCTTCCCGCGCTTCCTTTTTTTTTTTTT
TTTAATCCCTTTATTTATTTTTCTTTCCGATTTCTTCGTTTTTCACTTAATTTCCTTTGGTTATTTTAATTTCTCTTTTTACCTCG
CCATGCCCGCCGACACGGGCATGGAAAAACCCACCGCTTCCCCCATCGCGGGGGGCACCGGCCAGCGCCAGTCACACGCCGGACAAA
CCTCGGAGCGCCAGCGAGCACCGGAAGTCCTCCAAACCCATCATGGAGAAGCGACGGCGGGCGCGCATCAACGAGAGCCTGGGGCA
GCTGAAGATGCTCATCCTGGACGCGTTGAAAAAAGACAGCTCTCGGCACTCCAAGCTGGAGAAGGCGGACATCCTGGAGATGACGG
TGAAGCACCTGAGGAACCTCCAGAGGGCACAGATGGCAGCCGCTCTCAGCGCTGACCCCAGCGTGCTGGGCAAATACCGGGCTGGT
TTCAATGAGTGCATGAACGAGGTGACGAGGTTCCTGTCCACCTGCGAGGGGGTGAACGCCGACGTGCGAGCCCGCCTGCTGGGCCA
CCTCTCTGCTTGCCTGGGGCAGATCGTGGCCATGAACTACCTGCCACCCCCTCCTGCTGGCCAACCCGCACATCTGGCTCAACCCC
                          3'Outer Primer                    5'Inner Primer
TGCACGTCCAACTGCCTCCTACCACCACCGGAGCGGTGCCGGTGCCGTGCAAGCTGGAACCGACCGAAGCTCTCTCCCCTAAAGTC
      3'Inner Primer                                          5'Outer Primer
TATGGGGGTTTCCAGCTCGTCCCGGCCACTGACGGTCAATTCGCCTTCCTCATCCCCAACCCTGCGTTCCCACCTGGCTCTGGACC
GGTTATCCCACTCTACGCCAACGCCAACGTCCCAGTGTCCACAAGTGGAGGCTCTGGGAACGCGTCCACCACCCCATCAGCATCCC
CGGTGCAGGGCTTGACATCATTTGGGCACAGCGTGGTGCCGGCATCCCAGGCAGGCAGCCCTATTGCTGAGCGCAGGGAGGCCGTC
TGGCGGCCGTGGTAGCCGTCCCATCCGTCCCTCATCCCAAAAACTCTGCCTTGGCTCCCGTTGTATGRAGCCCCGTTGTGGTTGTT
GTATTTTCCTAAGCAGATTTCAACACGGGGAAAGGATTAAAAAGGACCTCGCTAGGAACCGTGCWATTTATTAATCTCAGAGGTTG
GGATCTCGATTTGTATTTTTACTCCTTTTGAAATGCCTTTATCTGAGATAYGTTTGTGTTTTTTTTTTGGTTGGTTTTTTNNTTTA
TTTTTAARGAGAAATAAATGAGAAATAGTTTTGTAACAGATAACTCCATCATGCCAAAGTTGTTTGTATCCCGTCGGTCCGTGCGT
GTGTTGCGTGTGCCTGTGTTGGAGACCTCCGAGAAAAAGAGAACAAAATGATACCGCAGGTGTTGAAATAAATTGCTCTTTGAAAT
AGAAAAAAAAAAA
```

D)

To see PA sites, click on the number below

| Gene_Symbol | Chromosome | Start | End | Location | Strand | Read_Median | Read_Mode | Reads_Supporting_Site |
|---|---|---|---|---|---|---|---|---|
| HES1 | chr21 | 2805636 | 2805643 | UTR3 | + | 2805641.0 | 2805641 | 1643 |
| HES1 | chr21 | 2805477 | 2805517 | UTR3 | + | 2805506.0 | 2805506 | 369 |

2 chicken whole detail entries

**Figure 2.4 - Chicken *hairy1* alternative transcripts.**
**A)** Representation of the tissues used for RACE experiments. Caudal portions of HH12 to HH13 embryos and forelimbs of HH20 to HH22 embryos were collected and pooled. **B)** Agarose gel electrophoresis results of 5'RACE and 3'RACE assays. A small *hairy1 (s-hairy1)* isoform that differs in the 5' UTR was identified. This isoform seems to be specific for the caudal tissue. In the 3'RACE, two distinct 3'UTRs were identified in both tissues analyzed. **C)** *hairy1* sequence identifying the 5'UTR (dark blue) and the 3'UTR (purple). The first ATG codon (red) encodes the first methionine of the larger transcript, corresponding to the canonical protein. The second and the third are putative translation initiation sites of *s-hairy1*. Underlined brown are the first ten nucleotides of s-hairy1; underlined green, last ten translated nucleotides; underlined light blue shows the last ten nucleotides of the smaller 3'UTR; underlined black, primers used in the RACE experiment. **D)** Alternative 3'UTRs retrieved from RNA sequencing data using genxpro's tool, APADB, (http://tools.genxpro.net/apadb/browse/chicken/whole/HES1/).

The smaller 5' transcript (which we coined *s-hairy1*) lacks the first exon and the beginning of the second exon (Figure 2.5 B). From sequence analysis alone, we speculate that it is difficult for this transcript to be translated. The first ATG is present at the beginning of the transcript (six nucleotides after the beginning of the transcript), making it highly improbable that the translation machinery will bind the *s-hairy1* and start the translation. However, 48 nucleotides after, there is a second in frame ATG (Figure 2.5 C). If translated into a protein starting from this ATG codon, s-Hairy1 would lack the basic domain and the part of the Helix-Loop-Helix (HLH) domain (Figure 2.5 C). For that reason, it would not be able to bind directly to the DNA. This is similar to the 5' alternative transcript of the mouse Hes3 (Hirata *et al.*, 2000). An alternative hypothesis, yet to be explored, is that *s-hairy1* is a long non-coding RNA (lncRNA).

By considering this data, we hypothesize that the alternative transcript may have a role in the EC finetuning and be, at least, partially responsible for the difference of EC pace between the somitogenesis and forelimb outgrowth.

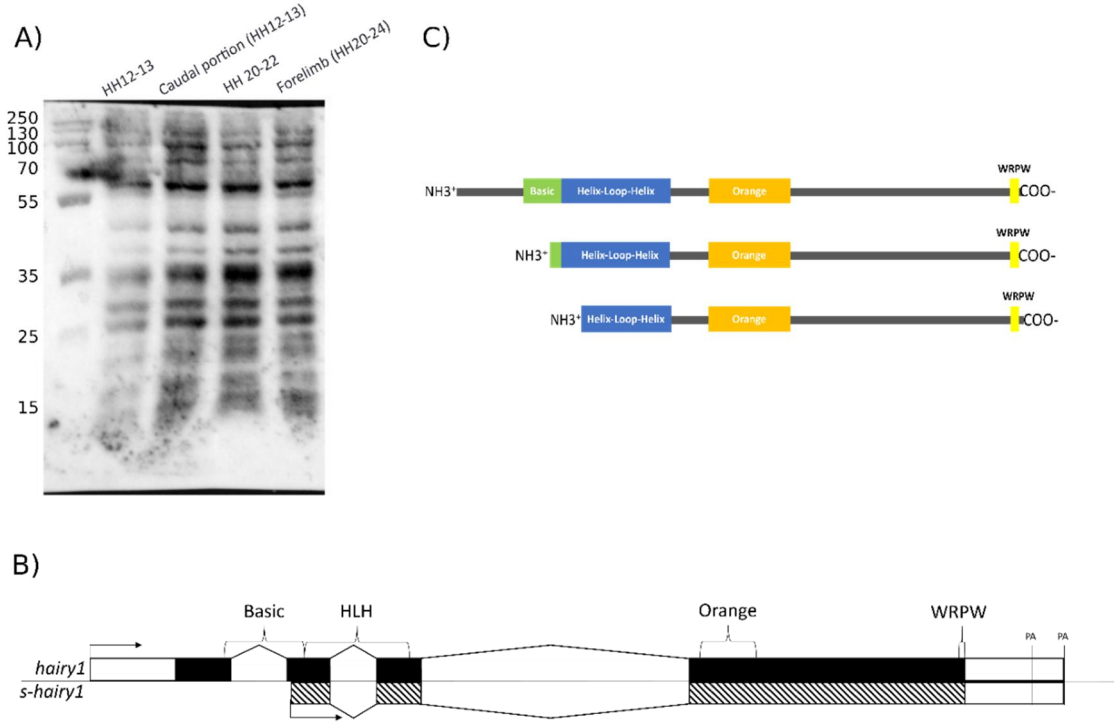



**Figure 2.5 - *hairy1* gene architecture and putative protein forms generated.**
**A)** Western-blot for Hairy1. It was used four different tissues: first lane, protein ladder; second lane, total protein extracts from HH12 to HH13 embryos; third lane, protein extracts from caudal portions of HH12 and HH13 embryos; fourth lane, protein extracts from whole embryos HH20 to HH22; fifth lane, protein extracts from forelimbs of HH20 to HH22 embryos. **B)** *hairy1* gene architecture: *hairy1 locus* is able to produce more than one transcript. It displays two distinct TSS and two polyadenylation signals. White boxes represent the RNA that is not translated;

Black boxes represent exons; Dashed boxes, possible translated RNA. **C)** Possible protein isoforms translated from *hairy1* (upper scheme) and *s-hairy1* (middle and lower schemes, depending on the ATG codon used).

### 2.3.2 Functional characterization of *hairy1* transcripts

To assess the functional role of *hairy1* transcripts, each transcript, *hairy1,* and *s-hairy1,* was overexpressed in the PSM progenitors in primitive streak stage embryos (HH4). Figure 2.3 (Material and Methods) illustrates the plasmids used for overexpression of *hairy1* and *s-hairy1*.

Electroporated embryos were cultured for 24 hours using the EASY culture technique (Chapman *et al.*, 2001). Only the embryos with a normal head phenotype, fluorescence in the paraxial mesoderm tissues, and absence of fluorescence in the head were collected and used for further experiments. To ensure that *hairy1* and *s-hairy1* were effectively overexpressed in the electroporated embryos, we performed whole-mount *in situ* RNA hybridization for *hairy1* using an antisense RNA probe capable of detecting both *hairy* and *s-hairy1* RNAs. In contrast to the empty vector, where we only detected the endogenous, non-perturbed expression of *hairy1* (Figure 2.6 A-C), in both *hairy1* and *s-hairy1* overexpressing embryos, ectopic expression was detected in the electroporated tissues (Figure 2.6 D-I), coincident with the presence of GFP signal (Figure 2. e, f, h, and i).

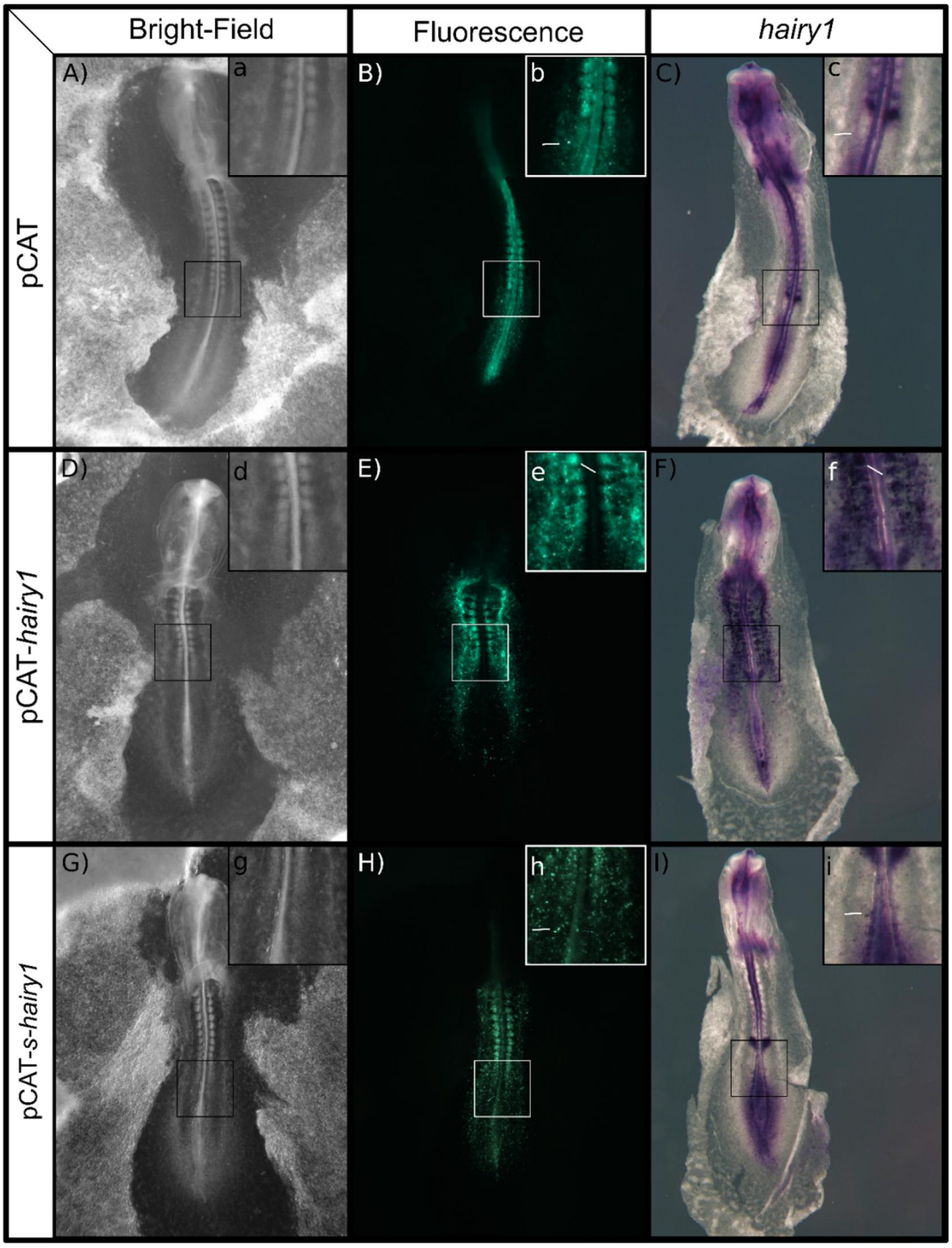



**Figure 2.6 - Tissue electroporation using pCAT-hairy1 and pCAT-*s*-hairy1 plasmids leads to overexpression of *hairy1* and *s-hairy1*, respectively.**
**A - C)** Electroporation of the empty vector (pCAT). A) Bright field imaging 24 hours after the electroporation; B) GFP expression 24 hours after electroporation; C) Whole-mount *in situ* hybridization for *hairy1* expression; **a** - **c**, magnifications of the area inside of the square in the whole embryo; White arrows, shows that the presence of GFP in b does not correspond to *hairy1* over-expression in c. **D - F)** Electroporation of a plasmid containing *hairy1* sequence (pCAT-hairy1). D) Bright field imaging 24 hours after the electroporation; E) GFP expression 24 hours after electroporation; F) Whole-mount *in situ* hybridization for *hairy1* expression; **d** - **f**, magnifications of the area inside of the square in the whole embryo; White arrows, shows that bright GFP in e corresponds to ectopic expression of *hairy1* in f. **G - I)** Electroporation of a plasmid containing s-*hairy1* (pCAT-s-hairy1). G) Bright field imaging 24 hours after the

electroporation; H) GFP expression 24 hours after electroporation; I) Whole-mount *in situ* hybridization for *hairy1* expression; **g** - **i**, magnifications of the area inside of the square in the whole embryo; White arrows, shows that the bright spots of GFP in e correspond to an ectopic expression of s-*hairy1* in f.

Embryos expressing GFP in the mesodermal precursors 5 hours after electroporation were selected to perform time-lapse imaging during 6 to 16 hours. This analysis evidenced that embryos overexpressing *hairy1* or *s-hairy1* have a delay in the initiation of body segmentation (Figure 2.7). The headfold length was used to temporally align the electroporated embryos since this tissue was not targeted by the electroporation procedure. Also, unpublished work from our laboratory showed that the headfold length has linear growth over time and can be reliably used to stage embryos over time (Maia-Fernandes and Pais-de-Azevedo *et al.*, manuscript in preparation).

The pCAT electroporated embryos developed as they would have *in ovo*, with the number of formed somites corresponding to the expected headfold length over time (Figure 2.7 A). However, the embryos electroporated with either pCAT-hairy1 or pCAT-s-hairy1 formed fewer somites: when they reached headfold lengths comparable to the control embryos with three somites formed, they still did not have any somite (Figure 2.7 B and C, first panels). When the overexpressing embryos had three somites, the empty vector already presented five somites, and when they formed the fourth somite, the control embryos formed the sixth somite. This data suggests that overexpression of *hairy1* or *s-hairy1* in the PSM precursors delays the beginning of somitogenesis. However, over time and in both conditions, segmentation is resumed with the same periodicity as the control embryos. This has previously been seen in our lab, but not for s-hairy1. The data suggest that *s-hairy1* is functional and has a role in the segmentation process and possibly in the embryo clock.

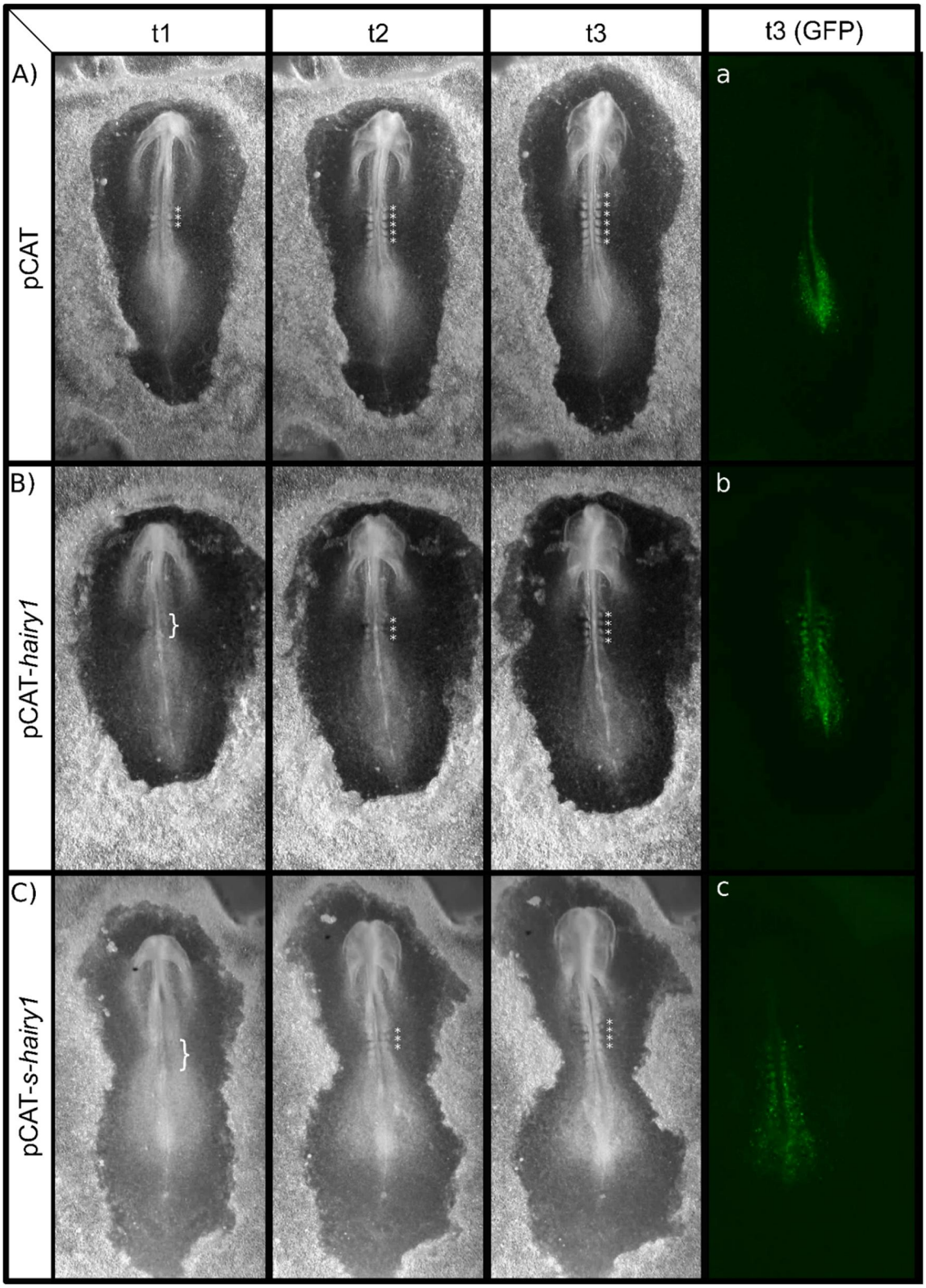



**Figure 2.7 - Overexpression of *hairy1* or *s-hairy* delay somitogenesis.**
**A-C)** Timelapse images of embryos electroporated with the empty vector (pCAT) (A), pCAT-*hairy1* (B) and pCAT-s-*hairy1*(C). **a-c** show the GFP localization in the third frame. The frames were chosen based on the length of the headfold. Asterisks show the somites in each frame; Brackets mark the tissue that should be segmented in comparison with the empty vector. The full videos are in the digital supplementary data.

We further questioned if only the segmentation was perturbed or if the development of the trunk was overall delayed. To clarify this question, several measurements were made in frames obtained with 1-hour intervals, as exemplified in Figure 2.8 A: 1 – Head-fold (Figure 2.8 C). 2 – From the most anterior part of the head until the end of the primitive streak (Figure 2.8 D); 3 – From the most anterior part of the head until Hensen's node (Notochord) (Figure 2.8 E); 4 – From Hensen's node until the end of the primitive streak (Figure 2.8 F); In all the graphs, the blue lines refer to embryos injected with the empty vector (pCAT), the pink lines represent the embryos electroporated with pCAT-*hairy1,* and in yellow, the embryos electroporated with pCAT-*s-hairy1* (Figure 2.8 B).

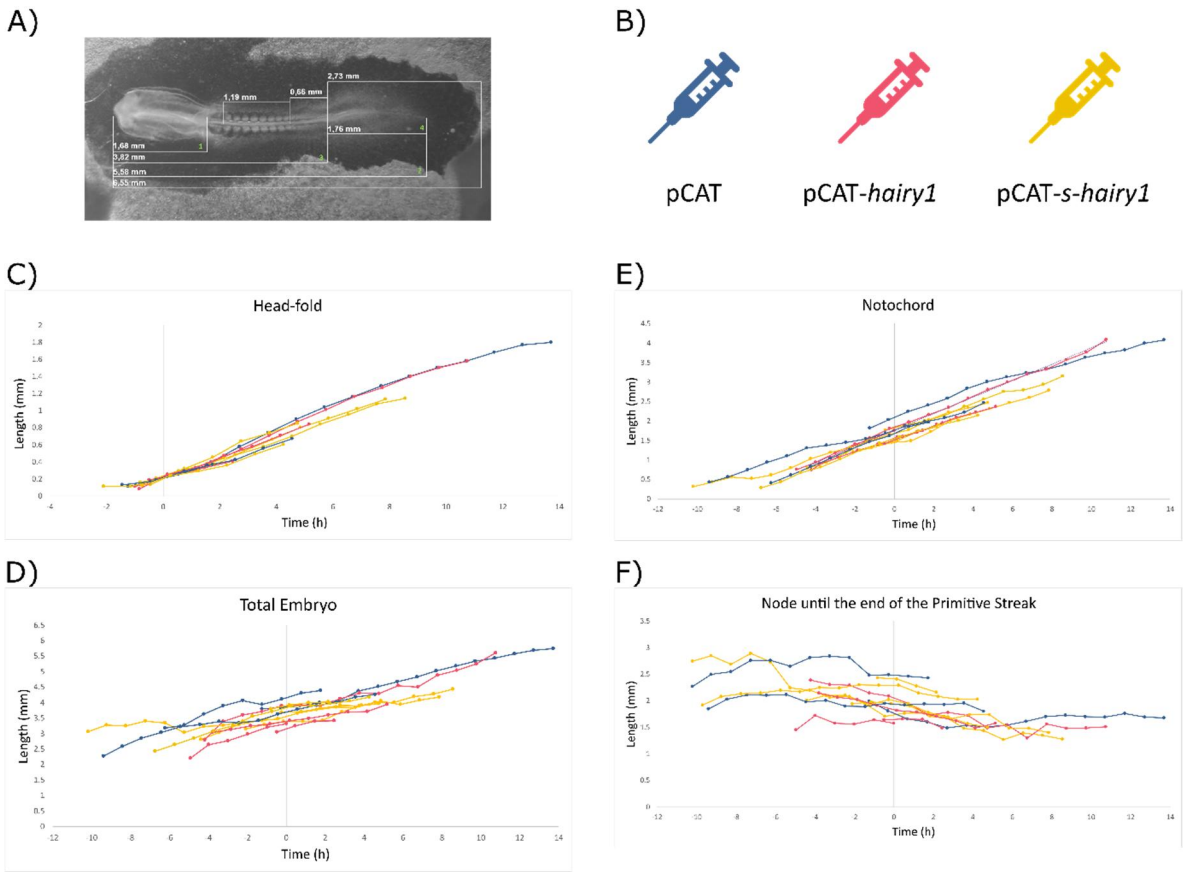



**Figure 2.8 - Measurements performed to characterize embryo tissue elongation over time.**
**A)** Example of one frame with all the measurements performed. **B)** Representation of the color code used in the graphs. blue: empty vector (pCAT); pink: *hairy1* overexpression (pCAT-hairy1); yellow: *s-hairy1* overexpression (pCAT-s-hairy1). **C)** Headfold length over time. Headfold measurements were used to establish the t=0 in all the graphs (t=0 is when the headfold has 0.22mm) (represented in A as 8). **D)** Whole embryo measurements. In each frame, embryos were measured from the top of the headfold until the end of the primitive streak (represented in A as 2). **E)** Measurements of the notochord. Embryos were measured from the top of the headfold until the Hensen's node (represented in A as 3). **F)** Measurement of the tissues that are in gastrulation, from the Hensen's node until the end of the primitive streak (represented in A as 5).

Since the segmentation in the *hairy1* and *s-hairy1* overexpressing embryos is delayed, we could not use the number of somites to align the different embryos. To overcome this problem, Maia-Fernandes and Pais-de-Azevedo showed that the head-fold has linear growth over time (data not published, manuscript in preparation). The same was reproduced in the time-lapse experiments performed in this work, independently of the constructs used (Figure 2.8 C). In order to align all the videos, the time 0 (t=0) was selected when the head-fold had 0,22 mm. This measure was set to be t=0 because it is the minimum head-fold length present in all embryos.

The whole embryo was measured from the most anterior part of the head until the end of the primitive streak and until the posterior limit of the area pellucida (Figure 2.8 D and data not shown). In both cases, the embryos overexpressing *hairy1* (pink curves) or *s-hairy1* (yellow curves) show a trend to be shorter over time in comparison to control embryos (blue curves). Although the number of embryos tested is small, this data is consistent with unpublished data from the laboratory, which suggests that the overexpression of *hairy1* promotes delay in trunk development. We further measured the embryos to pinpoint if a specific tissue can account for the delay. The notochord of the *hairy1* and *s-hairy1* overexpressing embryos was tendentially shorter over time. However, they show a similar growth rate as the control embryos. Lastly, we looked at the length of the primitive streak (measurement from the Hensen's node until the end of the primitive streak). This tissue was more variable than the other measured tissues, which greatly limited the possibility of taking any conclusions.

Finally, all the electroporated embryos were compared with non-treated embryos (Figure 2.9). This study included data previously obtained in the lab for non-treated embryos. When we look at both notochord and whole embryo elongation, the growth behavior of the embryos electroporated with the empty vector (n=6/6) closely resembles the non-electroporated embryos. This means that the electroporation *per se* does not significantly affect embryo growth. The *hairy1* and *s-hairy1* overexpressing embryos are tendentially smaller than the pCAT. Although we see a perturbation in notochord length, we identify a more significant elongation delay when we look at the whole embryo length. In the total embryo length, four *hairy1* overexpressing embryos (n=4/5) and three *s-hairy1* overexpressing embryos (n=3/5) are smaller than the pCAT and smaller than the non-electroporated embryos. This suggests that the most affected area is the more undifferentiated tissue, in the remaining primitive streak.

By closing analyzing the elongation curves of each embryo, we could find a maximum delay of nine hours for embryos overexpressing *hairy1*, seven hours for the embryos

overexpressing *s-hairy1*. By contrast, the maximum delay of the control embryos was of three hours. It is important to emphasize that only one control embryo has an overall delay superior to one hour, and in the experimental embryos, almost all showed a delay superior to three hours. Overall, our data show that overexpressing of *s-hairy1* has a similar effect to overexpressing the full-length *hairy1.*

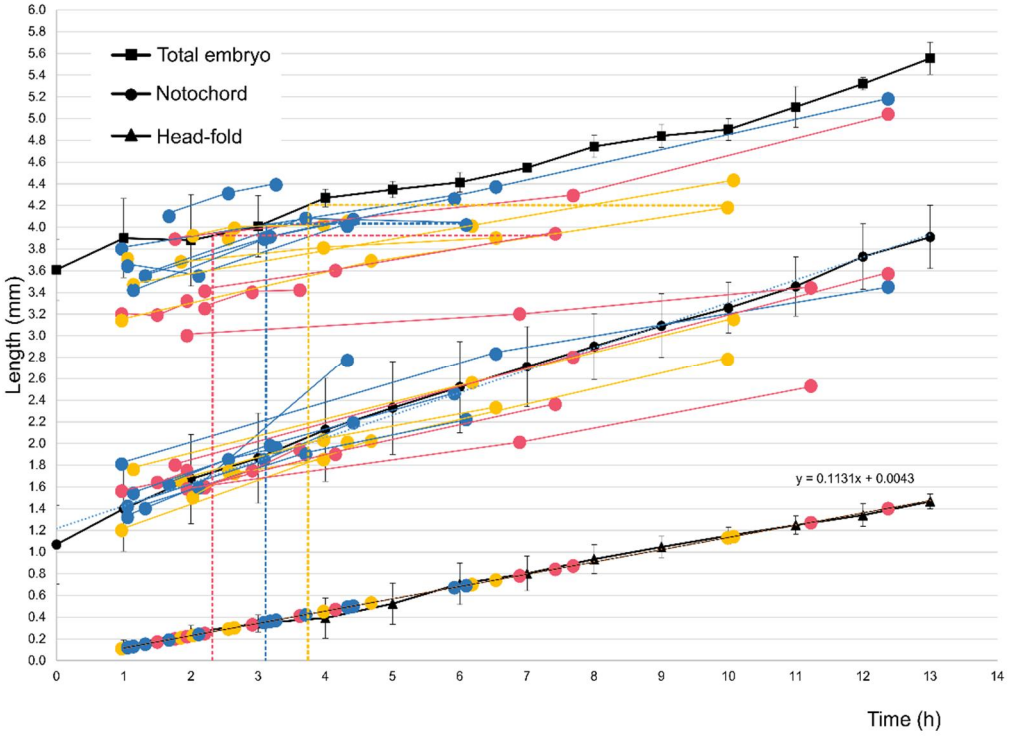



**Figure 2.9 – Comparison of anterior-posterior elongation of electroporated embryos with culture-only (WT) embryos.**
Headfold length was used as a *proxy* for developmental time in order to align the data of all embryos. Notochord and total embryo length were compared with WT embryos. Blue represents the embryos electroporated with the empty vector (pCAT), in pink embryos electroporated with pCAT-hairy1 and yellow embryos electroporated with pCAT-s-hairy1. The black lines represent the WT embryo growth over time. Dashed lines are an example of delay calculation of each condition.

As embryo segmentation takes place, each newly formed segment is assigned with an anterior-posterior identity. This is due to *Hox* cluster sequential activation during gastrulation in the epiblast cells, which then dictates the timing of cell ingression into the primitive streak (Iimura and Pourquie, 2006). Since we detected a delay in the growth of the embryos overexpressing *hairy1* or *s-hairy1,* we questioned if the *Hox* gene expression was also delayed, thus indicating a delay in gastrulation cell movements. We electroporated embryos at stage HH4 (1 day old), allowing them to grow for more than 24 hours and collected between stages HH10 to HH11 (Figure 2.10 A i, ii, iii). After evaluating that the electroporation was

successful by analysing the GFP expression (Figure 2.10 A i', ii' and iii'), we performed whole-mount *in situ* RNA hybridization to detect the expression of *HoxB9* (Figure 2.10 A i'', ii'' and iii''). If activation of *HoxB* expression was delayed, we expected that the anterior limit of *HoxB* expression in the PSM would be more posterior than in the control embryos.

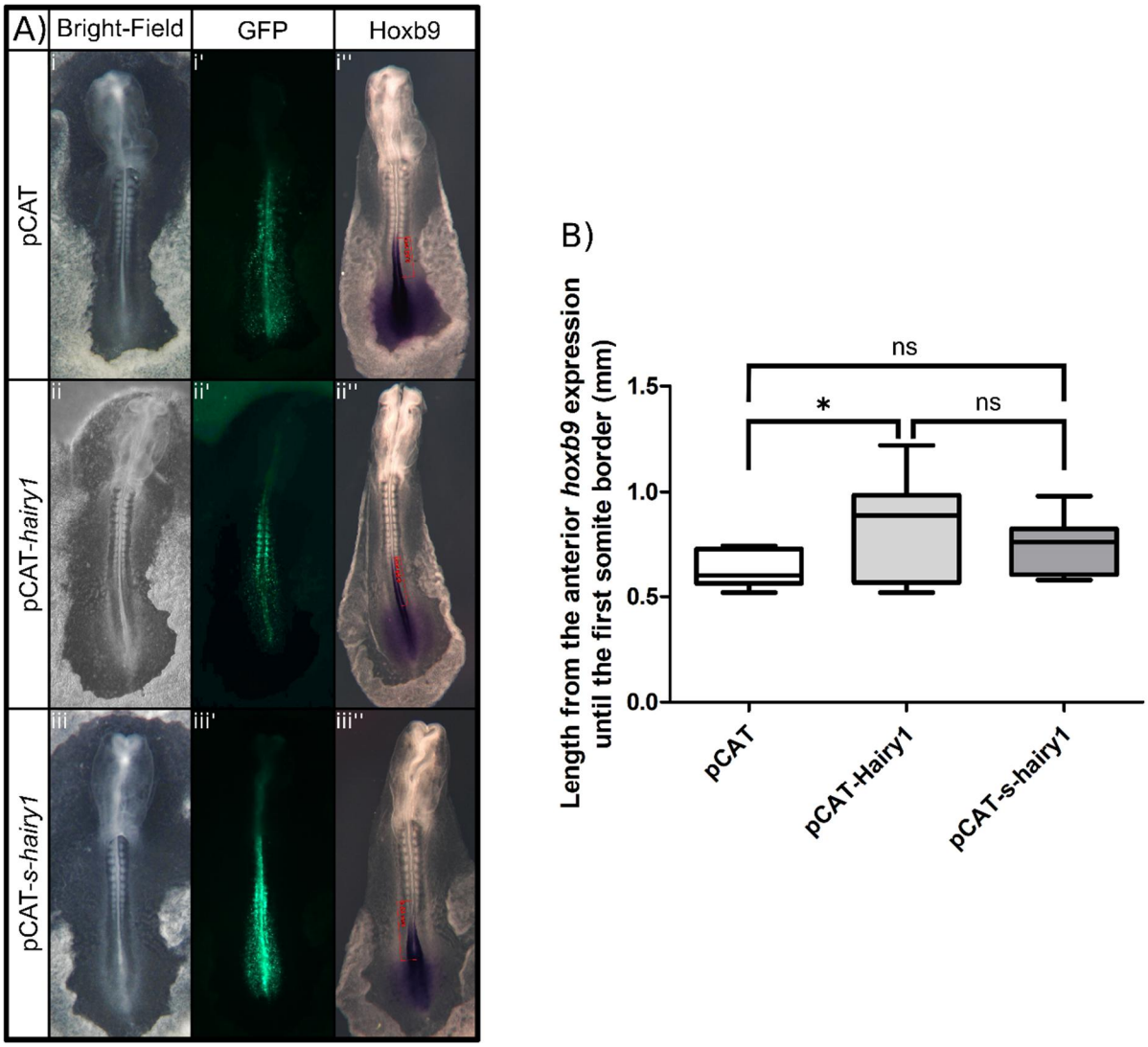



**Figure 2.10 - *HoxB9* expression in embryos overexpressing *hairy1 or s-hairy1* in the PSM.**
**A)** Representative images of embryos electroporated in each condition and respective *HoxB9* expression. **B)** Quantification of the length from the anterior limit of *HoxB9* expression in the PSM until the first somite border (mm). N=8 per condition. *Student t-test with a p-value < 0.05.

In the control embryos, *HoxB9* is expressed in the neural plate and neural tube until the second somite border, and it is also expressed in the paraxial and axial mesoderm in a posterior-to-anterior gradient until midway through the PSM (Figure 2.10 A i-i''). In embryos overexpressing *hairy1* or *s-hairy1*, the intensity of the graded *HoxB9* signal was significantly reduced (Figure 2.10 A ii-iii''). This supports the hypothesis of a delay in *Hoxb9* activation, since mRNA molecules have not accumulated to the same levels as in the control.

To further evaluate if *HoxB9* expression was delayed, we measured the distance between the last somite border and the anterior limit of *HoxB9* expression in the PSM (Figure 2.9 A, red lines). The average length in the control embryos was 0.63 mm, while it was 0.83 mm on the embryos overexpressing *hairy1* and 0.74 mm on the embryos overexpressing *s-hairy1* (Figure 2.10 B). Overall, in both cases, *HoxB9* gene expression is perturbed and our results support a putative delay in *HoxB9* activation during gastrulation. However, only *hairy1* overexpression resulted in statistically significant results (Figure 2.10 B), so we need to increase the number of embryos to confirm these results.

Unpublished work from the laboratory (manuscript under preparation) found that Hairy1 can bind the chromatin of the *HoxB* cluster. To better understand the impact of Hairy1 on *HoxB* cluster regulation, Chromatin immunoprecipitation (ChIP) experiments were performed. First, the protocol was optimized, in particular the sonication time required for proper chromatin fragmentation (DNA fragments from 200 to 600 base pairs, as shown in Figure 2.11 A). Then, an experiment was designed where stage HH8 embryos were bisected, and half was immediately fixed, and the other half was cultured for 30 minutes or 90 minutes (Figure 2.11 B). The chromatin was immunoprecipitated using an antibody against Hairy1 and, by qPCR, the enrichment of *HoxB9* promoter regions precipitated by the Hairy1 antibody was evaluated (Figure 2.11 C). The results evidenced that the binding of Hairy1 to the *HoxB9* promotor is highly dynamic over time. This data supports a model whereby cyclic Hairy1 formation/degradation may result in binding/dissociation to *HoxB* cluster promoters, thus timing their expression.

