

# **Investigating microRNA-target interactions during skeletal muscle development in chicken embryos**

**Camille J.M. VIAUT**

A thesis submitted for the degree of Doctor of Philosophy  
Ph.D. in Biomolecular Science

University of East Anglia  
School of Biological Sciences  
Norwich, United Kingdom

**May 2017**

## ABSTRACT

MicroRNAs (miRNAs), short non-coding RNAs, which act post-transcriptionally to regulate gene expression, are of widespread significance and have been implicated in many biological processes during development and disease, including muscle disease. In addition to the myomiRs, which are miRNAs highly enriched in striated muscles, recent advances in sequencing technology and bioinformatics led to the identification of a large number of miRNAs in vertebrates and other species. However, for many of these miRNAs specific roles, in particular during myogenesis, have not yet been determined.

Here, I investigated the potential functions of miR-128, confirmed an interaction with one of its candidate targets, Eya4, and looked at the impact of its knock-down on skeletal myogenesis in the chicken embryo.

The expression pattern of miR-128, as well as 22 other somitic miRNAs, were characterised by LNA *in situ* hybridisation (LNA ISH).

Eya4 was identified as a candidate ‘muscle’ target of miR-128 by computational analysis. Its expression pattern was characterised; miR-128 and Gga-Eya4 displayed similar profiles in developing somites. Using the miRanda algorithm potential miRNA binding sites were identified in the 3’ untranslated region (UTR) of other transcription factors, which along with Eya4 are members of the PAX-SIX-EYA-DACH (PSED) network (Six1/4, Eya1/2/3, and Dach1).

These miRNA/target interactions were examined *in vitro* and *in vivo*. Gga-Eya4 was confirmed as a target of miR-128 as well as miR-206 by luciferase reporter assays. MiR-128/Gga-Eya4 interaction was validated by RNA ISH and RT-qPCR after antagomiR (AM)-128 injection in chicken embryos. Knock-down of miR-128 resulted in a significant de-repression of Gga-Eya4 expression; an increase in Gga-Six4 and Gga-Pax3 expression was also observed, whereas Gga-MyoD1 expression was decreased.

With this project, using a combination of cell-based experiments and animal studies, I showed that miR-128 could play an important role in the regulation of skeletal myogenesis in the chicken embryo by targeting Gga-Eya4, a member of the PSED network.

**'Perfer et obdura'**

- Ovid -

**To my Mum**

# TABLE OF CONTENTS

<b>TITLE</b> .....	<b>1</b>
<b>ABSTRACT</b> .....	<b>2</b>
<b>TABLE OF CONTENTS</b> .....	<b>4</b>
<b>LIST OF FIGURES</b> .....	<b>9</b>
<b>LIST OF TABLES</b> .....	<b>11</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>12</b>
<b>CHAPTER 1: INTRODUCTION</b> .....	<b>14</b>
<b>1.1. Skeletal muscle development</b> .....	<b>14</b>
<b>1.1.1. The chicken embryo: a model system for skeletal myogenesis</b> .....	<b>14</b>
<b>1.1.2. Origin of skeletal muscle: from gastrula to somite</b> .....	<b>14</b>
<b>1.1.3. Somite formation and differentiation: from pre-somitic mesoderm (PSM) to myotome</b> .....	<b>16</b>
<b>a. Somitogenesis</b> .....	<b>16</b>
<b>b. Regulation of somitogenesis</b> .....	<b>18</b>
<b>c. Somite differentiation</b> .....	<b>19</b>
<b>d. Regulation of somite differentiation</b> .....	<b>22</b>
<b>1.1.4. Myogenesis</b> .....	<b>24</b>
<b>a. Cellular heterogeneity: from myogenic progenitor cells (MPCs) to myotube</b> .....	<b>24</b>
<b>b. Genetic networks regulating myogenesis</b> .....	<b>25</b>
<b>i) Myogenic regulatory factors (MRFs)</b> .....	<b>25</b>
<b>ii) PAX-SIX-EYA-DACH (PSED) network</b> .....	<b>27</b>
<b>1.2. MicroRNAs</b> .....	<b>31</b>
<b>1.2.1. Discovery of microRNAs</b> .....	<b>31</b>
<b>1.2.2. MicroRNA biogenesis</b> .....	<b>32</b>
<b>a. MicroRNA genes</b> .....	<b>32</b>
<b>b. Production: from gene to mature microRNA</b> .....	<b>34</b>
<b>1.2.3. RNA silencing</b> .....	<b>37</b>
<b>1.2.4. MicroRNAs and myogenesis</b> .....	<b>38</b>
<b>a. Dicer-dependent myogenesis</b> .....	<b>38</b>
<b>b. Muscle-specific microRNAs: myomiRs</b> .....	<b>39</b>

1.2.5. Identification of novel microRNAs .....	44
1.3. Research aims and objectives .....	45
<b>CHAPTER 2: MATERIALS AND METHODS .....</b>	<b>47</b>
2.1. Culture of chicken eggs .....	47
2.2. Embryo staging, harvesting and fixation .....	47
2.3. Embryo dehydration .....	47
2.4. Embryo rehydration .....	47
2.5. Whole-mount <i>in situ</i> hybridisation (WMISH) .....	48
2.5.1. LNA probe pre-absorption .....	48
2.5.2. <i>In situ</i> hybridisation protocol .....	48
2.5.3. <i>In situ</i> hybridisation – Buffers and solutions .....	49
2.5.4. Embedding, sectioning and imaging .....	51
2.6. Locked Nucleic Acid (LNA) probes for WMISH .....	52
2.7. Labelled RNA probe synthesis by PCR .....	52
2.7.1. DNA template preparation .....	52
2.7.2. Probe synthesis (Transcription Reaction) .....	53
2.7.3. Probe purification .....	54
2.8 RNA extraction from embryos and dissected somites .....	54
2.9. Reverse transcription .....	55
2.10. Amplification of cDNA by Polymerase Chain Reaction (PCR) .....	55
2.11. Agarose gel electrophoresis of DNA and RNA .....	57
2.12. Restriction digestion .....	57
2.13. Purification of PCR products and digested plasmid DNAs .....	57
2.14. Ligation of DNA into a plasmid vector .....	58
2.14.1. Ligation efficiency .....	58
2.14.2. Ligation into pGEM-T Easy vector .....	59
2.14.3. Ligation into pGL3-modified vector and pCAB vector .....	60
2.15. Preparation of DH5 $\alpha$ <i>Escherichia coli</i> ( <i>E. coli</i> ) competent cells .....	60
2.15.1. Solutions and Buffers .....	61
2.15.2. Transfection efficiency test .....	61
2.16. Plasmid transformation into DH5 $\alpha$ competent cells .....	62
2.16.1. Selection of transformants .....	62
2.16.2. Colony PCR .....	63
2.17. Plasmid isolation from bacterial cultures .....	63

2.18. Bioinformatics: databases, software and algorithms -----	65
2.19. MicroRNA target validation – <i>In vitro</i> interaction -----	65
2.19.1. Mutagenesis of inserts cloned into pGL3-modified vector -----	66
2.19.2. Cell culture -----	67
2.19.3. Cell transfection -----	68
2.19.4. Luciferase Reporter Assays -----	69
2.19.5. Normalisation of luciferase assays data -----	69
2.20. MicroRNA target validation – <i>In vivo</i> interaction -----	70
2.20.1. Microinjection -----	70
2.20.2. Harvesting of treated chicken embryos -----	71
2.20.3. RNA extraction from dissected somites – qPCR -----	71
2.20.4. Real-Time quantitative PCR (RT-qPCR) -----	72
2.20.5. Normalisation of RT-qPCR data -----	73
<b>CHAPTER 3: MicroRNA CHARACTERISATION AND TARGET IDENTIFICATION -----</b>	<b>74</b>
3.1. Introduction -----	74
MiR-128 -----	76
3.2. Results and discussion -----	78
3.2.1. Characterisation of microRNA expression patterns -----	78
a. MicroRNAs: conservation across species -----	79
b. MicroRNAs: expression patterns -----	81
i) New insights into the myomiRs -----	81
ii) Characterisation of a group of 16 somitic miRNAs-----	84
iii) Characterisation of miR-128 -----	89
3.2.2. Identification of microRNA targets -----	92
a. <i>Eya4</i> : a candidate target for miR-128 -----	92
b. Identification of microRNAs targeting PSED members -----	99
3.3. Conclusion -----	102
<b>CHAPTER 4: CHARACTERISATION OF MicroRNA TARGETS AND <i>in vitro</i> INTERACTION INVESTIGATION -----</b>	<b>104</b>
4.1. Introduction -----	104
<i>Eya4</i> and the other PSED members -----	105
4.2. Results and discussion -----	110