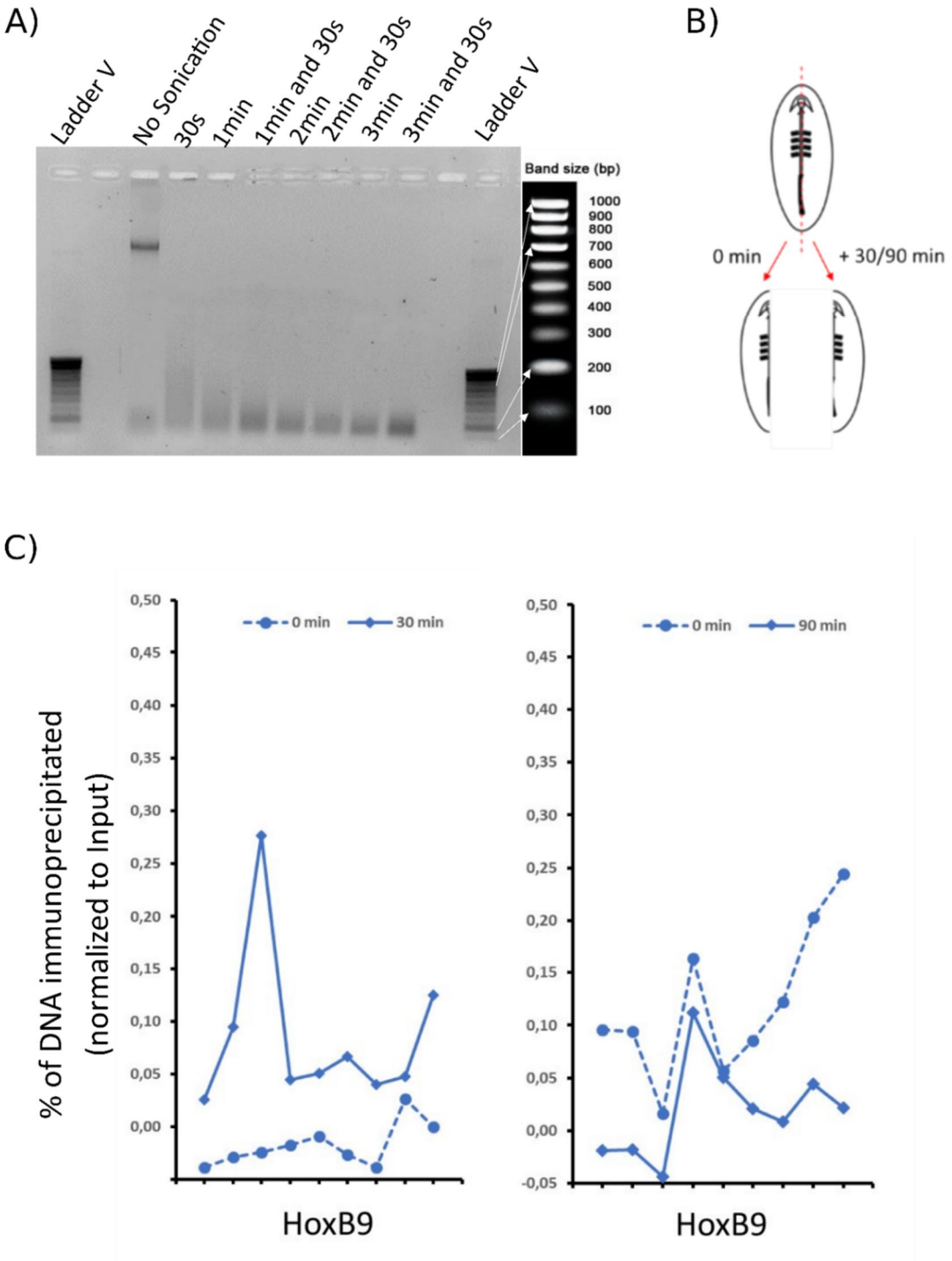



**Figure 2.11 - Evaluation of Hairy1 binding dynamics to *HoxB9* promoter using Chromatin Immunoprecipitation followed by quantitative PCR (ChIP-qPCR).**
**A)** Chromatin sonication optimization. Agarose gel electrophoresis analysis of the embryo chromatin fragment length upon different time of sonication. With 1 min of sonication, we got the desired chromatin fragment length (between 200 and 600 nucleotides). **B)** Experimental approach: half of the embryo was collected and fixed and the other half was cultured for 30 or 90 minutes. After that time, they were collected and fixed and ChIP-qPCR was performed. **C)** Dynamic Hairy1 binding to *HoxB9* promoter evidenced by ChIP-qPCR. The xx axis represents the HoxB9 promoter landscape (intervals of 1 kb). Dashed line represents the control half and the solid line represent the incubated half. Dot represents the average site of the amplified fragments.

### 2.3.3 Expanding the limb clock

The EC is closely related with PSM segmentation, however in recent years it has also been implicated in other tissues and developmental processes, such as central nervous system development and limb outgrowth (Pascoal *et al.*, 2007; Shimojo *et al.*, 2008). In the PSM, several genes that have a cyclic behavior are already known, which are synchronized and tightly regulated by notch signaling (Dequéant *et al.*, 2006; Jiang *et al.*, 2000; Krol *et al.*, 2011). During limb development, only *hairy2* was described to have cyclic expression, with a period four times longer than in the PSM (6h *vs* 90 min) (Pascoal *et al.*, 2007). Also, there is very little information regarding *hairy1* expression dynamics in this tissue. In the present topic, we intended to better understand the oscillatory network present in the distal mesenchymal tissue during the forelimb development.

During chicken somitogenesis, *dll1* has a cyclic expression behavior and, in cooperation with Notch1, it is responsible for local synchrony within the PSM (Jiang *et al.*, 2000). In the limb, however, we found that the expression of *dll1* is absent in the early stages of the limb bud outgrowth (Figure 2.12 A). Besides *delta*, there are other notch ligands, *serrate*, and in chicken, it was already found that *serrate1* and *serrate2* are both present during the limb bud outgrowth, however in different areas. *Serrate1* was described to be expressed in the distal forelimb mesenchyme, and *serrate2* is exclusively expressed in the AER (Myat *et al.*, 1996; Vargesson *et al.*, 1998). Since *serrate1* is present in the distal limb mesenchyme, we further investigated the possibility of its expression being cyclic as it happens in the PSM with *dll1* (Krol *et al.*, 2011). Contrarily to *dll1* expression, we found that *serrate1* is expressed in the entire forelimb mesenchyme of embryos at stage HH22. As the limb grows, the expression of *serrate1* is distalized, and at HH24, its expression is restricted to the distal-most portion (Figure 2.12 B).

To evaluate if *serrate1* has a cyclic expression, we opened a window on the eggshell, surgically removed the right limb, and immediately fixed it. We then closed the egg and let the embryo grow for more 1.5, 3, or 6 hours and afterward, we collected the left forelimb and processed both limbs for *in situ* hybridization. We found that *serrate1* expression presents only subtle variations over time, that do not recapitulate (Figure 2.12 B). Although *serrate1* expression at these stages is in the same domain of *hairy2* cyclic expression, the expression seems to be only dynamic, slightly changing as the tissue differentiates, but in a non-cyclic manner. In fact, Vargesson and colleagues described that the expression of *serrate1* at stages from HH27 onwards is limited to the interdigital tissue, and it may indicate a function in interdigital cell apoptosis, as was already described for mouse *jagged2* (Francis *et al.*, 2005; Vargesson *et al.*, 1998).

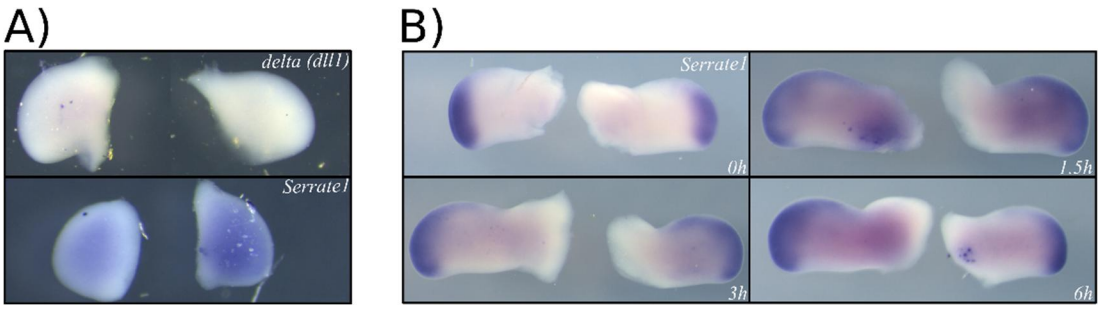



**Figure 2.12 - Expression of notch ligands *delta1* (*dll1*) and *serrate1* during forelimb development.**
**A)** Expression of *dll1* and *serrate1* at stage HH22. **B)** Expression of *serrate1* over time. The right limb from a four-day-old embryo (HH24) was collected at time 0h. The embryo was left to grow for anoter 1,5 hours, 3 hours, and 6 hours. After that, the left limb was collected. Gene expression was accessed by RNA *in situ* hybridization and both, right and left limbs, were processed together.

Regarding *hairy1* expression in the developing limb, Vasiliauskas and colleagues showed that misexpression of *hairy1* leads to delay in forelimb development (Vasiliauskas *et al.*, 2003), leading us to question the role of *hairy1* in forelimb outgrowth and if it is part of the limb clock with a cyclic expression. To be consistent with the known data for *hairy2*, we decided to perform this work at the same stages as (Pascoal *et al.*, 2007), between stages HH24 and 25. We first analyzed *hairy1* expression in the forelimb at stage HH24 (Figure 2.13). We identified three different patterns of *hairy1* expression. All three patterns showed expression in the posterior distal mesenchyme, near the ZPA. *Hairy1* expression only differs in the central and anterior distal mesenchyme. It could be present in the posterior distal mesenchyme and mostly absent in the anterior tissue (n=4/34 Figure 2.13 A). It could be present through all the distal mesenchyme (n=15/34 Figure 2.13 B). At last, we could find expression in the posterior and anterior distal mesenchyme, but partially or totally absent from the central mesenchyme (n=15/34 Figure 2.13 C). Additionally, *hairy1* presented a constant expression in the ventral side of the proximal limb region (Figure 2.13 white arrows).

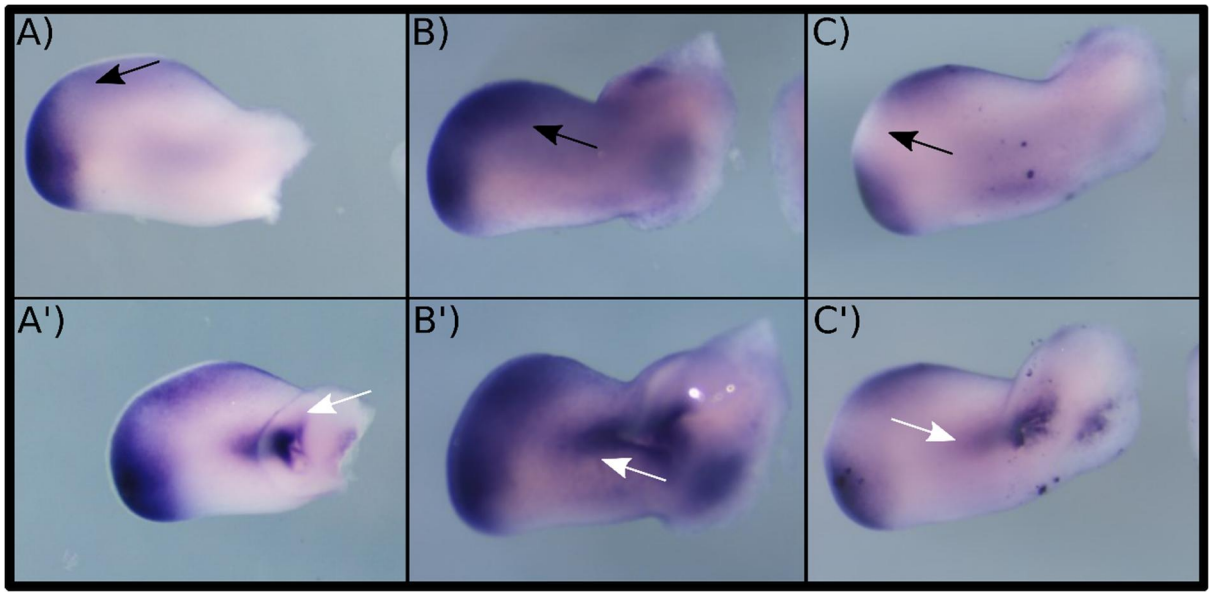



**Figure 2.13 - *hairy1* has a dynamic expression during limb outgrowth.**
We identified three main *hairy1* expression patterns in the distal mesenchyme: **A)** The expression is mainly present in the posterior distal limb bud and it is absent from the anterior distal limb (Black arrow); **B)** *hairy1* expression occupies all the distal mesenchyme, including the anterior distal area (Black arrow); **C)** *hairy1* is expressed in the more anterior and posterior distal mesenchyme but absent from the central distal mesenchyme (Black arrow). It also presents expression in the ventral central mesenchyme (White arrow). Top row (A-C): dorsal view; bottom row (A'-C'): ventral view.

We next evaluated if *hairy1* dynamic expression was also cyclic and its possible periodicity. To pursue this hypothesis, we fixed one forelimb and cultured the other for 1.5, 3, or 6 hours. To analyze the results, we compared the pattern of expression of the right forelimb (t=0) and the left forelimb (cultured for 1.5, 3, or 6 hours) (Figure 2.14). Not only does *hairy1* have a dynamic expression in the distal forelimb mesenchyme, the result obtained indicate that *hairy1* expression is cyclic, with a periodicity of 6 h (Figure 2.14 A). Limbs pairs collected 1.5 hours apart only presented identical expression patterns in 25% of the cases (n=3/12). Similar results were observed when the left forelimb was cultured for more three hours, with only 33.3% (n=6/18) of the forelimb pairs presenting an identical expression. However, when the left forelimb was allowed to grow for more six hours, 73.3% (n=11/15) of the limbs showed an identical expression pattern (Figure 2.14 B). These results suggest that *hairy1* expression also has a cyclic behavior, with a periodicity of 6 h, such as *hairy2*.

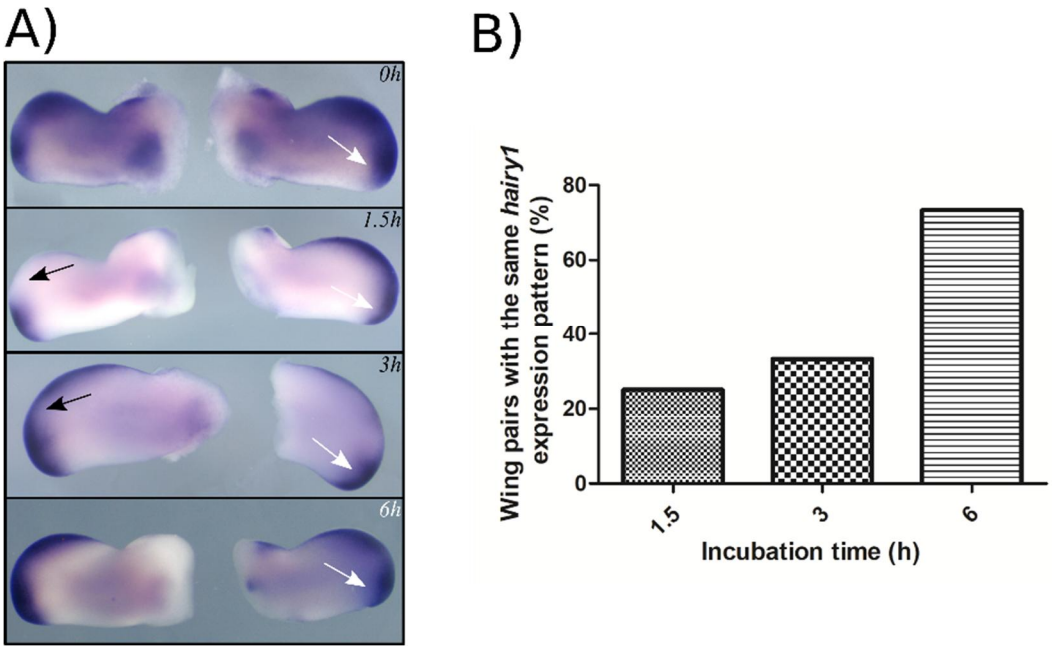



**Figure 2.14 - *hairy1* expression cycles with a 6 hours pace in the chick embryo forelimb. A)** Expression pattern of *hairy1*. The right limb from a four-day-old embryo (HH24) was collected at time 0h. The embryo was allowed to grow for more 1.5, 3, or 6 hours. After that, the left limb was collected. Gene expression was assessed by RNA *in situ* hybridization and both, right and left limbs, were hybridized together. Black arrows indicate the dynamic expression domain; White arrows, the constant expression domain. **B)** Quantification of forelimb pairs with identical *hairy1* expression pattern.

At stage HH24, the *hairy1* cyclic domain seems to be different from the cyclic domain presented by *hairy2* (Figure 2.15). At this stage, the *hairy2* cyclic domain is contained in the central distal mesenchyme, the wave of expression goes from the posterior distal mesenchyme to the central distal mesenchyme (progress zone), and in the end, the wave goes towards the central proximal mesenchyme. However, the *hairy1* wave of expression extends from the posterior distal mesenchyme towards the anterior distal mesenchyme (Figure 2.15).

We intend to extend our functional experiments to the forelimb. As we previously showed in topic 2.3.2 of this chapter, we promoted a delay in trunk development by overexpressing *hairy1* or *s-hairy1*. Now we wanted to see the impact of overexpressing both transcripts in the limb bud. Since *s-hairy1* is not endogenously expressed in the limb bud, can its expression perturb the correct patterning? By promoting its expression, can we change the dynamics of the embryo clock? Overexpression studies in the limb tissue is technically very challenging, so we first need to optimize the experimental technique, based on the work of (Ueda *et al.*, 2017). We injected the plasmid in the coelom under the right forelimb field of HH15 to 16 embryos. Then, by placing a thin negative electrode under the forelimb field and a positive on top, we gave an electric pulse, and the cells were electroporated (Figure 2.16 A). Afterward, we let the embryos grow for more four days and collected both forelimbs (Figure

2.16 B). The limbs that presented a strong GFP expression (Figure 2.16 C) were then stained with alcian blue to evaluate the cartilage segments (Figure 2.16 D). Having successfully obtained electroporated limb tissues, in the future we will measure the length and area of the limb segments to compare both forelimbs (one electroporated and the other as control) (Figure 2.16 C a and b).

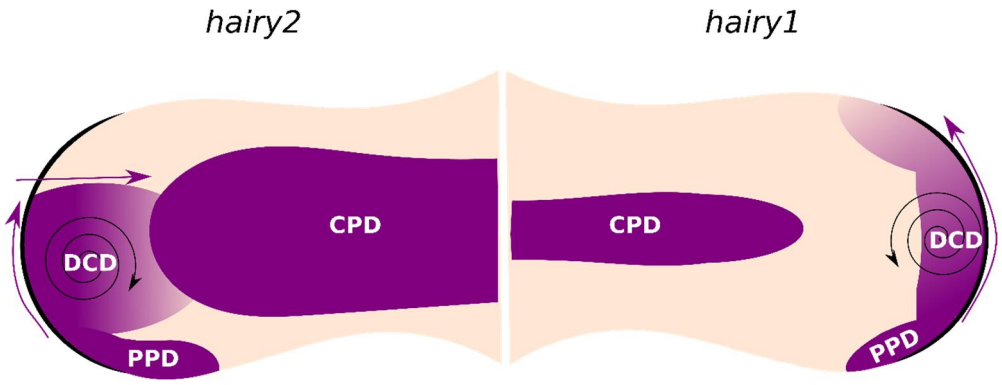



**Figure 2.15 - Domains of expression of *hairy2* and *hairy1* in an HH24 stage limb bud.** (Left) *hairy2* expression pattern presents two constant domains, the Posterior Positive Domain (PPD) and the Central Positive domain (CPD). Moreover, *hairy2* has a cyclic expression pattern in the Distal Cyclic Domain (DCD). (Right) *hairy1* expression pattern also has two constant positive domains, the PPD and the ventral CPD. However, the CPD is narrower and shorter in *hairy1*, and limited to the ventral side. It also presents a cyclic domain, but, comparatively with *hairy2*, it extends more anteriorly and does not expand to the central mesenchyme. Spirals: cyclic expression; Purple arrows indicate the direction of the wave of expression in the cyclic domains.

## 2.4 Chapter Discussion

### 2.4.1 *Gallus gallus hairy1 locus* produces multiple transcripts

The work described in the present chapter focused on the first characterized gene of the EC, the chicken *hairy1*. Our aim was to identify possible alternative transcripts that may explain the different paces of the embryo clock in the PSM (90 min) and the forelimb (6 h). We identified one 5' alternative transcript (*s-hairy1*) and one 3' alternative transcript. The alternative 3' transcript was expressed in both tissues, PSM and forelimb. However, *s-hairy1* was only present in the caudal portion tissues, and for the first time, we identified a *hairy1* transcript expressed in the PSM and absent in the distal mesenchyme of the forelimb.

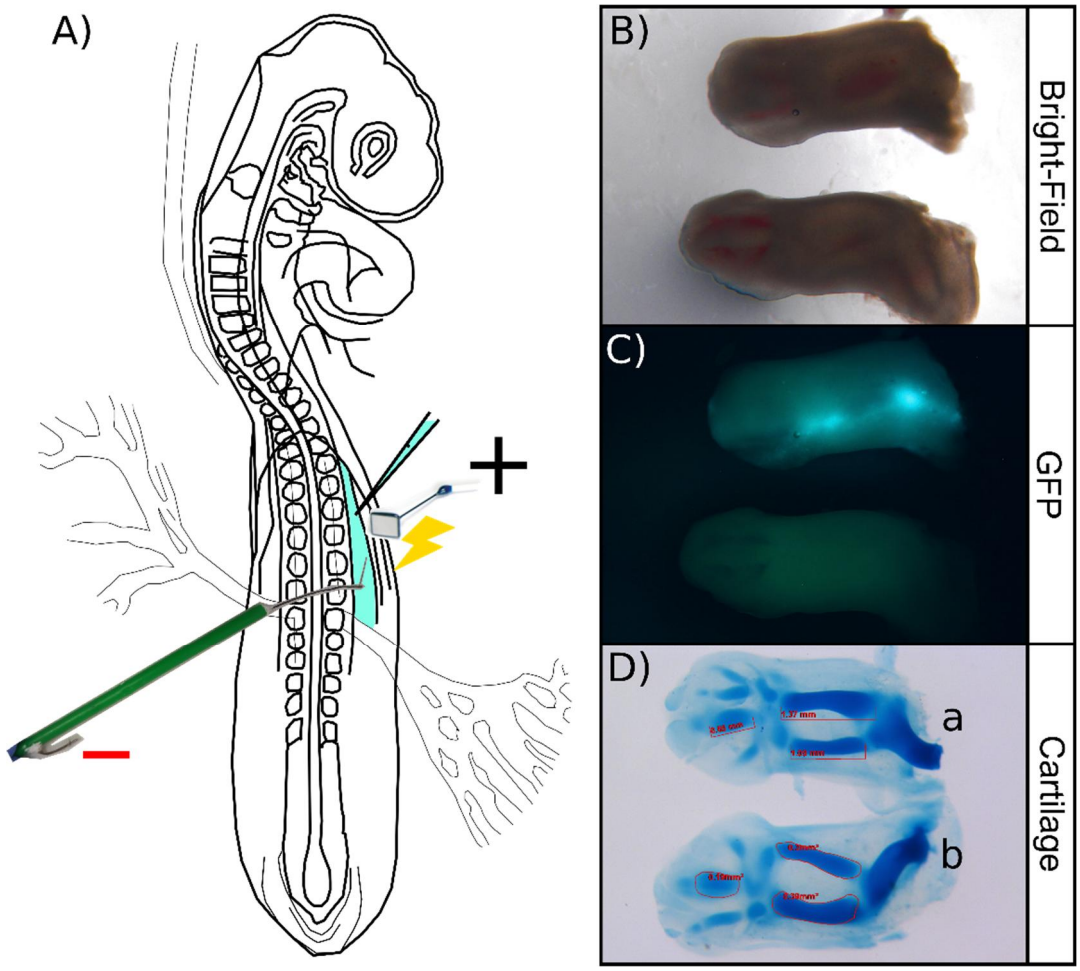



**Figure 2.16 - Optimization of electroporation for functional analysis of *hairy1* and *s-hairy1* in the forelimb.**
**A)** Schematic representation of the electroporation of a HH16 embryo. **B)** Six-day-old forelimbs collected four days after electroporation. The top forelimb was electroporated, and the bottom served as control. **C)** GFP expression, which shows where the plasmid was effectively electroporated and expressed. **D)** Alcian blue staining of the forelimbs to analyze the future bone elements. a and b show an example of the measurements of the length and the area of the bone.

Our first question was if it is translated or if it is a long non-coding RNA. Our western-blot results did not identify different Hairy1 protein species between the caudal portion and the forelimb, but they could exist, mainly because the antibody used in this experiment was highly unspecific, and further experimentation should be performed with a different antibody. The first ATG of *s-hairy1* is only six nucleotides from the 5'end, which makes it very difficult to be translated since the translational machinery does not have space to dock on the RNA.

However, if *s-hairy1* would be translated, it would give rise to a protein lacking part of the basic domain, essential to bind the DNA (Ohsako *et al.*, 1994).

Vasiliauskas and colleagues showed that an alternative splicing of *hairy1* produces a protein with fourteen extra amino acids in the basic domain, Hairy1B (Vasiliauskas *et al.*, 2003). Their data suggest that overexpression of this protein form has less impact than the overexpression of Hairy1A. They suggested that Hairy1B may have an impairment in its capability to bind the DNA. The same is possible to happen if the *s-hairy1* is translated from its first ATG. A more plausible hypothesis it that *s-hairy* could be translated from the second ATG onwards. This smaller protein would have a completely absent basic domain and a truncated *helix-loop-helix*, thus lacking DNA binding and protein dimerization capacity. A final alternative hypothesis is that it could be a functional long non-coding RNA.

### 2.4.2 *hairy1* or *s-hairy1* overexpression alters *HoxB9* expression and delays PSM segmentation

To functionally analyze *s-hairy1,* we overexpressed it in the prospective PSM tissue during gastrulation. We found that this overexpression led to a delayed segmentation start. This phenocopied the overexpression of *hairy1.* By comparing the headfold extension, we found that the *hairy1* or *s-hairy1* overexpression produced a general delay of one to two somites over time. In a non-published work in the lab, overexpression or downregulating *hairy1* equally led to a somitogenesis delay and overall delay in trunk development (Andrade *et al.*, unpublished). Although the overexpression is maintained until, at least, embryonic stage HH12-13 (when we stopped the experiment and collected the embryos in the present work), the delay seems to be only in the beginning, since the overexpressing embryos maintained the proportional difference of the number of somites compared with the control, suggesting that embryos manage to recover and/or use a compensatory mechanism. Further work should be conducted in order to evaluate if the formed somites are normal or if the embryos develop skeletal defects. More striking is the fact that *s-hairy1* also has the capacity of promoting the same defective phenotype as *hairy1*. Moreover, when we overexpressed *s-hairy1,* we found that the *HoxB9* expression limit was more posterior compared with the control. This means that not only the segmentation itself was perturbed, but also trunk identity pattering was delayed. We had a similar result when with *hairy1* overexpression.

We further analyzed the impact of the overexpression of *hairy1* and *s-hairy1* and found that the trunk of overexpressing embryos was shorter than the control. By making several measurements of the embryo, we concluded that the most affected portion, in both cases, was the primitive streak. The embryo clock starts operating during early gastrulation (Maia-

Fernandes and Jesus *et al.*, unpublished). Therefore, by affecting the embryo clock in these early stages, we may be perturbing the correct timing of cell differentiation or the motile capacity of the cells, or even perturbing cell division, which was already correlated with Notch signaling and the EC (Braunreiter and Cole, 2019; Delaune *et al.*, 2012; Kobayashi and Kageyama, 2010). More experimentation is needed to clarify the role of *s-hairy* and *hairy* in trunk elongation.

*s-hairy* overexpression results in the same phenotype as *hairy1*. Data from Andrade et al., shows that overexpression of the C-terminal domain of Hairy1 also promotes the same segmentation delay (Andrade *et al.*, unpublished). If, in fact, *s-hairy1* is able to produce a peptide, it may have the same function as the Hairy1 C-terminal alone. Hes proteins work in protein complexes (homodimer or heterodimer) (Kageyama *et al.*, 2007; Schroter *et al.*, 2012), and Hairy1 is able to bind to the mesoderm marker Brachyury through its c-terminal domain (Andrade *et al.*, unpublished). So, the putative S-Hairy1 peptide could be competing with Hairy1 for dimerization, thus having a dominant-negative effect. This is supported by the fact that down-regulation of *hairy1* also produces the same phenotypes (Andrade *et al.*, unpublished).

### 2.4.3 Hairy1 protein binds dynamically to the *HoxB9* promoter

We showed that Hairy1 protein is able to bind in a dynamic way to the chromatin *HoxB9* promoter. Although we did not observe a recapitulation of the binding profile after 90 min, it may be because the clock is faster during the formation of the first somites, as recently characterized in our lab (Maia-Fernandes and Jesus *et al.*, unpublished). This experiment is in the context of a more extensive work, where we found that Hairy1 dimerizes Brachyury, and together they have the ability to bind *Hox* promoter chromatin. Also, they show that the C-terminal domain of Hairy1 can heterodimerize with Brachyury, which means that the putative S-Hairy1 peptide also could dimerize with Brachyury and directly compete with Hairy1 for Brachyury dimerization, and consequently promoting a delay of the *Hox* opening (Andrade *et al.*, unpublished). Our results support a model whereby cyclic Hairy1 formation/degradation may result in binding/dissociation to *HoxB* cluster promoters, thus timing their expression.

### 2.4.4 *hairy1* has a cyclic expression in the distal forelimb

To date, *hairy2* was the only gene documented to have a cyclic behavior in the distal forelimb mesenchyme, with a pace of six hours (Pascoal *et al.*, 2007). Since the primary question that drove this thesis was: “How do the same molecular mechanisms work in the same organism but with different paces?”, we found it critical to deepen our knowledge of the

limb clock. For that reason, we decided to look at Notch ligands' expression and *hairy1* (which cycle in the PSM) (Krol *et al.*, 2011).

Regarding the Notch ligands, previously documented results support the data we obtained. *Delta-like1* expression is totally absent from the chicken forelimb mesenchyme at HH22, and in the mouse developing limb, *Jag1* is the primary Notch ligand, whereas *Dll1* is expressed at very low levels (Dong *et al.*, 2010). On the other hand, *serrate1*, another Notch ligand, is ubiquitously expressed in the HH22 forelimb. Afterward, *serrate1* expression is confined to the distal forelimb, where it might bind to Notch1 (also present in the distal mesenchyme) (Francis *et al.*, 2005; Vargesson *et al.*, 1998). However, since at stage HH27, *serrate1* is only expressed in the interdigital domain, we do not know what could drive cyclic expression since at least *hairy2* continues to have a cyclic expression.

*Lunatic fringe (lfng)* is absent during limb outgrowth (data from geisha repository (Darnell *et al.*, 2007)). The lack of *lfng* may be compensated by the expression of another fringe family protein, such as *radical fringe* (*rfng*), which is expressed in the AER and distal limb (Ha and Riddle, 2003). However, the mice limb skeleton is normal in the KO of the three mice fringe family genes (*Lfng, Rfng*, and *Mfng* (*Manic Fringe*)). So, the components of Notch signalling that might participate in the limb clock regulation remain unclear.

During this work, we found a second gene with an oscillatory behavior in the developing limb, *hairy1*. As *hairy2,* we found that *hairy1* has cyclic expression in the distal forelimb, between the stages HH23 and 25, with a six-hour pace. Although both genes' domains of cyclic expression are partially overlapping, we found that *hairy1* only cycles from the posterior to the anterior, and not to the central mesenchyme, as *hairy2* (Pascoal *et al.*, 2007). Since there is collinear *Hox* expression along the anterior to posterior distal forelimb (Abbasi, 2011), it is tempting to propose that the limb clock could also be important for digit patterning.

## 2.5 References


**Abbasi, A.A.** (2011). Evolution of vertebrate appendicular structures: Insight from genetic and palaeontological data. Dev Dyn *240*, 1005-1016. https://doi.org/10.1002/dvdy.22572.

**Braunreiter, K.M., and Cole, S.E.** (2019). A tale of two clocks: phosphorylation of NICD by CDKs links cell cycle and segmentation clock. EMBO reports *20*, e48247. https://doi.org/10.15252/embr.201948247.

**Chapman, S.C., Collignon, J., Schoenwolf, G.C., and Lumsden, A.** (2001). Improved method for chick whole-embryo culture using a filter paper carrier. Developmental Dynamics *220*, 284-289. https://doi.org/10.1002/1097-0177(20010301)220:3<284::aid-dvdy1102>3.0.co;2-5.

**Darnell, D.K., Kaur, S., Stanislaw, S., Davey, S., Konieczka, J.H., Yatskievych, T.A., and Antin, P.B.** (2007). GEISHA: an in situ hybridization gene expression resource for the chicken embryo. Cytogenet Genome Res *117*, 30-35. https://doi.org/10.1159/000103162.

**Delaune, E.A., François, P., Shih, N.P., and Amacher, S.L.** (2012). Single-cell-resolution imaging of the impact of Notch signaling and mitosis on segmentation clock dynamics. Dev Cell *23*, 995-1005. https://doi.org/10.1016/j.devcel.2012.09.009.

**Dequéant, M.-L., Glynn, E., Gaudenz, K., Wahl, M., Chen, J., Mushegian, A., and Pourquié, O.** (2006). A Complex Oscillating Network of Signaling Genes Underlies the Mouse Segmentation Clock. Science *314*, 1595-1598. https://doi.org/10.1126/science.1133141.

**Dong, Y., Jesse, A.M., Kohn, A., Gunnell, L.M., Honjo, T., Zuscik, M.J., O'Keefe, R.J., and Hilton, M.J.** (2010). RBPjkappa-dependent Notch signaling regulates mesenchymal progenitor cell proliferation and differentiation during skeletal development. Development *137*, 1461-1471. https://doi.org/10.1242/dev.042911.

**Francis, J.C., Radtke, F., and Logan, M.P.O.** (2005). Notch1 signals through Jagged2 to regulate apoptosis in the apical ectodermal ridge of the developing limb bud. Developmental Dynamics *234*, 1006-1015. https://doi.org/10.1002/dvdy.20590.

**Gilbert, F.S., and Barresi, M.J.F.** (2019). Developmental biology, 12th Edition (Sinauer Associates, Inc.).

**Gonzalez-Reyes, A., Elliott, H., and St Johnston, D.** (1997). Oocyte determination and the origin of polarity in Drosophila: the role of the spindle genes. Development *124*, 4927-4937. https://doi.org/10.1242/dev.124.24.4927

**Ha, A.S., and Riddle, R.D.** (2003). cBlimp-1 expression in chick limb bud development. Gene Expr Patterns *3*, 297-300. https://doi.org/10.1016/s1567-133x(03)00042-5.

**Hamburger, V., and Hamilton, H.L.** (1951). A SERIES OF NORMAL STAGES IN THE DEVELOPMENT OF THE CHICK EMBRYO. Journal of Morphology *88*, 49-&. https://doi.org/10.1002/jmor.1050880104.

**Henrique, D., Adam, J., Myat, A., Chitnis, A., Lewis, J., and Ish-Horowicz, D.** (1995). Expression of a Delta homologue in prospective neurons in the chick. Nature *375*, 787-790. https://doi.org/10.1038/375787a0

**Hirata, H., Ohtsuka, T., Bessho, Y., and Kageyama, R.** (2000). Generation of Structurally and Functionally Distinct Factors from the Basic Helix-Loop-Helix Gene Hes3 by Alternative First Exons. Journal of Biological Chemistry *275*, 19083-19089. https://doi.org/10.1074/jbc.M001075200.

**Iimura, T., and Pourquie, O.** (2006). Collinear activation of Hoxb genes during gastrulation is linked to mesoderm cell ingression. Nature *442*, 568-571. https://doi.org/10.1038/nature04838.

**Jiang, Y.J., Aerne, B.L., Smithers, L., Haddon, C., Ish-Horowicz, D., and Lewis, J.** (2000). Notch signalling and the synchronization of the somite segmentation clock. Nature *408*, 475-479. https://doi.org/10.1038/35044091.

**Jouve, C., Palmeirim, I., Henrique, D., Beckers, J., Gossler, A., Ish-Horowicz, D., and Pourquie, O.** (2000). Notch signalling is required for cyclic expression of the hairy-like gene HES1 in the presomitic mesoderm. Development *127*, 1421-1429. https://doi.org/10.1242/dev.127.7.1421

**Kageyama, R., Ohtsuka, T., and Kobayashi, T.** (2007). The Hes gene family: repressors and oscillators that orchestrate embryogenesis. Development *134*, 1243-1251. https://doi.org/10.1242/dev.000786.

**Kling, D.E., Lorenzo, H.K., Trbovich, A.M., Kinane, T.B., Donahoe, P.K., and Schnitzer, J.J.** (2002). MEK-1/2 inhibition reduces branching morphogenesis and causes mesenchymal cell apoptosis in fetal rat lungs. American Journal of Physiology-Lung Cellular and Molecular Physiology *282*, L370-L378. https://doi.org/10.1152/ajplung.00200.2001.

**Kobayashi, T., and Kageyama, R.** (2010). Hes1 regulates embryonic stem cell differentiation by suppressing Notch signaling. Genes Cells *15*, 689-698. https://doi.org/10.1111/j.1365-2443.2010.01413.x.

**Krol, A.J., Roellig, D., Dequeant, M.L., Tassy, O., Glynn, E., Hattem, G., Mushegian, A., Oates, A.C., and Pourquie, O.** (2011). Evolutionary plasticity of segmentation clock networks. Development *138*, 2783-2792. https://doi.org/10.1242/dev.063834.

**Larkin, A., Marygold, S.J., Antonazzo, G., Attrill, H., dos Santos, G., Garapati, P.V., Goodman, Joshua L., Gramates, L S., Millburn, G., Strelets, V.B., et al.** (2020). FlyBase: updates to the Drosophila melanogaster knowledge base. Nucleic Acids Research *49*, D899-D907. https://doi.org/10.1093/nar/gkaa1026.

**Müller, S., Rycak, L., Afonso-Grunz, F., Winter, P., Zawada, A.M., Damrath, E., Scheider, J., Schmäh, J., Koch, I., Kahl, G., and Rotter, B.** (2014). APADB: a database for alternative polyadenylation and microRNA regulation events. Database (Oxford) *2014*, bau076. https://doi.org/10.1093/database/bau076.

**Myat, A., Henrique, D., Ish-Horowicz, D., and Lewis, J.** (1996). A chick homologue of Serrate and its relationship with Notch and Delta homologues during central neurogenesis. Dev Biol *174*, 233-247. https://doi.org/10.1006/dbio.1996.0069.

**Nelson, C.E., Morgan, B.A., Burke, A.C., Laufer, E., DiMambro, E., Murtaugh, L.C., Gonzales, E., Tessarollo, L., Parada, L.F., and Tabin, C.** (1996). Analysis of Hox gene expression in the chick limb bud. Development *122*, 1449-1466. https://doi.org/10.1242/dev.122.5.1449

**Ohsako, S., Hyer, J., Panganiban, G., Oliver, I., and Caudy, M.** (1994). Hairy function as a DNA-binding helix-loop-helix repressor of Drosophila sensory organ formation. Genes & Development *8*, 2743-2755. https://doi.org/10.1101/gad.8.22.2743.

**Palmeirim, I., Henrique, D., Ish-Horowicz, D., and Pourquie, O.** (1997). Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. Cell *91*, 639-648. https://doi.org/10.1016/S0092-8674(00)80451-1

**Pascoal, S., Carvalho, C.R., Rodriguez-León, J., Delfini, M.-C., Duprez, D., Thorsteinsdóttir, S., and Palmeirim, I.** (2007). A Molecular Clock Operates During Chick Autopod Proximal-distal Outgrowth. Journal of Molecular Biology *368*, 303-309. http://dx.doi.org/10.1016/j.jmb.2007.01.089.

**Rushlow, C.A., Hogan, A., Pinchin, S.M., Howe, K.M., Lardelli, M., and Ish-Horowicz, D.** (1989). The Drosophila hairy protein acts in both segmentation and bristle patterning and shows homology to N-myc. Embo j *8*, 3095-3103. https://doi.org/10.1002/j.1460-2075.1989.tb08461.x

**Sakamoto, K., Yan, L., Imai, H., Takagi, M., Nabeshima, Y., Takeda, S., and Katsube, K.** (1997). Identification of a chick homologue of Fringe and C-Fringe 1: involvement in the neurogenesis and the somitogenesis. Biochem Biophys Res Commun *234*, 754-759. https://doi.org/10.1006/bbrc.1997.6652.