4.2.1. Validation of Eya4 as a target for miR-128 .....	110
a. Cloning of a fragment of the coding sequence of Gga-Eya4 .....	110
b. Characterisation of Gga-Eya4 .....	114
c. Cloning of a fragment of the 3' untranslated region (UTR) of Gga-Eya4 .....	118
d. Mutagenesis: Gga-Eya4 3'UTR mutants .....	119
e. Luciferase reporter assays: miRNA/target interaction <i>in vitro</i> .....	121
i) MiR-128 targets Gga-Eya4 3'UTR .....	123
ii) Potential synergism between miRNAs in the regulation of Gga-Eya4 expression .....	123
4.2.2. MicroRNA regulation of members of the PSED network, and their expression patterns .....	128
a. Cloning of 3'UTR fragments of PSED members and mutagenesis .....	128
b. Luciferase reporter assays: miRNA/target interaction <i>in vitro</i> .....	130
i) MiR-128 and miR-133a target Gga-Eya1 3'UTR .....	132
ii) MiR-133a targets Gga-Eya3 3'UTR .....	132
iii) Gga-Six1, Gga-Six4 and Gga-Dach1 3'UTRs, and microRNA regulation .....	132
c. Chicken Six1/4, Eya1-3, Dach1 expression patterns .....	135
i) Gga-Eya1, Gga-Eya2 and Gga-Eya3 .....	136
ii) Gga-Six1, Gga-Six4 and Gga-Dach1 .....	138
4.3. Conclusion .....	140
CHAPTER 5: MicroRNA/mRNA target interactions <i>in vivo</i> and their role(s) during myogenesis .....	142
5.1. Introduction .....	142
MicroRNAs, the PSED network and myogenesis .....	142
Functional experiments .....	145
5.2. Results and discussion .....	146
5.2.1. MiR-128 regulates myogenesis via its interaction with Gga-Eya4 .....	147
a. MiR-128 targets Gga-Eya4 <i>in vivo</i> .....	147
b. MiR-128, Gga-Eya4 and myogenesis .....	152
5.2.2. MiR-206/Gga-Eya4, miR-133/Gga-Eya1 and miR-133/Gga-Eya3 interactions <i>in vivo</i> .....	154
a. MiR-206/Gga-Eya4 interaction and myogenesis .....	154

b. MiR-133/Gga-Eya1 and miR-133/Gga-Eya3 interactions <i>in vivo</i> ---	156
5.2.3. Control experiments – Scrambled antagomiR (AM-scr) injections --	156
5.3. Conclusion -----	158
CHAPTER 6: DISCUSSION AND CONCLUSION -----	161
6.1. Summary and discussion -----	161
6.1.1. MicroRNA characterisation -----	161
MiR-128: myomiR? -----	163
6.1.2. MiR-128 target identification -----	165
MicroRNAs and regulation of target expression -----	165
6.1.3. MiR-128/Gga-Eya4 interaction -----	166
MiR-128 and Gga-Eya4 – overlapping expression -----	166
Gga-Eya4: co-regulation by miR-128 and miR-206? -----	169
6.1.4. MiR-128/Gga-Eya4 interaction, PSED network and myogenesis ----	170
MiR-128/Gga-Eya4 interaction and myogenesis -----	170
6.2. Future work -----	171
6.3. Conclusion -----	173
ABBREVIATIONS -----	174
PUBLICATION -----	178
REFERENCES -----	179
APPENDICES -----	203
APPENDIX I – Primers and antagomiRs sequences -----	203
APPENDIX II – MicroRNA expressions and Bioinformatics analysis -----	207
APPENDIX III – Sequences of coding and 3'UTR fragments cloned -----	219

## LIST OF FIGURES

<b>Figure 1.1: Somitogenesis in the chicken embryo</b> .....	<b>17</b>
<b>Figure 1.2: Formation of the early myotome in the somite</b> .....	<b>21</b>
<b>Figure 1.3: Regulation of myogenesis during chicken development</b> .....	<b>26</b>
<b>Figure 1.4: The biogenesis and function of miRNAs</b> .....	<b>33</b>
<b>Figure 1.5: MiR-1 and miR-206 contribute to the downregulation of Pax3 during myoblast differentiation</b> .....	<b>42</b>
<b>Figure 2.1: pGEM-T Easy vector map</b> .....	<b>59</b>
<b>Figure 2.2: pRL-TK vector map</b> .....	<b>68</b>
<b>Figure 3.1: Expression patterns of miR-1a and miR-1b</b> .....	<b>83</b>
<b>Figure 3.2: Expression pattern of miR-206</b> .....	<b>83</b>
<b>Figure 3.3: Expression patterns of miR-133 family members</b> .....	<b>85</b>
<b>Figure 3.4: Expression patterns of 16 selected miRNAs</b> .....	<b>87</b>
<b>Figure 3.5: Expression pattern of miR-128</b> .....	<b>91</b>
<b>Figure 4.1: Cloning of Gga-Eya4</b> .....	<b>111</b>
<b>Figure 4.2: Multiple protein sequences alignment of EYA4</b> .....	<b>115</b>
<b>Figure 4.3: Expression profile of Gga-Eya4</b> .....	<b>117</b>
<b>Figure 4.4: Mutagenesis strategy</b> .....	<b>120</b>
<b>Figure 4.5: Luciferase reporter assays – miR-128/Gga-Eya4 3' untranslated region (UTR) interaction</b> .....	<b>122</b>
<b>Figure 4.6: Luciferase reporter assays – miRNA combinations and interaction with Gga-Eya4 3' untranslated region (UTR)</b> .....	<b>124</b>
<b>Figure 4.7: Luciferase reporter assays – Gga-Eya1 and Gga-Eya3</b> .....	<b>131</b>
<b>Figure 4.8: Luciferase reporter assays – Gga-Six1 and Gga-Six4</b> .....	<b>133</b>
<b>Figure 4.9: Luciferase reporter assays – Gga-Dach1</b> .....	<b>134</b>
<b>Figure 4.10: Expression patterns (Part 1): Gga-Eya1; Gga-Eya2; Gga-Eya3</b> ..	<b>137</b>
<b>Figure 4.10: Expression patterns (Part 2): Gga-Six1; Gga-Six4; Gga-Dach1</b> --	<b>139</b>
<b>Figure 5.1: Functional experiments: miRNA loss-of-function – AntagomiR strategy</b> .....	<b>146</b>

<b>Figure 5.2: Expression patterns of Gga-Eya4, Gga-Six4, Gga-Pax3 and Gga-MyoD1 after antagomiR-128 injection</b> .....	<b>148</b>
<b>Figure 5.3: RT-qPCR results – PSED and MRF transcript levels in somites injected with antagomiR-128</b> .....	<b>150</b>
<b>Figure 5.4: RT-qPCR results – PSED and MRF transcript levels in somites injected with antagomiR-206</b> .....	<b>155</b>
<b>Figure 5.5: RT-qPCR results – PSED and MRF transcript levels in somites injected with scrambled antagomiR (AM-scr)</b> .....	<b>157</b>
<b>Figure 5.6: Proposed model – MiR-128/Eya4 interaction and potential effect on the PSED network and the MRFs</b> .....	<b>159</b>

<b>Appendix III List 1: Sequences of coding and 3'untranslated region (UTR) fragments cloned and used to generate RNA oligonucleotides in order to perform RNA <i>in situ</i> hybridisation (ISH)</b> .....	<b>219</b>
<b>Appendix III List 2: Sequences of 3' untranslated region (UTR) fragments cloned in order to perform luciferase reporter assays</b> .....	<b>221</b>