**Schroter, C., Ares, S., Morelli, L.G., Isakova, A., Hens, K., Soroldoni, D., Gajewski, M., Julicher, F., Maerkl, S.J., Deplancke, B., and Oates, A.C.** (2012). Topology and dynamics of the zebrafish segmentation clock core circuit. PLoS Biol *10*, e1001364. https://doi.org/10.1371/journal.pbio.1001364.

**Shimojo, H., Ohtsuka, T., and Kageyama, R.** (2008). Oscillations in Notch Signaling Regulate Maintenance of Neural Progenitors. Neuron *58*, 52-64. https://doi.org/10.1016/j.neuron.2008.02.014.

**Small, S., and Arnosti, D.N.** (2020). Transcriptional Enhancers in Drosophila. Genetics *216*, 1-26. https://doi.org/10.1534/genetics.120.301370.

**Ueda, S., Suzuki, T., and Tanaka, M.** (2017). Transgene Introduction into the Chick Limb Bud by Electroporation. Methods Mol Biol *1650*, 203-208. https://doi.org/10.1007/978-1-4939-7216-6_13.

**Vargesson, N., Patel, K., Lewis, J., and Tickle, C.** (1998). Expression patterns of Notch1, Serrate1, Serrate2 and Delta1 in tissues of the developing chick limb. Mech Dev *77*, 197-199. 10.1016/s0925-4773(98)00138-5. https://doi.org/10.1016/S0925-4773(98)00138-5

**Vasiliauskas, D., Laufer, E., and Stern, C.D.** (2003). A role for hairy1 in regulating chick limb bud growth. Developmental Biology *262*, 94-106. http://dx.doi.org/10.1016/S0012-1606(03)00360-9.

**Voiculescu, O., Papanayotou, C., and Stern, C.D.** (2008). Spatially and temporally controlled electroporation of early chick embryos. Nat Protoc *3*, 419-426. https://doi.org/10.1038/nprot.2008.10.

**Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., and Madden, T.L.** (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics *13*, 134. https://doi.org/10.1186/1471-2105-13-134.

# Chapter 3

## Relevance of Embryo Clock mRNA 3'UTR for gene expression oscillations in *Danio rerio*

The work described in this chapter was performed at the Department of Molecular and Cellular Biology of Harvard University, under the supervision of Professor Alexander F. Schier, supported by a Fulbright Research Grant fellowship

*This page is intentionally left blank*

## 3.1 Chapter Introduction

mRNA stability and translation are multifactorial mechanisms, in which the 3'UTR plays a crucial regulatory role. Depending on the factors that bind to the 3'UTR, it can lead to translation repression, poly(A) deadenylation and mRNA decay, and/or dictate mRNA subcellular localization (Figure 3.1) (Reviewed in (Berkovits and Mayr, 2015)).

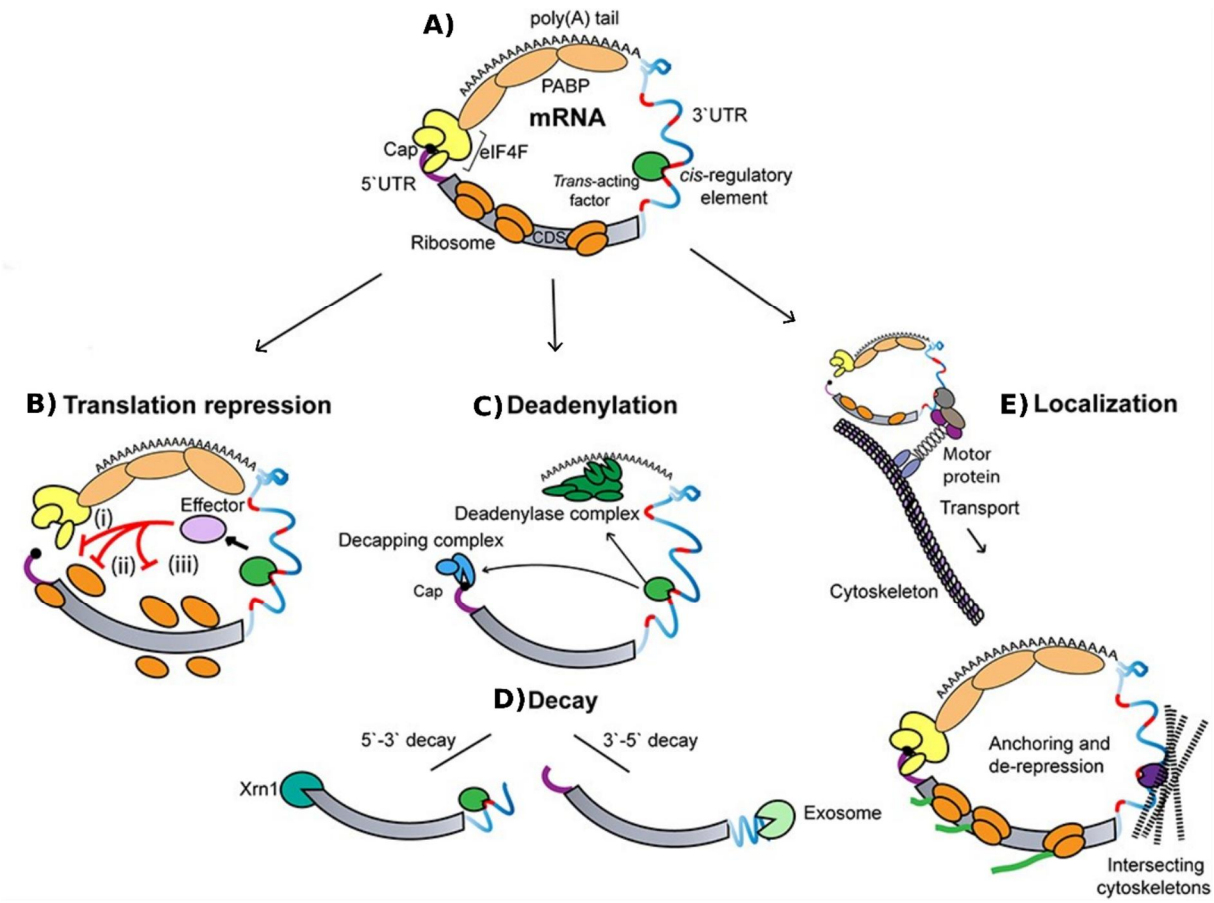



**Figure 3.1 - Overview of mRNA 3'UTR functions.**
**A)** General schematics of an eukaryotic mRNA. The 3' poly(A) interacts with the 5'Cap by poly(A) binding protein (PAPB) recognition of the eIF4F, which binds to the 5'Cap. **B)** The 3'UTR can lead to translation repression by (i) interference with the cap-binding complex, (ii) interference with the ribosome mRNA binding capacity, or (iii) by inhibiting ribosome-mediated elongation. **C-D)** The 3'UTR can lead to mRNA decay by promoting 3' deadenylation and/or recruiting decapping factors. **E)** Translationally inactive mRNA can be transported along the cytoskeleton (through RBPs bound to the 3'UTR) to a specific cell compartment/location and, anchored there, it is activated to be translated. Adapted from (Mayya and Duchaine, 2019).

The embryo clock negative feedback regulatory mechanism relies on fast mRNA turnover rates and the 3'UTRs of EC genes have been described to play an important role in gene expression oscillations. For instance, the specific 3'UTR of mouse *Lfng* and *Hes7* was sufficient to drive cyclic expression of an exogenous vector, recapitulating the expression pattern of the endogenous *Lfng* or *Hes7* genes, respectively (Nitanda *et al.*, 2014). Moreover, when Fujimuro and colleagues inserted a long piece of DNA in the *hes7* prospective 3'UTR, this not only disrupted *hes7* oscillations but also drastically reduced HES7 protein levels (Fujimuro *et al.*, 2014). In zebrafish, it was also demonstrated *her1, her7,* and *hes6* mRNAs are regulated by CCR4-NOT deadenylating complex. Their data also suggests that the 3'UTR of these genes is responsible for the recruitment of CCR4-NOT, regulating to their own degradation rate (Fujino *et al.*, 2018).

The work described in this chapter was performed to analyze the importance of the 3'UTR of three zebrafish EC genes, *her1*, *her7*, and *deltaC* (*dlc*). Ulitsky and colleagues found that both *her1* and *her7* have APA sites, resulting in several possible 3'UTR (Ulitsky *et al.*, 2012). Using CRSPR technology, we deleted the alternative distal APA site, and examined the subsequent resulting phenotype. In turn, *dlc* has a 3'UTR of around 1400 nucleotides with multiple putative miRNAs target sites. We divided the 3'UTR into three Regulatory Regions (RR) and deleted them using CRISPR technology. We further analyzed F2 embryos for phenotype. With this work, we intended to shed light on possible mechanisms of regulating the zebrafish EC through the mRNA 3'UTR.

### 3.1.1 Zebrafish as a powerful model in Developmental Biology

This thesis was performed using the chicken embryo as the main experimental model, which is a great model to study the embryonic clock and development in general (Stern, 2005). We can efficiently study gene expression and correlate it with biological processes. However, it has some drawbacks. It is still challenging to produce mutants and would require a lot of space to have a chicken facility. Instead, tissue electroporation or viral vectors are commonly employed to modulate gene expression. This results in chimeric tissues and in expression levels that do not mimic physiological conditions, hindering data interpretation.

Zebrafish is another important model used to study embryo development. It is small, hence it is possible to have a large number of animals growing in a limited physical space; they reproduce all year round, and each clutch can have hundreds of embryos; the embryos are transparent, which also helps to evaluate possible phenotypes and study organogenesis more efficiently; they reach sexual maturity in only three months (Figure 3.2). Most importantly for our work, there is a vast array of experimental techniques that allow genetic manipulation,

among them RNAi, morpholino experiments, RNA injection for overexpression, and gene modifications, such as knockouts and knock-ins (Reviewed in (Pathak and Barresi, 2020)).

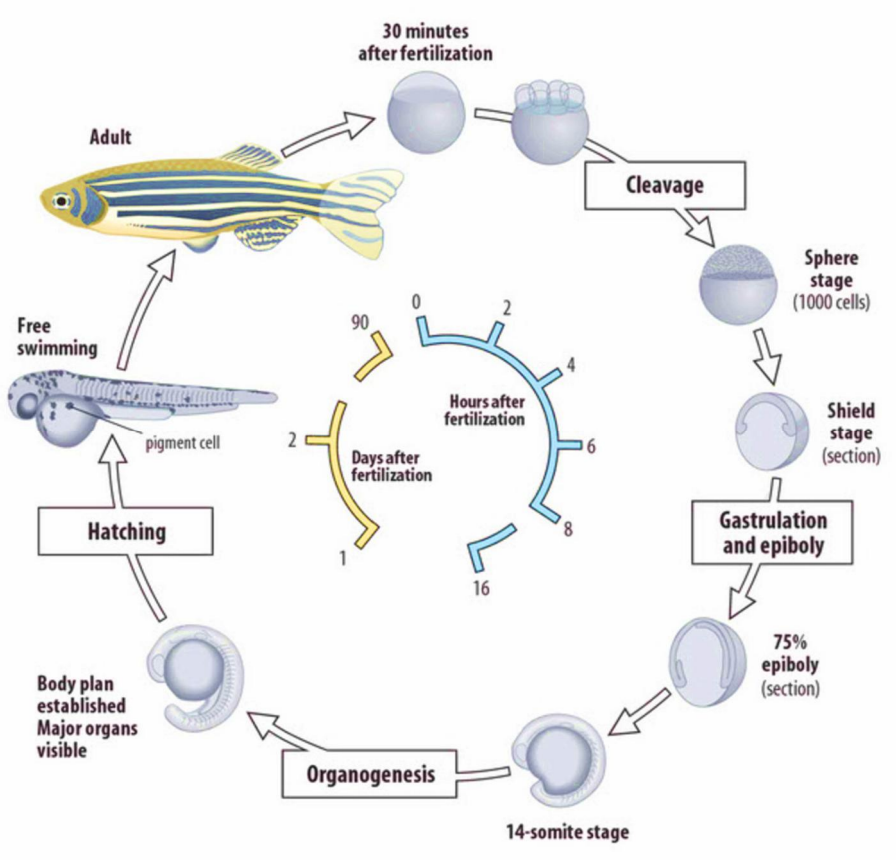



**Figure 3.2 - The *Danio rerio* life cycle.**
Schematic representation of the zebrafish life cycle, evidencing the key steps over time. A zebrafish embryo starts its gastrulation 6hpf and finishes it at around 10hpf, starting somitogenesis and ending at 1dpf. They hatch 2dpf and achieve sexual maturity in 90 days. A zebrafish can live up to five years. Adapted from (D'Costa and Shepherd, 2009)

One of the hypotheses regarding this thesis is that alternative polyadenylation (APA) sites could modulate mRNA stability, consequently impacting the pace of the embryo clock. Using chicken embryos and the techniques available it would be very difficult to test this hypothesis. Instead, we decided to work with *Danio rerio*.

## 3.2 Material and Methods

### Ethics statement

Experimentation involving live vertebrate animals was performed at the zebrafish facility of the Faculty of Arts & Sciences (HU/FAS), Harvard University. All experiments were approved by the Harvard University/Faculty of Arts & Sciences Standing Committee on the Use of Animals in Research & Teaching under Protocol number 25–08.

### Zebrafish strains and embryos

Zebrafish (*Danio rerio*) wild-type strains TL/AB and AB were used in this study. Animals were kept on cycles of 14 hours of light and 10 hours in the dark at 28ºC. Embryos were cultured in blue water (250 mg/L Instant Ocean salt, 1 mg/L methylene blue) and were staged according to (Kimmel *et al.*, 1995).

### RNA extraction and quality control

Total RNA was extracted using RNAqueous™-Micro Total RNA Isolation Kit (Invitrogen™ #AM1931) according to the manufacturer's instructions. RNA was resuspended in 20 µL of Elution Solution, and it was immediately quantified using NanoDrop 2000 (Thermo Scientific) and stored at -80ºC. RNA quality control was performed using Experion™ RNA StdSens Analysis Kit (BioRad #700-7103). Only samples with an RQI (RNA Quality Indicator) equal to or above nine were used for subsequent steps.

### Rapid Amplification of cDNA Ends (RACE)

5'RACE and 3'RACE were performed using FirstChoice™ RLM-RACE Kit (Invitrogen™ AM1700M) according to the manufacturer's instructions, using 10µg of total RNA as starting material. During the 5'RACE execution and after Calf-intestinal alkaline phosphatase (CIP) treatment, phenol:chloroform (Sigma-Aldrich #P2069) was used to stop the reaction and purify the Caped RNA. PCR was performed using *her7*-specific primers that were designed using primer BLAST (Ye *et al.*, 2012). Sequences of the primers used are in Table 3.1.

**Table 3.1 - Gene-specific PCR primers used for her7 PCR after RACE**

| | |
|---|---|
| her7_5'_Outer | 5'-GACCGAGCCTTCCAGTCCACC-3' |
| her7_5'_Inner | 5'-GGGCCTTGCAGGAGCAGAAGT-3' |
| her7_3'_Outer | 5'-ACCTGAAACTTCTGCTCCTGC-3' |
| her7_3'_Inner | 5'-ACCTGAAACTTCTGCTCCTGC-3' |

Outer and Inner PCRs were performed using the FirstChoice™ RLM-RACE Kit guidelines and were performed using DreamTaq DNA Polymerase (Thermo Scientific™ #EP0701) and run in a C1000 Touch™ Thermal Cycler (Bio-Rad #1851196). 5µL of Outer and Inner PCR were run in 2% agarose gel with the DNA GreenSafe Premium (Nzytech #MB13201). A ChemiDoc XRS+ Imaging system (Bio-Rad #1708265) acquired gel images.

**Bioinformatics queries**

A bioinformatics search for annotated alternative transcripts produced by *her1*, *her7,* and *dlc* was performed using UCSC genome browser http://genome.ucsc.edu (Kent *et al.*, 2002). Zebrafish (Danio rerio) Zv9 (GCA_000002035.2) and Genome Reference Consortium Zebrafish Build 11 assemblies were used. To identify alternative 3'UTRs, TargetScan Fish Release 6.2 (http://www.targetscan.org/fish_62/) (Agarwal et al., 2015; Ulitsky *et al.*, 2012) was also used. TargetScan Fish Release 6.2 was used to identify miRNA that could target *her1*, *her7*, and *dlc* (Agarwal *et al.*, 2015).

**CRISPR/Cas9 mediated mutagenesis**

Gene-specific guide RNA (gRNA) sequences were designed using CHOPCHOP (https://chopchop.cbu.uib.no/) (Montague *et al.*, 2014) (Table 3.2). sgRNAs were synthesized by performing a ligation reaction using T4 DNA polymerase (NEB #M0203) between the gene-specific oligo and the constant oligo (AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC) and subsequent transcription using MEGAscript T7 Kit (Ambion #AM1334) (Gagnon *et al.*, 2014). CRISPR-Cas9 mutagenesis was conducted as described by (Gagnon *et al.*, 2014). ~1nL of mutagenesis mixture, consisting of sgRNAs (300ng/µL), Cas9 protein (600ng/µL) and phenol red injection dye (Sigma-Aldrich #P3532), was injected into the cytoplasm of one-cell stage TL/AB wild-type embryos.

**Genotyping**

Generation F0 embryos (Figure 3.3) were collected 24-30 hours after injection to evaluate the success of the mutation event by genotyping.

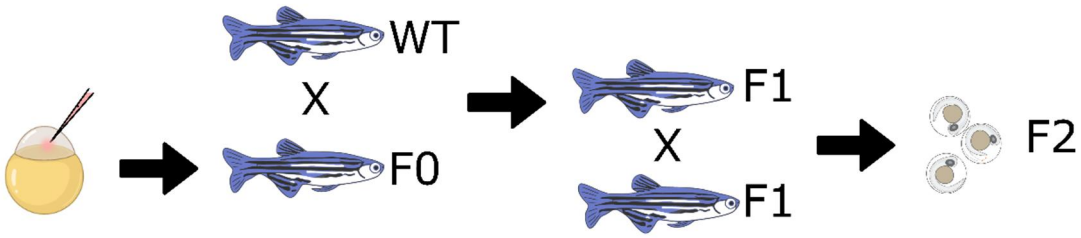



**Figure 3.3 - Embryo cross and mutant lines production.**
One cell stage embryos were injected with the mutation mix, generating F0 embryos. F0 embryos grew for three months and were crossed with wild-type fish (WT) to generate heterozygous fish (generation F1). Two F1 siblings were crossed to generate generation F2. F2 embryos were collected for whole-mount in situ hybridization.

Caudal tips of adult F0 and F1 fish were collected for genotyping. Genomic DNA was extracted by using HotSHOT protocol (Meeker *et al.*, 2007) and PCR was performed using PCR Master Mix (2X) (Thermo Scientific™ # K0171) and the primer pairs in Table 3.3 in a C1000 Touch™ Thermal Cycler (Bio-Rad #1851196). 3µL of PCR product was analysed in 2% agarose gel stained with UltraPure™ Ethidium Bromide (Invitrogen™ # 15585011). A ChemiDoc XRS+ Imaging system (Bio-Rad #1708265) acquired gel images.

**Table 3.2 - Gene-sepecific oligos to generate sgRNAs used in CRISPR-Cas9 mutagenesis.**

| sgRNA_Name | Sequence (5' to 3') |
|---|---|
| sgRNA_her7_PA3_1 | TAATACGACTCACTATAGGTATAAATCTGAACATGAAGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_her7_PA3_2 | TAATACGACTCACTATAGTATAAATCTGAACATGAAAGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_her7_PA2&3_1 | TAATACGACTCACTATAGATCCAGTACAGATCATGCAGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_her7_PA2&3_2 | TAATACGACTCACTATAGCTCCTTGCATGATCTGTACGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_her7_END_1 | TAATACGACTCACTATAGTGTCGGGAACACATAACAAGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_her7_END_2 | TAATACGACTCACTATAGGGAGTGGCATGATGTTTGCGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_her1_PA2_1 | TAATACGACTCACTATAGTTAGATACTTCAAGAAGTAGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_her1_PA2_2 | TAATACGACTCACTATAGGACTCGTTTCTTTACATTTGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_her1_END_1 | TAATACGACTCACTATAGCCATTAATCTTCAAGAAGTGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_her1_END_2 | TAATACGACTCACTATAGGACTCGTTTCTTTACAGTTGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_dlc_RR1_1 | TAATACGACTCACTATAGTCAGGCCCACTGGTGTTTGGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_dlc_RR1_2 | TAATACGACTCACTATAGATAGCAGAATTCTTCAATCGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_dlc_RR2_1 | TAATACGACTCACTATAGAGAGAGTTAGATGAACTTCGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_dlc_RR2_2 | TAATACGACTCACTATAGCATAGACACAAGAACTCTCGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_dlc_RR3_1 | TAATACGACTCACTATAGCTTTTCGACAGTAGATGTTGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_dlc_RR3_2 | TAATACGACTCACTATAGGAGGCACTTAATTGAGATTGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_dlc_END_1 | TAATACGACTCACTATAGATGCAATGTAAATTATTCGGTTTTAGAGCTAGAAATAGCAAG |
| sgRNA_dlc_END_2 | TAATACGACTCACTATAGTGCAATGTAAATTATTCGGGTTTTAGAGCTAGAAATAGCAAG |

### Whole-mount *in situ* hybridization

Embryos were fixed overnight in a fresh 4% formaldehyde (Sigma-Aldrich # 47608) solution at 4ºC. Whole-mount *in situ* hybridization was performed as described by (Thisse and Thisse, 2008). Digoxigenin-labeled RNA probes were synthesized by *in vitro* transcription of linearized using DIG RNA 119abelling kit (Roche #11175025910). Used probes: *her1* (Muller *et al.*, 1996); *her7* and *dlc* (Oates and Ho, 2002).

**Table 3.3 - PCR primers for *her1*, *her7* and *dlc* genotyping**

| Primer pair name | Forward (5' to 3') | Reverse (5' to 3') |
|---|---|---|
| Her1_pair_1 | CTGTTTTGCTTGGAGTTGGAGGT | TGAGTCCTGCAGTCACATTCT |
| Her7_pair_1 | AAAACACTGTCTCCAAAGCAGGAC | ACCGACCGATGTTTTGGAGG |
| Her7_pair_2 | CGTCGTCCACACCACAAAAACACT | AGACTGGAAAGCTGAGGTGTGAG |
| Dlc_pair_1 | CTCAGACCCATTTTCCATCCTCTT | GCATAACGTGGATGTGTTTCCGTA |
| Dlc_pair_2 | CAGACCCATTTTCCATCCTCTTAAT | GTTCAAGTAATCCTCAGTTTCTGTT |

### Imaging

Embryos obtained by whole-mount *in situ* hybridization were dehydrated in methanol and transparentized using a benzyl benzoate:benzyl alcohol (BBBA) solution. Afterward, they were imaged using Zeiss Axio Imager.Z1 microscope coupled to a Zeiss Axiocam 305 color 5MP camera.

## 3.3 Results

### 3.3.1 Characterization of the mRNAs produced by embryo clock genes in zebrafish

To understand the relevance of the mRNA 3'UTR for clock gene expression oscillations, we ablated different portions of the 3'UTRs from zebrafish EC genes, namely *her1*, *her7*, and *dlc* (Holley *et al.*, 2000; Oates and Ho, 2002).

According to the UCSC genome browser, *her7* (Ensembl gene ENSDARG00000017917 / RefSeq ID: NM_131609) is present in chromosome 5 from the genomic position 71711885 to 71713200. It can produce two transcripts, one that is translated into Her7 protein (Ensembl transcript ENSDART00000017849 and Ensembl protein ENSDARP00000006074, respectively) and a smaller, non-coding transcript (Ensembl transcript ENSDART00000146654) (Figure 3.4). Additionally, this gene has three

polyadenylation (PA) signals for the coding transcript (ENSDART00000017849). These three PA sites originate different 3'UTRs, with 365, 255, and 167 nucleotides (Ulitsky *et al.*, 2012). TargetScan (Agarwal *et al.*, 2015) identified four conserved putative miRNAs target sites in against *her7* transcripts. Three of them are specific to the longer 3'UTR, and one is shared between all possible 3'UTRs.

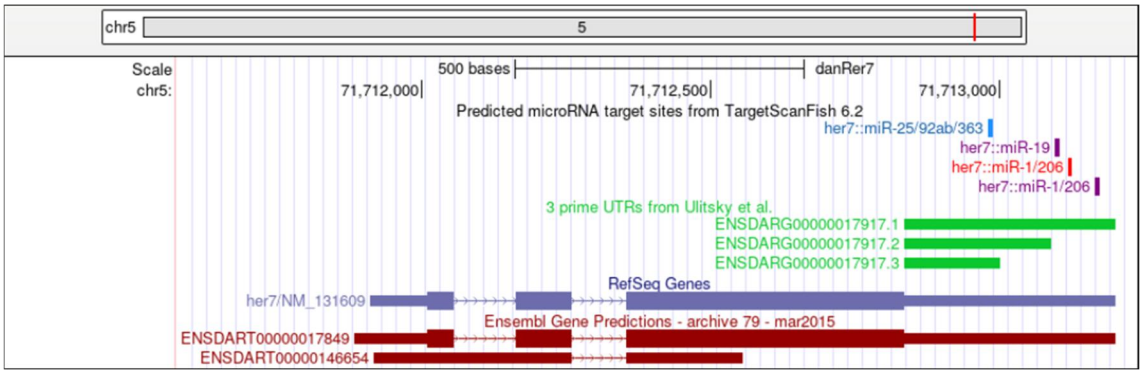



**Figure 3.4 - UCSC genome browser view of *her7 locus.***
*Her7* can produce two transcripts with different transcription start sites (Red). The longer transcript can present three different 3'UTRs (green) (http://genome.ucsc.edu/cgi-bin/hgTracks?db=danRer7&lastVirtModeType=default&lastVirtModeExtraState=&virtModeType=default&virtMode=0&nonVirtPosition=&position=chr5%3A71711716%2D71713227&hgsid=1132772285_gWau8g9M4oMboKadtFOLKuwcxHCF) (Kent *et al.*, 2002).

Regarding *her1* gene (Ensembl gene ENSDARG00000014722 / RefSeq ID: NM_131078), it is located in chromosome 5 between the genomic position 71692761 to 71699778. *her1* only presents one transcript in the UCSC genome browser (Ensembl transcript ENSDART00000016307), responsible for producing Her1 protein (Ensembl protein ENSDARP00000027076). According to (Ulitsky *et al.*, 2012), *her1* has four PA signals, producing four different 3'UTRs. Two of them are longer, with 1334 and 1276 nucleotides, and two are shorter, with 723 and 660 nucleotides. TargetScan identifies nine miRNA target sites in the *her1* 3'UTR. Eight are present in all isoforms, and one (for miR-30) is present only in the two longer 3'UTRs (Figure 3.5).

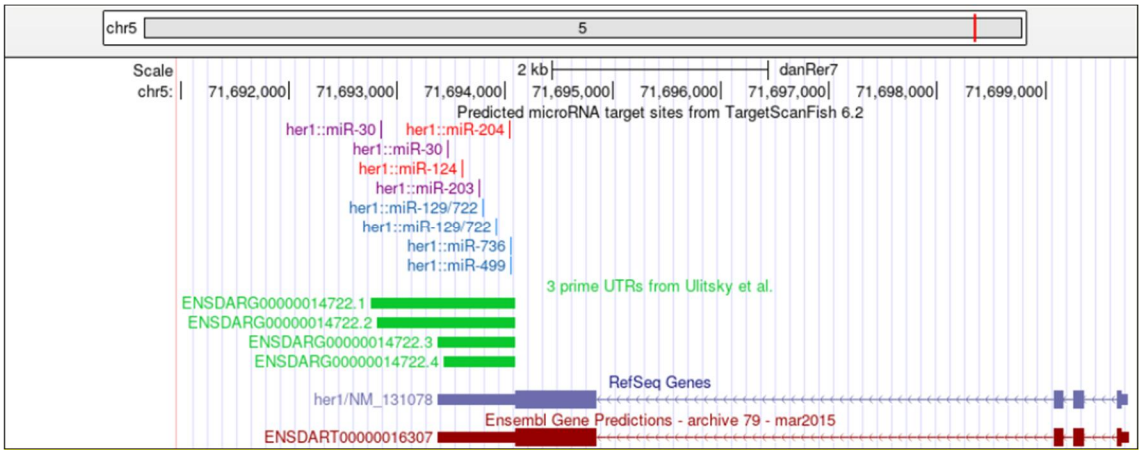



**Figure 3.5 - UCSC genome browser view of *her1 locus.***
*Her1* is only transcribed through one TSS (Red). However, it can present four different 3'UTRs (green). (http://genome.ucsc.edu/cgi-bin/hgTracks?db=danRer7&lastVirtModeType=default&lastVirtModeExtraState=&virtModeType=default&virtMode=0&nonVirtPosition=&position=chr5%3A71692181%2D71700856&hgsid=1132812745_UkfThk9NuzAMBTrw6dMJftGIsIqT) (Kent *et al.*, 2002)

The last gene analyzed was *deltaC* (Ensembl gene ENSDARG00000002336 / RefSeq ID: NM_130944), which is located in chromosome 15 between the genomic position 19425090 to 19429944. This gene is potentially transcribed into two different RNA identities. One coding transcript (Ensembl transcript ENSDART00000018514), which is translated into DeltaC protein (Ensembl protein ENSDARP00000018643) and a smaller non-coding RNA (Ensembl transcript ENSDART00000148359). According to (Ulitsky *et al.*, 2012), the coding RNA only has one PA site, which produces a 3'UTR of 1326 nucleotides. TargetScan found 20 miRNA target sites capable of targeting this transcript. No information was found regarding the smaller non-coding transcript (Figure 3.6).

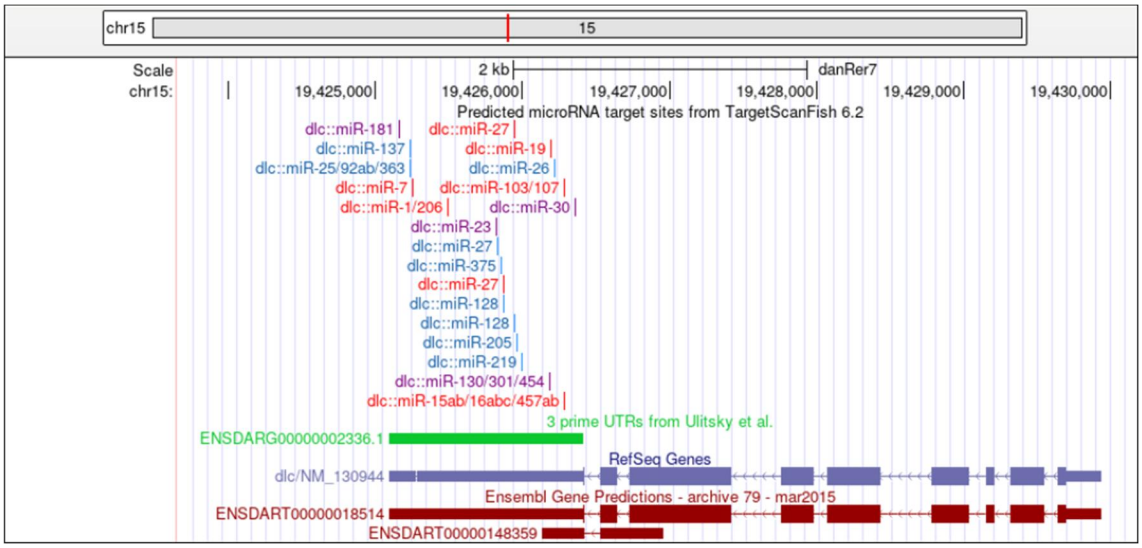



**Figure 3.6 - UCSC genome browser view of *dlc locus.***
*dlc* can produce two different transcripts, with different transcription start sites (Red). Moreover, ENSDART00000018514 transcript only presents one possible 3'UTRs (green). (http://genome.ucsc.edu/cgi-bin/hgTracks?db=danRer7&lastVirtModeType=default&lastVirtModeExtraState=&virtModeType=de0fault&virtMode=0&nonVirtPosition=&position=chr15%3A19424108%2D19430742&hgsid=1132814329_ZWubWUCEusAsqjFFEOIqDoHS0Osy) (Kent *et al.*, 2002).

The gene structure of *her7* is similar to its chicken homologue *hairy1*. Both genes have more the one polyadenylation site, three in *her7* and two in *hairy1.* More interestingly, both can produce one smaller transcript with a start site after the coding transcript (Figure 3.7 A). *s-her7* is a non-coding RNA and since both are EC core genes, it would be interesting to explore if their regulation and functional role is similar and, with that unveil common mechanisms in the vertebrate EC regulation.

The data above regarding alternative transcripts produced by *her7* is based on *in silico* annotations or on experiments performed on embryos that are not in somitogenesis stages. For this reason, we performed Rapid Amplification of cDNA Ends (RACE) assays to experimentally identify the transcripts produced by *her7* in zebrafish at somitogenesis stages, more precisely on 18 hpf embryos. By 5'RACE, two amplicons with ~260 and ~350 nucleotides were obtained, corresponding to two putative transcription start sites (TSS) (Figure 3.7 B). The longer amplicon obtained is a putative novel transcript, with a transcription start site ~100 nucleotides upstream of the described protein-coding transcript. The shorter 5'RACE amplicon is compatible with the canonical transcript (ENSDART00000017849). We could not identify the smaller transcript (ENSDART00000146654), which would produce an expected amplicon of 223 nucleotides.

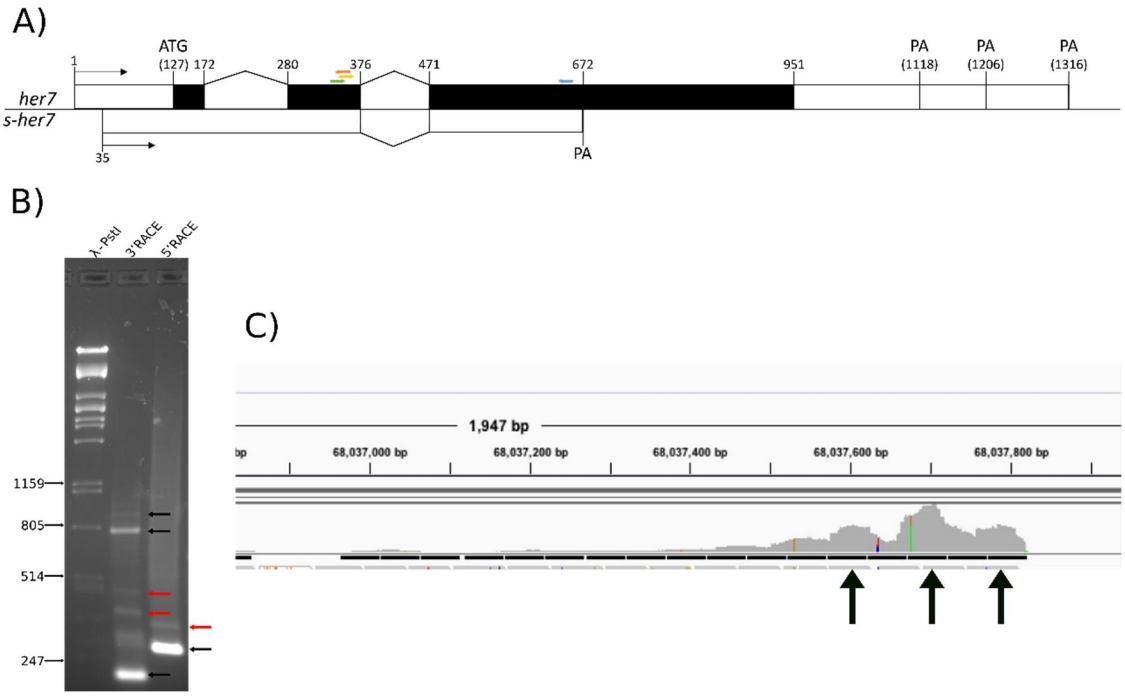



**Figure 3.7 - *her7* transcript characterization during somitogenesis stages.**
**A)** *her7 locus* and encoded transcripts according to UCSC Genome Browser on Zebrafish Jul. 2010 (Zv9/danRer7) Assembly (Kent *et al.*, 2002). The start site of *her7* canonical transcript is nucleotide number 1. It can have three APA sites. The proximal PA gives rise to a transcript of 905 nucleotides, the middle one gives rise to a transcript of 993 nucleotides, and the distal PA a transcript of 1103 nucleotides. The transcription of *s-her7* starts 35 nucleotides after the canonical transcript and has only one PA site that originates a transcript of 542 nucleotides. Black arrows indicate the transcription start sites (TSS); PA indicates the polyadenylation sites; Colored arrows indicate the primers used for the RACE experiment: Blue – 5'RACE Outer primer; Red – 5' RACE Inner primer; Green – 3'RACE Outer primer; Yellow – 3'RACE Inner Primer. **B)** RACE experiments using RNA from 18hpf embryos (somitogenesis stage). The molecular-weight size marker in the first lane, λ phage DNA digested with the restriction enzyme *PstI*. Black arrows evidence bands that resemble expected RNA molecules, and the Red arrows evidence new possible mRNA molecules, new putative PA sites, or TSS. **C)** Analysis of the single-cell RNA sequencing data showing the read count of the 3'end of *her7.* Black arrows evidence the peaks of expression of each PA site (all three peaks are compatible with the (Ulitsky *et al.*, 2012) data) (Farrell *et al.*, 2018).

Regarding the 3'RACE, we obtained fragments compatible with the existence of six potential polyadenylation sites (Figure 3.7 B). Two of the amplicons are compatible with PA sites already described by (Ulitsky *et al.*, 2012), with an expected 869 (ENSDARG00000017917.1) and 759 (ENSDARG00000017917.2) nucleotides, respectively. One of the Ulisky 3'UTRs was not detected (ENSDARG00000017917.3), which would have produced an amplicon with 691 nucleotides. We observed three additional amplicons with sizes ranging from 300 to 500 nucleotides that could potentially originate three undescribed 3'UTRs. We also identified an amplicon compatible with the smaller non-coding transcript (ENSDART00000146654), with an expected size of 227 nucleotides. Finally, by analyzing

single-cell RNA sequencing data from early embryos (from bud stage to six somites stage) (Farrell *et al.*, 2018), we found an enrichment in PA usages compatible with the three 3'UTR published by (Ulitsky *et al.*, 2012) (Figure 3.7 C, black arrows).

The discrepancies between RACE and single-cell data results may be explained because of the stages of the embryos used for both experiments. However, since the single-cell data was compatible with the results obtained by Ulitsky and colleagues, we decided to consider the same three PA sites for further experiments. Furthermore, the three smaller PA sites only detected by RACE, would terminate the transcription before the end of the *her7*'s ORF. If we promoted any deletion in these three PA sites, we would truncate Her7 protein and compromise its function.

### 3.3.2 Analysis of the functional role of *her7* and *her1* 3' UTRs

To test our hypothesis that alternative 3'UTR is involved in the EC dynamics of the zebrafish, we induced genomic deletions in the transcribed DNA that originates the *her7*'s 3'UTRs. We performed two deletions using the CRISPR technique, one of ~70 bp and another of ~165 bp. With the 70 bp deletion, we intended to remove the distal 3'UTR, allowing *her7* to be transcribed only with the 255 (ENSDARG00000017917.2) or 167nt (ENSDARG00000017917.3) 3'UTRs. The second deletion of 165bp only allows *her7* to produce the more proximal 3'UTR of 167nt (ENSDARG00000017917.3) (Figure 3.8).