## LIST OF TABLES

<b>Table 3.1: Summary – Characterisation of miRNAs during chicken development --</b> -----	<b>80</b>
<b>Table 3.2: MiR-128 targets identification – GOTERM_BP_1 results-----</b>	<b>94</b>
<b>Table 3.3: MiR-128 targets identification – GOTERM_BP_FAT results -----</b>	<b>94</b>
<b>Table 3.4: MiR-128 targets identification – KEGG_PATHWAY results -----</b>	<b>94</b>
<b>Table 3.5: MiR-128 targets identification – g:GOSl results-----</b>	<b>96</b>
<b>Table 3.6: MiR-128 targets identification – UP_TISSUE results -----</b>	<b>96</b>
<b>Table 3.7: MiR-128 targets identification – MiRanda analysis-----</b>	<b>100</b>
<b>Table 4.1: Percentage identity and similarity between EYA protein sequences-</b>	<b>113</b>
<b>Table 4.2: Summary – Luciferase reporter assays -----</b>	<b>129</b>
<b>Appendix I Table 1: Primers to generate <i>in situ</i> hybridisation (ISH) probes --</b>	<b>203</b>
<b>Appendix I Table 2: Primer sequences for PCR amplification of the 3' untranslated region (UTR) of PSED members -----</b>	<b>203</b>
<b>Appendix I Table 3: Mutagenesis primers -----</b>	<b>204</b>
<b>Appendix I Table 4: Gga-Eya4 Full-length primers -----</b>	<b>205</b>
<b>Appendix I Table 5: Other primers for sequencing and mutagenesis -----</b>	<b>205</b>
<b>Appendix I Table 6: AntagomiR sequences-----</b>	<b>206</b>
<b>Appendix I Table 7: qPCR primers -----</b>	<b>206</b>
<b>Appendix II Table 1: MiRNA sequences used to run the miRanda algorithm- -</b>	<b>207</b>
<b>Appendix II Table 2: Summary of miRNA expression in different tissues in chicken embryo -----</b>	<b>208</b>
<b>Appendix II Table 3: Somitic miRNAs – Bioinformatics analysis -----</b>	<b>208</b>
<b>Appendix III Table 1: <i>In vitro</i> transcription conditions for the synthesis of the RNA oligonucleotides used to perform RNA ISH -----</b>	<b>224</b>

## ACKNOWLEDGEMENTS

Firstly, I would like to express my gratitude to my supervisor Prof. Andrea Münsterberg for giving me the opportunity to work in her laboratory and for her support and guidance throughout my project.

I would also like to thank my secondary supervisors, Dr Grant Wheeler and Prof. Tamas Dalmay, for their support, challenging questions and input during our meetings.

I would like to thank Dr Tim Grocott for sharing his invaluable knowledge with me and for his precious help with everything. An incredible source of inspiration!

Many thanks to Dr Dominic McCormick, Dr Tracey Swinger and Dr Linh Le for sharing their knowledge and helping me with cell culture and luciferase assays.

I would like to thank Dr Simon Moxon for helping me with the bioinformatics part of my project and Dr Paul Thomas for his assistance with bioimaging.

I would like to acknowledge all of my fellow lab members, past and present, for making my PhD journey enjoyable. Thanks to the ‘chick team’: Abdul, Melissa, Danielle, Jan, Alice, Tim1, Estefanía, Geoff, Johannes, Tim2, Wenhui, Kasia, Marlene, Niki, James, Connie; and to the ‘frog team’: Inès, Marta, Angels, Katy, Nicole, Ayisha, Amanda, Chris, Vicky, Adam, Kim. Also thanks to Amy and Andy for their assistance around the lab and to all of the visitor PhD students and project/summer students for putting up with me.

Many thanks to the Norwich Biosciences Doctoral Training Partnership programme for funding this project and also to The Company of Biologists for giving me an insight into the science publishing world during my 3-month internship.

I would like to thank Dr Karine Massé for introducing me to Developmental Biology and greatly contributing to my career path choice by giving me the opportunity as an undergraduate student to work in her laboratory.

I would like to thank all of my friends for their unwavering support and the fantastic moments we shared together. Thank you Tim, James, Kadri, Ben, Liz, Jan! My life in Norwich would not have been the same without you! A very special thank you goes to Sophie and Marta for moral support and for always being by my side, especially in the toughest situations.

I would also like to thank Sarah, Jon, Molly and Tim for introducing me to the proper British culture, for helping me with my English, and for having been like a second family to me during the first year of my PhD. Thanks to Stephen and the dancers of the Knightshift Dance Company for sharing the same passion for dancing.

Finally, I would like to thank my family and especially my Mum, my Dad and my Grandma for their endless love and support throughout my life and my studies. Many thanks to my parents for always believing in me, for their generosity and their understanding. Thank you so much to my lovely Grandma for everything she has done for me over the years. Her contribution to my life and education is priceless.

Merci du fond du cœur ! Merci pour tout !

# CHAPTER 1: INTRODUCTION

## 1.1. Skeletal muscle development

### 1.1.1. The chicken embryo: a model system for skeletal myogenesis

The chicken embryo has a long history as a major system in developmental biology (Stern 2005; Hirst & Marcelle 2015).

Readily available and easy to incubate, embryo development can be directly observed by cutting a small window in the egg shell. The establishment of a staging atlas by Hamburger and Hamilton in 1951, allowed specific developmental landmarks to be seen and correlated with experimental manipulations of development (Hamburger & Hamilton 1992). The relative transparency of the embryos, and the regular formation of pairs of somites, transient metameric structures, along the anterior-posterior axis, allowed to accurately stage the embryos.

The external development of chicken embryos allows for convenient experimental manipulations at specific embryonic stages, such as, for example, grafting and lineage tracing (Ordahl & Le Douarin 1992). In the last few years, the classical approaches have been enriched by major technical advances, such as the development of new methods for gain- and loss-of function analysis (*in vivo* electroporation) (Nakamura & Funahashi 2012); and the completion of the first draft of the sequence of its genome (Hillier et al. 2004). In December 2015, the newest chicken genome, version Gallus\_gallus-5.0 (Galgal5; GCA\_000002315.3) was released; it was sequenced and assembled using varied sequencing technologies, including Sanger, Illumina and 454 (International Chicken Genome Sequencing Consortium).

With its rapid external development, its relative transparency, and an easy access to its somites, as well as its genome having been sequenced, the chicken embryo constitutes a model of choice for studying skeletal myogenesis.

### 1.1.2. Origin of skeletal muscle in vertebrates: from gastrula to somite

The vertebrate skeletal musculature is a complex and heterogeneous organ system serving multiple functions in the organism.

During early embryonic development, skeletal muscles are generated by a series of distinct morphogenetic events extremely well regulated (reviewed in Musumeci et al. 2015). Embryogenesis, the process by which the embryo forms and develops, is a succession of several phases: cleavage, gastrulation, neurulation and organogenesis.

The first phase, cleavage, is characterised by a series of cell divisions (mitoses). The different cells derived from cleavage are called blastomeres and form the blastula. The single-layered blastula is then reorganised into a three-layered structure, the gastrula.

Gastrulation is a very important phase which, in the chicken, starts with the formation of the primitive streak and the determination of the body axes (reviewed in Bénazéraf & Pourquié 2013). The gastrulation process results from the integration of cell proliferation, differentiation and migration of thousands of cells. Large-scale flows of cells from the epiblast – cells at the posterior edge of the upper layer of the *area pellucida* – migrate into the midline of the embryo to form the primitive streak. Then, these prospective mesodermal and endodermal cells undergo an epithelial-to-mesenchymal transition (EMT) and migrate beneath the remaining epiblast cells, the future ectoderm (Nakaya & Sheng 2008). At the end of the gastrulation, three germ layers are formed: the ectoderm, the endoderm and, localised in between them, the mesoderm (Solnica-Krezel & Sepich 2012).

Gastrulation is followed by organogenesis, where each layer will give rise to specific tissues and organs in the developing embryo.

Ectoderm (most external layer) is composed of three parts: external ectoderm, also known as surface ectoderm, the neural crest and the neural tube (formed during neurulation). These ectodermal structures will differentiate and form the epidermis, the skin appendages, the lens, melanocytes and the nervous tissues.

Endoderm (most internal layer) will form the epithelial lining of most of the organs related to the digestive and respiratory systems (lungs, digestive tract and annexe organs (liver, pancreas)).

Mesoderm (middle layer), is composed of the lateral plate mesoderm, intermediate mesoderm, paraxial and axial mesoderm; it will give rise to the heart, blood cells, the notochord, bone and cartilage, kidney and most of the smooth, cardiac and skeletal muscle of the body (Solnica-Krezel 2005).

The paraxial mesoderm, localised on both sides of the neural tube and the notochord, is composed of an anterior part, the cephalic paraxial mesoderm, and a posterior part, the pre-somitic mesoderm (PSM). The PSM will progressively segment to give rise to the somites (Christ & Ordahl 1995).

### **1.1.3. Somite formation and differentiation: from PSM to myotome**

#### **a. Somitogenesis**

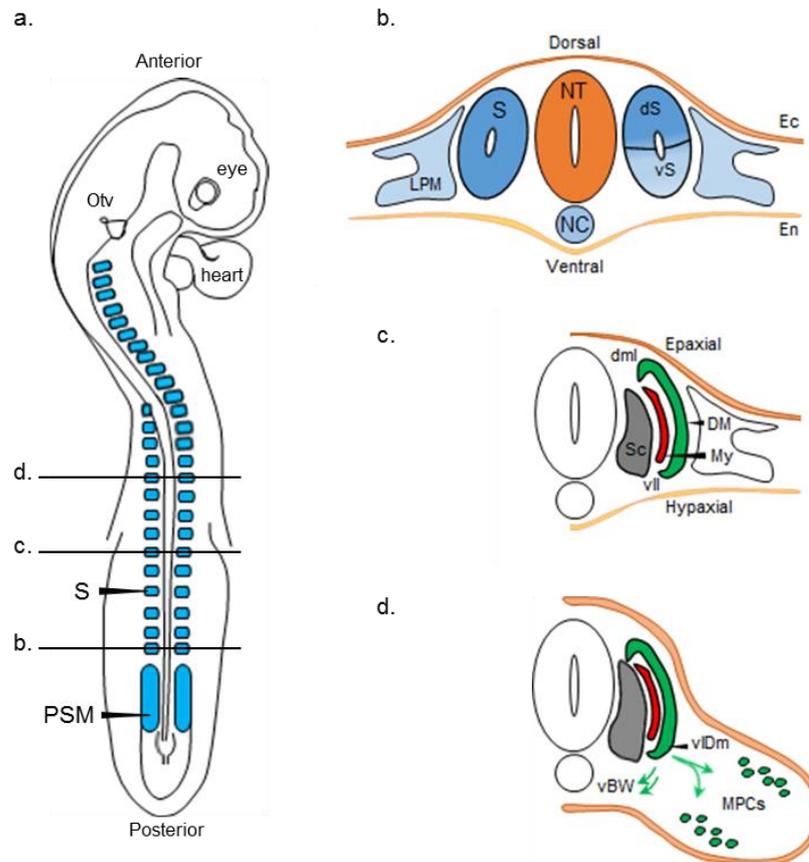
In vertebrates, most of the axial skeleton and all skeletal muscles of the body, with the exception of the craniofacial bone and head muscles, are derived from the somites (Christ & Ordahl 1995). Somites, transient metameric structures, are generated by segmentation from the PSM (**Fig. 1.1**).

Before the somites form, the paraxial mesoderm of vertebrate is segmented into somitomeres. They form along the length of the embryo during gastrulation, in a strict anterior to posterior order, and appear in bilateral pairs (Meier 1982; Jacobson 1988). Unlike the somitomeres in the head which remain contiguous, the somitomeres in the trunk and tail, gradually condense and epithelialise to become somites.

Each newly formed somite undergoes mesenchymal-to-epithelial transition (MET) and results in the formation of an epithelial ball of columnar cells enveloping mesenchymal cells within a central cavity, the somitocoel (Fig. 1.1b). Each somite is surrounded by extracellular matrix (ECM) components establishing important connections with adjacent structures.

Somitogenesis is a sequential, bilateral, directional and a periodic process. The first somite pair forms directly posterior to the otic vesicle region (future ear; Fig. 1.1a) (Hinsch & Hamilton 1956; Huang et al. 1997). From this moment, a new pair of somites forms sequentially, adding new segments, on both sides of the median line of the chicken embryo, along the anterior-posterior axis to the caudal tip as the embryo extends.

While the first formed (or oldest) somite is located at the anterior tip of the trunk paraxial mesoderm, the last produced (youngest) somite is located more posteriorly.



**Fig. 1.1: Somitogenesis in the chicken embryo.** (a) Schematic of HH15 chicken embryo showing the position of the 25 pairs of somites along the Anterior-Posterior axis (according to the Hamburger and Hamilton classification (Hamburger & Hamilton 1992)). (b-d) Transverse sections (indicated by straight lines in (a)) at different levels (posterior (b), intermediate (c), and anterior (d)) in the embryo show the evolution of the somite. Originating from the mesoderm (b), the somite forms the dermomyotome, sclerotome and the myotome (c). (d) At the flank-limb junction, cells from the dermomyotome, and the myotome migrate to give rise to most of the muscles of the body (vBW) and limbs (MPCs). DM: dermomyotome (dorsomedial lip (dml), ventrolateral lip (vll)); ventrolateral dermomyotome (vIDM)); Ec: ectoderm; En: endoderm; LPM: lateral plate mesoderm; MPCs: myogenic progenitor cells; My: myotome; NC: notochord; NT: neural tube; Otv: otic vesicle; PSM: pre-somitic mesoderm; S: somite (dorsal (dS) and ventral (vS) somite); Sc: sclerotome; vBW: ventral body wall. Adapted from Mok & Sweetman 2011.

In the chicken, a total of 52 pairs of somites are going to be generated, with a pair formed every 90 minutes at 38-39°C, during the first five days of embryonic development (Dale & Pourquié 2000). The total number of somites (zebrafish: 30; mouse: 65; human: 42), as well as the time needed to form a new somite (zebrafish: 30 minutes; mouse: 2 hours; human: 4-5 hours) are species-dependent.

### **b. Regulation of somitogenesis**

Somites originate at regular and cyclic species-specific intervals. This ability of the paraxial mesoderm can be based on a molecular oscillator, at least theoretically. Existence of such an oscillator had been predicted in a model called the ‘Clock and wavefront’ model (Cooke & Zeeman 1976), but it was only in 1997 that components of this oscillator were identified and called ‘segmentation clock’ (Palmeirim et al. 1997).

Temporal periodicity is regulated by expression of ‘oscillating genes’ (‘clock’) and by gradients of signal molecules providing a ‘wave’ motion. The number of somites is established during the initial stages of paraxial mesoderm production. Somites appear exactly at the same time bilaterally in the embryo, and the clock for production of the first pair of somites is defined when cells enter the PSM (Palmeirim et al. 2008). This process appears to be under the control of multiple signalling gradients involving the WNT, NOTCH, fibroblast growth factor (FGF), and retinoic acid (RA) pathways (reviewed in Ozbudak & Pourquié 2008; Aulehla & Pourquié 2008; Aulehla & Pourquié 2010).

The PSM can be divided in two regions that differ not only in terms of gene expression patterns, but also in morphology of the PSM cells.

In the caudal two-third of the PSM, high FGF activity is believed to keep cells in a mesenchymal, undifferentiated state and oscillatory expression of segmentation clock genes occurs.

Expression of these oscillating genes appears cyclically in PSM cells, at defined intervals. Although the origin of this periodic ‘clock’ remains unclear and highly discussed (Aulehla & Pourquié 2008), several gradients seem to play a role into the somitic segmentation. FGF and WNT proteins are produced at regular intervals in the most posterior portion of the PSM, and RA is produced by the newly formed somites in the anterior region.

A gradient of RA exists going in the anterior-to-posterior direction, while the gradient of FGF and WNT proteins is going in the posterior-to-anterior direction (Aulehla & Pourquié 2010). The region where these two gradients meet is called the ‘determination front’, or ‘wavefront’, and corresponds to the region where expression of oscillating genes is temporary segregated and remains active to form a new pair of somites (Del Corral & Storey 2004; Dubrulle & Pourquié 2004).

As PSM cells leave the posterior immature region, crossing the determination front, and enter the anterior third of the PSM, several changes take place. The first signs of morphological formation occur when the most peripheral PSM cells undergo an EMT (Nakaya et al. 2004). At this point, the somite boundaries are specified and formed through the activation by the NOTCH pathway of Hairy and Lunatic Fringe (Lfng). Hairy, via the activation of Eph/ephrin proteins and their receptors (Palmeirim et al. 1997; Jouve et al. 2000), and Lfng (Dale et al. 2003) facilitate the somite boundary formation, the detachment of a new somite from the others, and contribute to the establishment of the somite anterior-posterior polarisation.

Recent studies questioned the role of this ‘clock and wavefront’ mechanism in somite formation, and suggested that somites are self-organising structures. Palmeirim *et al.* demonstrated that the molecular segmentation of the anterior part of the PSM is an intrinsic property and that no signal coming from neighbouring tissues is required (Palmeirim et al. 1997). In addition, Dias *et al.* showed, in ectopic somite experiments, that cyclic expression of clock genes, as well as waves and gradients, are not necessary for somite formation (Dias et al. 2014); however ‘the clock’ appears to be required for normal subdivision of the somites into anterior and posterior halves (Stern & Piatkowska 2015).