As a preliminary approach to verify if these deletions impacted the gene expression patterns, whole-mount *in situ* hybridization was performed on generation F2 embryos. According to a Mendelian distribution, it was expected that 25% of the clutch was WT, 50% heterozygous, and 25% homozygous for each deletion. Around 25 embryos were analyzed per condition, and in all cases, the different characteristic patterns of expression of *her7* were observed (Figure 3.9). In the WT, *her7* expression is cyclic and limited to the PSM tissue (Oates and Ho, 2002). Each cycle of expression starts in the tailbud, gradually the expression is extended until the anterior PSM. As the expression extends to the anterior PSM, it starts to go down in the tailbud. Over time it stabilizes and forms a stripe in the determination front. As the stripe "moves" anteriorly towards a somite fate, it starts to disappear, and when a new somite forms, the expression stripe disappears (Oates and Ho, 2002). Occasionally, embryos show two stripes. When the posterior stripe forms, the anterior stripe cells have not been integrated yet into a somite, and *her7* expression is still present.

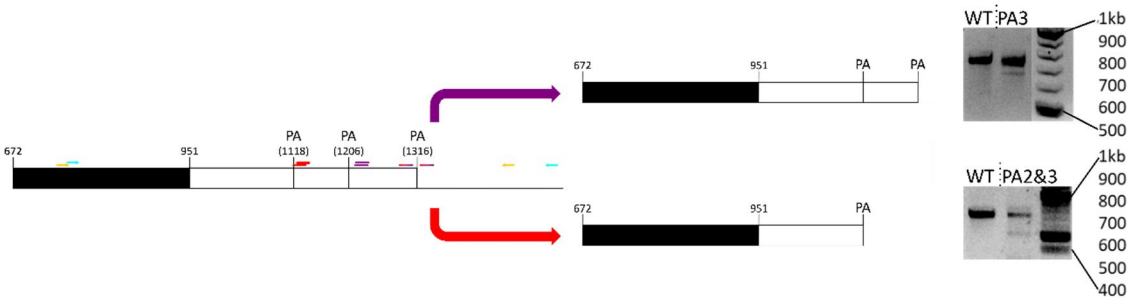



**Figure 3.8 - Schematics and genotyping of 3'UTR deletions from *her7.***
Two deletions were made in the *her7 locus* using CRISPR technology. The first deletion removed the more distal 3'UTR (Purple arrow) (deletion of ~70 nucleotides). The second deletion removed the medial and distal PA, and consequently, the produced mRNAs only could use the proximal PA (Red Arrow), (deletion of ~165 nucleotides). On the right, PCR genotyping results are shown, evidencing heterozygous embryos harboring the mutations. Red and purple bars in the WT representation of the gene indicate the gRNA position for CRISPR. Arrows in yellow and light blue indicate the primer pairs used for genotyping.

Figure 3.9 A shows examples of WT *her7* expression between 10 to 18 hpf. The *her7-ΔPA3* F2 embryos presented similar expression patterns as found in the WT clutch, which suggests that the absence of the transcript containing the longest 3'UTR is dispensable for *her7* oscillations. However, the expression seems more intense in the posterior PSM from 16hpf onwards than the WT (Figure 3.9 B black arrowheads). Interestingly, the stripes of the majority of the F2 embryos display irregularities in the stripe's borders, and they are not as sharp as the WT stripes (Figure 3.9 B white arrowheads).

The same approach was used to evaluate the effects of the deletion of the mid and distal PA sites. No significant expression alterations were found in any of the embryos (*her7-ΔPA2&3*). However, there are important observations in the expression of *her7* that may indicate a possible role of these alternative 3'UTRs. Namely, the bands are not as sharp as the ones of the WT. Several embryos had a smear-like effect in *her7* expression (Figure 3.9 C, red arrows). Moreover, some of the embryos presented an asymmetric expression (Figure 3.9 C, bottom panel). In both cases, the phenotype did not follow Mendelian segregation.

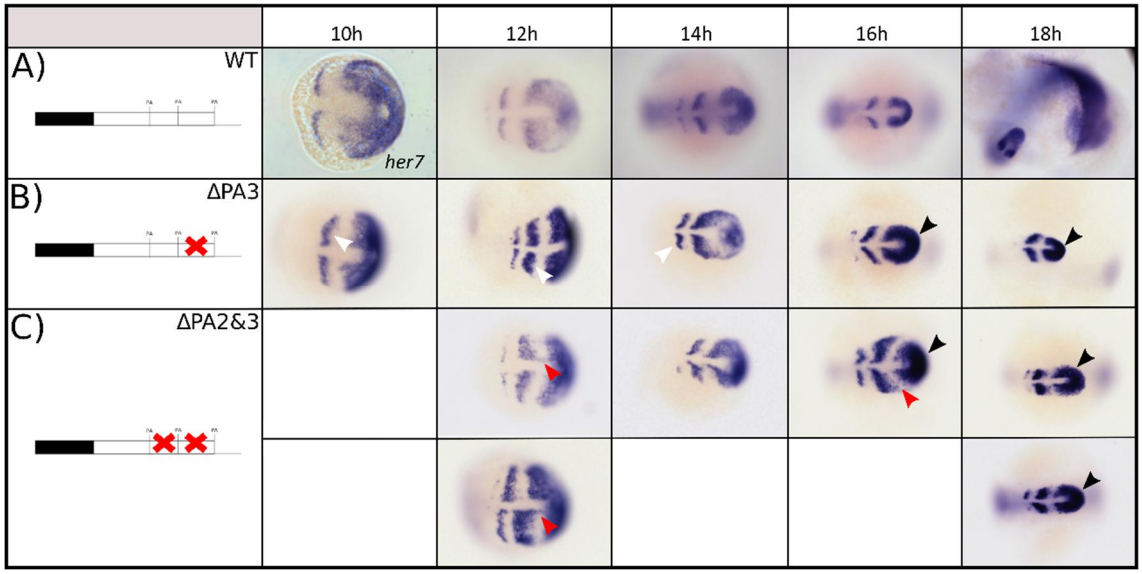



**Figure 3.9 - Whole-mount *in situ* hybridization for *her7* expression on *wild-type* (A) and F2 generation embryos carrying ΔPA3 (B) or ΔPA2&3 (C) deletions.**
In B) and C), generation F2 was obtained by crossing two F1 heterozygous siblings. Representative results of ~25 embryos *per* time point (10, 12, 14, 16, and 18 hpf), *per* condition are shown. C) Upper panel shows embryos with symmetric expression and the bottom panel, asymmetric expression. White arrowheads indicate stripe irregularities; Black arrowheads indicate more intense labeling compared to WT embryos in A); Red arrowheads show a smear-like expression in the F2 embryos.

It was already demonstrated that interfering with the notch pathway could lead to asymmetric expression of *her* genes in zebrafish. Echeverri and Oates demonstrated that knockdown of the Notch target gene Su(H) leads to asymmetric expression of *her* genes (Echeverri and Oates, 2007). This may suggest that alterations in the mRNA stability, even though not interfering with the oscillation behavior, could lead to alterations in delta-notch signaling with consequent problems in the symmetry of clock gene expression.

Another well-described HES family gene in zebrafish somitogenesis is *her1.* According to (Ulitsky *et al.*, 2012), this gene has four possible PA sites, two proximal (ENSDARG00000014722.4 and ENSDARG00000014722.3) generating 3'UTRs with 660 and 723 nucleotides, respectively, and two more distal (ENSDARG00000014722.2 and ENSDARG00000014722.1), generating 3'UTRs with 1276 and 1334 nucleotides, respectively (Figure 3.10 A). To delete the 3' region posterior to the proximal 3'UTRs, we designed gRNAs able to target the 3'UTR coding region after the second PA site (Figure 3.10 B, Blue bars) and gRNAs that target close to the fourth PA site (Figure 3.10 C, Red bars). By using the combination of these two sets of gRNAs, we generated zebrafish embryos harboring a deletion of around 450 nucleotides, *her1-ΔPA2* (Figure 3.10 B, Electrophoresis gel).

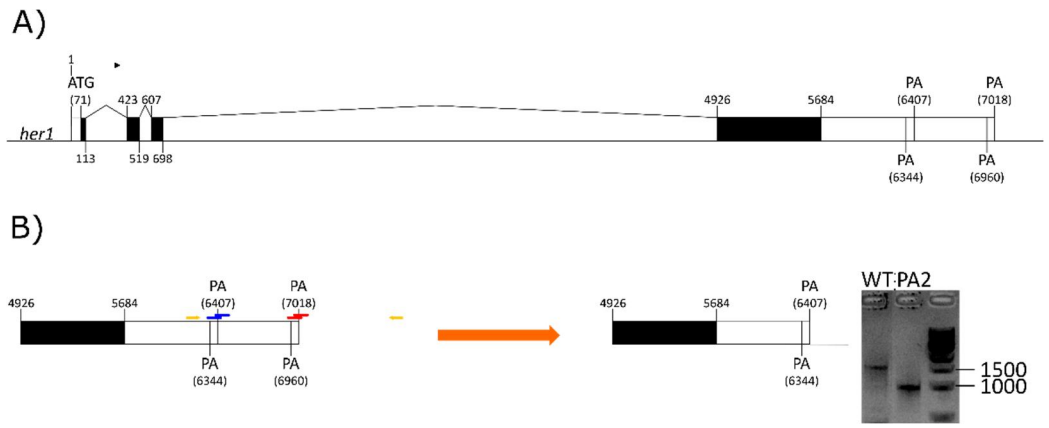



**Figure 3.10 - Structure of *her1 locus* and 3'UTR mutant generation.**
**A)** *her1 locus* according to UCSC Genome Browser on Zebrafish Jul. 2010 (Zv9/danRer7) assembly (Kent *et al.*, 2002). The start site of h*er1* transcript is nucleotide number 1. It can have four different APA sites. Two more proximal PAS give rise to transcripts of 1718 and 1781 nucleotides and two more distal PAS that originate transcripts of 2334 and 2392 nucleotides. Black arrows indicate the transcription start sites (TSS); PA indicates the polyadenylation sites. **B)** Schematics and genotyping of embryos carrying 3'UTR deletion from *her1.* A single deletion in the *her1 locus* was made using CRISPR technology that removed the two more distal 3'UTR (deletion of ~600 nucleotides). On the right, genotyping results of an homozygous embryo harboring the deletion. Blue and red bars indicate the gRNA position for CRISPR. Arrows in yellow indicate the position of the primers used for genotyping.

Once the injected fish achieved sexual maturity, carriers of the ΔPA2 deletion were selected and crossed with WT fish (generation F1). Once the F1 generation achieved sexual maturity, siblings were crossed (generation F2). F2 embryos were collected from 10hpf to 18hpf in intervals of two hours. As a preliminary approach, we performed whole-mount *in situ* RNA hybridization to assess *her1* expression and check any major expression alteration.

The expression dynamics of *her1* closely resembles the expression dynamics of *her7.* The expression of *her1* is activated in the tailbud, through cell-cell interaction, the expression propagates towards the anterior PSM. As the expression anteriorizes, more posterior cells stop expressing it, leading to the emergence of stripes of expression in the anterior PSM (Oates and Ho, 2002). Once the cells integrate a somite, they definitely stop expressing *her1* (Muller *et al.*, 1996) (Figure 3.11 A).

Since the F2 embryos are a cross between two heterozygous siblings, we expected to have an heterogeneous population containing 25% of WT embryos, 50% of heterozygous embryos, and 25% homozygous for the deletion ΔPA2. In this case, we did not observe any alteration to the expression bands' shape, neither any gross modification to morphological development (Figure 3.11 B). In later stages of development, the expression seems more intense (Figure 3.11 B black arrowhead). However, more in-depth experiments should be performed in order to confirm this result, since the domains of expression do not seem to be

affected. As seen in the case of *her7* ΔPA2&3 mutants, *her1* Δ*PA2* also presented in some sporadic cases an asymmetric expression (Figure 3.11 B, bottom panel). This strengthens the hypothesis that 3'UTR of cyclic genes is required for the symmetry of expression between the left and right PSM.

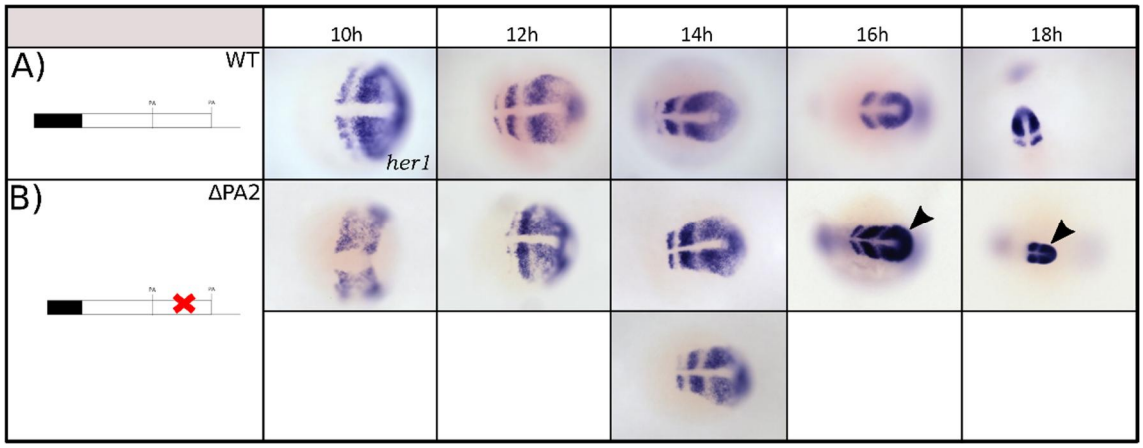



**Figure 3.11 - Whole-mount *in situ* hybridization for *her1* expression detection in *wild-type* (A) and F2 generation embryos carrying ΔPA2 deletion (B).**
In B), generation F2 was obtained by crossing two F1 heterozygous siblings. Representative results of ~25 embryos *per* time point (10, 12, 14, 16, and 18 hpf), *per* condition are shown. B) Upper panel shows embryos with symmetric expression and the bottom panel, asymmetric expression. Black arrowheads indicate more intense labeling compared to WT embryos in A).

### 3.3.3 Analysis of the functional role of deltaC 3' UTR

From the three EC best-described genes, the only one that does not present any alternative PA site is *deltaC* (*dlc*). However, it has a very large 3'UTR, with around 1400 base pairs (Figure 3.6; Figure 3.12 A) (Ulitsky *et al.*, 2012). Here, there are multiple miRNAs target sites predicted by TargetScan (Figure 3.12 B) (Agarwal *et al.*, 2015). We wondered if miRNAs had any implication on the embryonic clock *dlc* oscillations. And, if so, which are the miRNAs involved. Since there are too many miRNA target sites in *deltaC* 3'UTR, we decided to divide it into three main regulatory regions (Figure 3.12 B red squares). Then we asked if any of these single regions or combinations had any relevance to deltaC mRNA stability and, consequently, if it impacted the embryonic clock.

Figure 12 C shows six deletions made using CRISPR technology on the *dlc* 3'UTR. We made three single deletions ΔRR1, ΔRR2, and ΔRR3, two double deletions (ΔRR1&2 and ΔRR2&3), and one triple deletion (ΔRR1,2&3). We designed four pairs of gRNAs, and a combination of two pairs was used to generate each of the deletions (Figure 3.12 C, colored bars represented in the WT *locus* scheme).

Firstly, we generated mutants harboring deletions of a single RR. *Dlc*-ΔRR1 was generated by deleting ~320 nucleotides from the proximal 3'UTR region (using the gRNAs depicted in red and green bars in Figure 3.12 C). The *dlc* transcript generated by this mutant is expressed with a 3'UTR that only has RR2 and RR3 (*dlc*-ΔRR1 3'UTR structure is depicted in Figure 3.12 C, red/green arrow). *Dlc*-ΔRR2 was generated by deleting ~420 nucleotides from the mid-region (using the gRNAs depicted in green and purple bars in Figure 3.12 C). The *dlc* of this mutant is expressed with a 3'UTR that only has RR1 and RR3 (*dlc*-ΔRR2 3'UTR structure is depicted in Figure 3.12 C, green/purple arrow). *Dlc*-ΔRR3 was generated by deleting ~500 nucleotides from the distal 3'UTR region (using the gRNAs depicted in purple and yellow bars in Figure 3.12 C). The *dlc* of this mutant is expressed with a 3'UTR that only has RR1 and RR2. (*dlc*-ΔRR3 3'UTR structure is depicted in Figure 3.12 C, green/purple arrow).

Afterward, we deleted combinations of RR. *Dlc*-ΔRR1&2 was generated by deleting ~900 nucleotides that include the proximal and mid 3'UTR regions (using the gRNAs depicted in red and purple bars in Figure 3.12 C). The *dlc* of this mutant is expressed with a 3'UTR that only has RR3 (*dlc*-ΔRR1&2 3'UTR structure is depicted in Figure 3.12 C, red/purple arrow). *Dlc*-ΔRR2&3 was generated by deleting ~950 nucleotides from the mid and distal prospective 3'UTR genomic sequence (using the gRNAs depicted in green and yellow bars in Figure 3.12 C). The *dlc* of this mutant is expressed with a 3'UTR that only has RR1 (*dlc*-ΔRR2&3 3'UTR structure is depicted in Figure 3.12 C, green/yellow arrow). *Dlc*-ΔRR1,2&3 was generated by deleting ~1300 nucleotides from the 3'UTR region (using the gRNAs depicted in red and yellow bars in Figure 3.12 C). The *dlc* of this mutant is expressed with a 3'UTR that only comprises a few nucleotides after the STOP codon and a few nucleotides comprising the PA signal and PA site (*dlc*-ΔRR1,2&3 3'UTR structure is depicted in Figure 3.12 C, red/yellow arrow).

Injected fish were genotyped for the presence of the mutations (Figure 3.12 C, electrophoresis gels). It is important to evidence that we could find and isolate individual adult fish carriers of each deletion, except the *dlc*-ΔRR1,2&3. For this last experiment, only WT, non-carriers survived until adulthood. The lethality of this deletion could mean that a minimum length of the 3'UTR is required for proper embryo development and survival. During the genotyping, we found that the bands were relatively shorter than expected for this mutant. This may indicate that the deletion was more extensive than what was expected and deleted the PA signal and site, or the end of the ORF. For these reasons, we did not study this mutant further.

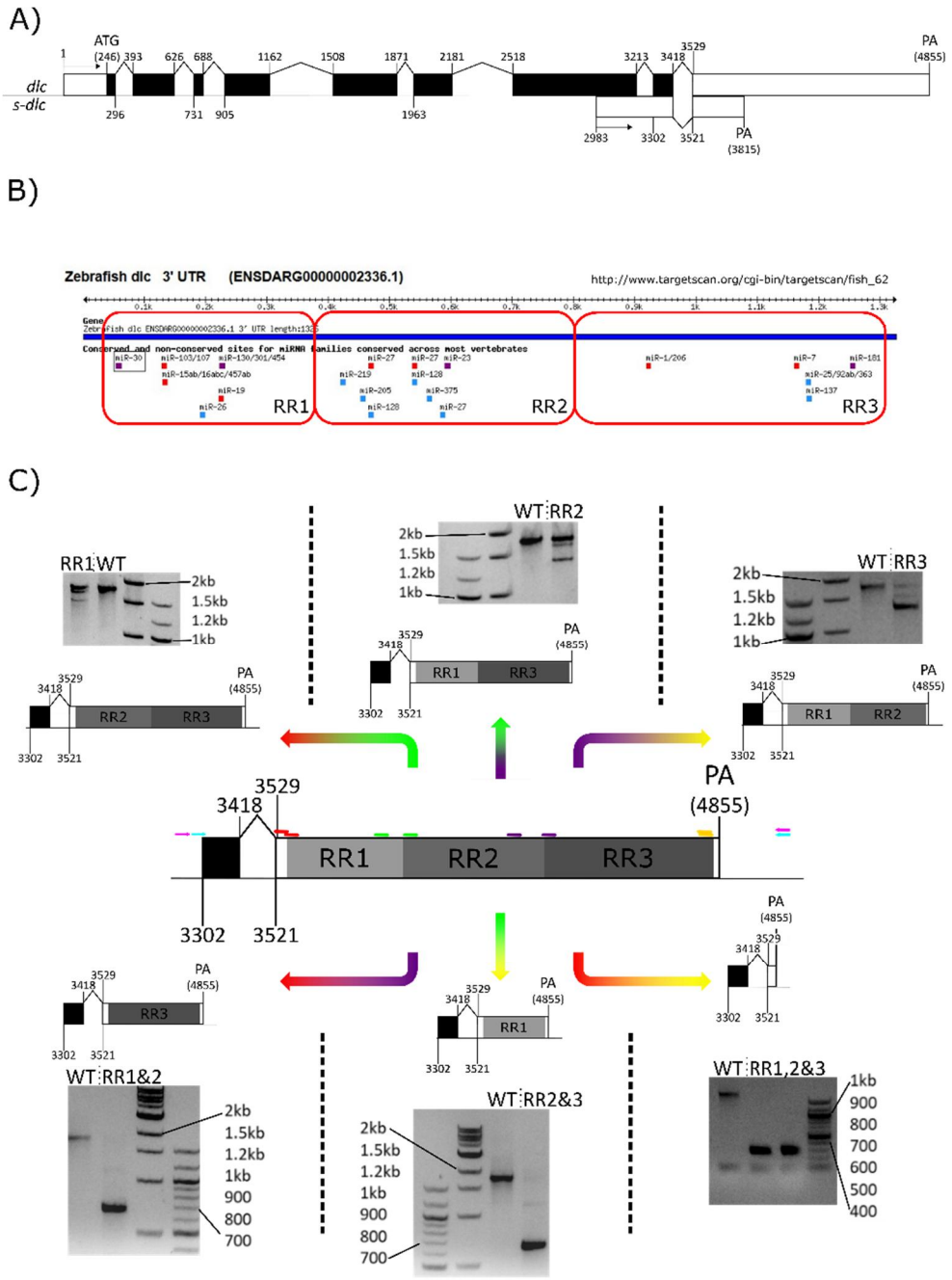



**Figure 3.12 - Structure of *deltaC locus* and 3'UTR mutant generation.**
**A)** *dlc locus* according to UCSC Genome Browser on Zebrafish Jul. 2010 (Zv9/danRer7) assembly (Kent *et al.*, 2002). The start site of *dlc* canonical transcript is nucleotide number 1 and only has one PA site, which generates a transcript of 3567 nucleotides. The transcription of the smaller *s-dlc* starts 2983 nucleotides after the canonical transcript and has only one PA site, originating a transcript of 728 nucleotides. Black arrows indicate the transcription start sites (TSS); PA indicate the polyadenylation site. **B)** Three regulatory regions were identified in the *dlc* 3'UTR . Each region harbors a cluster of miRNAs targets sites (http://www.targetscan.org/cgi-bin/targetscan/fish_62/view_gene.cgi?taxid=7955&rs=ENSDARG00000002336.1&showcnc=0&shownc=0&showncf=). **C)** Representation of the six deletions made in the *dlc locus* using CRISPR technology. In the first deletion, Regulatory Region 1 (RR1) was removed (Red/Green arrow showing a deletion of ~320 nucleotides). The second deletion removed the RR2 (Green/Purple Arrow showing the *locus* with a deletion of ~420 nucleotides). The third deletion

removed RR3 (Purple/Yellow Arrow showing the *locus* with deletion of around 500 nucleotides). The fourth deletion both RR1 and RR2 (RR1&2) (Red/Purple Arrow showing the *locus* with deletion of around 900 nucleotides). The fifth deletion removed the Regulatory Regions 2 and 3 (RR2&3) (Green/Yellow Arrow showing the *locus* with deletion of around 950 nucleotides). The sixth deletion removed RR1,2&3 (Red/Yellow Arrow showing the *locus* with deletion of around 1300 nucleotides). For each deletion represented, there is an example gel of genotyping results of heterozygous or homozygous embryos harboring the mutations. Red, green, purple, and yellow bars in the WT representation of the gene indicate the gRNA position used for CRISPR. Arrows in pink and light blue indicate the primer pairs used for genotyping.

We analyzed the expression pattern of *dlc* in the mutants and tried to identify expression defects. We considered an expression defect if the expression pattern was significantly different from the WT, and not only more intense, since it could be only the consequence of different times of the whole-mount *in situ* hybridization staining. The WT expression pattern of *deltaC* is characterized by a wave-like expression in the posterior PSM, as *her1* and *her7,* when the expression wave goes more anteriorly, it forms an isolated stripe of expression. According to the clock phase, it can show a single or two strong bands of expression, and after that, it stays in basal levels of expression in the somites (Jiang *et al.*, 2000; Smithers *et al.*, 2000) (Figure 3.13).

By deleting the regulatory regions of the 3'UTR, and consequently decreasing the number of miRNA target sites available, we expected that the mRNA would be stabilized and, consequently, the pattern of expression would present alterations. We evaluated at *dlc* expression by whole-mount *in situ* hybridization, in all mutants from 10hpf to 18hpf, except for the RR2&3, which we only analyzed from 12 hours onwards due to very few embryos obtained *per* clutch. In each condition, 25 F2 generation embryos were used, resulting from crosses of heterozygous siblings. In each clutch, we expected to have a heterogeneous population containing 25% of WT embryos, 50% of heterozygous embryos, and 25% homozygous for each deletion (Figure 3.13).

When we deleted the Regulatory Region 1 (RR1), Regulatory Region 2 (RR2), or both at the same time (RR1&2), no prominent effect was detected. The expression pattern of *dlc* closely resembles the one of the WT (Figure 3.13 A). However, small changes in the expression pattern in these three mutants were detected. Firstly, we detected an increased intensity of the labelling in *dlc*-ΔRR1 and *dlc*-ΔRR2 (Figure 3.13 B and C, black arrowheads). However, the shape and expression domains resembled the WT counterparts. The same was not detected with the same frequency in the *dlc*-ΔRR1&2 mutant. Expression intensity was overall identical to *dlc*-ΔRR1 and *dlc*-ΔRR2 mutants (Figure 3.13 D). We found some occasional irregularities in the stripe shape at the beginning of segmentation, *dlc* seems to have a more leaky expression instead of a sharp stripe expression like in the WT (Figure 3.13 D, white and red arrowheads).

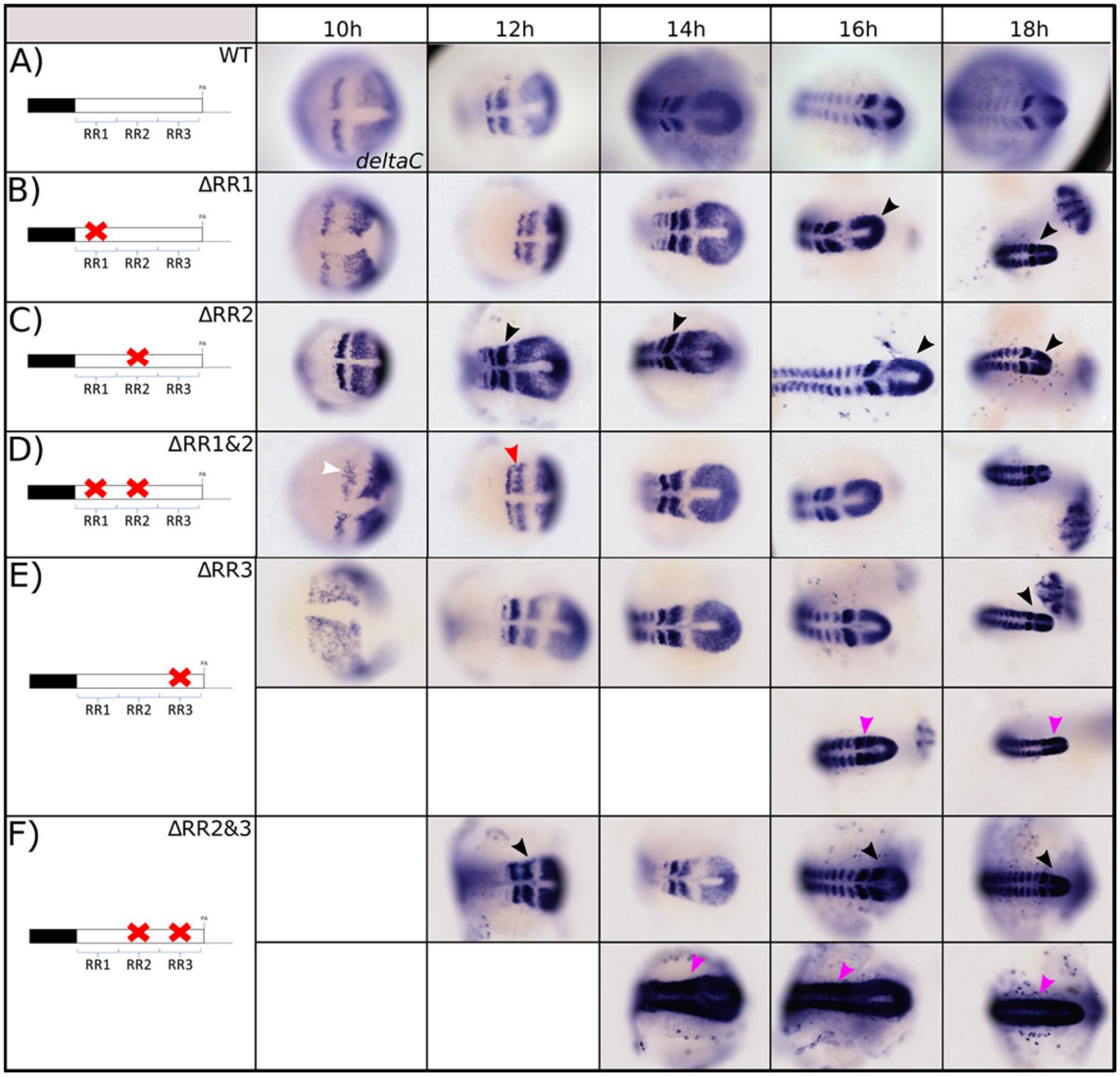



**Figure 3.13 - Whole-mount *in situ* hybridization for *dlc* expression in *wild-type* (A) and F2 generation embryos carrying deletions in the 3'UTR: ΔRR1 (B), ΔRR2 (C), ΔRR1&2 (D), ΔRR3 l or ΔRR2&3 (F).**
In B-F), generation F2 was obtained by crossing two F1 heterozygous siblings. Representative results of ~25 embryos *per* time point (10, 12, 14, 16, and 18 hpf), *per* condition are shown. White Arrowheads indicate stripe irregularities; Black arrowheads indicate more intense labeling compared to WT embryos in A); In E) and F), upper panels show embryos with normal expression stripes on the anterior PSM, and the bottom panels show embryos with increased expression resulting in loss of stripes. Red arrowheads show a smear-like expression in the F2 embryos. Pink arrowheads show an increased *dlc* expression and consequent loss of characteristic stripes in the anterior PSM on F2 embryos.

On the other hand, when the Regulatory region 3 (RR3) was taken, it was possible to see a clear phenotype (Figure 3.13 E, bottom panel). During the first 14 hours of development, most embryos seemed to have a WT expression pattern. From 16 hours onwards, 18% of the

embryos presented a strong *dlc* expression throughout all PSM (N=9/50 embryos from 16 and 18hpf). These embryos showed a continuous expression in all the PSM, and the somites also have an increased expression (Figure 3.13 E, pink arrowheads). It is important to mention that in this case, it is still possible to see *dlc* expression in the individual somites (Figure 3.13 E, bottom panel).

The cumulative deletion of the Regulatory Region 2 and 3 (RR2&3) had an even more evidenced mRNA stabilization phenotype. Overall, the staining was stronger in almost all analyzed embryos, even compared with *dlc*-ΔRR1 and *dlc*-ΔRR2 (Figure 3.13 B). From 14 hpf onwards, in 24,14% (N=14/58 embryos analyzed in all three time points) the expression of *dlc* becomes is so strong that it is impossible to distinguish any pattern. The *dlc* expression was continuously detected from the most posterior part of the tailbud until the anterior formed somites, including in the intersomitc space (Figure 3.1F E, pink arrowheads). However, besides the intersomitic boundary, this strong expression is only present in domains where it was expressed in the WT embryos. There is no expression in the notochord and neural tube, as well in most of the head. Only the five most anterior somites have a compartmentalized expression. This suggests that the transcripts are considerably more stable and maintained highly expressed even after the cells were integrated into a somite (Figure 3.13 F, bottom panel). This preliminary result strongly suggests an important role for the *dlc* 3'UTR in mRNA turnover and, consequently, for embryo clock oscillations.

By perturbing the embryo clock, it was expected that somitogenesis would be also perturbed. To evaluate somite morphology, whole-mount *in situ* hybridization for *xirp2a* (Xin actin-binding repeat-containing 2a) was performed. *Xirp2a* is only expressed in the myotendinous junctions, making it a suitable marker to visualize somite borders (Figure 3.14 A, left image). The total number of somites formed were counted, and somite length was measured of both anterior somites (somites 9 to 13) (Figure 3.14 A, red lines on the right image) and posterior somites (somites 20 to 24) (Figure 3.14 A, purple lines on the right image).

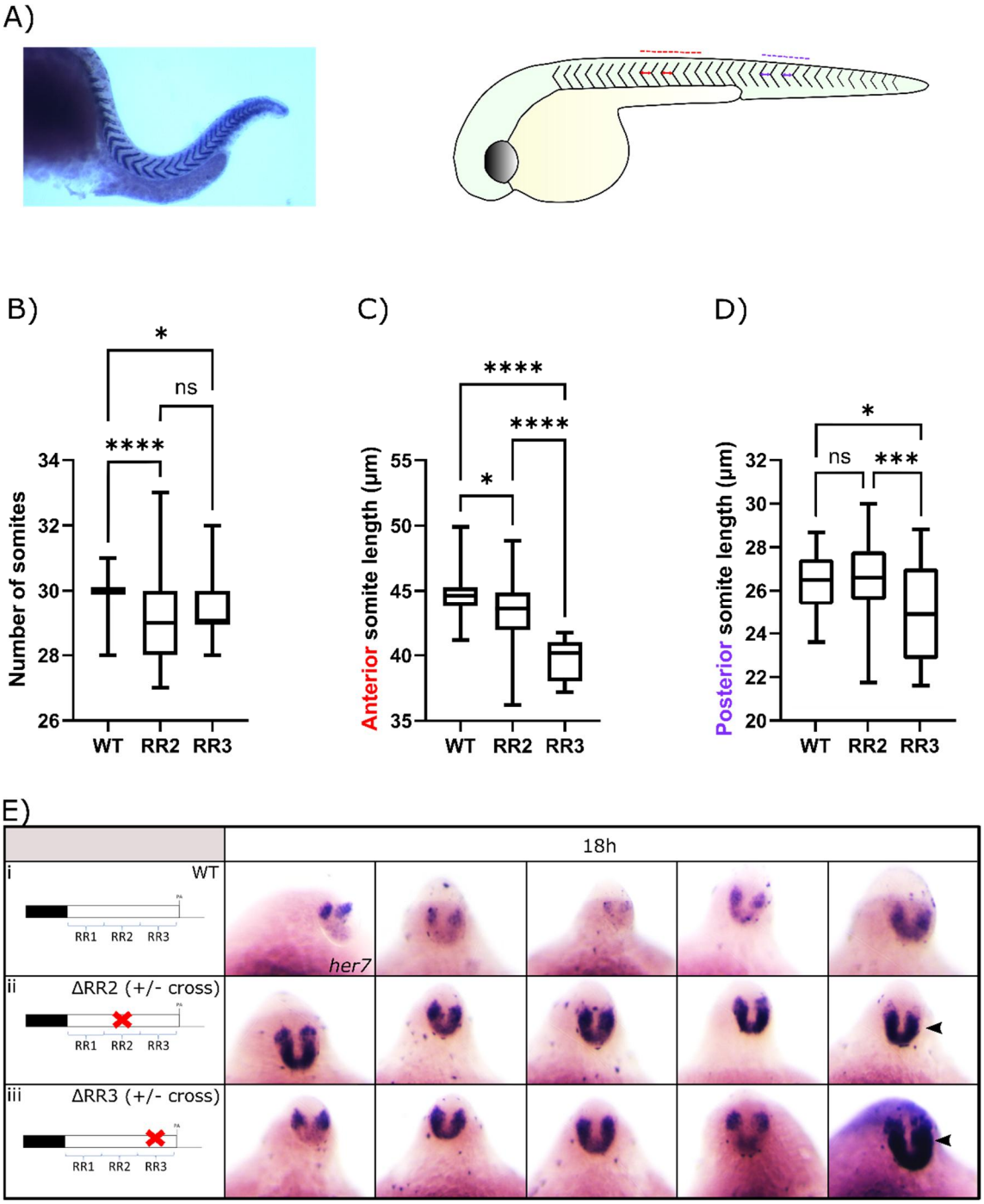



**Figure 3.14 - Characterization of the *dlc* ΔRR2 and ΔRR3 mutants.**
**A)** Whole-mount *in situ* hybridization for *xirp2a* expression (intersomitic gene marker) on a *wild-type* embryo (Left). Representation of how somite measurements were performed, indicating the anterior (Red) and posterior (Purple) somites measured. **B)** Total number of somites of 24hpf WT, ΔRR2, or ΔRR3 embryos. **C)** Length of the anterior somites of 24hpf WT, ΔRR2, or ΔRR3 embryos. **D)** Length of the posterior somites of 24hpf WT, ΔRR2, or ΔRR3 embryos. **E)** Whole-mount *in situ* hybridization for *her7* expression on *wild-type* (i) and embryos carrying ΔRR2 (ii), ΔRR3 (iii) deletions. Black arrowheads indicate more intense labeling compared to WT embryos (summary of work performed by A. Tavares in collaboration with G. Carraco).

Regarding somite integrity, both the *dlc*-ΔRR2 and the *dlc*-ΔRR3 mutants showed normal somite shape and boundaries. However, both mutants formed fewer somites in total than the WT (Figure 3.14 B). According to the clock and wavefront model (Cooke and Zeeman, 1976) this could suggest that the clock is oscillating with a larger period (smaller periodicity), which would result in bigger somites. Hence, we also measured the length of the somites. The five more anterior somite (from somite 9 to 13) and five more posterior ones (from somite 20 to 24) were measured. The WT presented an average of 44.5 μm, while both mutants presented decreased somite length. The *dlc*-ΔRR2 mutant had anterior somites with around 43 μm, and the ΔRR3 mutant had approximately 40 μm (Figure 3.14 C). At the level of the most posterior somites, only the dlc-ΔRR3 had a significant difference in somite length compared with the WT. Both WT and the *dlc*-ΔRR2 mutants had, on average, the same somite length of around 26.5 μm, and the *dlc*-ΔRR3 posterior somites had on average 25 μm (Figure 3.14 D).

Since the somitogenesis was perturbed, we decided to evaluate another EC gene expression, besides the expression of the mutated gene *deltaC*. For that, we evaluated the expression of *her7* at 18hpf, the time when stabilized *dlc* expression was more evident. For this experiment, heterozygous fish carrying the ΔRR2 or ΔRR3 deletions were crossed, and the resulting embryos were collected and fixed 18 hpf. After that, whole-mount *in situ* hybridization was performed for *her7* expression (Figure 3.14 E). In both mutants, there is an evidently increased intensity of *her7* expression in comparison with the WT. Surprisingly, even *dlc*-ΔRR2 where the preliminary results showed only minor differences in *dlc* expression, has *her7* with an increased expression. This suggests that stabilization of *dlc* expression leads to an overexpression of *her7.* Although both mutants presented an increased *her7* expression, it seems to be more evident in the *dlc*-ΔRR3 mutant. In some cases, the expression was detected across all the PSM (Figure 3.14 E ii, black arrowheads). An increased expression of *dlc* and, consequently, an over-activation of Notch signaling may underlie the observations described above.

## 3.4 Chapter Discussion

3'UTRs are, by excellence, mRNA regulatory regions. They can harbor multiple targets sites for proteins or miRNAs that have an impact on RNA stability, translation or even mRNA cellular localization (Berkovits and Mayr, 2015; Mayya and Duchaine, 2019). To better study these regions, and their impact on the embryo clock, we generated several deletions on the 3'UTR of the three best characterized zebrafish clock genes, *her1*, *her7*, and *deltaC* (*dlc*) (Holley *et al.*, 2000; Jiang *et al.*, 2000; Oates and Ho, 2002).