### **c. Somite differentiation**

Once formed, somites rapidly differentiate and develop into three distinct cellular compartments: sclerotome, dermomyotome, and myotome (Fig. 1.1) (Christ & Ordahl 1995; Christ et al. 2007).

During early maturation, the ventral portion of the somite undergoes an EMT resulting in the formation of the sclerotome (Fig. 1.1b-c).

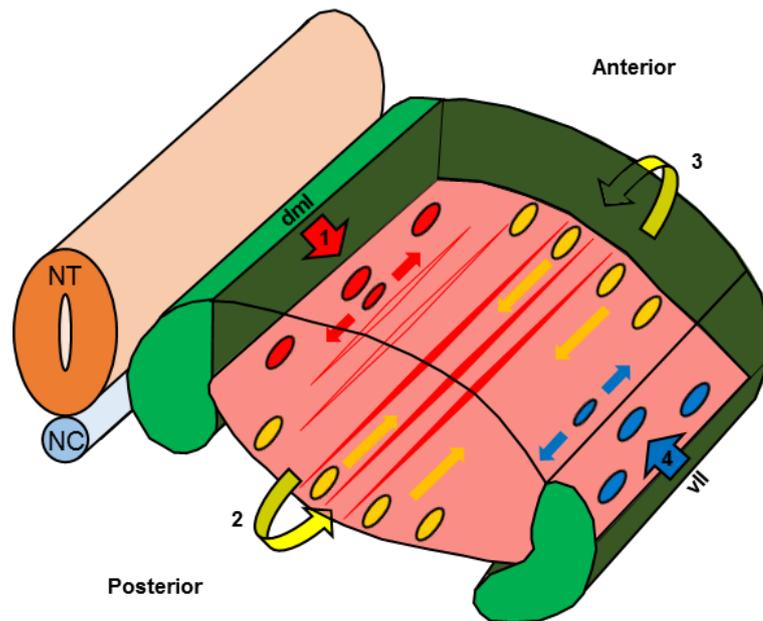
The sclerotome, along with mesenchymal cells from the somitocoel, will contribute to the formation of most of the axial skeleton, ribs, cartilage and connective tissues (syndetome), and bones (vertebrae) (Scaal & Christ 2004; Brent et al. 2003; Baykal & Korkusuz 2016).

The most dorsal part of the somite remains epithelial and is referred to as the dermomyotome (Fig. 1.1b-c). The dermomyotome then further sub-divides into a medial portion, the epaxial domain, and a lateral portion, the hypaxial domain. The epaxial domain will give rise to the axial musculature and skeleton, while the hypaxial domain will contribute to the muscles of the limbs and body walls (Fig. 1.1d). The central part of the dorsal layer of the dermomyotome (also called dermatome) will give rise to the muscles and dermis of the back (Ordahl & Le Douarin 1992; Baykal & Korkusuz 2016). The formation of the sclerotome and the dermomyotome define the dorsoventral axis of the somite.

The myotome forms in between the dermomyotome and the sclerotome (Fig. 1.1b-c), and involves two sequential steps (Fig. 1.2) (Gros et al. 2004). In a first step, myotome expands only from the translocation of dermomyotomal cells through the dorsomedial lip (dml). Through this process, older myocytes are displaced laterally by newer ones arising at the dml, resulting in an incremental myotome growth; cells elongate bidirectionally in the anterior-posterior axis of the embryo to form full-sized myocytes. In a second step, myocytes produced from the posterior, the anterior, and the ventrolateral (vll) borders enter the myotome, in a phase that combines incremental growth at the dml and vll and coherent growth at the posterior and anterior borders. In the process of coherent growth, myocytes elongate unidirectionally and the relative position of the progenitors within the dermomyotome is maintained with their progeny in the myotome (Denetclaw et al. 1997; Ordahl et al. 2001).

The cells originating from the dorsomedial and ventrolateral lips of the dermomyotome contribute exclusively to the epaxial and hypaxial domain of the myotome, respectively, whereas the cells from the anterior and posterior borders populate both mediolateral domains (Gros et al. 2004; Manceau et al. 2008).

At limb levels, cells from the ventrolateral edges of the dermomyotome lose their epithelial characteristics and migrate into the limb buds.



**Fig. 1.2: Formation of the early myotome in the somite.** Cells present at the four epithelial borders of the dermomyotome contribute to the myotome formation at different rates: Cells emanating from the dorsomedial lip (dml) translocate under the dermomyotome and elongate bidirectionally in the anterior-posterior axis of the embryo to form full-sized myocytes (1); then myocytes arise sequentially from the posterior border (2), from the anterior border (3), and finally from the ventrolateral lip (vll) (4). NC: notochord; NT: neural tube. Adapted from Gros *et al.* 2004.

The myotome is the first skeletal muscle structure to be formed during development. The epaxial myotome will give rise to the neck and deep back muscles, while the hypaxial myotome will give rise to the intercostal muscles, body wall muscles, trunk and limb muscles (Fig. 1.1d) (Kaehn et al. 1988; Baykal & Korkusuz 2016).

#### **d. Regulation of somite differentiation**

Although mesoderm is specified in a very early stage of embryogenesis along the anterior-posterior axis, determination of the fate of cells in each somite occurs only when the somite is completely formed (Brand-Saberi et al. 1996).

The differentiation of a somite into sclerotome, dermomyotome and myotome depends on interactions with surrounding tissues and is regulated by extrinsic molecular signals from the dorsal neural tube and surface ectoderm (WNT proteins), the lateral plate mesoderm (bone morphogenetic proteins (BMP)), and notochord and floor plate of the neural tube (Sonic hedgehog homolog (SHH) and Noggin proteins (Christ et al. 2007; Yusuf & Brand-Saberi 2006)).

The ventromedial portion of the somite undergoes an EMT. The notochord and the floor plate of the neural tube produce and secrete SHH and Noggin proteins, which are able to re-programme cells; they will lose their epithelial characteristics (de-epithelialisation) and revert to a mesenchymal-like fate (Fan & Tessier-Lavigne 1994; Brand-Saberi et al. 1993). These mesenchymal cells then lose the expression of N-cadherin and become mobile (Sosic et al. 1997). They migrate ventrally to form the sclerotome, start to express transcription factors, such as Pax1, necessary for differentiation into cartilage, and form the vertebrae and a large part of each rib (Barnes et al. 1996).

The dermomyotome cells in the dorsal part of the somite remain epithelial. This is mainly due to WNT signals coming from dorsal neural tube and surface ectoderm. Formation of the medial half (epaxial dermomyotome) is attributed to Wnt1/3a (dorsal neural tube) (Munsterberg et al. 1995; Ikeya & Takada 1998); and that of the lateral half (hypaxial dermomyotome) is influenced by Wnt6 (ectoderm) (Fan et al. 1997; Dietrich et al. 1997; Schubert et al. 2002). Dermomyotomal cells express Pax3 and Pax7 (Scaal & Christ 2004).

The hypaxial domain is also defined by BMP signals (Bmp4, member of the TGF $\beta$ -superfamily) from the lateral plate mesoderm (Cheng et al. 2004; Pourquié et al. 1996).

Soon after the establishment of the dorsal and ventral compartments of the somite, the myotome starts to form in between these two structures.

The dorsal WNT signals act synergistically with the ventral SHH signal to promote its formation (Dietrich et al. 1997). Based on somite explant studies, it was proposed that Wnt1/3a, from the dorsal neural tube, in combination with low levels of SHH, from the notochord and ventral neural tube, could induce the myogenic differentiation in the epaxial and then hypaxial domain of the dermomyotome (Munsterberg et al. 1995).

Other important players in myotome formation are BMP and Notch signalling. While the medial half of the myotome is under the regulation of WNT and SHH activity to form the epaxial domain, the lateral half receives BMP signals (Bmp4) and Notch signals (Delta1) from the lateral plate mesoderm (Pourquié et al. 1995; Pourquié et al. 1996; Dietrich et al. 1998; Hirst & Marcelle 2015). Pourquié *et al.* showed that Bmp4 is responsible for maintaining the undifferentiated state of prospective hypaxial muscle and hence counteracts the differentiation-inducing activity of the neural tube (WNT signalling) (Pourquié et al. 1996). Limb muscles originate from cells that have migrated as undifferentiated precursors from the somites. Amthor *et al.* showed, by performing bead experiments, that a dose-dependent response of myogenic cells to BMP may spatially coordinate their correct positioning and growth in the limbs (Amthor et al. 1998).

In addition, Marcelle *et al.* observed that Noggin, expressed in the dorsomedial somite, which antagonises BMP activity, could play a role in regulating BMP patterning of the somite (Marcelle et al. 1997). They showed that BMP is required for Wnt1/3a expression in the dorsal neural tube, which in turn can promote muscle differentiation (Munsterberg et al. 1995); on the other hand, they observed that ectopic expression of Bmp4 in the paraxial mesoderm resulted in the inhibition of myotome formation (Marcelle et al. 1997).

WNT proteins induce the expression of myogenic regulatory factors (MRFs), such as MYF5 and MYOD1, indicating the beginning of the myogenesis.

#### **1.1.4. Myogenesis**

##### **a. Cellular heterogeneity: from myogenic progenitor cells to myotube**

Formation of skeletal muscle – myogenesis – is a process allowing differentiation of mesenchymal cells into myoblasts, which proliferate, exit from the cell cycle and fuse together to form multinuclear structures, called myotubes, expressing the characteristic proteins of muscle tissue. Myogenesis starts in the dermomyotome and requires the commitment of a pool of cells into the skeletal muscle lineage (Fig. 1.3).

The dermomyotome is composed of a mixture of dermal and myogenic (MPCs) progenitor cells, which subsequently are going to give rise to the dermatome and the myotome. The first molecular markers characterising myogenic precursors, in the dermomyotome, are the paired-box transcription factors PAX3 and PAX7 (Kassar-Duchossoy et al. 2005; Relaix et al. 2005). Their activation results from WNT signals (Wnt1/3a/6) from the overlying surface ectoderm (Otto et al. 2006). PAX3/7 label proliferating myoblasts in the dermomyotome, where they form a regulatory network with other factors, such as SIX, EYA and DACH proteins, to initiate the myogenesis programme (Heanue et al. 1999). PAX3/7 also support the proliferation and survival of myoblast before differentiation (Buckingham & Relaix 2007), as long as growth factors, especially FGFs, are available.

As the dermomyotome matures, PAX3-expressing myoblasts migrate from the dorsomedial and ventrolateral lips (Galli et al. 2008). Myogenic differentiation starts when myoblasts delaminate from the edges of the dermomyotome and migrate ventrally to form the primary myotome, exclusively composed of post-mitotic myocytes. The myoblasts exit from the cell cycle and start to express the myogenic determination genes, MYF5 and MYOD1 (Ordahl et al. 2001; Gros et al. 2004). This process is associated with the downregulation of PAX3, in part regulated by SHH signals from the notochord and floor plate (Williams & Ordahl 1994; Goulding et al. 1994; Johnson et al. 1994). Once in the myotome, myocytes receive signals from the neural tube; Wnt11 expression in the dml is essential to orient myocyte elongation, all parallel and aligned along the anterior-posterior axis of the embryo (Gros et al. 2009).

The second wave of myogenesis involves fusion of myoblasts. They no longer proliferate and begin to secrete fibronectin in the ECM (Menko & Boettiger 1987).

The signal provided by this adhesion between myoblasts and fibronectin promotes the differentiation of the myoblasts into muscle cells. They align to form chains and fuse to give rise to myotubes destined to become skeletal muscle fibres. When myoblasts become able to fuse, myogenin, a myogenic basic helix-loop-helix (bHLH) protein, starts to be expressed (Bergstrom & Tapscott 2001). The latest step in myotube growth and differentiation involves the increased synthesis of contractile proteins, such as skeletal muscle actin, and myosin heavy and light chains.

At limb level, a fraction of myoblasts from the hypaxial dermomyotome and hypaxial myotome delaminate and migrate into the forming limb buds. These migrating hypaxial myogenic precursor cells express Pax3 and ladybird gene, Lbx1 (Williams & Ordahl 1994; Pourquié et al. 1995; Dietrich et al. 1998); Lbx1 being exclusively expressed in this sub-population of lateral somite cells. When they reach their final destination, they initiate skeletal muscle differentiation programme to generate the future limb muscles (Cinnamon et al. 1999).

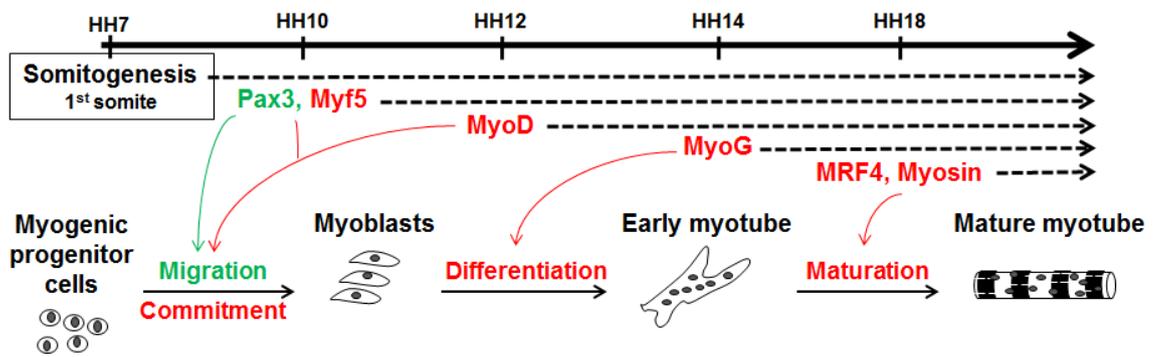
## **b. Genetic networks regulating myogenesis**

### **i) Myogenic regulatory factors**

The formation of skeletal muscle is under the control of the MYOD family of myogenic regulatory factors (MRFs) (Fig. 1.3).

The MRFs are bHLH domain-containing transcription factors. There are four MYOD family members. Myogenic factor 5 (MYF5) and Myogenic differentiation (MYOD) are primary, because they are required for determination of the myoblasts. Myogenin (MYOG) and MYF6 (also called MRF4) are secondary; they regulate terminal differentiation by activating transcription of genes encoding for specific muscle proteins (Bergstrom & Tapscott 2001).

While SHH and WNT signalling pathways are necessary to form the epaxial myotome, they are also required to induce the activation of the primary MRFs, MYF5 and MYOD. SHH, from the notochord and the ventral neural tube (floor plate), and Wnt1/3a, from the dorsal neural tube, contribute to the activation of MYF5. Wnt6, from the surface ectoderm, participate in the activation of MYOD (Munsterberg et al. 1995; Dietrich et al. 1997).



**Fig. 1.3: Regulation of myogenesis during chicken development.** Each step leading the myogenic progenitor cells to differentiate and mature myotubes are controlled by Pax3 and the myogenic regulatory factors (MRFs). Myf5: myogenic factor 5; MyoD: myogenic differentiation; MRF4: myogenic regulatory factor 4 (or Myf6); MyoG: myogenin.

In the dermomyotome, MPCs become committed to the skeletal muscle lineage once they express MYF5 and MYOD. To enter the myotome, these cells need to inhibit the BMP signals. SHH and Wnt1/3a lead to the activation of Noggin in the dorsomedial lip of the dermomyotome, stimulating conversion of dermomyotomal precursor tissue into differentiating myotome. Then the expression of the terminal differentiation genes, required for the fusion of myocytes and the formation of myotubes, is performed by both MRF4 and MYOG (Bentzinger et al. 2012).

In chicken, Myf5 is the earliest marker of determined muscle cells expressed, closely followed by MyoD (Mok & Sweetman 2011; Berti et al. 2015). In mice deficient for Myf5 and MyoD (Myf5<sup>-/-</sup>/MyoD<sup>-/-</sup>), most of the myogenic programme is severely affected with embryos failing to develop any skeletal muscle (Rudnicki et al. 1993). The MyoG knockout has an equally severe phenotype with perinatal death; the expression of several differentiation markers, such as myosin heavy chain and MRF4, appears to be reduced, whereas MyoD levels were normal (Nabeshima et al. 1993). While myoblasts are formed there is a complete absence of functional skeletal muscle supporting the idea that MyoG regulates the later stages of myogenic differentiation, whilst Myf5 and MyoD (and in some case MRF4) are involved in the process of determination (Moncaut et al. 2013).

## ii) PAX-SIX-EYA-DACH network

The PAX-SIX-EYA-DACH (PSED) network, composed of *eyeless*, *eyes absent*, *sine oculis* and *dachshund* genes, has been described for the first time in *Drosophila* where it plays an important role during eye development (Chen et al. 1997; Pignoni et al. 1997).

Vertebrate homologues for these genes have been described and grouped into the PAX (paired-homeobox; PAX1-9), SIX (sine oculis-related homeobox; SIX1-6), EYA (eyes absent-related homeobox; EYA1-4) and DACH (dachshund-related homeobox; DACH1, 2) multigene families (Hanson 2001; Relaix & Buckingham 1999).