### 3.4.1 *dlc* 3'UTR is required for the cyclic expression of EC genes in *Danio rerio*

In this study, we showed that the *dlc* 3'URT is important to control mRNA expression levels, most likely through regulation of mRNA stability. When we deleted the RR2 or RR3, somite segmentation was affected, and *her7* expression was altered. However, we did not test if the EC periodicity or the segmentation time were also altered. Liao and colleagues found that EC oscillations were accelerated by 6.4% when they introduced 50 extra copies of *deltaD* in the genome, another notch ligand important for the segmentation clock (Liao *et al.*, 2016). In our *dlc*-ΔRR2 and *dlc*-ΔRR3 mutants, although the embryo clock pace was not directly measured, the somites are smaller, so this could suggest that the EC would also be accelerated in both backgrounds. However, the total number of somites formed was diminished, suggesting that there could also be a partial impairment of the molecular mechanisms involved in embryo axial elongation. The same experiments should also be conducted with the *dlc*-ΔRR2&3 mutant where the *dlc* expression was even higher to help elucidate these and other pertinent questions.

Rapid *dlc* mRNA turnover is regulated through Pncr2 and Pumilio responsive elements present in the terminal part of the 3'UTR of *dlc* (Gallagher *et al.*, 2017; Tietz *et al.*, 2020). Our results support these findings, showing that the RR3 of the *dlc* 3'UTR is crucial for mRNA stability. Moreover, we also detected alterations in somite number and size and obtained altered expression levels of *her7* in the *dlc*-ΔRR2 mutant. This suggests that *dlc* instability is not only driven by Pncr2 and Pumilio but also by other factors. Since the *dlc* 3'UTR RR2 is rich in miRNA target sites and also contains an AU-rich element (ARE), they may also be involved in mRNA stability regulation. This could explain why the *dlc*-ΔRR2&3 has an even stronger stabilization than the other two mutants.

Interestingly, by perturbing the expression of dlc, we also affected the expression of her7. We detected an increased expression of her7 in the PSM in both dlc-ΔRR2 and dlc-ΔRR3 mutants. It was already described that downregulation or mutation of dlc leads to downregulation of her7, leading to EC perturbation (Oates and Ho, 2002). In our study, by doing the opposite, overexpressing dlc, we were able to overexpress her7. her genes are downstream targets of Notch/Delta signalling (Wahi *et al.*, 2016), so a constant activation of Notch by Delta1 overexpression may explain this observation.

### 3.4.2 *Danio rerio her1* and *her7* alternative 3'UTRs are dispensable for gene expression oscillations

Recently *her1* 3'UTR was shown to be crucial for the instability of the mRNA molecule by harboring a Pumilio response element and an ARE (Tietz *et al.*, 2020). Tietz and colleges studied the proximal 3'UTR (Tietz *et al.*, 2020) and not the distal part, which we deleted and did not observe alterations in gene expression dynamics. The distal 3'UTR (deleted in this work) also has a Pumilio response element (PRE) and an AU-rich element (ARE), so one PRE and ARE seem to be sufficient to recruit Pnrc2 and the mRNA decay machinery and promote cyclic expression (Fujino *et al.*, 2018; Gallagher *et al.*, 2017; Tietz *et al.*, 2020). This possibly means that in the case of the *her1,* the constant part of the 3'UTR is more important for mRNA stability than the distal alternative APA described (Ulitsky *et al.*, 2012). *her7* also has one PRE and one ARE, which are localized in the proximal 3'UTR, and so, our two mutants, *her7*-ΔPA3 and *her7*-ΔPA2&3 maintained both elements. This may explain how they maintained their cyclic expression in our study.

Symmetric expression of the clock genes requires Retinoic Acid (RA) signaling, which acts as a buffer, preventing the Left-Right signals from reaching the PSM and influencing the EC and somitogenesis (Brend and Holley, 2009). In fact, when *raldh2* expression is impaired, the PSM segments asymmetrically (Kawakami *et al.*, 2005; Lourenço and Saúde, 2010; Sirbu and Duester, 2006; Vermot *et al.*, 2005; Vermot and Pourquié, 2005). Moreover, when Su(H) (Suppressor of hairless) is downregulated, misregulation of *cyp26a1* and asymmetric EC genes expression is observed (Echeverri and Oates, 2007). Curiously, in two of our mutants, *her1*-ΔPA2 and *her7*-ΔPA2&3 we occasionally found asymmetric expression patterns. This suggests a possible role for the 3'UTR in the protection from the left-right signals. To date, there is no evidence supporting a role of the 3'UTR of clock genes in maintaining the symmetry of expression in the PSM. RA has been described to promote RNA stability mediated by binding of the cellular retinoic acid-binding protein 2 (CRABP2), a protein known to transport RA from the cytosol to the nuclear retinoic acid receptor (RAR), to the RNA-binding protein HuR (Vreeland *et al.*, 2014). So, a possible explanation of our observations is that HuR may bind to the deleted 3'UTR regions. This is a very interesting hypothesis that remains to be experimentally tested.

## 3.5 References


**Agarwal, V., Bell, G.W., Nam, J.-W., and Bartel, D.P.** (2015). Predicting effective microRNA target sites in mammalian mRNAs. eLife *4*, e05005. https://doi.org/10.7554/eLife.05005.

**Berkovits, B.D., and Mayr, C.** (2015). Alternative 3′ UTRs act as scaffolds to regulate membrane protein localization. Nature *522*, 363-367. https://doi.org/10.1038/nature14321.

**Brend, T., and Holley, S.A.** (2009). Balancing segmentation and laterality during vertebrate development. Semin Cell Dev Biol *20*, 472-478. https://doi.org/10.1016/j.semcdb.2008.11.009.

**Cooke, J., and Zeeman, E.C.** (1976). A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. Journal of Theoretical Biology *58*, 455-476. http://dx.doi.org/10.1016/S0022-5193(76)80131-2.

**D'Costa, A., and Shepherd, I.T.** (2009). Zebrafish development and genetics: introducing undergraduates to developmental biology and genetics in a large introductory laboratory class. Zebrafish *6*, 169-177. https://doi.org/10.1089/zeb.2008.0562.

**Echeverri, K., and Oates, A.C.** (2007). Coordination of symmetric cyclic gene expression during somitogenesis by Suppressor of Hairless involves regulation of retinoic acid catabolism. Dev Biol *301*, 388-403. https://doi.org/10.1016/j.ydbio.2006.10.003.

**Farrell, J.A., Wang, Y., Riesenfeld, S.J., Shekhar, K., Regev, A., and Schier, A.F.** (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. Science *360*. https://doi.org/10.1126/science.aar3131.

**Fujimuro, T., Matsui, T., Nitanda, Y., Matta, T., Sakumura, Y., Saito, M., Kohno, K., Nakahata, Y., and Bessho, Y.** (2014). Hes7 3'UTR is required for somite segmentation function. Sci Rep *4*, 6462. https://doi.org/10.1038/srep06462.

**Fujino, Y., Yamada, K., Sugaya, C., Ooka, Y., Ovara, H., Ban, H., Akama, K., Otosaka, S., Kinoshita, H., Yamasu, K., et al.** (2018). Deadenylation by the CCR4-NOT complex contributes to the turnover of hairy-related mRNAs in the zebrafish segmentation clock. FEBS Lett *592*, 3388-3398. https://doi.org/10.1002/1873-3468.13261.

**Gagnon, J.A., Valen, E., Thyme, S.B., Huang, P., Akhmetova, L., Pauli, A., Montague, T.G., Zimmerman, S., Richter, C., and Schier, A.F.** (2014). Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. PloS One *9*, e98186. https://doi.org/10.1371/journal.pone.0098186.

**Gallagher, T.L., Tietz, K.T., Morrow, Z.T., McCammon, J.M., Goldrich, M.L., Derr, N.L., and Amacher, S.L.** (2017). Pnrc2 regulates 3'UTR-mediated decay of segmentation clock-associated transcripts during zebrafish segmentation. Developmental Biology *429*, 225-239. https://doi.org/10.1016/j.ydbio.2017.06.024.

**Holley, S.A., Geisler, R., and Nüsslein-Volhard, C.** (2000). Control of her1 expression during zebrafish somitogenesis by a delta-dependent oscillator and an independent wave-front activity. Genes Dev *14*, 1678-1690. https://doi.org/10.1101/gad.14.13.1678

**Jiang, Y.J., Aerne, B.L., Smithers, L., Haddon, C., Ish-Horowicz, D., and Lewis, J.** (2000). Notch signalling and the synchronization of the somite segmentation clock. Nature *408*, 475-479. https://doi.org/10.1038/35044091.

**Kawakami, Y., Raya, Á., Raya, R.M., Rodríguez-Esteban, C., and Belmonte, J.C.I.** (2005). Retinoic acid signalling links left–right asymmetric patterning and bilaterally symmetric somitogenesis in the zebrafish embryo. Nature *435*, 165-171. https://doi.org/10.1038/nature03512.

**Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., Haussler, and David** (2002). The Human Genome Browser at UCSC. Genome Research *12*, 996-1006. https://doi.org/10.1101/gr.229102.

**Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B., and Schilling, T.F.** (1995). Stages of embryonic development of the zebrafish. Dev Dyn *203*, 253-310. https://doi.org/10.1002/aja.1002030302.

**Liao, B.K., Jorg, D.J., and Oates, A.C.** (2016). Faster embryonic segmentation through elevated Delta-Notch signalling. *7*, 11861. https://doi.org/10.1038/ncomms11861.

**Lourenço, R., and Saúde, L.** (2010). Symmetry OUT, Asymmetry IN. Symmetry *2*, 1033-1054. https://doi.org/10.3390/sym2021033

**Mayya, V.K., and Duchaine, T.F**. (2019). Ciphers and Executioners: How 3′-Untranslated Regions Determine the Fate of Messenger RNAs. Frontiers in Genetics *10*. https://doi.org/10.3389/fgene.2019.00006.

**Meeker, N.D., Hutchinson, S.A., Ho, L., and Trede, N.S.** (2007). Method for isolation of PCR-ready genomic DNA from zebrafish tissues. Biotechniques *43*, 610, 612, 614. https://doi.org/10.2144/000112619.

**Montague, T.G., Cruz, J.M., Gagnon, J.A., Church, G.M., and Valen, E.** (2014). CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Res *42*, W401-407. https://doi.org/10.1093/nar/gku410.

**Muller, M., Weizsacker, E., and Campos-Ortega, J.A.** (1996). Expression domains of a zebrafish homologue of the Drosophila pair-rule gene hairy correspond to primordia of alternating somites. Development *122*, 2071-2078. https://doi.org/10.1242/dev.122.7.2071.

**Nitanda, Y., Matsui, T., Matta, T., Higami, A., Kohno, K., Nakahata, Y., and Bessho, Y.** (2014). 3'-UTR-dependent regulation of mRNA turnover is critical for differential distribution patterns of cyclic gene mRNAs. FEBS J *281*, 146-156. https://doi.org/10.1111/febs.12582.

**Oates, A.C., and Ho, R.K.** (2002). Hairy/E(spl)-related (Her) genes are central components of the segmentation oscillator and display redundancy with the Delta/Notch signaling pathway in the formation of anterior segmental boundaries in the zebrafish. Development *129*, 2929-2946. https://doi.org/10.1242/dev.129.12.2929

**Pathak, N.H., and Barresi, M.J.F.** (2020). Chapter 45 – Zebrafish as a Model to Understand Vertebrate Development. In The Zebrafish in Biomedical Research, S.C. Cartner, J.S. Eisen, S.C. Farmer, K.J. Guillemin, M.L. Kent, and G.E. Sanders, eds. (Academic Press), pp. 559-591. https://doi.org/10.1016/B978-0-12-812431-4.00045-2.

**Sirbu, I.O., and Duester, G.** (2006). Retinoic-acid signalling in node ectoderm and posterior neural plate directs left-right patterning of somitic mesoderm. Nat Cell Biol *8*, 271-277. https://doi.org/10.1038/ncb1374.

**Smithers, L., Haddon, C., Jiang, Y.-J., and Lewis, J.** (2000). Sequence and embryonic expression of deltaC in the zebrafish. Mechanisms of Development *90*, 119-123. https://doi.org/10.1016/S0925-4773(99)00231-2.

**Stern, C.D.** (2005). The Chick: A Great Model System Becomes Even Greater. Developmental Cell *8*, 9-17. https://doi.org/10.1016/j.devcel.2004.11.018.

**Thisse, C., and Thisse, B.** (2008). High-resolution in situ hybridization to whole-mount zebrafish embryos. Nat Protoc *3*, 59-69. https://doi.org/10.1038/nprot.2007.514.

**Tietz, K.T., Gallagher, T.L., Mannings, M.C., Morrow, Z.T., Derr, N.L., and Amacher, S.L.** (2020). Pumilio response and AU-rich elements drive rapid decay of Pnrc2-regulated cyclic gene transcripts. Developmental Biology *462*, 129-140. https://doi.org/10.1016/j.ydbio.2020.03.017.

**Ulitsky, I., Shkumatava, A., Jan, C.H., Subtelny, A.O., Koppstein, D., Bell, G.W., Sive, H., and Bartel, D.P.** (2012). Extensive alternative polyadenylation during zebrafish development. Genome Res *22*, 2054-2066. https://doi.org/10.1101/gr.139733.112.

**Vermot, J., Llamas, J.G., Fraulob, V., Niederreither, K., Chambon, P., and Dollé, P.** (2005). Retinoic Acid Controls the Bilateral Symmetry of Somite Formation in the Mouse Embryo. Science *308*, 563-566. https://doi.org/10.1126/science.1108363.

**Vermot, J., and Pourquié, O.** (2005). Retinoic acid coordinates somitogenesis and left-right patterning in vertebrate embryos. Nature *435*, 215-220. https://doi.org/10.1038/nature03488.

**Vreeland, A.C., Yu, S., Levi, L., Rossetto, D.d.B., and Noy, N.** (2014). Transcript Stabilization by the RNA-Binding Protein HuR Is Regulated by Cellular Retinoic Acid-Binding Protein 2. Mol Cell Biol *34*, 2135-2146. https://doi.org/10.1128/MCB.00281-14.

**Wahi, K., Bochter, M.S., and Cole, S.E.** (2016). The many roles of Notch signaling during vertebrate somitogenesis. Seminars in Cell & Developmental Biology *49*, 68-75. https://doi.org/10.1016/j.semcdb.2014.11.010.

**Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., and Madden, T.L.** (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics *13*, 134. https://doi.org/10.1186/1471-2105-13-134.

# Chapter 4

## Regulation of the embryo clock by miRNAs

*This page is intentionally left blank*

## 4.1 Chapter Introduction

The EC is characterized by a negative feedback loop mechanism underlying gene expression oscillations. This mechanism relies on the ability of some transcription factors to repress their own transcription (Bessho *et al.*, 2003; Bessho and Kageyama, 2003). It is finetuned by delays, such as delay in transcription, nuclear exportation, translation, and mRNA and protein decay (Saga, 2012). To maintain their cyclic expression, the activation and repression of these genes needs to be tightly regulated. Small perturbations in the EC dynamics can lead to a complete loss of gene expression oscillations (Reviewed in Section 1.4).

Besides being conserved among vertebrates, operating in the PSM with a species-specific periodicity, the EC was already reported in the development of other tissues, and with tissue-specific periodicities. Pascoal and colleagues found that *hairy2* has a cyclic expression in the distal forelimb during its outgrowth with a six-hour period, instead of 90 minutes as in the PSM (Pascoal *et al.*, 2007). So, how are different periods of gene expression oscillations generated?

micro-RNAs (miRNA) are small molecules containing ~23 nucleotides, which target mRNA 3'UTRs leading to translation repression and/or mRNA degradation (Bartel, 2018). Riley et al. described that micro-RNA (miRNA) mir-125a-5p controls the pace of the chicken EC by targeting the cyclically expressed gene, *lunatic fringe* (*lfng*). When miR-125a-5p was misexpressed (overexpressed or downregulated), or its target on *lfng* was protected from its action, the expression of another EC gene, *hairy1,* was continuous and static through all PSM (Riley *et al.*, 2013). Also, miR-9 leads to mouse *hes1* and zebrafish *her6* mRNA degradation, which is crucial for the maintenance of *hes1* and *her6* expression oscillations in neural stem cells (Bonev *et al.*, 2012; Soto *et al.*, 2020). Based on this information, we elaborated the following hypothesis: miRNA species in the different tissues account for tissue-specific mRNA half-lives and, consequently, different EC paces.

Since knowledge of miRNAs that may have a role in the EC is scarce, we performed high throughput miRNA sequencing (miRNA-Seq) in order to find known and unknown miRNAs that could be related to the EC and which could explain the different EC periodicity in chicken development. Three chick embryo tissues were analyzed: undetermined PSM (U-PSM), determined PSM (D-PSM), and the limb distal cyclic domain (DCD). As PSM cells pass the determination front, cyclic gene expression progressively stabilizes. For this reason, we divided the PSM into determined and undetermined tissues, and we analysed them separately by miRNA-Seq. This miRNA-Seq approach resulted in two scientific manuscripts (one currently under review), which are the main text of the present chapter.

## 4.2 “gga-miRNOME, a microRNA-sequencing dataset from chick embryonic tissues” (*manuscript in press)*

manuscripttrackingsystem scientific data

tracking system home | author instructions | reviewer instructions | help | logout | journal home

**Detailed Status Information**

| | |
|---|---|
| **Manuscript #** | SDATA-21-00646B |
| **Current Revision #** | 2 |
| **Other Versions** | SDATA-21-00646 SDATA-21-00646A |
| **Submission Date** | 15th December 21 |
| **Current Stage** | |
| **Title** | gga-miRNOME, a microRNA-sequencing dataset from chick embryonic tissues |
| **Manuscript Type** | Data Descriptor |
| **Corresponding Author** | Dr Raquel Andrade (rgandrade@ualg.pt) (Universidade do Algarve) |
| **Contributing Authors** | Dr Isabel Duarte , Dr Gil Carraco , Dr Nayara Azevedo , Dr Vladimir Benes |
| **Authorship** | Yes |
| **Abstract** | MicroRNAs (miRNAs) are small non-coding RNA molecules, with sizes ranging from 18 to 25 nucleotides, which are key players in gene expression regulation. These molecules play an important role in fine-tuning early vertebrate embryo development. However, there are scarce publicly available miRNA datasets from non-mammal embryos, such as the chicken (Gallus gallus), which is a classical model system to study vertebrate embryogenesis. Here, we performed microRNA-sequencing to characterize the early stages of trunk and limb development in the chick embryo. For this, we profiled three chick embryonic tissues, namely, Undetermined Presomitic Mesoderm (PSM_U), Determined Presomitic Mesoderm (PSM_D) and Forelimb Distal Cyclic Domain (DCD). We identified 926 known miRNAs, and 1,141 novel candidate miRNAs, which nearly duplicates the number of Gallus gallus entries in the miRBase database. These data will greatly benefit the avian research community, particularly by highlighting new miRNAs potentially involved in the regulation of early vertebrate embryo development, that can be prioritized for further experimental testing. |
| **Editorial Board Member** | Assigned |
| **Subject Terms** | Biological sciences/Molecular biology/Non-coding RNAs/miRNAs<br>Biological sciences/Developmental biology/Experimental organisms/Non-model organisms |
| **Competing Interests Policy** | There is **NO** Competing Interest. |
| **Human Subjects Question** | No |
| **Competing Interests Policy** | No competing interest declared |
| **Applicable Funding Source** | Fundação para a Ciência e Tecnologia (FCT), Portugal [Andrade]<br>Fundação para a Ciência e Tecnologia (FCT), Portugal [Duarte]<br>Fundação para a Ciência e Tecnologia (FCT), Portugal [Carraco] |
| **Publication Agreements** | 1 of 1 tasks completed.<br>**Article Processing Charge:** completed by Raquel (P) Andrade on 21st December 21 |

| Stage | Start Date | End Date | Approximate Duration |
|---|---|---|---|
| | 15th December 21 | | |
| Decision Made | 15th December 21 | | |
| Awaiting Decision | 15th December 21 | | |
| Reviewer Assignment | 15th December 21 | | |
| Editor Assignment | 13th December 21 | | |
| Author Approved Converted Files | 13th December 21 | | |
| Preliminary Manuscript Data Submitted | 13th December 21 | | |

eJournalPress

tracking system home | privacy policy | cookie policy | manage cookies

# gga-miRNOME, a microRNA-sequencing dataset from chick embryonic tissues

Isabel Duarte[1,2,*], Gil Carraco[1,3,4,*], Nayara T. D. de Azevedo[5], Vladimir Benes[5], Raquel P. Andrade[1,4,6,¥]

1. Faculdade de Medicina e Ciências Biomédicas (FMCB), Universidade do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal
2. Center for Health Technology and Services Research (CINTESIS), Polo da Universidade do Algarve, 8005-139 Faro, Portugal
3. ProRegeM-PhD Program in Mechanisms of Disease and Regenerative Medicine, Faro, Portugal
4. ABC-RI, Algarve Biomedical Center Research Institute, Faro, Portugal
5. Genomics Core Facility, EMBL, Heidelberg, Germany
6. Champalimaud Research Program, Champalimaud Center for the Unknown, Lisbon, Portugal

* These authors contributed equally to this work

¥Corresponding author: Raquel P. Andrade (rgandrade@ualg.pt)

## Abstract

MicroRNAs (miRNAs) are small non-coding RNA molecules, with sizes ranging from 18 to 25 nucleotides, which are key players in gene expression regulation. These molecules play an important role in fine-tuning early vertebrate embryo development. However, there are scarce publicly available miRNA datasets from non-mammal organisms, such as the chicken (*Gallus gallus*), which is a classical model system to study vertebrate embryogenesis. Here, we performed microRNA-sequencing to characterize the early stages of trunk and limb development in the chick embryo. For this, we profiled three chick embryonic tissues, namely, Undetermined Presomitic Mesoderm (PSM_U), Determined Presomitic Mesoderm (PSM_D) and Forelimb Distal Cyclic Domain (DCD). We identified 926 known miRNAs, and 1,141 novel candidate miRNAs, which nearly duplicates the number of *Gallus gallus* entries in the miRBase database. These data will greatly benefit the avian research community, particularly by highlighting new miRNAs potentially involved in the regulation of early vertebrate embryo development, that can be prioritized for further experimental testing.

## Background & Summary

MicroRNAs (miRNAs) are small, single-stranded RNAs with sizes ranging from 18 to 25 nucleotides that are involved in gene expression regulation. This is achieved via post-transcriptional silencing of complementary messenger RNA (mRNA) targets by repression of translation and/or mRNA degradation[1]. miRNAs were initially called small temporal RNAs (stRNAs), since they were first described as essential for proper developmental stage transition in the *C. elegans* life cycle[2]. Today they are recognized to act as gatekeepers of developmental time in many other systems, by mediating cell proliferation-to-differentiation transitions[3].

The canonical pathway of miRNA biogenesis starts with the transcription of a primary miRNA (pri-miRNA) by RNA polymerase II. The pri-miRNA forms a hairpin that is recognized by DGCR8, which recruits a Class 2 ribonuclease III enzyme, Drosha. This enzyme cleaves the RNA releasing the hairpin, called precursor miRNA (pre-miRNA), which is then exported to the cytoplasm via the Exportin-5 transporter. Here, it is recognized by a second Class 2 ribonuclease III enzyme, Dicer, that cleaves the loop from the hairpin releasing a small double-stranded RNA. One of the strands binds to an Argonaute protein from the RNA-induced silencing complex (RISC), while the other is degraded. At this point, the mature miRNA selectively recognizes and binds to the 3' untranslated region (3'UTR) of its target mRNA through a small 2-7 nucleotide seed region, leading to RISC-mediated mRNA degradation and/or translational repression[1].

An essential step in addressing miRNA-mediated regulation of gene expression is to identify and quantify the miRNAs present in the biological system of interest. High throughput miRNA profiling studies have identified thousands of miRNAs in Human and mouse samples[4]. However, this effort has been lagging behind in other model organisms, hindering the elucidation of their role in these systems. This is the case of the chicken (*Gallus gallus*) embryo, a well-established model for studying human embryogenesis due to its extraordinary molecular and morphological similarities in the early stages of development, alongside the ease of experimental manipulation it offers[5]. It was in the chicken embryo that the molecular embryonic clock (EC) underlying the periodic formation of vertebrae precursors was first described[6]. EC genes present cyclic expression maintained by negative feedback regulation in the posterior undetermined presomitic mesoderm (PSM), which gradually slows in the anterior PSM and halts in the segmented somites[7,8]. The periodicity of gene expression oscillations in the PSM is species-specific but can also differ in different tissues of the same organism. Namely, *hairy2* gene expression oscillates with a periodicity of 90 min in the chick PSM and 6h in the distal cyclic domain (DCD) of the developing forelimb bud[9].

mRNA instability is essential for EC cycles of expression and there is evidence of a miRNA-dependent regulation of EC gene oscillations[10]. Namely, miR-125a-5p is required for cyclic *LFNG* expression in the chick PSM[11] and miR-9 drives *Hes1* oscillations in mouse neural progenitor cells[12]. Additionally, we previously showed that the genes encoding the enzymatic machinery for miRNA biogenesis are expressed in both the chick PSM and forelimb bud[13], tissues where the EC is oscillating.

A thorough characterization of the role of small RNAs in chick embryo development and in the regulation of the EC has been hampered by the scarcity of miRNA expression datasets in embryonic tissues of this model system. To overcome this limitation, we performed a miRNA profiling analysis (miRNA-Seq) of three different tissues of the developing chick embryo (Figure 1a,b). Namely, two regions of the PSM - Undetermined Presomitic Mesoderm (PSM_U) and Determined Presomitic Mesoderm (PSM_D) - and the Forelimb Distal Cyclic Domain (Limb). We report the identification of 926 known miRNAs, and 1,141 candidate novel miRNAs, not previously described in chicken. Accordingly, we believe that this will be an invaluable data resource for the research community studying miRNA-mediated gene expression in early vertebrate development, particularly in the chick embryo.

## Methods

### Embryos

Fertilized *Gallus gallus* eggs (Pintobar, Portugal) were incubated at 38ºC in a humidified atmosphere for two or four days to obtain embryos in stages HH12-13 and HH20-22[14], respectively.

### Sample collection

Presomitic mesoderm (PSM) tissues were isolated from embryos in stages HH12-13. To obtain these samples, embryos were collected from 48h-incubated eggs, placed in a petri dish containing phosphate buffer saline (PBS) solution and staged according to Hamburger and Hamilton[14]. Only the embryos in stages HH12-13 were selected for further use. The embryos were then placed ventral side up in PBS and 4 µL of Pancreatin (25 mg/mL) (Sigma #8049-47-6) was added to the surface of the embryo. After 3 to 5 minutes, pancreatin was inactivated with goat serum (Gibco #16210-072). The mesoderm located on either side of the neural tube was isolated from all surrounding tissues and divided into determined PSM (PSM_D, upper one-third portion) and undetermined PSM (PSM_U, caudal two-thirds) (Figure

1a,b). Due to the extraordinarily small size of these tissues, 20 pairs of PSM portions were pooled together for RNA extraction from each biological sample. The samples were snap frozen in liquid nitrogen and stored at -80ºC.

Distal Cyclic Domain (DCD Limb) tissues were isolated from embryos at stages HH20-22 (Figure 1a,b). Embryos were collected from fertilized eggs incubated for four days, placed in PBS and staged according to Hamburger and Hamilton[14]. Only the embryos in stages HH20-22 were selected for further use. The limb tissue (distal medial portion of the forelimb bud) was manually dissected using forceps. 20 DCD Limb pairs were pooled together for each sample, snap frozen in liquid nitrogen and stored at -80ºC.

**RNA extraction**

Biological samples were defrosted on ice. Total RNA was extracted using TRIzol Reagent (Invitrogen #15596-018) according to the manufacturer's instructions with slight adaptations, namely, the aqueous phase from the first step of extraction was washed once with Phenol:Chloroform (Sigma #P2069) and then with Chloroform:Isoamyl Alcohol (24:1). The aqueous phase was recovered using Phase lock Gel Heavy (5Prime #2302830) and RNA was precipitated by addition of 1/10 volume of 3 M sodium acetate, 2.5 volumes of 100% ethanol and 3 µL per mL of Linear Acrylamide (Ambion #AM9520). After one hour at -80ºC, the RNA was precipitated by centrifugation at 14,000 rpm for 30 minutes at 4ºC. The pellet was washed with 70% ethanol and centrifuged for 15 minutes at 4ºC, briefly air-dried and resuspended in 50 µL of MilliQ (Merck Millipore) purified water. The samples were quantified using NanoDrop 2000 (Thermo Scientific) and stored at -80ºC.

**RNA quality control**

A first-round of quality control was performed by Reverse Transcription-PCR. 100ng of RNA was reverse transcribed using iScript™ cDNA Synthesis Kit (BioRad #1708890). Subsequent PCR for GAPDH was done using DreamTaq DNA Polymerase (Thermo Scientific™ #EP0701). In a second instance, RNA quality control was performed using Experion™ RNA StdSens Analysis Kit (BioRad #700-7103) (Table 1). Only samples with an RQI (RNA Quality Indicator) equal to or above 8.5 were sent for sequencing.

**Library preparation and miRNA-sequencing**

The sequencing libraries were prepared using the NEBNext Multiplex Small RNA Library Prep Set for Illumina (NEB #E7300S/L Version 5.0), starting with 150 ng of total RNA

as input. As a first step in the protocol, adaptors ligate directly to the small RNA fragments containing 5' phosphate and 3' OH, followed by cDNA generation and PCR amplification. 15 cycles of amplification were performed using specific SR primers for Illumina and index primer of choice for each sample (according to NEB #E7300S/L Version 5.0 protocol).

Size distribution of the final library was assessed on Bioanalyzer (Agilent Technologies) with a DNA High Sensitivity kit (Agilent Technologies #5067-4626), and concentration was measured with Qubit® DNA High Sensitivity kit (Life Technologies #Q32854) in Qubit® 2.0 Flurometer (Life Technologies). Individual libraries that passed the QC step were pooled equimolarly in a 9-plex, and final pool was purified with SPRI select beads at a 1.3x bead ratio (Beckman Coulter #B23319). Pool was loaded to a single lane of an Illumina HiSeq2000 sequencing instrument (Illumina Inc.) at 6 pM concentration and was sequenced in 50 bp single-read mode[26]. Library construction and sequencing were performed at EMBL's GeneCore facility in Heidelberg, Germany.

**Quality control of sequencing reads**

Sequencing reads were firstly evaluated using FastQC (version 0.11.5)[15] to verify the overall read quality of each sample. One library (PSM_U2), from undetermined PSM, did not pass the quality control step, mainly due to its small library size, leading to its removal from further analyses (Figure 2a).

**Quantification and Normalization of annotated miRNAs**

For the remaining 8 samples that passed the quality control, annotated microRNA read counts were obtained using the Chimira software (version 1.5)[16]. Briefly, the pipeline implemented in Chimira for miRNA-seq analysis comprises the following steps: firstly, the sequences are cleaned, trimmed, and size selected to remove adapters and low quality microRNA reads. Next, the reads passing the previous filters are mapped to *Gallus gallus* hairpin sequences present in miRBase (release 22)[4] using BLASTn[17] allowing up to two mismatches. Finally, a count-based miRNA expression dataset is generated[27] and normalized across all samples using DESeq2[18]. Further data validation, visualization, and statistical analyses were conducted using the normalized log2 expression data.

**Detection, Quantification and Normalization of novel miRNAs**

Detection of novel miRNAs was performed using the Mirnovo tool (v1.0)[19], which is a machine learning algorithm that predicts novel miRNAs by analysing structural features of

miRNA precursor hairpin sequences gathered directly from small RNA-Sequencing data. Briefly, the Mirnovo pipeline entails the following steps: (i) adapter removal followed by sequence de-duplication; (ii) the tallied sequences then enter a series of clustering steps, followed by cluster refinement to obtain consensus sequences; (iii) the prediction step identifies known and novel miRNAs; and (iv) the final step, aligns the consensus sequences from all miRNAs (known and novel) to the reference genome. This was done by selecting the most stable hairpins (scored by Delta G free energy) found in a 90-nucleotide window around the consensus sequences followed by genomic feature calculation[19].

The specific parameters used for Mirnovo were: *Gallus gallus* input species, using the Universal prediction model (since there are no models specifically trained for chicken), length filter between 16 and 28 nucleotides, minimum read depth of 5, minimum variants 1, and initial clustering using an alignment identity threshold of 0.9 (vsearch-id parameter).

The candidate novel miRNAs were then quantified and normalized using Chimira[16] with the Mirnovo extension. The analysis was performed as described above for the annotated miRNAs, with the difference that the custom hairpin FASTA files output from Mirnovo for each sample were uploaded together with the corresponding raw FASTQ files.

Data resulting from this identification and quantification (i.e. hairpin sequences, genomic location, and normalized counts) is freely available[28].

### microRNA expression profiling

Using customized R scripts (R version 3.6)[20], we conducted quality control analysis, and briefly inspected the profile of annotated and novel microRNA expression in each embryo tissue. For this we used R packages for data visualization, namely, *Tidyverse*[21], *UpSetR*[22], *Patchwork*[23], and *plot3D*[24].

## Data Records

All sequencing data has been deposited in the ArrayExpress data repository[25] with accession number E-MTAB-8176[26]. This dataset consists of 8 microRNA expression raw data files in fastq format. Detailed experimental procedures and data analysis are also available there.

Processed data (in tabular text format) containing the log2 normalized counts of the sequencing reads for annotated miRNAs has been deposited in Figshare[27]. Similarly, the list

of predicted novel miRNAs, with sequence, and log2 normalized counts is available in Figshare[28].

All the sequencing data and the normalized miRNA expression counts are open. The R code used for the exploratory data analysis and visualizations are also freely available for consultation in Figshare[29].

## Technical Validation

### Quality control of microRNA-Seq data

The quality control of the raw sequencing reads was performed using FastQC[15] to assess overall read quality and flag potentially poor-quality samples. All samples except one, passed the QC metrics performed by FastQC[15]. The poor-quality sample presented a variable PHRED score distribution across the read length (Figure 2a), as well as a very low total number of reads (244,296 reads compared to 3 million average reads in the other samples), hinting that the sequencing step was faulty, possibly due to sample degradation prior to or during library preparation. Accordingly, this sample was removed from further data analyses.

Chimira, the software used to quantify the miRNA expression, also performs quality control for the samples, namely read length distribution after trimming, nucleotide distribution per position, and GC content ratio at each position. These were all manually inspected (to ensure that biases and outlier sequences were not present) before accepting the output miRNA quantification values.

### Validation of experimental design strategy

Since each tissue sample included pools of 20 embryos, expression variation was expected between the biological replicates for the same tissue. Accordingly, to validate our experimental design and check for sample coherence between replicates, we evaluated sample variance via a Principal Component Analysis (PCA).

The first and second components (17.5% and 12.6% explained variance, respectively) can only distinguish between the Limb and the PSM, but by adding the third component (11.2%) the distinction between determined and undetermined PSM becomes apparent (Figure 2b). This shows, as expected, that the differences between Limb and PSM, two distinct tissues, are more extensive than the differences between determined and undetermined PSM, two molecular states of the same tissue. Albeit more subtle, such differences within the PSM

are visible in the dataset, therefore validating the samples collected for our experimental design (Figure 1a).

**Performance measures for novel miRNA predictions**

The quality metrics reported by Mirnovo for the novel miRNAs predicted for *Gallus gallus* show an overall good scoring for all samples, as seen in the ROC curves reported for the Random Forest algorithm applied (Figure 3).

As shown in Table 2, the method is highly specific (>95% of true negative identification), despite not being very sensitive (circa 50% true positive identification). This means that although many new miRNAs might be missed, the ones reported should be regarded as highly reliable. These results mirror the fact that the prediction had to be run using a general animal model, given the lack of specific models trained with chicken miRNAs.

This method identified circa 50% novel candidate miRNAs as shown by the novel prediction values presented in Table 2. This represents the addition of 1,141 new candidate miRNAs to the previously existing 1,232 mature miRNAs in the miRBase database for *Gallus gallus*, further granting relevance to this dataset.

**Validation of expression profiling**

To validate the read normalization and quantification steps, we evaluated the read distribution before and after normalization, and briefly compared the miRNA expression profile between the three tissues. Known and novel miRNA datasets were independently evaluated, since the analysis was conducted separately, and each dataset represents a different resource for the community.

**miRNA-seq Reads distribution and normalization**

The distribution of the total number of reads (Figure 4a,c) shows that there are some differences between the replicates before read count normalization, particularly for the Determined PSM tissue in the known miRNAs set. This fact is most likely a reflection of the embryo pooling strategy that might be contributing asymmetrically to the total miRNA amount present in each sample. After normalization (Figure 4b,d), the distributions become more balanced between replicas, and therefore amenable for further expression comparisons between tissues. As expected, the miRNA expression distribution for all three tissues is positively skewed (even after log2 transformation), showing a long tail to the right.

### miRNA Expression profile in the different tissues

Looking at the miRNA expression profile in the three tissues helps with uncovering possible experimental errors; for example, large asymmetries in the diversity of miRNAs found for each tissue could indicate faulty sequencing, or a total overlap of miRNA identities between tissues could indicate mislabelling or inadequate experimental design.

The top-20 most expressed miRNAs (Figure 5a,c) are found in all three tissues, with roughly comparable distributions in both annotated and novel miRNAs. Additionally, the intersection plot (Figure 5b,d) clearly shows that the majority of miRNAs (637 in known miRNAs and 849 in novel miRNAs) are found in all three tissues. Importantly, each tissue presents exclusive miRNAs, namely 71 in Limb, 35 in determined PSM, and 8 in undetermined PSM for known miRNAs (Figure 5b); and 51 in Limb, 41 in determined PSM, and 7 in undetermined PSM for novel miRNAs (Figure 5d), showing that each sample is sufficiently different from the others, allowing for proper differentiation between tissues.

## Code Availability

Technical validation and data visualization was performed in RStudio (Version 1.1.463)[30], using R (version 3.6)[20], and Bioconductor (version 3.9)[31], with packages tidyverse (version 1.3.1)[21], UpSetR (version 1.4.0)[22], patchwork (version 1.1.1)[23], plot3D (1.3)[24]. The R code used for these analyses, in the form of an annotated R notebook, is freely available in Figshare[27]. Additional software tools used to analyse this miRNA-seq dataset were the following: FastQC (version 0.11.5)[15], Chimira (version 1.5)[16], and Mirnovo (version 1.0)[19] as described in the Methods section.

## Usage Notes

The bioinformatics analysis described here made use of freely available software tools commonly used by the research community (Figure 1c). There are alternative miRNA-seq analysis pipelines equally applicable to the FASTQ reads from *Gallus gallus*[28], for example, miRDeep2[32], QuickMIRSeq[33] and sRNAnalyzer[34]. For a recently published miRNA-seq analysis protocol, see Potla et al.[35], discussing available individual tools for each step: (i) quality control (adaptor trimming, read quality/length filtering); (ii) read mapping; (iii) annotation (using miRBase); (iv) quantification; and optionally (v) detection of novel miRNAs.

These data can equally be used to seek the complete small RNA'ome, using for example the recently developed platform coMpSRA that is reported to identify and quantify diverse RNA molecule types, including miRNA, piRNA, snRNA, snoRNA, tRNA, and circRNA[36].

The miRNA expression data herein reported[27] will be useful to study gene regulation in the early phases of vertebrate embryo development, for example by performing differential expression and target gene annotation analyses. Some considerations should be taken into account for downstream analyses. Namely, the RNA was extracted from pools of 20 dissected tissues meaning that each sample represents an heterogenous mixture of individuals, whose variability is present in the data. This is even more relevant if we consider that oscillations of clock gene expression occur in the tissues analysed. Thus, care should be taken when using such static sample datasets to contrast tissues with dynamical gene expression. Additionally, some SNPs can potentially interfere with the successful mapping of some miRNA transcripts that might have been discarded, and therefore cause an underrepresentation of expression for those miRNAs. Finally, for differential expression studies comprising the PSM_U tissue, since this group comprises only two replicates, the comparison will have lower statistical power to detect small effect sizes. Accordingly, appropriate statistical techniques should be applied to deal with this limitation.