In vertebrates, the functions of this gene network are not restricted to eye formation and the PSED network plays key regulatory roles in the development of numerous organs and tissues such as kidney, ear and muscle (Relaix & Buckingham 1999).

Notably, this network leads to the activation of the MRF genes, placing it upstream of the genetic regulatory cascade that directs dermomyotomal progenitors toward the myogenic lineage.

#### PAX transcription factors:

The balance between stem cell self-renewal and progression into a differentiation programme is of critical importance for tissue growth and regeneration.

During skeletal muscle development, each muscle contains a pool of resident stem cells that can either differentiate into muscle fibres or remain as proliferating progenitors. These cells express two related paired-homeobox transcription factors, PAX3 and PAX7, that are essential for ensuring the myogenic potential and survival of progenitors in embryonic (PAX3) and adult (PAX3 and PAX7) myogenesis (Buckingham & Relaix 2007; Buckingham & Relaix 2015).

PAX3 and PAX7 are important regulators of muscle development and are upstream of myogenic genes in somites, limb muscles and satellite cells.

PAX3, initially globally expressed throughout the somite, subsequently becomes restricted to the dermomyotome, and then to the epaxial and hypaxial dermomyotome, is finally downregulated as progenitor cells enter myogenesis. In Pax3 mouse mutant (Pax3<sup>-/-</sup>), somitogenesis is affected, with abnormal myotome formation, trunk muscle defects, and a complete absence of limb muscles (Bober et al. 1994; Goulding et al. 1994). PAX7 is strongly expressed in the central dermomyotome. In Pax7 mutant (Pax7<sup>-/-</sup>), skeletal muscle forms normally in the developing embryo (Mansouri et al. 1996). However, in mice lacking both Pax3 and Pax7 (Pax3<sup>-/-</sup>/Pax7<sup>-/-</sup>), major defects in myogenesis occur, suggesting that together these genes are required for normal muscle development (Relaix et al. 2005). It has been shown in studies performed on mouse mutants and overexpression in chicken embryos, that Pax3/7 activate and control the expression of the MRF genes, such as Myf5 and MyoD (Williams & Ordahl 1994; Bajard et al. 2006).

The expression of PAX3/7 is regulated by the activity of members of the SIX, EYA and DACH families (Heanue et al. 1999).

### SIX, EYA and DACH transcription factors:

SIX homeodomain transcription factors, with EYA and DACH cofactors, have also been implicated in the initiation of myogenesis (Heanue et al. 1999). Similar to PAX3 and PAX7, SIX1/4, EYA1/2/4, and DACH1/2 have been shown to synergistically regulate myogenesis and play a key role in the migration of myogenic precursors. The first indication of an upstream function in myogenesis came from experiments in the chicken embryo where ectopic expression of Six1 and Eya resulted in the activation of Pax3 and the myogenic regulatory genes (Heanue et al. 1999). Since then, analyses in mouse mutants have provided insight into the complex roles of Six, Eya, and Dach.

SIX1 and SIX4 are currently considered to be the apex of the genetics cascade that directs dermomyotomal progenitors toward the myogenic lineage. SIX family proteins are transcription factors characterised by the presence of two conserved domains, a homeodomain (HD) that binds to DNA, and an amino-terminal SIX domain (SD) that interacts with coactivators (EYA) or corepressors (DACH, Groucho (GRO)) of transcription (Kumar 2009).

EYA proteins are characterised by the EYA domain (ED), located in their C-terminal region, responsible for the interaction of EYA with other proteins, including SIX and DACH (Li et al. 2003). Recent works have shown that EYA proteins contain both threonine and tyrosine phosphatase activities, placing them as unique co-transcription factor phosphatases (Sano & Nagata 2011; Rayapureddi et al. 2003; Tootle et al. 2003). It is proposed that this activity inhibits DACH corepressor function. EYA function also involves the recruitment of RNA polymerase II and coactivators such as CREB-binding protein (CBP), or corepressor such as histone deacetylase (HDAC), to the SIX complex (Jemc & Rebay 2007; Spitz et al. 1998; Li et al. 2003). These findings highlight a dual activity for EYA proteins both in the cytosol and in the nucleus.

SIX proteins (SIX1 and SIX4) bind to and translocate EYA proteins (EYA1, EYA2), sometimes associated with DACH proteins, to the nucleus, where they act as cofactors to activate SIX target genes, such as PAX3, MYOD, MRF4, and MYOG (Grifone et al. 2005).

The presence of EYA proteins in these complexes converts SIX and DACH, which are repressors or weak activators of transcription, into strong transcriptional activators.

In mouse, *Six1*, *Six4*, *Eya1*, *Eya2*, *Dach1* and *Dach2*, are expressed in somite, in the dermomyotome and subsequently in Pax3-positive myogenic progenitors.

Unlike Pax3 and Pax7, these factors are also present in differentiated skeletal muscle. In the dermomyotome, *Eya1* and *Eya2* are mainly expressed in the epaxial and hypaxial domains, after the initial onset of epaxial myogenesis.

The critical role of SIX/EYA in myogenesis is revealed by the phenotype of *Six1/Six4* and *Eya1/Eya2* double mutants (Grifone et al. 2005; Grifone et al. 2007), which are more severe than the single mutants, with loss of all muscles derived from the hypaxial dermomyotome, including limb and many trunk muscles. Epaxial myogenesis, leading to the formation of the back muscles, is not affected. In these *Eya* or *Six* double mouse mutants, Pax3 expression is lost in the hypaxial dermomyotome, with an absence of progenitor cell migration and cell death.

A second feature of the double mutants, is a pronounced downregulation of the myogenic regulatory genes (*Myf5*, *MyoD*, and *Mrf4*, but also *MyoG*), observed from the time when *Six/Eya* complex would normally be active (Giordani et al. 2007; Relaix et al. 2013). *Six1/4/Myf5(Mrf4)* mouse mutants do not activate *MyoD* and do not form skeletal muscle in the trunk and limbs (Relaix et al. 2013). This resembles the phenotype of *Pax3/Myf5(Mrf4)* mutants (Tajbakhsh et al. 1997).

These observations suggest that SIX/EYA complex can act upstream of the PAX genes, by regulating PAX3 expression, but also downstream of the PAX genes, by directly targeting and regulating the expression of some of the MRFs.

Similar experiments have been carried out for the DACH proteins, which are classified as negative SIX regulators, although they do not have any identified binding site. Mice with a knockout of *Dach1* or *Dach2* die quickly after birth, but limb development does not seem to be affected (Davis et al. 2001; Davis et al. 2006). This suggests potential overlapping function(s) of the two DACH members in this tissue; or that their activity as repressors of SIX, via EYA, in the SIX/EYA complex, might be only important in some cases.

Myogenesis is a complex process with different levels of regulation. During the last decade, a further layer of complexity has been added with the discovery of microRNAs.

## **1.2. MicroRNAs**

MicroRNAs (miRNAs or miRs) are a class of recently identified small non-coding RNAs that regulate gene expression post-transcriptionally (Bartel 2004; reviewed in Goljanek-Whysall et al. 2012; Ha & Kim 2014). In just two decades, miRNAs have been shown to play important roles in many biological processes, including cell proliferation, apoptosis, and differentiation.

### **1.2.1. Discovery of microRNAs**

The first two miRNAs, *lin-4* and *let-7*, were originally identified in the nematode *Caenorhabditis elegans* (*C. elegans*) as small non-coding RNAs required for the temporal regulation of larval development (Lee et al. 1993; Reinhart et al. 2000).

*lin-4* was proved not to encode for a protein but to encode for a 22-nucleotide non-coding RNA that is partially complementary to a conserved site located in the 3' untranslated region (UTR) of the *lin-14* messenger RNA (mRNA) (Lee et al. 1993; Wightman et al. 1991). When this complex is formed, *lin-14* is downregulated to allow the developmental transition from the first to the second larval stage (Ruvkun & Giusto 1989).

The discovery of *lin-4* and its target-specific translational inhibition activity has highlighted a new mechanism of gene regulation during the development. Few years later, a second miRNA, *let-7*, was discovered (Reinhart et al. 2000).

*Let-7* encodes a temporally regulated 21-nucleotide small RNA that controls the transition from the fourth stage to the adult stage during *C. elegans* development (Reinhart et al. 2000). Similar to *lin-4*, *let-7* performs its function by binding to the 3'UTR of *lin-14*, but also *lin-28*, *lin-41* and *lin-57*, and inhibits their translation (Vella et al. 2004; Abrahante et al. 2003).