Since the chicken genome annotation is not yet up-to-par with the annotations from other vertebrate genomes, most chicken miRNAs deposited in databases are not yet experimentally validated, and their target genes are based mostly on chicken-specific computational predictions. This study opens the door for new findings specific for birds, and for validation of known vertebrate miRNAs and their respective target genes. Finally, the predicted novel miRNAs[28] represent an invaluable resource for the avian research community looking to experimentally validate novel candidate miRNAs acting in early vertebrate development capable of regulating their gene of interest. Additionally, these data coupled with transcriptomics data for the same tissues can help uncover potential regulatory modules active in early vertebrate embryogenesis.

## Acknowledgements

This study was supported by the Portuguese Fundação para a Ciência e Tecnologia (FCT) grant PTDC/BEX-BID/5410/2014 to RPA and ID and by the Município de Loulé. GC was supported by the FCT scholarship SFRH/BD/101609/2014.

## Author contributions

I.D. performed data analysis, managed, and archived the data, and drafted the manuscript. G.C. collected the samples, performed the RNA extraction, and drafted the manuscript. N.T.D.A. and V.B. acquired the data. R.P.A. conceived the study, designed the experiments, coordinated the project, and drafted the manuscript. All authors read and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Figures

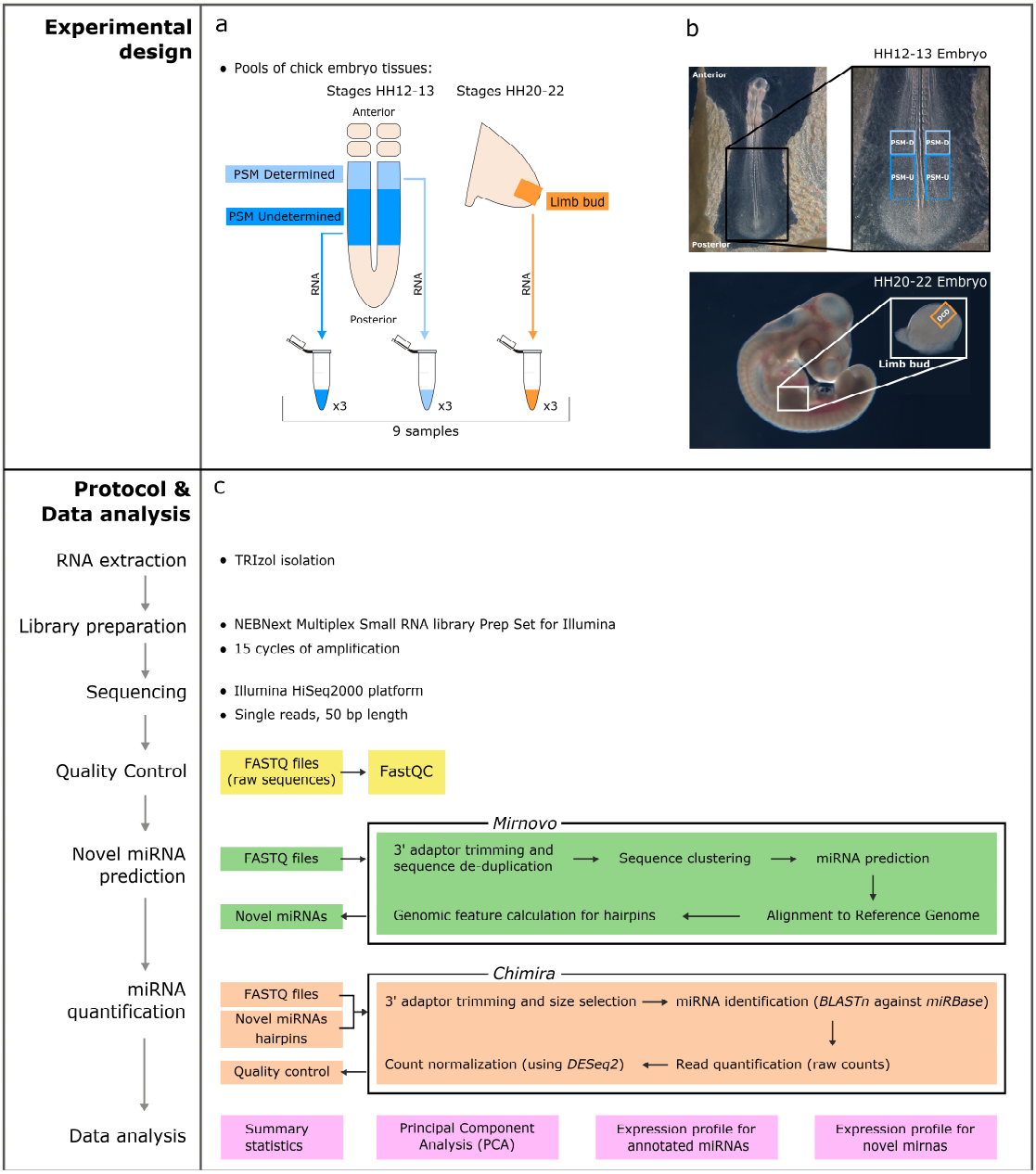



**Figure 1: Experimental design, protocol overview, and data analysis workflow.** (**a**,**b**) Overview of the experimental design, showing the sampling sites of the chick embryonic tissues collected. (**c**) Pipeline for annotated miRNA-seq data analysis and novel miRNA prediction. PSM_D: determined Presomitic Mesoderm (PSM); PSM_U: undetermined PSM; DCD: Limb Distal Cyclic Domain.

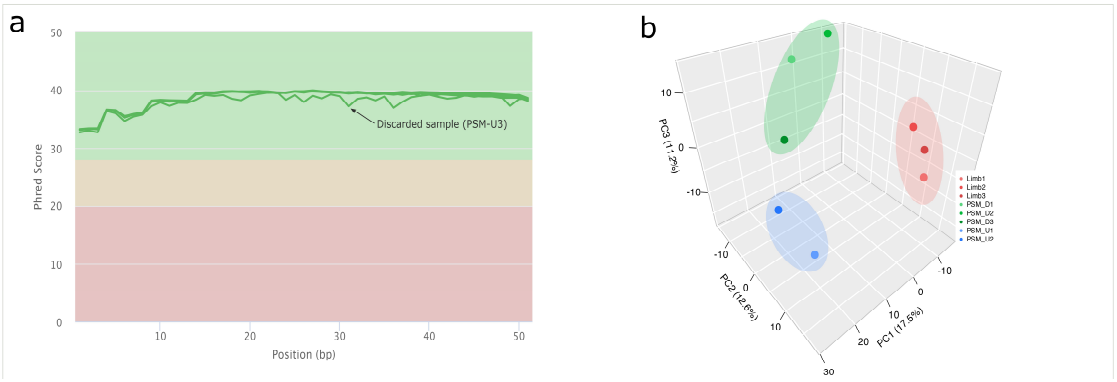



**Figure 2: miRNA-seq Quality Control and experimental design validation.** (**a**) Raw sequencing reads were evaluated with FastQC[15]. (**b**) Principal Component Analysis (PCA) showing the overall variance between samples.

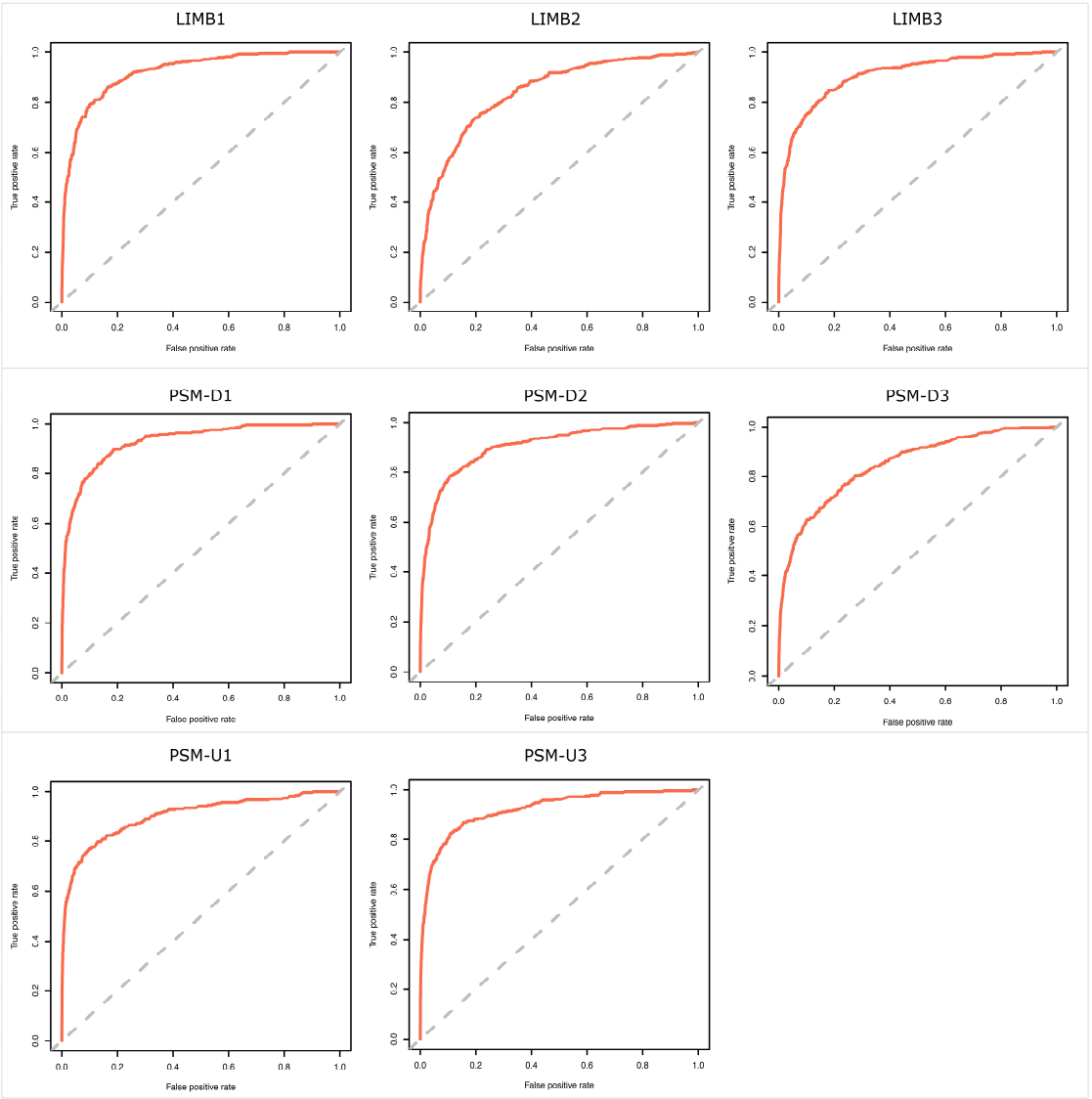



**Figure 3: Novel miRNA prediction quality.** ROC curves for the random forest algorithm applied to each tissue sample.

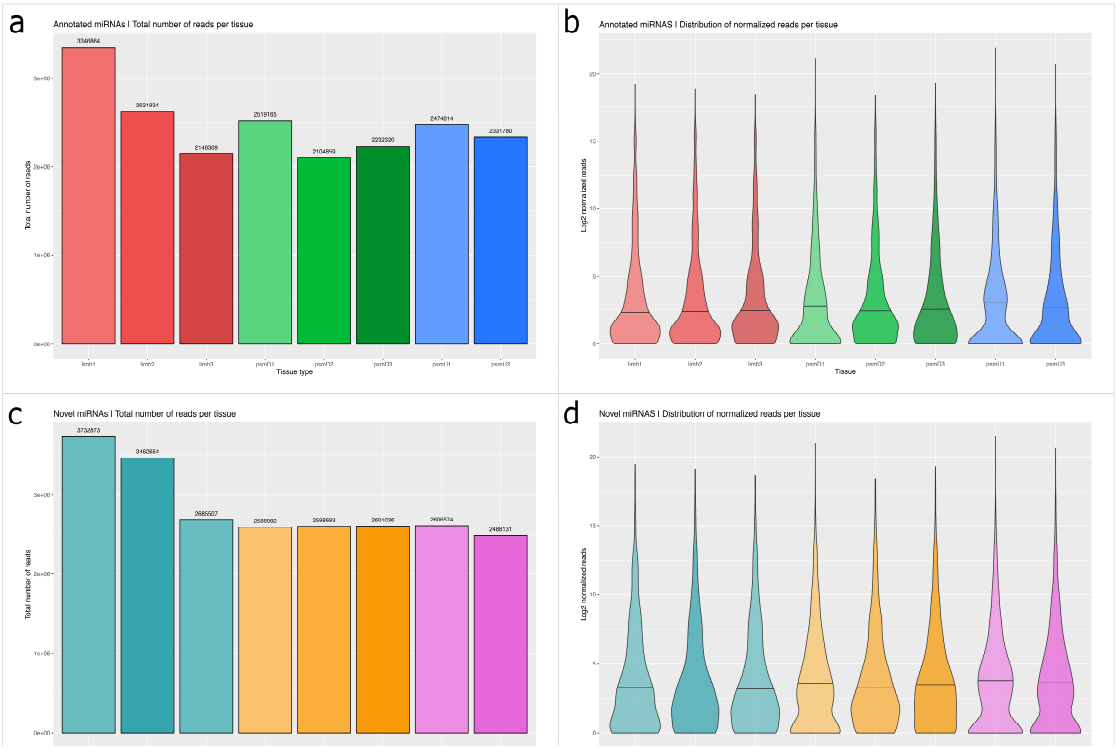



**Figure 4: Total counts and distribution of normalized expression per tissue, for annotated (a-b) and novel (c-d) miRNAs.** Total number of reads for (**a**) annotated miRNAs, and (**c**) novel predicted miRNAs. Distribution of normalized read counts per tissue, in (**b**) known, and (**d**) novel miRNAs.

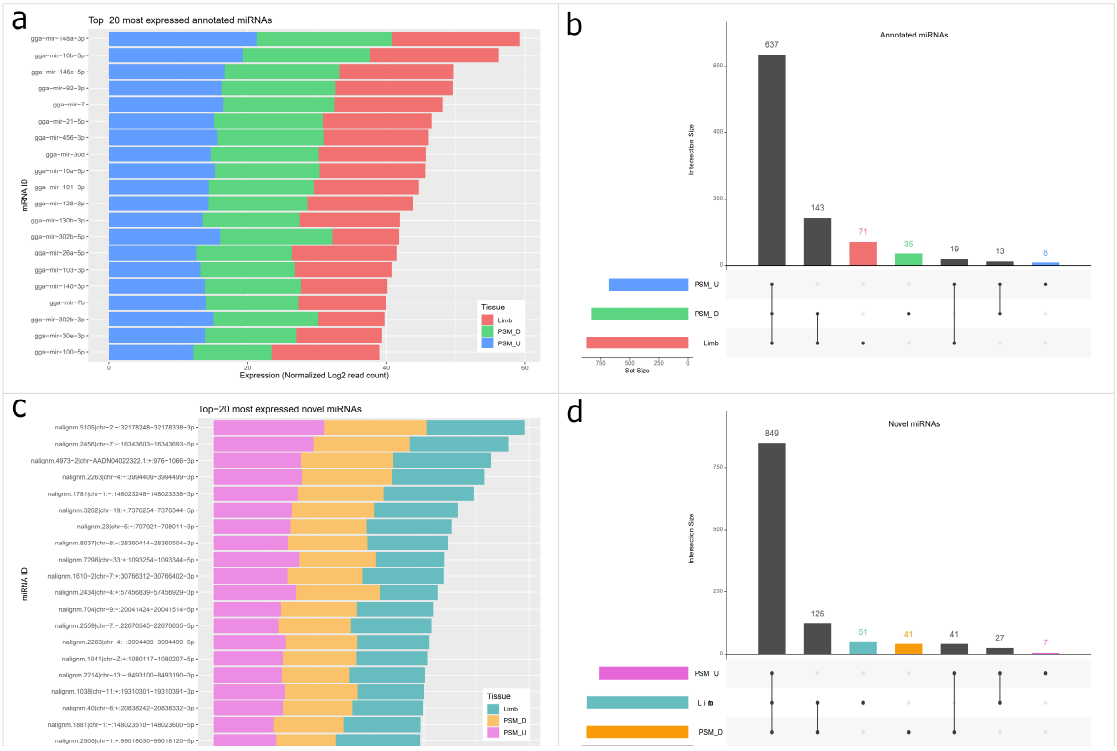



**Figure 5: Expression profiling and overlap between the three tissues for annotated (a-b) and novel (c-d) miRNAs.** (**a,c**) Top 20 highly expressed miRNAs per tissue. (**b,d**) Intersection between miRNAs expressed in each tissue.

## Tables

**Table 1: Total RNA quality control.**

| Sample ID | Tissue | Hamburger Hamilton Stages (HH) | Concentration (ng/µL) | RQI |
|---|---|---|---|---|
| Limb1 | Forelimb Distal Cyclic Domain | HH20-22 | 515.72 | 9.5 |
| Limb2 | | | 1,477.60 | 9.9 |
| Limb3 | | | 742.8 | 9.7 |
| PSM_D1 | Determined Presomitic Mesoderm | HH12-13 | 48.00 | 9.3 |
| PSM_D2 | | | 63.41 | 10.0 |
| PSM_D3 | | | 38.03 | 9.3 |
| PSM_U1 | Undetermined Presomitic Mesoderm | HH12-13 | 52.10 | 8.5 |
| PSM_U2 | | | 39.24 | 8.7 |
| PSM_U3 | | | 61.50 | 9.2 |

**Table 2: Performance measures from novel miRNA prediction using the Mirnovo algorithm.**

| | | Performance Measures | | | Predicted miRNAs |
|---|---|---|---|---|---|
| Sample ID | Tissue | Sensitivity (%) TP/P | Specificity (%) TN/N | Precision (%)* TP/(TP+FP) | Novel predictions (%) FP/(TP+FP) |
| Limb1 | Forelimb Distal Cyclic Domain | 59.18 | 96.35 | 52.25 | 47.75 |
| Limb2 | | 44.06 | 95.00 | 44.53 | 55.47 |
| Limb3 | | 58.40 | 96.40 | 55.35 | 44.65 |
| PSM_D1 | Determined Presomitic Mesoderm | 59.40 | 97.24 | 52.21 | 47.79 |
| PSM_D2 | | 52.03 | 97.19 | 52.56 | 47.44 |
| PSM_D3 | | 36.79 | 98.21 | 39.12 | 60.88 |
| PSM_U1 | Undetermined Presomitic Mesoderm | 52.31 | 98.80 | 49.33 | 50.67 |
| PSM_U3 | | 54.36 | 98.10 | 47.23 | 52.77 |

*The precision metric reported by Mirnovo represents the proportion of known miRNAs (annotated in miRBAse) relative to the total known (True Positives) plus novel predicted (False Positives)[19].

## References


1. Bartel, D. P. Metazoan MicroRNAs. *Cell* **173**, 20-51 (2018).

2. Lee, R., Feinbaum, R. & Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843-854 (1993).

3. Ambros, V. MicroRNAs and developmental timing. *Current Opinion in Genetics & Development* **21**, 511-517 (2011).

4. Kozomara, A., Birgaoanu, M. & Griffiths-Jones, S. miRBase: from microRNA sequences to function. *Nucleic Acids Research* **47**, D155-D162 (2018).

5. Stern, C. The ChickA Great Model System Becomes Even Greater. *Developmental Cell* **8**, 9-17 (2005).

6. Palmeirim, I., Henrique, D., Ish-Horowicz, D. & Pourquié, O. Avian hairy Gene Expression Identifies a Molecular Clock Linked to Vertebrate Segmentation and Somitogenesis. *Cell* **91**, 639-648 (1997).

7. Oates, A., Morelli, L. & Ares, S. Patterning embryos with oscillations: structure, function and dynamics of the vertebrate segmentation clock. *Development* **139**, 625-639 (2012).

8. Shih, N., François, P., Delaune, E. & Amacher, S. Dynamics of the slowing segmentation clock reveal alternating two-segment periodicity. *Development* **142**, 1785-1793 (2015).

9. Sheeba, C., Andrade, R. & Palmeirim, I. Joint interpretation of AER/FGF and ZPA/SHH over time and space underlies hairy2 expression in the chick limb. *Biology Open* **1**, 1102-1110 (2012).

10. Jing, B., et al. Dynamic properties of the segmentation clock mediated by microRNA. *Int. J. Clin. Exp. Pathol*. **8**, 196-206 (2015).

11. Riley, M., Bochter, M., Wahi, K., Nuovo, G. & Cole, S. mir-125a-5p-Mediated Regulation of Lfng Is Essential for the Avian Segmentation Clock. *Developmental Cell* **24**, 554-561 (2013).

12. Bonev, B., Stanley, P. & Papalopulu, N. MicroRNA-9 Modulates Hes1 Ultradian Oscillations by Forming a Double-Negative Feedback Loop. *Cell Reports* **2**, 10-18 (2012).

13. Carraco, G., Gonçalves, A., Serra, C. & Andrade, R. MicroRNA processing machinery in the developing chick embryo. *Gene Expression Patterns* **16**, 114-121 (2014).

14. Hamburger, V. & Hamilton, H. A series of normal stages in the development of the chick embryo. *Journal of Morphology* **88**, 49-92 (1951).

15. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (2016)

16. Vitsios, D. & Enright, A. Chimira: analysis of small RNA sequencing data and microRNA modifications: Fig. 1. *Bioinformatics* **31**, 3365-3367 (2015).

17. Boratyn, G. et al. BLAST: a more efficient report with usability improvements. *Nucleic Acids Research* **41**, W29-W33 (2013).

18. Love, M., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, (2014).

19. Vitsios, D. et al. Mirnovo: genome-free prediction of microRNAs from small RNA sequencing data and single-cells using decision forests. *Nucleic Acids Research* **45**, e177-e177 (2017).

20. R Core Team. R: The R Project for Statistical Computing. *R-project.org* https://www.R-project.org/ (2017)

21. Wickham, H. et al. Welcome to the Tidyverse. *Journal of Open Source Software* **4**, 1686 (2019).

22. Gehlenborg, N. UpSetR: A More Scalable Alternative to Venn and Euler Diagrams for Visualizing Intersecting Sets. R package version 1.4.0. https://CRAN.R-project.org/package=UpSetR (2019)

23. Pedersen, T. L. patchwork: The Composer of Plots. R package version 1.1.1. https://CRAN.R-project.org/package=patchwork (2020)

24. Soetaert, K. plot3D: Plotting Multi-Dimensional Data. R package version 1.3. https://CRAN.R-project.org/package=plot3D (2019)

25. Athar, A. et al. ArrayExpress update – from bulk to single-cell expression data. *Nucleic Acids Research* **47**, D711-D715 (2018).

26. Carraco, G., Duarte, I. & Andrade R. P. microRNA-Seq of *Gallus gallus* embryo tissues: Undetermined Presomitic Mesoderm (PSM), Determined PSM, and Limb bud. *ArrayExpress* https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-8176 (2021).

27. Duarte, I., Carraco, G. & Andrade R. P. gga_mirnOME | microRNA-seq | miRNA Expression dataset from chick embryonic tissues. *Figshare* https://doi.org/10.6084/m9.figshare.14706867 (2021).

28. Duarte, I., Carraco, G. & Andrade R. P. gga_mirnOME | microRNA-seq | Novel Predicted miRNAs and Expression values from chick embryonic tissues. *Figshare* https://doi.org/10.6084/m9.figshare.14901102 (2021).

29. Duarte, I., Carraco, G. & Andrade R. P. gga_mirnOME | R notebook | miRNA Expression data analysis. *Figshare* https://doi.org/10.6084/m9.figshare.14706891 (2021).

30. RStudio Team. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA http://www.rstudio.com/ (2015)

31. Bioconductor. *Bioconductor.org*. https://www.bioconductor.org/ (2019)

32. Friedländer, M., Mackowiak, S., Li, N., Chen, W. & Rajewsky, N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Research* **40**, 37-52 (2011).

33. Zhao, S. et al. QuickMIRSeq: a pipeline for quick and accurate quantification of both known miRNAs and isomiRs by jointly processing multiple samples from microRNA sequencing. *BMC Bioinformatics* **18**, (2017).

34. Wu, X. et al. sRNAnalyzer—a flexible and customizable small RNA sequencing data analysis pipeline. *Nucleic Acids Research* **45**, 12140-12151 (2017).

35. Potla, P., Ali, S. & Kapoor, M. A bioinformatics approach to microRNA-sequencing analysis. *Osteoarthritis and Cartilage Open* **3**, 100131 (2021).

36. Li, J. et al. COMPSRA: a COMprehensive Platform for Small RNA-Seq data Analysis. *Scientific Reports* **10**, (2020).

## 4.3 "A Genome-wide approach to the identification of miRNAs in temporal control of the chicken Embryo Clock" *(manuscript in preparation)*

Gil Carraco[1,2,3,¥], Isabel Duarte[2,4,¥], Raquel P. Andrade[1,2,5,*]

1. ABC-RI, Algarve Biomedical Center Research Institute, Faro, Portugal
2. Faculdade de Medicina e Ciências Biomédicas (FMCB), Universidade do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal
3. ProRegeM-PhD Program in Mechanisms of Disease and Regenerative Medicine, Faro, Portugal
4. Center for Health Technology and Services Research (CINTESIS), Polo da Universidade do Algarve, 8005-139 Faro, Portugal
5. Champalimaud Research Program, Champalimaud Center for the Unknown, Lisbon, Portugal.

¥ Equal contribution

*Corresponding author: Raquel P. Andrade (rgandrade@ualg.pt)

### Introduction

Embryo development is a four-dimensional process in which the fourth dimension is time. Embryonic developmental timing is critical for correct tissue or organ development and positioning and is clearly evidenced in rhythmic segmentation of the vertebrate body (somitogenesis). Underlying somitogenesis, an embryo clock (EC) dictates the pace of somite formation, having been first described in 1997 by Palmeirim and colleagues (Palmeirim *et al.*, 1997). In their work, the authors found that chicken *hairy1* (*hes4*, NCBI NM_001005848.3) has a cyclic expression in the presomitic mesoderm (PSM). Each cycle of expression takes 90 minutes, in synchrony with the formation of a new pair of somites (Palmeirim *et al.*, 1997). Nowadays, it is known that several genes from several gene pathways have this cyclic behavior in the PSM, and the mechanism is conserved in all vertebrates (Dequéant *et al.*, 2006; Krol *et al.*, 2011).

The pace of both the EC and somitogenesis is species-specific. For instance, the EC pace in chicken is 90 minutes, 120 in mice, and approximately 5 hours in humans. This is not due to genomic sequence differences; instead, different biochemical reaction speeds may explain the difference in the pace of the different species (Matsuda *et al.*, 2020). Matsuda et al. found that when the mouse *hes7* is expressed in human PSM cells, the periodicity of its

oscillations is comparable to the endogenous human *hes7*, and vice versa. Moreover, species-specific mRNA degradation rates underlied this phenomenon (Matsuda *et al.*, 2020).

Besides an EC in the PSM for somite formation, it was shown that there is also an EC operating in the chicken forelimb outgrowth (Pascoal *et al.*, 2007). Namely, *hairy2* expression cycles in the distal forelimb, but with a 6-hour periodicity instead of 90 minutes, like in the PSM (Pascoal *et al.*, 2007). However, the mechanism that underlies tissue-specific EC periodicities in the same organism remains elusive.

MicroRNAs (miRNAs) are 18 to 25 nucleotides long, single-stranded RNA molecules that are implicated in gene expression regulation. Their canonical mechanism of action is by post-transcriptional silencing of their specific target mRNA, by repression of translation and/or mRNA degradation (Bartel, 2018). miRNAs were firstly discovered in *C. elegans* development and were essential for the correct timing of larval stage transition, and because of that, they were called small temporal RNAs (stRNAs) (Lee *et al.*, 1993; Wightman *et al.*, 1993). A role for miRNAs in EC mRNA degradation was previously proposed (Jing *et al.*, 2015; Xie *et al.*, 2007). Furthermore, it was shown that miR-9 regulates the oscillatory expression of *hes1* and *her6* in the mouse and zebrafish neuronal progenitors, respectively (Bonev *et al.*, 2012; Soto *et al.*, 2020). In chicken, the canonical miRNA processing machinery is present in the tissues where the EC operates, namely PSM and forelimb bud (Carraco *et al.*, 2014). Also, in chicken, miR-125a-5p targets *lunatic fringe* (*lfng*) mRNA, and its absence impairs the EC and consequently somitogenesis (Riley *et al.*, 2013). Hence, we hypothesize that miRNAs could participate in the regulation of the strikingly different EC paces observed in the PSM and the forelimb.

Here, we performed a genome-wide characterization of the miRNA expression landscape in three key embryonic tissues: undetermined PSM (U-PSM), determined PSM (D-PSM), and the forelimb distal cyclic domain (DCD). For this, we used newly obtained miRNA-sequencing (miR-Seq) datasets (Duarte and Carraco *et al.*, Scientific Data, *in press*). Following qPCR validation of the sequencing data, differential expression analysis of the miRNAs revealed candidate EC-regulatory miRNAs in the Limb and PSM. Next, we searched for the target genes of the differentially expressed miRNAs using TargetScan (Agarwal *et al.*, 2015), and crossed this list of target genes with the list of downregulated genes found in the transcriptomics microarray datasets published for the PSM (Krol *et al.*, 2011; Oginuma *et al.*, 2017) and the forelimb (Anderson *et al.*, 2016). Finally, we conducted a gene functional enrichment analysis of the downregulated targets of differentially expressed miRNAs to search for candidate regulatory processes that could be differently active between the PSM and the limb. With this, we generated a list of miRNAs and respective candidate target genes to be further studied, which could be responsible for the different paces of the embryonic clock.

## Materials and Methods

### Embryo sample collection

Fertilized *Gallus gallus* eggs (Pintobar, Portugal) were incubated at 38ºC in a humidified atmosphere for two or four days to obtain embryos in stages HH12-13 and HH20-22, respectively (Hamburger and Hamilton, 1951).

The embryo tissue samples used for microRNA sequencing were collected and processed as previously described (Duarte and Carraco et al., Scientific Data, *in press*). Briefly, presomitic mesoderm (PSM) tissues were isolated from embryos in stages HH12-13. Embryos were collected from fertilized eggs incubated for 48h, and the mesoderm located on either side of the neural tube was isolated from all surrounding tissues and divided into determined PSM (PSM_D, upper one-third portion) and undetermined PSM (PSM_U, caudal two-thirds). 20 pairs of PSM portions were pooled together for each of the three biological replicates per tissue. Distal Cyclic Domain (DCD Limb) tissues were isolated from embryos at stages HH20-22. Embryos were collected from fertilized eggs incubated for 4 days, placed in PBS and the limb tissue (distal medial portion of the forelimb bud) was manually dissected using forceps. 20 DCD Limb pairs were pooled together for each of the three biological samples. Next, total RNA was extracted, processed, and the samples were quantified using NanoDrop 2000 (Thermo Scientific), and quality control was performed (Experion™ RNA StdSens Analysis Kit (BioRad #700-7103). Only samples with an RQI (RNA Quality Indicator) equal to or above 8.5 were used.

### miRNA sequencing dataset

The miRNA-seq dataset was previously obtained by Duarte and Carraco et al., (*in press*). Briefly, the sequencing reads were first quality controlled to remove low quality samples using FastQC (version 0.11.5) (Babraham Bioinformatics – FastQC A Quality Control tool for High Throughput Sequence Data., 2016), leading to the removal of one library from undetermined PSM (PSM_U2). For the remaining 8 samples, annotated microRNA read counts were obtained using the Chimira software (version 1.5)(Vitsios and Enright, 2015), yielding normalized log2 expression data which was used for the remaining analysis reported here.

### microRNA-seq differential expression

We conducted a differential expression analysis using the R package *DESeq2* (version 1.20)(Love *et al.*, 2014) using default parameters, for the following comparisons: *PSM_U vs*

*Limb*, *PSM_D vs Limb*, *PSM_U vs PSM_D*, and *PSM_TOTAL vs Limb* (where *PSM_*TOTAL corresponds to using all samples from *PSM_*U and *PSM_*D as PSM replicates, allowing the study of the overall differential expression between PSM and Limb). From each comparison, we ranked the miRNAs with larger differences between tissues by fold change (FC), using as a metric of significance the adjusted p-value < 0.05 (p-value adjusted for multiple testing).

**Target gene prediction**

To annotate the potential target genes from the differentially expressed miRNAs, we used the target gene prediction data from TargetScan (release version 7.2) (Agarwal *et al.*, 2015). For this, we selected the database entries conserved in *gga* (*Gallus gallus*) from the Predicted targets database file from TargetScan, yielding a total of 47,057 unique mirna::gene pairs. Then we intersected these with the differentially expressed miRNA ids, resulting in 1,350 pairs annotated in PSM_U vs PSM_D; 28,193 pairs in PSM_U vs Limb; 14,974 pairs in PSM_D vs Limb; and 29,142 pairs in PSM_TOTAL vs Limb. The full annotation table is presented as supplementary material.

It should be noted that the gene ids retrieved in this step are the human orthologues and not *Gallus gallus* gene ids (not present in the TargetScan database). Accordingly, the subsequent functional analysis of the predicted target genes uses the human gene ids.

**Transcriptomics meta-analysis | Data description**

Three *Gallus gallus* microarray transcriptomics datasets were gathered from GEO and ArrayExpress, namely accessions GSE75798 (Oginuma *et al*., 2017) and E- MTAB-406 (Krol *et al*., 2011) from presomitic mesoderm (PSM); and E-MTAB-4048 (Anderson *et al*., 2016) from the limb tissue.

All 3 datasets pertained to tissues collected during somitogenesis, when the embryo clock is known to be active, and all used the same microarray chip – Affymetrix Chicken Genome Array (microarray platform description: GPL3213 in NCBI, and A-AFFY-103 in ArrayExpress). Since all three original datasets contained microarrays from several chicken embryo parts in different developmental stages, and even from different model organisms, we selected for meta-analysis only the samples that matched our experimental design. Accordingly, from the PSM GSE75798 dataset were selected the 15 arrays corresponding to right and left PSM regions, from embryos in the HH12 stage. From the PSM E-MTAB-4048 dataset were selected the 18 chicken arrays from the right PSM at HH12 stage. Finally, from

the limb dataset (E-MTAB-4048) we selected only the 7 samples pertaining to the limb in stage HH20 (which is the same HH stage used in the miRNA-seq dataset).

**Transcriptomics meta-analysis | Data pre-processing**

The .CEL files from all the samples gathered were retrieved and analysed using customized R scripts (R version 4.0.2), using the following packages: *affy* (version 1.66.0), *FrozenChicken* (version 1.0), *frma* (version 1.40.0), *AnnotationDbi* (version 1.50.3), *chicken.db* (version 3.2.3), and *genefilter* (version 1.70.0), and *tidyverse* (version 1.3.1).

The data pre-processing steps were: quality control using *ArrayQualityMetrics* package (version 3.42.0), followed by data normalization using frozen-RMA (fRMA) making use of the *FrozenChicken* Rdata package (containing the frozen vectors required for the joint pre-processing of datasets coming from different experimental settings (Duarte *et al.*, 2021)). Then we removed duplicated probe sets for the same gene using the *genefilter* package (annotation-based filtering only, retaining the highest Inter-Quartile-Range probe for each gene).

**Transcriptomics meta-analysis | Differential expression analysis**

The expression matrix obtained from the transcriptomics data pre-processing was then used for differential expression analysis to find the list of down-regulated genes in each tissue. Using the package *limma* (version 3.44.3) we calculated the differential expression between PSM and Limb. Then we extracted a list of down-regulated genes in PSM when compared with Limb, and vice-versa. We used adjusted p-value < 0.05 as significant. This list of down-regulated genes was subsequently used to narrow-down the list of potential target genes from the up-regulated miRNAs found by miRNA-seq.

**Functional enrichment of down-regulated target genes from up-regulated miRNAs**

After selecting only the down-regulated target genes from up-regulated miRNAs, we conducted a functional enrichment analysis to gather more insights into the potential roles of the discovered miRNAs. Using the *gprofiler2* (version 0.2.0) package we searched for enriched GO (Molecular Function, Biological Process, and Cellular Component) (Mi *et al.*, 2021) and REACTOME categories (Jassal *et al.*, 2020) for PSM and Limb target genes.

**Validation of miRNA-Seq expression by RT-qPCR**

Three pools of each tissue were collected, and RNA was extracted as described above (three biological replicates). Extracted RNA was reverse transcribed into cDNA, in duplicate or triplicate, using miRCURY® LNA® RT Kit (Qiagen) according to the manufacturer's instructions. Subsequent qPCR was performed using miRCURY® LNA® miRNA Custom PCR Panels and miRCURY® LNA® SYBR® Green PCR Kit (Qiagen). The qPCR specificities were used according to the manufacturer's instructions. The miRNA qPCR LNA® probes were designed by Qiagen for the selected miRNAs and panels were customized at https://geneglobe.qiagen.com/pt/customize/pcr/.

**Validation of miRNA expression by qPCR | Data analysis**

The R data package *gga.mirnas.qpcr* containing the data from the qPCR validation of the 27 miRNAs selected from the miRNA-seq analysis was developed, and will be freely available in GitHub upon publication (https://github.com/iduarte/gga.mirnas.qpcr). Briefly, the original files exported from the CFX Manager (Version 3.1.1517.0823) software were used as a starting point for the data analysis conducted in R (version 4.0.2). The following data processing steps were undertaken: (1) Finding the exponential phase of the amplification curves (by visual inspection); (2) Defining the fluorescence threshold that intersects all curves in the exponential phase (for this dataset we used 100 fluorescence arbitrary units); (3) Determination of quantification cycle (Cq) via interpolation of the amplification curves, followed by intersection with the fluorescence threshold; (4) Cq normalization with interplate calibrator (to allow the comparison of the results between plates); (5) Summarization of Cq values by calculating the mean of the technical replicates per biological replicate; and finally (6) Conversion of Cq values to Log2 expression (since Cq values are inversely correlated with expression, these must be reversed to represent Log2-expression values, and an arbitrary constant can be added to make all the values positive and easier to compare. Here we used the following transformation: max(Cq) – Cq = Log2 Expression).

**Validation of miRNA expression by qPCR | Differential expression analysis**

A differential expression analysis was conducted for the qPCR results using the *limma* (version 3.44.3) package. The following contrasts were calculated: PSM_U versus PSM_D, PSM_U versus Limb, and PSM_D versus Limb. Statistical significance was considered for adjusted p-values < 0.05.

### Correlation between qPCR and miRNA-seq results

To evaluate the extent of agreement between the expression values obtained by miRNA-seq and by qPCR, we conducted a Pearson correlation analysis between both results using the *ggpubr* (version 0.4.0) package.

## Results and Discussion

### miRNA quantification and differential expression

We previously performed a genome-wide characterization of the miRNAs present in multiple tissues of the chicken embryo where the EC oscillates with different periodicities (Duarte and Carraco *et al.*, Scientific Data *in press*). EC periodicity is 90 min in the undetermined PSM (U-PSM) (Palmeirim *et al.*, 1997), is progressively decelerated in the determined PSM (D-PSM) (Dubrulle *et al.*, 2001), and is 6 hours in the forelimb distal cyclic domain (DCD) (Pascoal *et al.*, 2007). A total of 926 miRNAs previously annotated in miRBase (Kozomara *et al.*, 2019) were identified, as well as 1141 putative novel miRNAs (Duarte and Carraco *et al.*, Scientific Data *in press*). In this work, we further analysed the 926 annotated miRNAs.

The top 10 miRNAs expressed among all tissues were gga-miR-148a-3p, gga-miR-10b-5p, gga-miR-92-3p, gga-miR-146c-3p, gga-miR-7, gga-miR-21-5p, gga-miR-30d, gga-miR-302b-5p, gga-miR-456-3p and gga-miR-10a-5p (Figure 1a). Interestingly, the most abundant miRNA, gga-miR-148a-3p, is predicted to target *delta1* mRNA by TargetScan (Agarwal *et al.*, 2015). Several other top 10 expressed miRNAs are predicted to target genes involved in Notch, Wnt and FGF signalling pathways, which are critical regulators of EC operation.

From the 926 miRNAs identified, 637 are present in the three tissues (Figure 1b). We also found that 71 miRNAs are only found in the DCD, 35 in the D-PSM, and eight in the U-PSM. When both PSM portions (U-PSM and D-PSM) are grouped together (TotalPSM), we have 56 miRNAs present in the PSM but not in the limb DCD (Figure 1b). Since many miRNAs are found at very low levels, this information could be misleading, since others may not be detected due to lack of experimental sensitivity. In order to avoid this and shorten our list of candidates, we performed pairwise differential expression analysis, namely U-PSM vs D-PSM, U-PSM vs DCD, D-PSM vs DCD, and TotalPSM vs DCD (Figure 1c). It was considered that a miRNA was differentially expressed when the adjusted p-value>0.05. In total, we found 191 miRNAs that are differentially expressed among the tissues.

**Putative EC-regulatory differentially expressed miRNAs**

Since we are interested in the embryo clock, we selected differentially expressed miRNAs that could target EC-related transcripts for further validation using RT-qPCR. We have looked within the three pathways mainly involved in somitogenesis timing and regulation (Notch, Wnt, and FGF), more precisely to the genes expressed periodically in the PSM (Bailey and Dale; Dequéant *et al.*, 2006; Krol *et al.*, 2011).

By crossing the differentially expressed miRNAs with TargetScan (http://www.targetscan.org/vert_72/), we have searched for miRNAs that could target transcripts produced by Notch signaling genes *notch1, delta1, hairy1, hairy2, nrarp, lunatic fringe,* and *hey2*. Among these genes, we only found miRNAs that are able to target *notch1, dll1,* and *hey2* transcripts. *Notch1* can be targeted by gga-miR-367, gga-miR-34c-5p, gga-miR-200b-3p and gga-miR-429-3p. *Dll1* can be targeted by gga-miR-130a-3p, gga-miR-34c-5p, and gga-148a-3p. *Hey2* can be targeted by gga-miR-144-3p, gga-miR-181a-5p, gga-miR-181b-5p, gga-miR-31-3p, gga-miR-190a-5p, and gga-miR-138-1-3p (Table 1). Regarding the Wnt pathway, we have looked at the genes *gpr177*, *axin2* and *brachyury*. Two of these three, *axin2* and *brachyury*, have target sites for miRNAs that are differentially expressed in our data set. *Axin2* can be targeted by gga-miR-144-3p, gga-miR-205a, gga-miR-205b, gga-miR-221-3p, and gga-miR-138-1-3p. *Brachyury* can be targeted by miR-219a. Lastly, we search for miRNAs that could target mRNAs from FGF pathway genes with cyclic expression, more precisely, *dusp2, dusp6, snail1, snail2,* and *fgf3.* We identified miRNAs that could target *dusp6*, namely gga-miR-181a-5p, gga-miR-181b-5p, gga-miR-367 and gga-miR-125b-5p, and *snail2,* that is targeted by gga-miR-138-1-3p.

Besides genes that have a cyclic expression in the PSM, we have looked for miRNAs that could target genes that are expressed in a gradient in the PSM. We search for miRNAs that could target mRNAs produced by *wnt3a, wnt5a, wnt8a, raldh2, cyp26a,* and *fgf8*. We only found one miRNA, gga-miR-302a, that could target *wnt5a* mRNA. The lack of differential miRNAs able to target these genes may happen because of two reasons. First, it may indicate that miRNAs are not important for establishing gradients in the PSM or the limb outgrowth. Second, the genes involved in generating gradients in the PSM are roughly the same in the PSM and the limb (Reviewed in (Sheeba *et al.,* 2016)). Consequently, putative miRNAs that could be regulating these genes may be present but not differentially expressed.

We may have missed some miRNA-mRNA pairs due to the incomplete chicken genome annotation. For instance, miR-9 was shown to target *M. musculus hes1* mRNA and *D. rerio her6* mRNA (Bonev *et al.*, 2012; Soto *et al.*, 2020). However, by using TargetScan, we did not find any interaction of gga-miR-9 with chicken *hairy1* or *hairy2* transcripts. However,

after manually searching for the presence of a target site for gga-miR-9, we found that, in fact, there is a sequence element in the 3'UTR of *hairy2* mRNA that is compatible with gga-miR-9-5p targeting (Figure 2). For this reason and because they are differentially expressed, gga-miR-9-3p and gga-miR-9-5p were also included for further study.

Additionally, six more miRNAs were included for validation. Gga-miR-125a-5p was selected because it was already shown to target EC genes in the PSM (Riley *et al.*, 2013). gga-miR146a-5p, gga-miR-222b-5p, and gga-miR-19b-5p were selected as possible qPCR normalizers, as determined by geNorm analysis of the miR-Seq data (Vandesompele *et al.,* 2002). Moreover, in zebrafish, it was found that dre-miR-19 targets cyp26a1, helping to create a gradient of Retinoic Acid (RA) in the embryo (Franzosa *et al.,* 2013). Finally, gga.miR-21-5p and gga-miR-30d were selected because they are in the top 10 most expressed miRNAs (Figure 1a) (Table 1).

Figure 3 presents the expression levels of the miRNAs in Table 1, as obtained by miRNA-Seq. RT-qPCR was performed for experimental validation in freshly-extracted RNA samples (Figure 4). The miRNA-Seq results were reliably validated using RT-qPCR, as seen in Figure 4. The expression levels obtained by each experimental technique were equivalent in the vast majority of the miRNAs assayed. In some cases, the expression levels obtained by RT-qPCR are higher, most probably due to the technique's higher sensitivity (Sup. Fig. 3). The exception was gga-miRNA-148a-3p, where miRNA-seq detected higher levels.

Having validated the sequencing data, we further performed differential expression analysis of the selected miRNA in each pair of tissues using the qPCR data obtained (Figure 5). We found four miRNAs that are upregulated in the PSM in comparison with the forelimb DCD expression, namely gga-miR138-1-3p, gga-miR-302a, gga-miR-367, and gga-miR-9-3p. Since the miRNAs are important negative regulators of mRNA stability, they may have a role in the faster pace of the EC in the PSM in comparison with the limb. Besides gga-miR-9-3p, the other three miRNAs have, as predicted targets, genes with an oscillatory expression in the PSM (Table 1). Furthermore, gga-miR-302 is predicted to target *wnt5a*, a gene expressed in a gradient in the PSM (Baranski *et al.,* 2000). Direct functional analysis should be performed in order to find the relevance of these three miRNAs, especially testing the pairs miRNA::EC target gene. It would be interesting also to conduct a functional enrichment analysis on gga-miR-9-3p targets to find possible roles for this miRNA in the PSM.

We found eight genes upregulated in the limb DCD in comparison with the PSM, namely gga-let7a-5p, gga-miR-125b, gga-miR-144-3p, gga-miR200b-3p, gga-miR-205a, gga-miR-205b miR-221-3p and miR-429-3p. Strikingly five of these miRNAs, gga-miR200b-3p, gga-miR-205a, gga-miR-205b miR-221-3p and miR-429-3p, are predicted to target *notch1* or

*axin2.* Both of these genes are exclusively expressed in the AER and downregulated in the limb mesenchyme (Eckei *et al.,* 2016; Hayashi *et al.*, 1996). Thus, we hypothesize a possible role for these miRNAs in the definition of these two genes' expression by inhibiting their expression in the limb mesenchyme.

Although gga-miR-9-5p miRNA is expressed in all tissues (Figures 3 and 4), we did not find a significant differential expression in the tissues analyzed. However, since *hairy2* is one of its possible targets, the functional analysis of this pair should be conducted.

Surprisingly, we were not able to detect any expression of the gga-miR-125a-5p, which was previously described as having a role in the EC (Riley *et al.*, 2013). In mice, the same group showed that mmu-miR-125a-5p is not necessary for the mouse embryo clock. However, mutations in Lunatic Fringe target sites for the miR-125 family lead to a non-oscillatory behaviour in the caudal PSM (Wahi *et al.*, 2017). Thus, in chicken, by manipulating the endogenous gga-miR-125a-5p, they may have impaired the ability of other miRNAs to target *lfng,* causing an arrest of EC oscillations. Further experiments should be conducted in order to clarify the existence of this miRNA in the chicken embryo.

**miRNA target genes functional analysis**

To further our insights into the potential functions of the differentially expressed miRNAs, we conducted a functional classification of the predicted target genes. Since each miRNA can potentially target several hundreds of different genes, most of which will not be actively transcribed in the tissues and stages under study, we proceeded with combining our list of target genes with a list of down-regulated genes obtained from a transcriptomics meta-analysis of PSM and Limb. For this we queried and gathered appropriate microarray data from publicly available repositories. These data were then processed, followed by a differential expression evaluation between PSM and Limb, using the adjusted p-value < 0.05 as statistically significant. To narrow-down this list of genes, we applied two stringent cut-off criteria to obtain a high-confidence list of differentially regulated genes, namely Adjusted P-value < 1E-5, and |Log2 Fold Change| > 2.

By comparing the expression levels of the miRNAs present in the PSM tissue as a whole and the distal limb, we identified 21 miRNAs that were strongly upregulated in the totalPSM relatively to the DCD, and 27 that were downregulated (Figure 6a). When comparing PSM-U vs PSM-D data with the aforementioned parameters, only three differentially expressed miRNAs were identified: gga-miR-196-5p is upregulated, whereas gga-miR-216b and gga-miR-217-5p are downregulated in the PSM-U (Figure 6b). It is not surprising that the difference between these two tissues is smaller than when comparing PSM with limb tissue, since both

belong to the unsegmented paraxial mesoderm and may even have existed some experimental overlapping when the tissue was collected. Nevertheless, the difference between the two PSM portions may be significant because in the D-PSM, the cells are already committed towards somitic fate (Dubrulle *et al.*, 2001).

Next, we performed a transcriptomics differential expression analysis between the PSM and Forelimb, using previously published microarray expression data from these tissues (Krol *et al.*, 2011; Oginuma *et al.*, 2017; Anderson *et al.*, 2016) (Figure 7). miRNAs negatively regulate gene expression. By crossing our differentially expressed miRNAs with TargetScan (http://www.targetscan.org/vert_72/) and posteriorly with the negatively differentially expressed genes (Figure 7), we found that the 48 differentially expressed miRNAs (21 in the PSM and 27 in the forelimb) have the ability to target 2439 downregulated genes in the PSM and 952 downregulated genes in the limb, respectively.

To further understand the miRNA's role and determine how they may impact the EC, we did a GO (Mi *et al.*, 2021) and REACTOME (Jassal *et al.*, 2020) analysis of target genes (Figure 8). By Go enrichment and Reactome analysis, we found an over-representation of "RNA polymerase" and "RNA transcription" categories in PSM, which are absent or less relevant in the limb. Moreover, "regulation of RNA metabolic process" is over-represented in the PSM and absent in the limb. This may suggest a role for miRNAs controlling the gene expression in the PSM by targeting genes related to transcription and regulating the RNA metabolism, which could be important for fast gene oscillations. We also found an over-representation in the "regulation of macromolecules biosynthesis". Overall, the mRNA and protein synthesis are extremely dynamic in the PSM, partially due to the EC, which requires a fast production and degradation. On the other hand, we found an over-representation of "WNT signaling pathway", a category absent in the PSM. Although the WNT pathway is present and is important for the correct patterning of both tissues, it seems that miRNA-mediated regulation of WNT is more relevant in the limb context.

This approach produced a valuable list of genes that are enriched in each one of the categories identified, that represent putative targets for the differentially expressed miRNAs between PSM and Limb.

Altogether, by undertaking a genome-wide approach to analyze differentially expressed miRNAs and their putative target genes in chick embryonic tissues with distinct EC dynamics, the present work has identified multiple miRNAs that could be regulating EC mRNA degradation and/or translation in a tissue-specific manner. These represent promising candidates as tissue-specific regulators of EC dynamics.

**Acknowledgements**

This study was supported by the Portuguese Fundação para a Ciência e Tecnologia (FCT) grants PTDC/BEX-BID/5410/2014 to RPA. GC was supported by the FCT scholarship SFRH/BD/101609/2014.

**Author contributions**

G.C. collected the samples, performed the RNA extraction, designed the RT-qPCR, and drafted the manuscript. I.D. performed data analysis, performed GO enrichment analysis, and drafted the manuscript. R.P.A. conceived the study, designed the experiments, coordinated the project, performed RT-qPCR and drafted the manuscript. All authors read and approved the final manuscript.

**Competing interests**

The authors declare no competing interests.

**Figures**

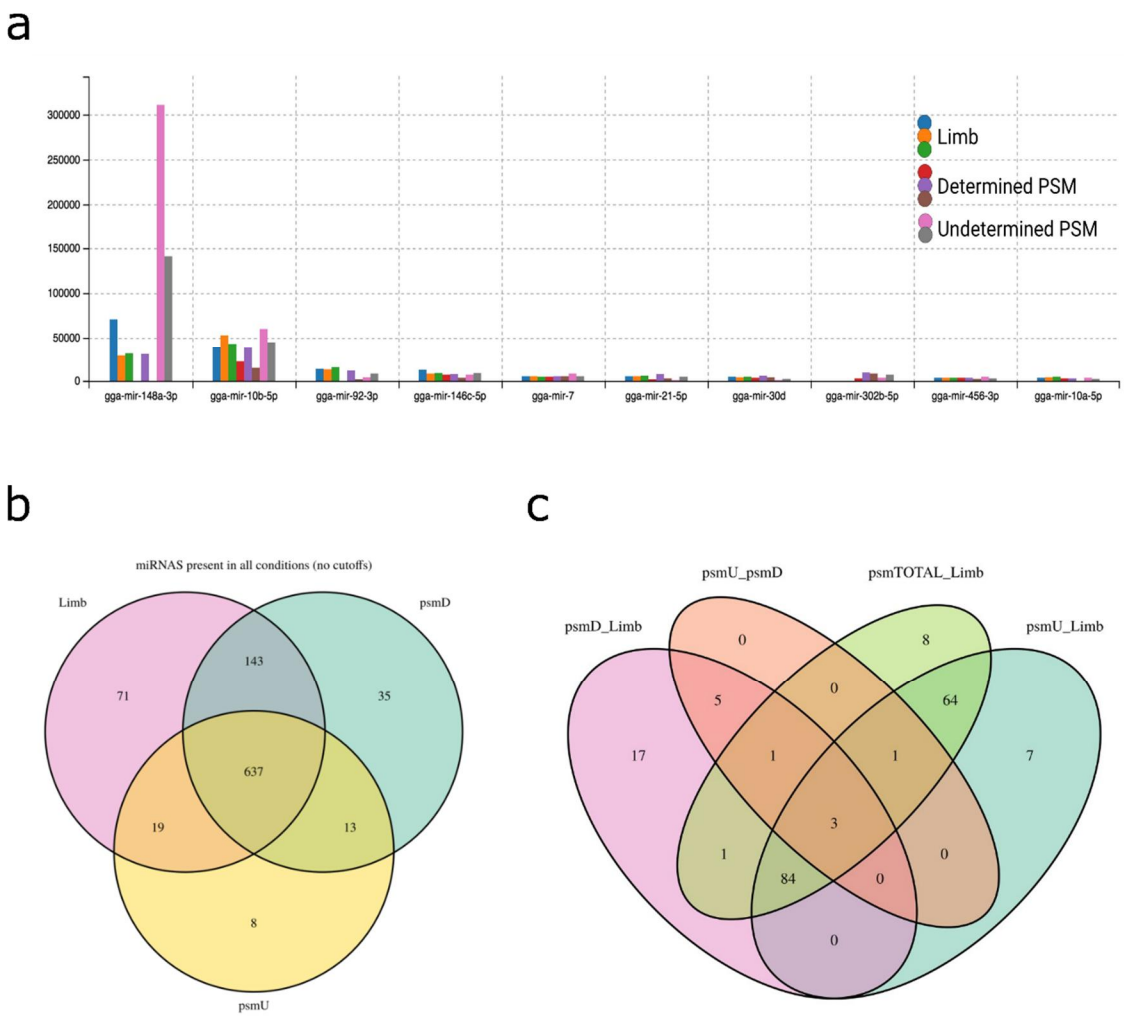



**Figure 1**. **Identification and differential miRNA expression among the three tissues. A** – Total miRNA counts of the top 10 expressed miRNAs among all samples. **B** – All 926 known miRNAs identified in the miRNA datasets and their distribution among the three studied tissues. **C** – Tissue pairwise differentially expressed miRNAs. The TotalPSM was generated by pooling together U-PSM and D-PSM datasets. Adjusted p-value < 0.05 was set as cut-off.

5'-TGAGGGTGCCAGGCACCGGACTGGACTGAGCACCGCGCCCCGTGCTGCTGGCACGACTGTACATAGGTTATA
CGTTCATATTGGATTGTGCCTTTGTACCATAGCACTCCCTCCGACAGTGTTTCGTGTTTTACACGAGATCTTTTT
TTCTCCCTTCCTTCCCCTTCCTTATGTGAC**ACCAAAGA**TCGGGATTTTTTTTTTTTGAATGCTCTTAAAATAAAA
GAGTCTCTTCTGTTTTGGAAGGCTTTTGGAGATA-3'

**Figure 2. *hairy2* 3'UTR sequence and gga-miR-9-5p target site (in bold).**

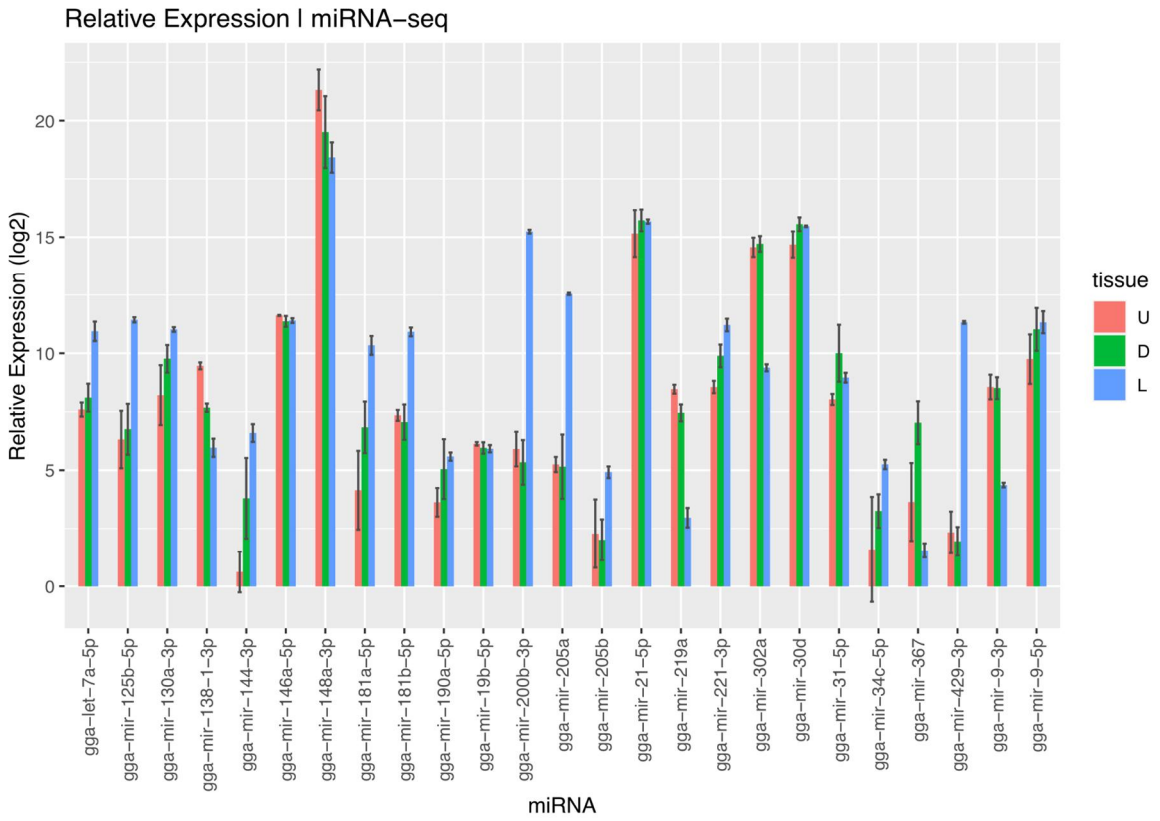



**Figure 3. Levels of expression of selected miRNAs detected by miRNA-Seq.** The miRNA expression levels of the undetermined PSM (U) are represented in red, the determined PSM (D) in green, and the limb DCD (L) in blue.

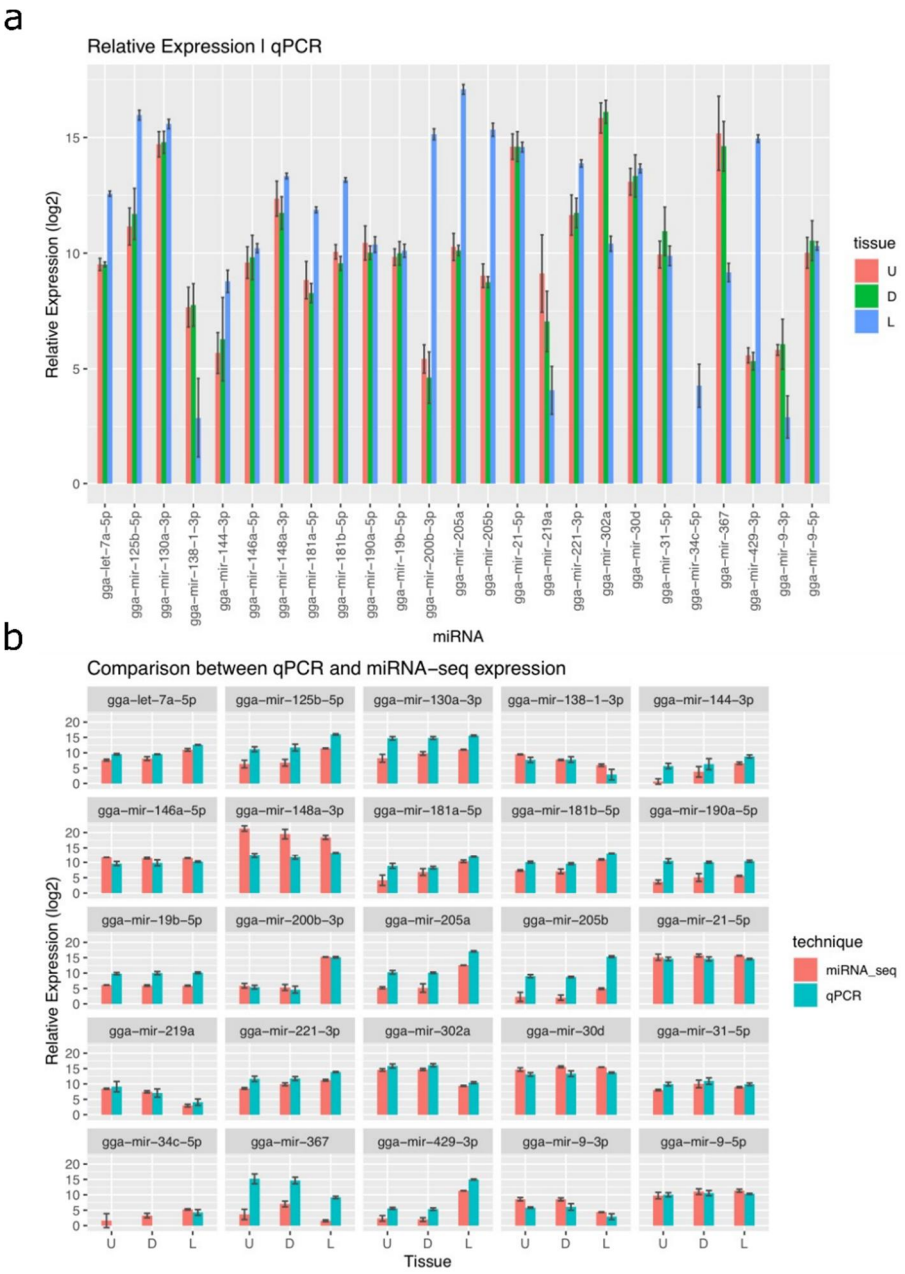



**Figure 4. Validation of expression levels of selected miRNAs RT-qPCR. A)** Relative expression levels obtained by RT-qPCR. The miRNA expression levels of the undetermined PSM (U) are represented in red, the determined PSM (D) in green, and the limb DCD (L) in blue. **B)** Direct comparison of the expression results obtained with both techniques per miRNA.

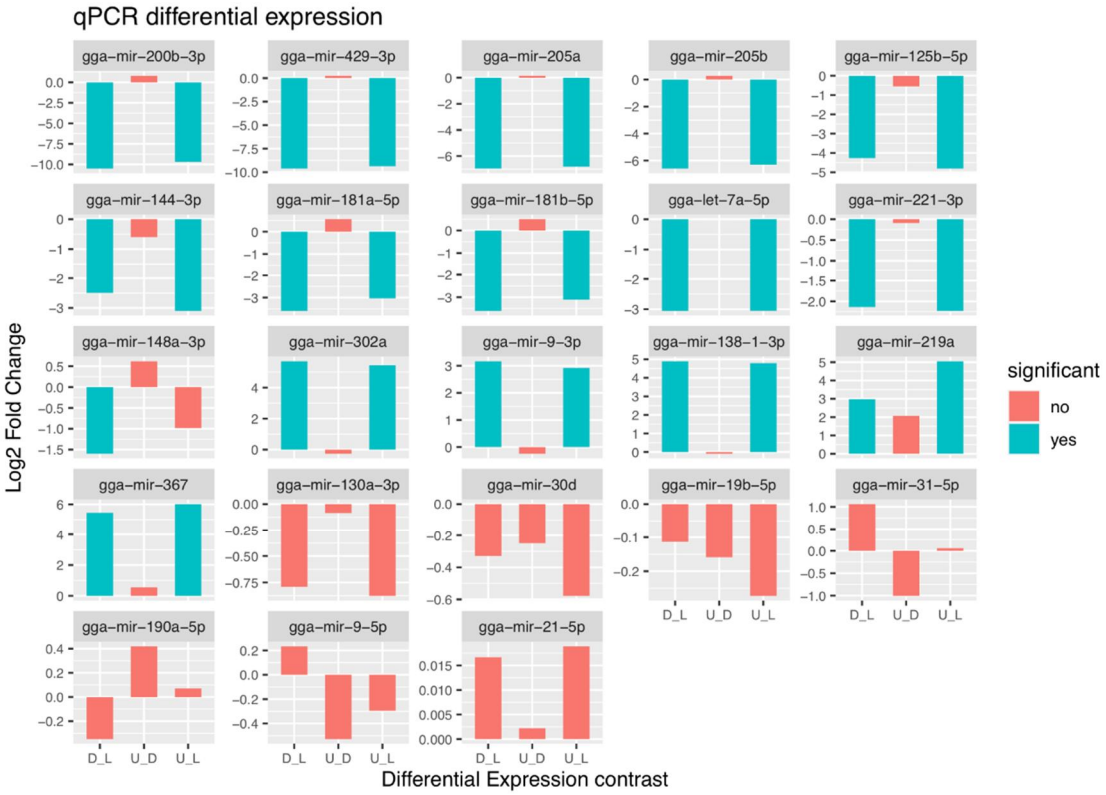



**Figure 5. Differential expression results obtained by qPCR.** D_L: PSM-D vs DCD; U_D: PSM-U vs PSM-D; U_L: PSM-U vs DCD. Significance is considered for P value <= 0.05.

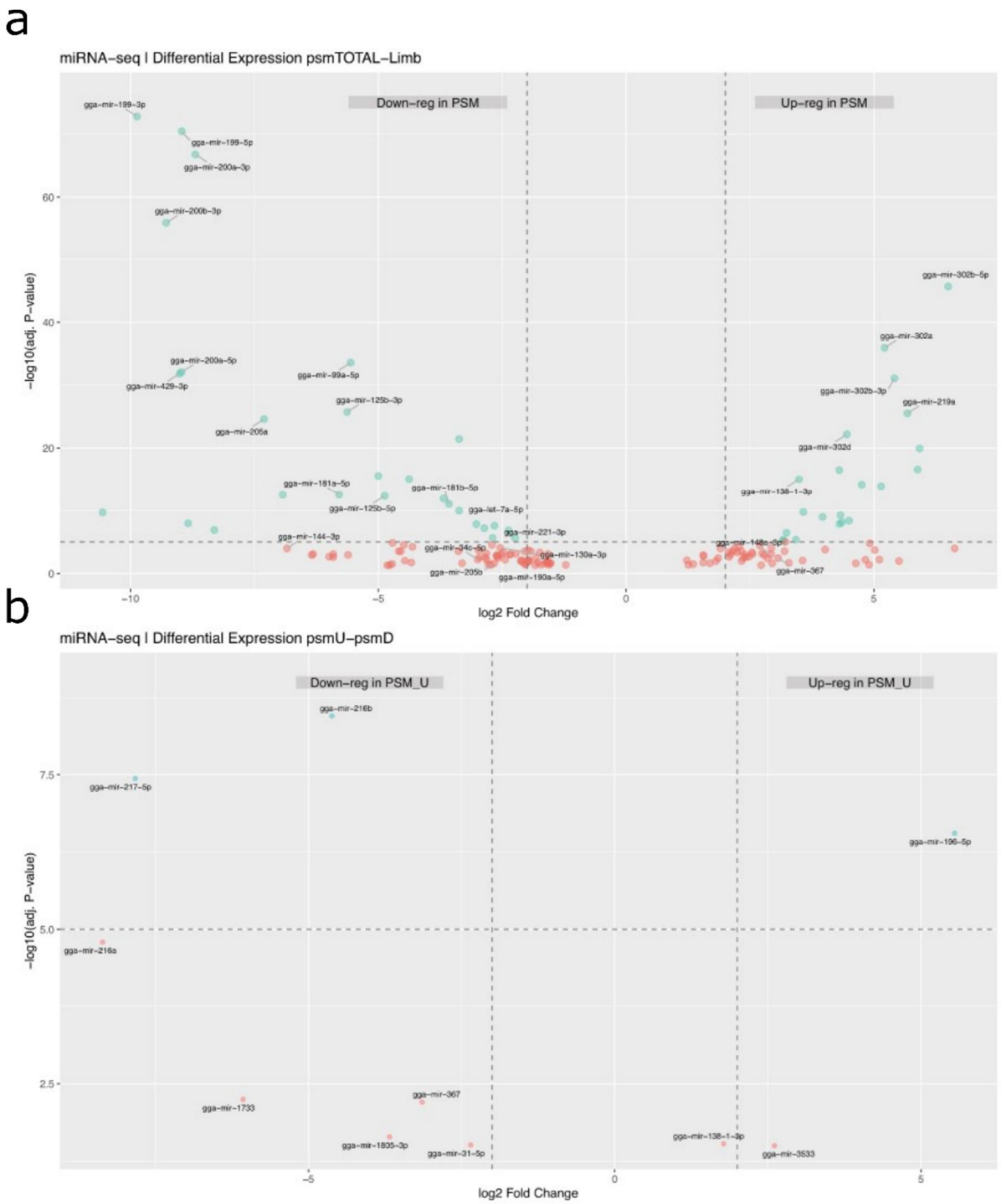



**Figure 6. Analysis of the levels of differentially expressed miRNAs. A** – Volcano plot showing differentially expressed miRNAs between psmTOTAL vs Limb. **B** – Volcano plot showing differentially expressed miRNAs between psmU vs psmD. Labeled in green are miRNAs with adjusted p-value < $10^{-5}$ and |Log2 Fold Change| >2.

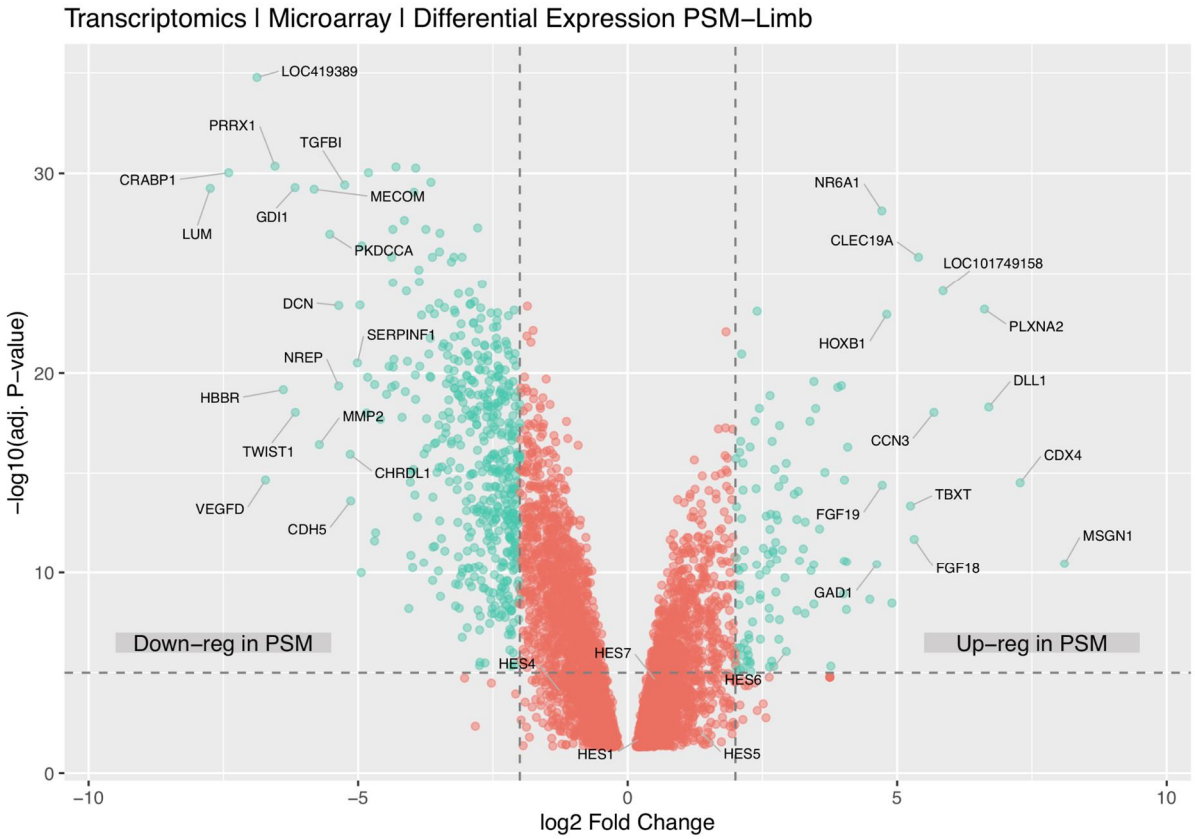



**Figure 7. Transcriptomics differential expression between the PSM and Forelimb.** Volcano plot showing the differentially expressed genes. The PSM gene expression microarrays datasets were obtained from (Krol *et al.*, 2011; Oginuma *et al.*, 2017), and Limb datasets were obtained from (Anderson *et al.*, 2016). Labeled in green are the genes that follow the cutoff adjusted p-value < $10^{-5}$ and |log2fold change| >2.

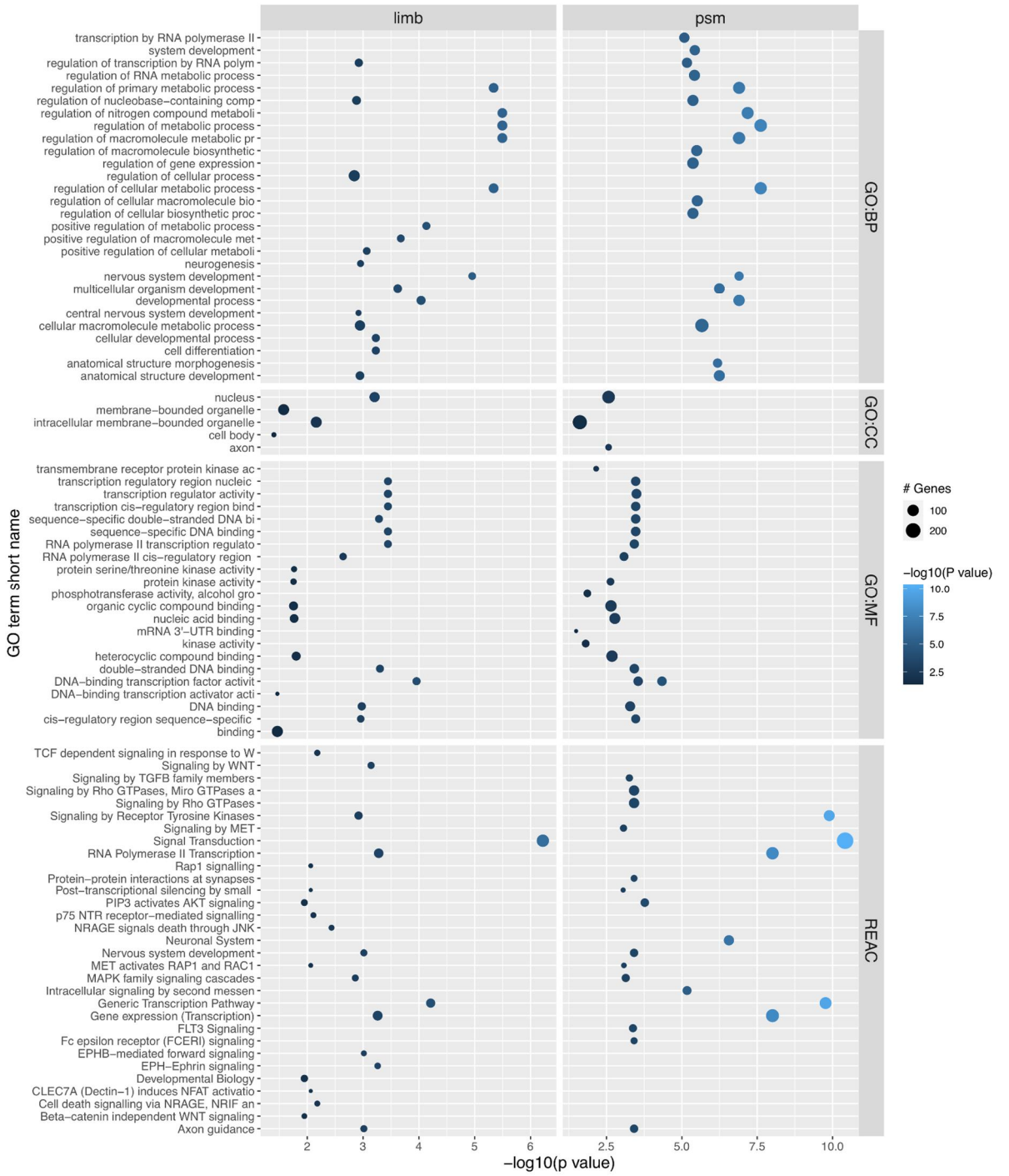
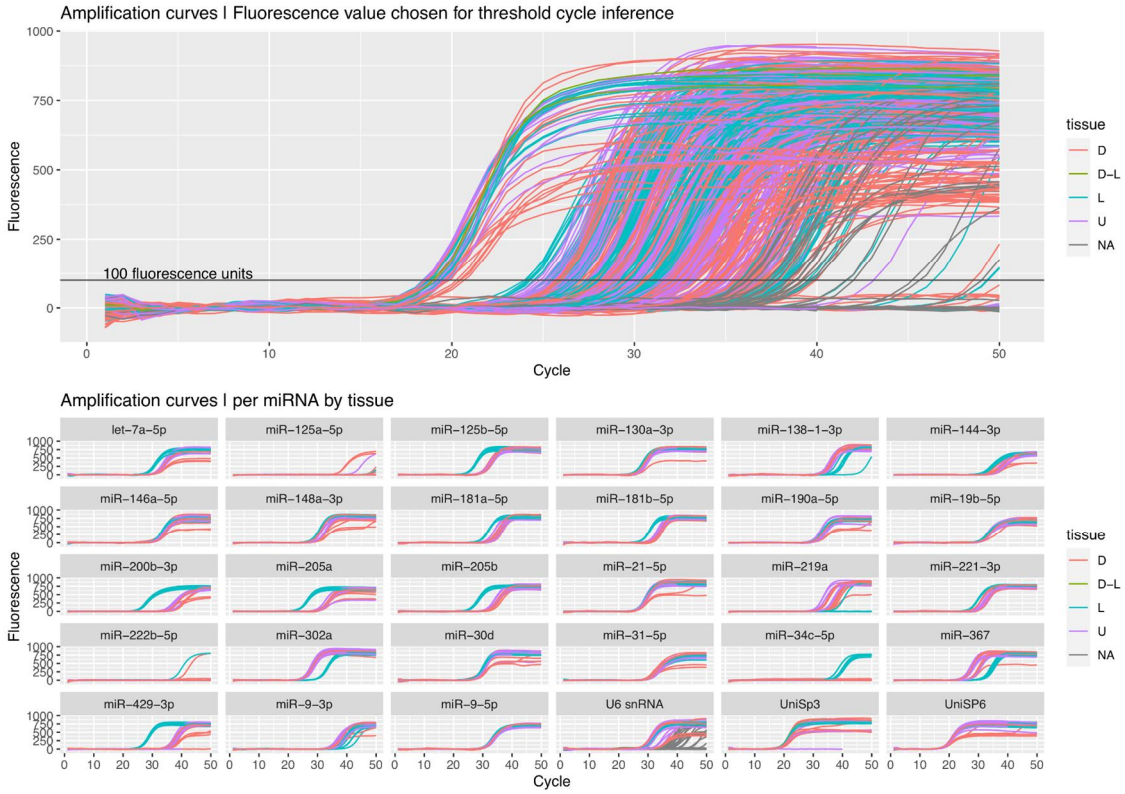
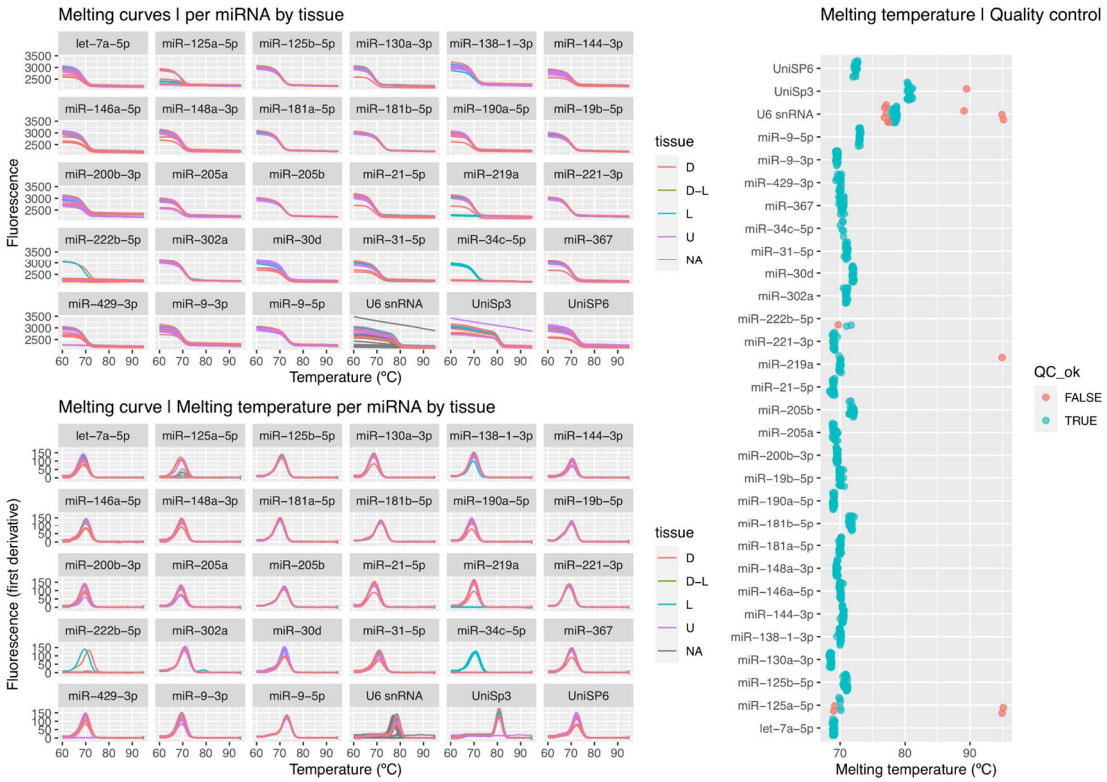


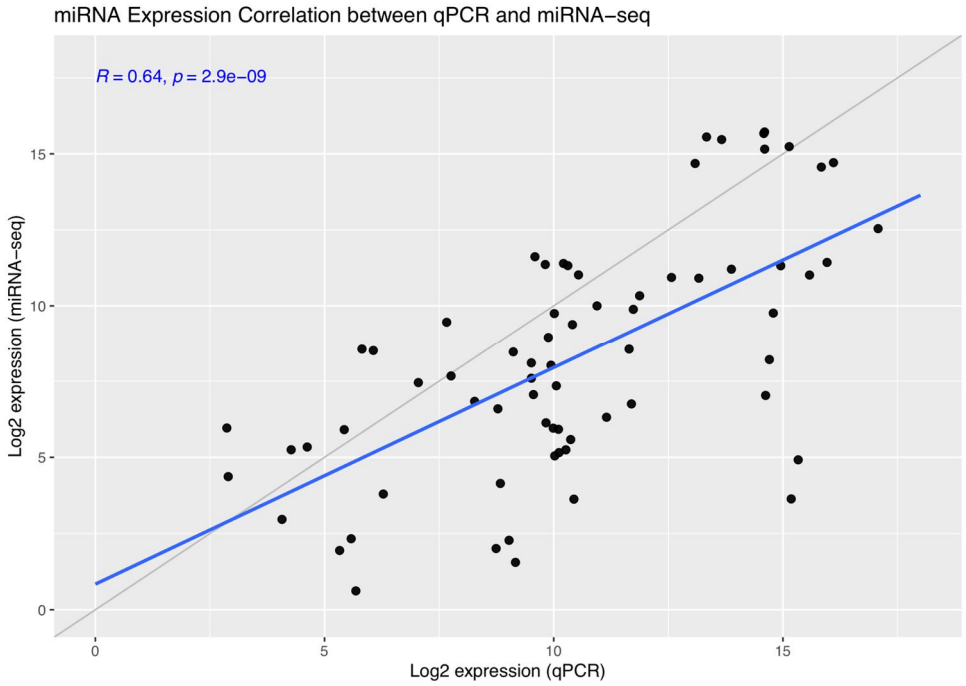

**Figure 8**. **miRNA putative targets GO enrichment analysis and Reactome analysis.** It is represented from the top to the bottom of the figure: GO Biological processes, GO Cellular component, GO Molecular functions and Reactome pathways.




**Supplementary Figure 1. Amplification qPCR curves.** The fluorescence threshold chosen (100 fluorescence units) intersects all reliable curves in their exponential phase. Individual curves, per miRNA and colored by tissue are shown for amplification success assessment.




**Supplementary Figure 2. qPCR Melting curve analysis.** Quality control was performed by evaluating the melting curves for each amplicon. The melting temperature was estimated for each miRNA in order to discard reactions that show evidence of more than one amplicon being amplified in the same reaction well. Accordingly, we discarded samples that displayed a melting temperature at least 1 degree Celsius away from the median value of all other replicates from the same amplicon (red dots shown).




**Supplementary Figure 3. Correlation between the miRNA expression values measured by miRNA-sequencing and by qPCR.** The miRNA-sequencing technique seems to under-estimate the magnitude of miRNA expression, when compared with qPCR, as shown by the slope of the fitted line (tilted towards the qPCR axis). However, the two measurements are moderately correlated (R = 0.64). The gray line indicates perfect correlation.

## Tables

**Table 1 – miRNAs with putative regulatory roles on genes required for EC operation.** RT-qPCR validation was performed. The selection was based on whether they were differentially expressed (criterium: adjusted p-value < 0.05) and/or their ability to target EC-related gene mRNAs.

| miRNA | EC related target genes | Differentialy expressed tissues |
|---|---|---|
| gga-miR-144-3p | *axin2; hey2* | PSM-D vs DCD; PSM-U vs DCD; PSM-Total vs DCD |
| gga-miR-205a | *axin2* | PSM-D vs DCD; PSM-U vs DCD; PSM-Total vs DCD |
| gga-miR-205b | *axin2* | PSM-D vs DCD; PSM-U vs DCD; PSM-Total vs DCD |
| gga-miR-221-3p | *axin2* | PSM-D vs DCD; PSM-U vs DCD; PSM-Total vs DCD |
| gga-miR-181a-5p | *dusp6; hey2;shh* | PSM-D vs DCD; PSM-U vs DCD; PSM-Total vs DCD |
| gga-miR-181b-5p | *dusp6; hey2* | PSM-D vs DCD; PSM-U vs DCD; PSM-Total vs DCD |
| gga-miR-367 | *dusp6; notch1* | PSM-D vs DCD; PSM-Total vs DCD; PSM-U vs PSM-D |
| gga-miR-125b-5p | *dusp6* | PSM-D vs DCD; PSM-U vs DCD; PSM-Total vs DCD |
| gga-miR-31-5p | *hey2* | PSM-D vs DCD; PSM-U vs PSM-D |
| gga-miR-190a-5p | *hey2* | PSM-U vs DCD; PSM-Total vs DCD |
| gga-miR-130a-3p | *dll1* | PSM-D vs DCD; PSM-U vs DCD; PSM-Total vs DCD |
| gga-miR-34c-5p | *dll1; notch1* | PSM-D vs DCD; PSM-U vs DCD; PSM-Total vs DCD |
| gga-miR-200b-3p | *notch1* | PSM-D vs DCD; PSM-U vs DCD; PSM-Total vs DCD |
| gga-miR-429-3p | *notch1* | PSM-D vs DCD; PSM-U vs DCD; PSM-Total vs DCD |
| gga-miR-219a | *brachyury* | PSM-D vs DCD; PSM-U vs DCD; PSM-Total vs DCD |
| gga-miR-302a | *wnt5a;Snail2* | PSM-D vs DCD; PSM-U vs DCD; PSM-Total vs DCD |
| gga-miR-138-1-3p | *axin2; snail2; hey2* | PSM-D vs DCD; PSM-U vs DCD; PSM-Total vs DCD; PSM-U vs PSM-D |
| gga-miR-9-3p | | PSM-D vs DCD; PSM-U vs DCD; PSM-Total vs DCD |
| gga-miR-9-5p | *hairy2** | PSM-U vs DCD |
| gga-miR-125a-5p** | *lunatic fringe* | Non detected in the mir-seq data |
| gga-miR-148a-3p*** | *dll1* | PSM-U vs DCD; PSM-Total vs DCD |
| gga-miR-21-5p*** | | Non-differentially expressed |
| gga-miR-30d*** | *snail1* | Non-differentially expressed |
| gga-miR-222b-5p**** | *dusp6* | Non-differentially expressed |
| gga-miR-19b-5p**** | *shh* | Non-differentially expressed |
| gga-miR-146a-5p**** | | Non-differentially expressed |
| gga-let-7a-5p | | PSM-D vs DCD; PSM-U vs DCD; PSM-Total vs DCD |

* Target manually found by sequence comparation.

** Publish to play a role on chicken somitogenesis, but not detected on our miR-seq data (Riley et al., 2013).

*** highly expressed miRNAs.

**** Possible miRNAs to be used has qPCR normalizers, according to our mir-seq data (Vandesompele et al., 2002).

## References


**Agarwal, V., Bell, G.W., Nam, J.-W., and Bartel, D.P.** (2015). Predicting effective microRNA target sites in mammalian mRNAs. eLife *4*, e05005. https://doi.org/10.7554/eLife.05005.

**Anderson, C., Khan, M.A.F., Wong, F., Solovieva, T., Oliveira, N.M.M., Baldock, R.A., Tickle, C., Burt, D.W., and Stern, C.D.** (2016). A strategy to discover new organizers identifies a putative heart organizer. Nature Communications *7*, 12656. https://doi.org/10.1038/ncomms12656.

**Babraham Bioinformatics – FastQC A Quality Control tool for High Throughput Sequence Data.** (2016). https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

**Bailey, C., and Dale, K.** (2015) Somitogenesis in Vertebrate Development. In eLS, pp. 1-15. https://doi.org/10.1002/9780470015902.a0003820.pub2.

**Baranski, M., Berdougo, E., Sandler, J.S., Darnell, D.K., and Burrus, L.W.** (2000). The dynamic expression pattern of frzb-1 suggests multiple roles in chick development. Dev Biol 217, 25-41. https://doi.org/10.1006/dbio.1999.9516.

**Bartel, D.P.** (2018). Metazoan MicroRNAs. Cell *173*, 20-51. https://doi.org/10.1016/j.cell.2018.03.006.

**Bonev, B., Stanley, P., and Papalopulu, N.** (2012). MicroRNA-9 Modulates Hes1 Ultradian Oscillations by Forming a Double-Negative Feedback Loop. Cell Reports *2*, 10-18. http://dx.doi.org/10.1016/j.celrep.2012.05.017.

**Carraco, G., Gonçalves, A.N., Serra, C., and Andrade, R.P.** (2014). MicroRNA processing machinery in the developing chick embryo. Gene Expression Patterns *16*, 114-121. http://dx.doi.org/10.1016/j.gep.2014.09.002.

**Dequéant, M.-L., Glynn, E., Gaudenz, K., Wahl, M., Chen, J., Mushegian, A., and Pourquié, O.** (2006). A Complex Oscillating Network of Signaling Genes Underlies the Mouse Segmentation Clock. Science *314*, 1595-1598. https://doi.org/10.1126/science.1133141.

**Duarte, I., Carraco, G., Azevedo, N.T.D., Benes, V., Andrade, R.P.** gga-miRNOME, A microRNA-sequencing dataset from chick embryonic tissues. Scientific Data (under review)3

**Duarte, I., Liber, M., Magno, R., and Andrade, R.P.** (2021). FrozenChicken: Promoting the meta-analysis of chicken microarray data. bioRxiv, 2021.2002.2025.432894. https://doi.org/10.1101/2021.02.25.432894.

**Dubrulle, J., McGrew, M.J., and Pourquié, O.** (2001). FGF Signaling Controls Somite Boundary Position and Regulates Segmentation Clock Control of Spatiotemporal Hox Gene Activation. Cell 106, 219-232. https://doi.org/10.1016/S0092-8674(01)00437-8.

**Eckei, G., Boing, M., Brand-Saberi, B., and Morosan-Puopolo, G.** (2016). Expression Pattern of Axin2 During Chicken Development. PloS One 11, e0163610. https://doi.org/10.1371/journal.pone.0163610.

**Franzosa, J.A., Bugel, S.M., Tal, T.L., La Du, J.K., Tilton, S.C., Waters, K.M., and Tanguay, R.L.** (2013). Retinoic acid-dependent regulation of miR-19 expression elicits vertebrate axis defects. Faseb j *27*, 4866-4876. https://doi.org/10.1096/fj.12-225524.

**Hamburger, V., and Hamilton, H.L.** (1951). A SERIES OF NORMAL STAGES IN THE DEVELOPMENT OF THE CHICK EMBRYO. Journal of Morphology *88*, 49-&. https://doi.org/10.1002/jmor.1050880104.

**Hayashi, H., Mochii, M., Kodama, R., Hamada, Y., Mizuno, N., Eguchi, G., and Tachi, C.** (1996). Isolation of a novel chick homolog of Serrate and its coexpression with C-Notch-1 in chick development. Int J Dev Biol 40, 1089-1096.

**Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., Sidiropoulos, K., Cook, J., Gillespie, M., Haw, R., et al.** (2020). The reactome pathway knowledgebase. Nucleic Acids Res 48, D498-d503. https://doi.org/10.1093/nar/gkz1031.

**Jing, B., Yuan, J., Yin, Z., Lv, C., Lu, S., Xiong, H., Tang, H., Ye, G., and Shi, F.** (2015). Dynamic properties of the segmentation clock mediated by microRNA. Int J Clin Exp Pathol *8*, 196-206.

**Kozomara, A., Birgaoanu, M., and Griffiths-Jones, S.** (2019). 187iRbase: from microRNA sequences to function. Nucleic Acids Res *47*, D155-D162. https://doi.org/10.1093/nar/gky1141.

**Krol, A.J., Roellig, D., Dequeant, M.L., Tassy, O., Glynn, E., Hattem, G., Mushegian, A., Oates, A.C., and Pourquie, O.** (2011). Evolutionary plasticity of segmentation clock networks. Development *138*, 2783-2792. https://doi.org/10.1242/dev.063834.

**Lee, R.C., Feinbaum, R.L., and Ambros, V.** (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854. http://dx.doi.org/10.1016/0092-8674(93)90529-Y.

**Love, M.I., Huber, W., and Anders, S.** (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

**Matsuda, M., Hayashi, H., Garcia-Ojalvo, J., Yoshioka-Kobayashi, K., Kageyama, R., Yamanaka, Y., Ikeya, M., Toguchida, J., Alev, C., and Ebisuya, M.** (2020). Species-specific segmentation clock periods are due to differential biochemical reaction speeds. Science *369*, 1450-1455. https://doi.org/10.1126/science.aba7668.

**Mi, H., Ebert, D., Muruganujan, A., Mills, C., Albou, L.P., Mushayamaha, T., and Thomas, P.D.** (2021). PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. Nucleic Acids Res *49*, D394-d403. https://doi.org/10.1093/nar/gkaa1106.

**Oginuma, M., Moncuquet, P., Xiong, F., Karoly, E., Chal, J., Guevorkian, K., and Pourquie, O.** (2017). A Gradient of Glycolytic Activity Coordinates FGF and Wnt Signaling during Elongation of the Body Axis in Amniote Embryos. Dev Cell *40*, 342-353 e310. https://doi.org/10.1016/j.devcel.2017.02.001.

**Palmeirim, I., Henrique, D., Ish-Horowicz, D., and Pourquie, O.** (1997). Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. Cell *91*, 639-648. https://doi.org/10.1016/S0092-8674(00)80451-1.

**Pascoal, S., Carvalho, C.R., Rodriguez-León, J., Delfini, M.-C., Duprez, D., Thorsteinsdóttir, S., and Palmeirim, I.** (2007). A Molecular Clock Operates During Chick Autopod Proximal-distal Outgrowth. Journal of Molecular Biology *368*, 303-309. http://dx.doi.org/10.1016/j.jmb.2007.01.089.

**Riley, M.F., Bochter, M.S., Wahi, K., Nuovo, G.J., and Cole, S.E.** (2013). Mir-125a-5p-mediated regulation of Lfng is essential for the avian segmentation clock. Dev Cell *24*, 554-561. https://doi.org/10.1016/j.devcel.2013.01.024.

**Sheeba, C.J., Andrade, R.P., and Palmeirim, I.** (2016). Mechanisms of vertebrate embryo segmentation: Common themes in trunk and limb development. Seminars in Cell & Developmental Biology. http://dx.doi.org/10.1016/j.semcdb.2016.01.010.

**Soto, X., Biga, V., Kursawe, J., Lea, R., Doostdar, P., Thomas, R., and Papalopulu, N.** (2020). Dynamic properties of noise and Her6 levels are optimized by miR-9, allowing the decoding of the Her6 oscillator. The EMBO Journal *39*, e103558. https://doi.org/10.15252/embj.2019103558.

**Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F.** (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biology 3, research0034.0031. https://doi.org/10.1186/gb-2002-3-7-research0034.

**Vitsios, D.M., and Enright, A.J.** (2015). Chimira: analysis of small RNA sequencing data and microRNA modifications. Bioinformatics *31*, 3365-3367. https://doi.org/10.1093/bioinformatics/btv380.

**Wahi, K., Friesen, S., Coppola, V., and Cole, S.E.** (2017). Putative binding sites for mir-125 family miRNAs in the mouse Lfng 3'UTR affect transcript expression in the segmentation clock, but mir-125a-5p is dispensable for normal somitogenesis. Dev Dyn 246, 740-748. https://doi.org/10.1002/dvdy.24552.

**Wightman, B., Ha, I., and Ruvkun, G.** (1993). Post-transcriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell *75*, 855-862. https://doi.org/10.1016/0092-8674(93)90530-4.

**Xie, Z.-R., Yang, H.-T., Liu, W.-C., and Hwang, M.-J.** (2007). The role of microRNA in the delayed negative feedback regulation of gene expression. Biochemical and Biophysical Research Communications *358*, 722-726. http://dx.doi.org/10.1016/j.bbrc.2007.04.207.

## 4.4 Chapter Discussion

Since the discovery of the miRNAs, they have been implicated in the timing of embryo development (Lee *et al.*, 1993; Wightman *et al.*, 1993). Moreover, they are mRNA negative regulators that could lead to rapid mRNA decay (Bartel, 2018), which is a key feature of EC mRNAs. In Chapter 4 of the present thesis we started to explore the possibility that, besides being regulators of the embryo clock (Bonev *et al.*, 2012; Riley *et al.*, 2013; Soto *et al.*, 2020), miRNAs could explain the different EC periodicities between the PSM and distal forelimb. Also, we wanted to find if miRNAs could be responsible for the arrest of cyclic expression in the anterior PSM. For that, we isolated determined PSM (D-PSM), undetermined PSM (U-PSM), and the forelimb distal cyclic domain (DCD), extracted total RNA, and analyzed by miRNA sequencing. The results obtained are described in two manuscripts presented.

In the first article, “gga-miRNOME, A microRNA-sequencing dataset from chick embryonic tissues”, we described the data sets generated from the miRNA-Seq approach. In this work, we found 926 known miRNAs and 1141 novel putative miRNAs, almost doubling the number of miRNAs annotated in the miRBase for *Gallus* gallus (1232) (Kozomara *et al.*, 2019). This represents an invaluable resource for the avian research community looking to experimentally validate novel candidate miRNAs acting in early vertebrate development.

In the second publication, “A Genome-wide Approach to the Identification of miRNAs in Temporal Control of the Chicken Embryo Clock”, we explored the 926 known miRNAs found in our dataset and tried to find potential new EC-related miRNAs and validate them. By miRNA differential expression analysis, we found that 191 miRNAs are differentially expressed between the three tissues. From this list, we selected 20 miRNAs capable of targeting EC genes and validated their expression by miRNA-qPCR. In a parallel approach, the putative target genes of the differentially expressed miRNAs were identified and then crossed with differentially expressed genes from the tissues under analysis, using previously published microarray datasets. Functional annotation of the resulting miRNA list was performed revealing multiple cellular and molecular processes that are under tissue-specific miRNA regulation.

Although we could identify and validate a list of miRNAs that are able to target EC clock genes and consequently regulate it, this study has some limitations. 1) By using this methodology, we were able to reduce and select only a few miRNAs that could be relevant for the EC. However, if a miRNA was not detected by miRNA-Seq in one of the tissues, it was not considered for subsequent experiments. 2) The *Gallus gallus* miRNome is still poorly annotated, as we could find nearly the same amount of potential new miRNAs as the number already annotated miRNAs. Since we only considered the known miRNAs for the second part of the work, we are potentially losing half of our candidate miRNAs. 3) We used TargerScan

(Agarwal *et al.*, 2015) to annotate potential targets for the differentially expressed miRNAs. Since *G. gallus* is not present on TargetScan, it only retrieves the genes that have conserved target sites with the human orthologs, thus limiting the potential to find *Gallus gallus* miRNA::gene pairs. 4) The EC in forelimb development is still poorly explored. Only *hairy1* and *hairy2* were reported to have a cyclic expression (this thesis and Pascoal *et al.*, 2007).

Following this work, functional experiments are needed to functionally validate some of the more promising miRNAs as regulators of the EC and potentially explain the different periodicity in different tissues.

## 4.5 References

**Agarwal, V., Bell, G.W., Nam, J.-W., and Bartel, D.P.** (2015). Predicting effective microRNA target sites in mammalian mRNAs. eLife 4, e05005. https://doi.org/10.7554/eLife.05005.

**Bartel, D.P.** (2018). Metazoan MicroRNAs. Cell 173, 20-51. https://doi.org/10.1016/j.cell.2018.03.006.

**Bessho, Y., Hirata, H., Masamizu, Y., and Kageyama, R**. (2003). Periodic repression by the bHLH factor Hes7 is an essential mechanism for the somite segmentation clock. Genes Dev 17, 1451-1456. https://doi.org/10.1101/gad.1092303.

**Bessho, Y., and Kageyama, R.** (2003). Oscillations, clocks and segmentation. Current Opinion in Genetics & Development 13, 379-384. https://doi.org/10.1016/S0959-437X(03)00083-2.

**Bonev, B., Stanley, P., and Papalopulu, N.** (2012). MicroRNA-9 Modulates Hes1 Ultradian Oscillations by Forming a Double-Negative Feedback Loop. Cell Reports 2, 10-18. http://dx.doi.org/10.1016/j.celrep.2012.05.017.

**Kozomara, A., Birgaoanu, M., and Griffiths-Jones, S.** (2019). miRBase: from microRNA sequences to function. Nucleic Acids Res 47, D155-D162. https://doi.org/10.1093/nar/gky1141.

**Lee, R.C., Feinbaum, R.L., and Ambros, V.** (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75, 843-854. http://dx.doi.org/10.1016/0092-8674(93)90529-Y.

**Pascoal, S., Carvalho, C.R., Rodriguez-León, J., Delfini, M.-C., Duprez, D., Thorsteinsdóttir, S., and Palmeirim, I.** (2007). A Molecular Clock Operates During Chick Autopod Proximal-distal Outgrowth. Journal of Molecular Biology 368, 303-309. http://dx.doi.org/10.1016/j.jmb.2007.01.089.

**Riley, M.F., Bochter, M.S., Wahi, K., Nuovo, G.J., and Cole, S.E.** (2013). Mir-125a-5p-mediated regulation of Lfng is essential for the avian segmentation clock. Dev Cell 24, 554-561. https://doi.org/10.1016/j.devcel.2013.01.024

**Saga, Y.** (2012). The synchrony and cyclicity of developmental events. Cold Spring Harb Perspect Biol 4. https://doi.org/10.1101/cshperspect.a008201

**Soto, X., Biga, V., Kursawe, J., Lea, R., Doostdar, P., Thomas, R., and Papalopulu, N.** (2020). Dynamic properties of noise and Her6 levels are optimized by miR-9, allowing the decoding of the Her6 oscillator. The EMBO Journal 39, e103558. https://doi.org/10.15252/embj.2019103558.

**Wightman, B., Ha, I., and Ruvkun, G.** (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell 75, 855-862. https://doi.org/10.1016/0092-8674(93)90530-4.

*This page is intentionally left blank*

# Chapter 5

## Final Considerations

*This page is intentionally left blank*

## 5.1 General overview

The precise timing of the processes that drive embryo development is critical for proper embryo development. Several processes, such as somitogenesis, rely on a clock-like molecular mechanism that dictates when a new pair of somites is produced and even possibly gives identity to the segments (Cordes *et al.*, 2004; Palmeirim *et al.*, 1997; Pourquie, 2001; Zákány *et al.*, 2001). Besides somitogenesis, other developmental processes may rely on the embryo clock (EC), such as limb outgrowth and neurogenesis (Aulehla and Pourquie, 2008; Pascoal *et al.*, 2007; Shimojo *et al.*, 2008). Several genes, for example, *hairy2,* present cycles of gene expression in the chicken PSM tissue with a periodicity of 90 minutes (Jouve *et al.*, 2000; Krol *et al.*, 2011). However, *hairy2* cycles in the chicken limb with a periodicity of six hours (Pascoal *et al.*, 2007). This made us question how the embryo clock periodicity is fine-tuned in the different tissues for the proper developmental timing of each structure. The scientific community has been trying to answer this question mostly by analyzing mutants in EC-related genes. However, even slight changes to the EC generally culminate in a total arrest of the oscillations (Choorapoikayil *et al.*, 2012; Fujimuro *et al.*, 2014; Lleras Forero *et al.*, 2018; Niwa *et al.*, 2011; Shimojo *et al.*, 2016; Takashima *et al.*, 2011). This makes it even more difficult to explain how the EC could have different paces within the same organism.

mRNA processing, translation, and decay have an essential role in the finetuning of the embryo clock (Saga, 2012). Hence, we hypothesized that alternative transcription or different mRNA stability regulators may account for the different periods of oscillation of the EC in the same embryo. In order to gain further insight on this matter, we pursued two hypotheses that gave rise to three core chapters of the present thesis: 1) Alternative transcripts finetune the EC and 2) miRNA species regulate the different pace of the EC in different tissues. Regarding hypothesis 1), in Chapter 1 we focused on the identification and functional relevance of the expression of *hairy1* transcripts produced by alternative transcription start sites. Moreover, we sought to identify novel genes belonging to the EC in limb outgrowth. In Chapter 2, we analyzed the 3'UTR of the zebrafish core EC genes, *her1, her7,* and *dlc.* As for hypothesis 2), in Chapter 4 we thoroughly characterized the chicken miRNOME in the PSM and developing limb using a genomewide approach.

## 5.2 Main Conclusions

The main goal of this thesis was to gain insights on the EC in tissue where it operates with different paces, namely the PSM and forelimb bud (Palmeirim *et al*., 1997; Pascoal *et al.,* 2007), which could explain the different *tempos* of the EC.

- In Chapter 2, we explored the functional role of an alternative *hairy1* transcript produced by an alternative TSS and that is expressed exclusively in the caudal portion of the embryo and not in the forelimb*, s-hairy1.* We found that overexpressing *s-hairy1* results in shortening of the primitive streak, and in an overall delay of embryo growth. Moreover, we found a delay in the beginning of somitogenesis by overexpressing *hairy1*, which is accompanied by a delay in *HoxB* gene expression. This phenocopied what is observed upon overexpression of *hairy1*. These results suggest that *s-hairy1* could have overlapping functions with the canonical *hairy1* transcript. We also found that the Hairy1 protein is able to bind the *HoxB9* promoter chromatin dynamically, suggesting that the periodicity of Hairy1 oscillations could be timing the aperture of the *HoxB* cluster.
- In Chapter 2, we described for the first time cyclic *hairy1* expression in the distal forelimb with a six-hour pace. This finding significantly strengthens the relevance of the operation of an EC in limb bud outgrowth and patterning.
- In Chapter 3, we evaluated the potential impact of the 3'UTR on the EC. For that, using CRISPR technology, we manipulated the 3'UTR of the three core zebrafish EC genes, *her1, her7,* and *dlc.* Regarding the alternative *her1* and *her7* 3'UTRs, we found evidence suggesting that it may have a role in ensuring Left-Right symmetry in EC oscillations.
- Regarding *dlC* 3'UTR, we found that RR3 is required for *dlc* mRNA instability, possibly because it harbors two PRE and one ARE, which were found to be important regulators of the zebrafish clock genes (Tietz *et al.*, 2020).
- *dlc*-ΔRR3 deletion and *dlc*-ΔRR2 (to less extent) show defects in somitogenesis and altered expression of *her7*. This may explain the intense phenotype of the *dlc*-ΔRR2&3. When both regions are deleted, it results in a constant overexpression, from the anterior somites until the posterior end of the PSM.
- In Chapter 4, we unveiled a dataset of 1141 putative new chicken miRNAs (almost the same amount annotated in miRBase (Kozomara *et al.*, 2019)), that have the potential to be involved in multiple development mechanisms, including the embryo clock. Moreover, we identified differentially expressed miRNAs in the PSM and limb bud that are inversely correlated with the expression of genes belonging to multiple biological processes.

Overall, with this work, we shed light on multiple putative mechanisms involving RNA molecules that could be in temporal control of EC oscillations.

## 5.3 Future Perspectives

There is still a long way to go to unveil the gene expression mechanisms that dictate the precise timing of the embryo clock. Our progress opened more questions than retrieved answers.

- We identified the alternative transcript *s-hairy1* however, there are still fundamental questions that remain unanswered about this transcript. First, the nature of the transcript must be clarified, if it is a long non-coding RNA or a protein-coding RNA. If it is protein-coding, what is the sequence of this putative protein? A new antibody against chicken Hairy1 is currently being optimized in the laboratory for western-blot analysis, and with it, we can more precisely find the Hairy1 protein forms in the PSM and forelimb tissues. If we find new Hairy1 isoforms, mass-spectrometry analysis would be useful to confirm the production of s-*hairy1.* Meanwhile, an *in vitro* translation assay could be informative. Also, more overexpression assays are needed to observe how the EC responds to the overexpression in both tissues.
- We found that *s-hairy* has a role in somitogenesis. However, the impact of the ectopic expression of this transcript in the forelimb remains elusive. As such, functional studies on the forelimb must be conducted.
- Further studies are also needed in the zebrafish mutants. Particularly the *dlc*-ΔRR2&3. We need to analyze how impaired the EC is in these mutants and how somitogenesis is affected. Moreover, more restrictive mutants are needed, such as mutants lacking PRE and ARE motifs, a single mutation, or combined mutations in PRE and ARE motifs in different locations in the 3'UTR. This way, we can pinpoint the relevance of each motif for the EC.
- By miRNA-sequencing and RT-qPCR we found interesting miRNAs::target genes pairs that are expressed in the limb mesenchyme, especially the miRNAs that seem to be "protecting" the mesenchymal forelimb cells from *notch1* and *axin2,* restricting their expression to the AER. Also, we found miRNAs that are enriched in the PSM that potentially can target EC-related genes. Further work in these pairs may unveil a possible role of the miRNAs on the definition of the *tempo* of the EC.
- GO analysis was performed for the targets of each one of the miRNAs that showed enrichment in the PSM for RNA transcription and macromolecules biosynthesis regulation. It would be interesting to reverse the analysis and find which are the specific genes that are upregulated to find new potential the miRNA::target gene pairs that could explain the difference in EC pace.
- LNA (Locked nucleic acid) whole-mount *in situ* hybridization should be performed to analyze the distribution of specific miRNAs' expression. Furthermore, functional

experiments are already in preparation to assess the putative regulation of *hairy2* by gga-miR-9-5p. We intend to electroporate morpholinos and make a target protection assay, in both the PSM and distal forelimb. With that, we intend to find if this miRNA has s role in the difference of the paces of the EC between the limb and PSM.

- Finally, we generated a dataset of new putative *Gallus gallus* miRNAs. A more in-depth study should be performed to validate them and explore the possibility of finding new regulators of the embryo clock.

## 5.4 References


**Aulehla, A., and Pourquie, O.** (2008). Oscillating signaling pathways during embryonic development. Curr Opin Cell Biol *20*, 632-637.

**Choorapoikayil, S., Willems, B., Strohle, P., and Gajewski, M.** (2012). Analysis of her1 and her7 mutants reveals a spatio temporal separation of the somite clock module. PLoS One *7*, e39073. https://doi.org/10.1371/journal.pone.0039073.

**Cordes, R., Schuster-Gossler, K., Serth, K., and Gossler, A.** (2004). Specification of vertebral identity is coupled to Notch signalling and the segmentation clock. Development *131*, 1221-1233. https://doi.org/10.1242/dev.01030.

**Fujimuro, T., Matsui, T., Nitanda, Y., Matta, T., Sakumura, Y., Saito, M., Kohno, K., Nakahata, Y., and Bessho, Y.** (2014). Hes7 3'UTR is required for somite segmentation function. Sci Rep *4*, 6462. https://doi.org/10.1038/srep06462.

**Jouve, C., Palmeirim, I., Henrique, D., Beckers, J., Gossler, A., Ish-Horowicz, D., and Pourquie, O.** (2000). Notch signalling is required for cyclic expression of the hairy-like gene HES1 in the presomitic mesoderm. Development *127*, 1421-1429. https://doi.org/10.1242/dev.127.7.1421

**Kozomara, A., Birgaoanu, M., and Griffiths-Jones, S.** (2019). miRBase: from microRNA sequences to function. Nucleic Acids Res *47*, D155-D162. https://doi.org/10.1093/nar/gky1141.

**Krol, A.J., Roellig, D., Dequeant, M.L., Tassy, O., Glynn, E., Hattem, G., Mushegian, A., Oates, A.C., and Pourquie, O.** (2011). Evolutionary plasticity of segmentation clock networks. Development *138*, 2783-2792. https://doi.org/10.1242/dev.063834.

**Lleras Forero, L., Narayanan, R., Huitema, L.F., VanBergen, M., Apschner, A., Peterson-Maduro, J., Logister, I., Valentin, G., Morelli, L.G., Oates, A.C., and Schulte-Merker, S.** (2018). Segmentation of the zebrafish axial skeleton relies on notochord sheath cells and not on the segmentation clock. Elife *7*. https://doi.org/10.7554/eLife.33843.

**Niwa, Y., Shimojo, H., Isomura, A., Gonzalez, A., Miyachi, H., and Kageyama, R.** (2011). Different types of oscillations in Notch and Fgf signaling regulate the spatiotemporal periodicity of somitogenesis. Genes Dev *25*, 1115-1120. https://doi.org/10.1101/gad.2035311.

**Palmeirim, I., Henrique, D., Ish-Horowicz, D., and Pourquie, O.** (1997). Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. Cell *91*, 639-648. https://doi.org/10.1016/S0092-8674(00)80451-1

**Pascoal, S., Carvalho, C.R., Rodriguez-León, J., Delfini, M.-C., Duprez, D., Thorsteinsdóttir, S., and Palmeirim, I.** (2007). A Molecular Clock Operates During Chick Autopod Proximal-distal Outgrowth. Journal of Molecular Biology *368*, 303-309. http://dx.doi.org/10.1016/j.jmb.2007.01.089.

**Pourquie, O.** (2001). The vertebrate segmentation clock. J Anat *199*, 169-175. https://doi.org/10.1046/j.1469-7580.2001.19910169.x.

**Saga, Y.** (2012). The synchrony and cyclicity of developmental events. Cold Spring Harb Perspect Biol *4*. https://doi.org/10.1101/cshperspect.a008201

**Shimojo, H., Isomura, A., Ohtsuka, T., Kori, H., Miyachi, H., and Kageyama, R.** (2016). Oscillatory control of Delta-like1 in cell interactions regulates dynamic gene expression and tissue morphogenesis. Genes Dev *30*, 102-116. https://doi.org/10.1101/gad.270785.115.

**Shimojo, H., Ohtsuka, T., and Kageyama, R.** (2008). Oscillations in Notch Signaling Regulate Maintenance of Neural Progenitors. Neuron *58*, 52-64. https://doi.org/10.1016/j.neuron.2008.02.014.

**Takashima, Y., Ohtsuka, T., González, A., Miyachi, H., and Kageyama, R.** (2011). Intronic delay is essential for oscillatory expression in the segmentation clock. Proceedings of the National Academy of Sciences *108*, 3300-3305. https://doi.org/10.1073/pnas.1014418108.

**Tietz, K.T., Gallagher, T.L., Mannings, M.C., Morrow, Z.T., Derr, N.L., and Amacher, S.L.** (2020). Pumilio response and AU-rich elements drive rapid decay of Pnrc2-regulated cyclic gene transcripts. Developmental Biology *462*, 129-140. https://doi.org/10.1016/j.ydbio.2020.03.017.

**Zákány, J., Kmita, M., Alarcon, P., de la Pompa, J.L., and Duboule, D.** (2001). Localized and transient transcription of Hox genes suggests a link between patterning and the segmentation clock. Cell *106*, 207-217. https://doi.org/10.1016/s0092-8674(01)00436-6.