This finding suggested not only that miRNAs had the capability of controlling developmental timing, but also that their biological function was an efficient, fast and cell-economic way of post-transcriptional regulation, since one miRNA could inhibit a variety of independent genes (Reinhart et al. 2000).

The identification of let-7 not only provided another example of developmental regulation by small RNAs, but also raised the possibility that such RNAs might be present in species other than nematode. While lin-4 appeared to be specific to worm, both let-7 and let-41 are evolutionary conserved with homologous detected from worm to human (Pasquinelli et al. 2000). This extensive conservation strongly indicated a more general role for these small RNAs in developmental regulation (He & Hannon 2004; Lagos-Quintana et al. 2001).

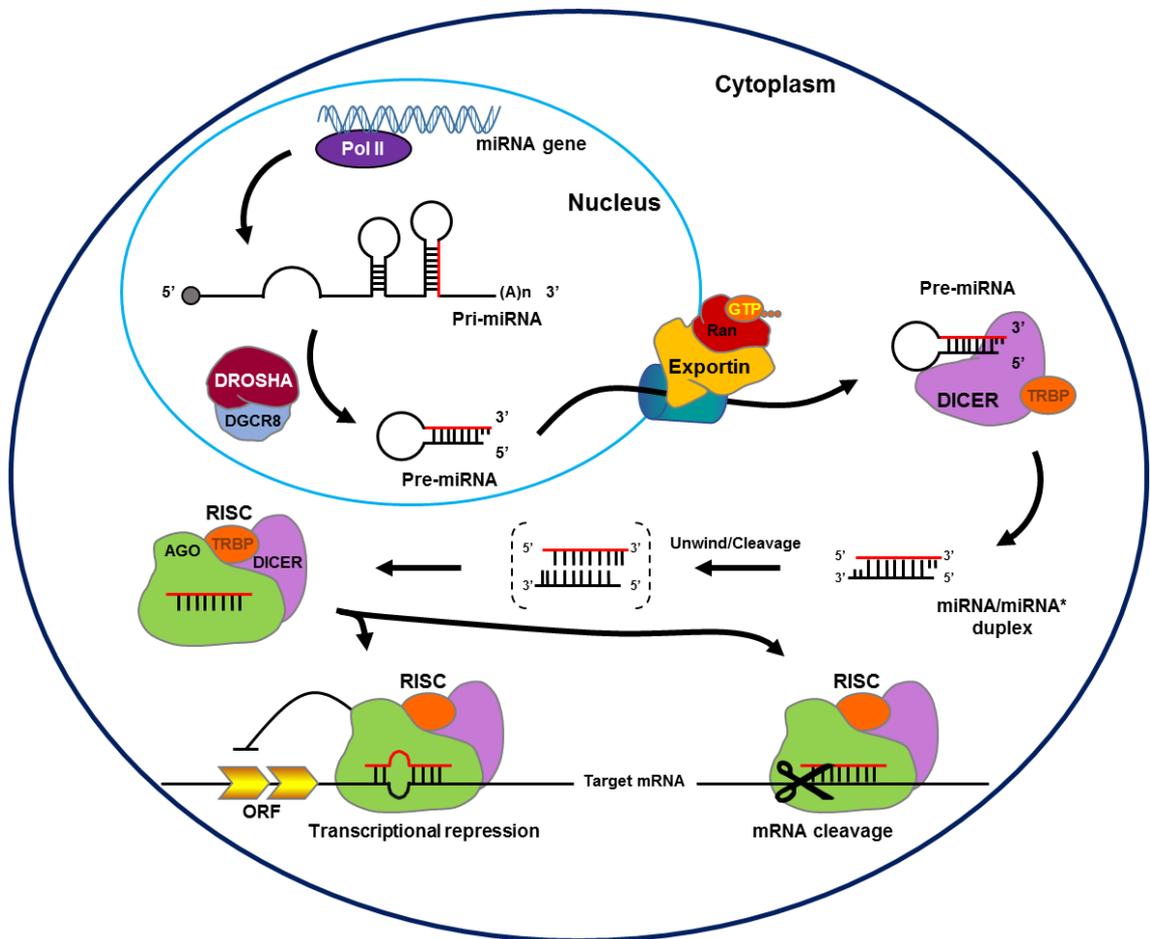
Since then, hundreds of miRNAs have been identified. Their discovery adds a new dimension to our understanding of large and complex gene regulatory networks.

### **1.2.2. MicroRNA biogenesis**

#### **a. MicroRNA genes**

MicroRNA genes constitute one of the most abundant gene families, and are widely distributed in animals, plants, protists and viruses (Griffiths-Jones et al. 2008).

MiRNA genes are located within various genetic context. In animals, most of the known miRNAs are located in introns of protein-coding genes and long non-coding transcripts; a small percentage of miRNAs are encoded by exonic regions (Rodriguez 2004). Often, several miRNA loci are in close proximity to each other, constituting a polycistronic transcription unit (Lee et al. 2002). The miRNAs in the same cluster are generally co-transcribed and have similar tissue expression profiles (coordinated cluster), however some clusters contain miRNAs with uncoordinated expression profiles. The predicted transcription start sites of such clusters being exclusively located upstream of the first miRNA this means that these clusters will be transcribed as single transcription units. The difference of tissue expression profiles of uncoordinated miRNAs suggests a post-transcriptional regulation of this processing.



**Fig. 1.4: The biogenesis and function of miRNAs.** MiRNA genes are transcribed by RNA polymerase II (Pol II) to form a capped and polyadenylated primary transcript (pri-miRNA). Pri-miRNA is cleaved by DROSHA/DGCR8 to form a hairpin-like precursor miRNA (pre-miRNA). Pre-miRNA is exported to the cytoplasm by Exportin 5 where it is further cleaved by DICER/TRBP to form a duplex that contains guide and passenger strands. One of the strands is then incorporated into RNA-induced silencing complex (RISC). MiRNA-loaded RISC can then recognise its target mRNA leading to transcriptional inhibition and/or degradation of the mRNA. ORF: open reading frame.

In addition, the precise location of their promoters have not yet been mapped for most miRNA genes. For some, which reside in the introns of protein-coding genes, they can share the promoter of their host gene. However, it has been shown that about one-third of intronic miRNAs are transcribed independently of their host gene (Monteys et al. 2010).

### **b. Production: from gene to mature microRNA**

The different steps involved in the biogenesis of mature miRNAs are summarised in Fig. 1.4 (also see Ha & Kim 2014; Gargalionis & Basdra 2013; Kim et al. 2016).

In animals, miRNAs are initially transcribed in the nucleus as long, capped and polyadenylated primary transcripts (pri-miRNAs), by RNA polymerase II (Lee et al. 2004; Cai et al. 2004).

These long non-coding pri-miRNAs, which are composed of one or several hairpin stem loop structures, where the double-stranded portion is partially complementary, are incorporated into the microprocessor complex. This complex contains a nuclear RNase III enzyme, called Drosha, and its co-factor, a protein encoded by DiGeorge Critical Region 8 (DGCR8). After cleavage by Drosha, 60-80 nucleotide long imperfect stem-loop or hairpin structures with a 3' two-nucleotide overhang, called precursor microRNAs (pre-miRNAs) are released (Lee et al. 2003; Gregory et al. 2004; Denli et al. 2004; Han et al. 2004). The basal junction, between single-stranded RNA and double-stranded RNA, has been shown to be a major reference point in determining the cleavage site (Han et al. 2006; Zeng et al. 2005); so is the apical junction, linked to the terminal loop of these pri-miRNAs, for optimal efficiency and accuracy of Drosha processing (Zeng et al. 2005; Ma et al. 2013).

The efficiency of Drosha-mediated processing is crucial for determining miRNA abundance. Multiple mechanisms exist to control the expression level, activity and specificity of Drosha.

For example, Drosha and DGCR8 auto-regulate each other; DGCR8 stabilises Drosha through protein-protein interactions, whereas Drosha destabilises DGCR8 mRNA by cleaving it at a hairpin in the second exon.

Post-translational modification can also regulate the protein stability, nuclear localisation and processing activity of the microprocessor complex (phosphorylation and acetylation of Drosha; phosphorylation and deacetylation of DGCR8). Also, it has been shown that Drosha-mediated processing can be controlled specifically by RNA-binding proteins that selectively interact with Drosha and/or certain pri-miRNAs (Finnegan & Pasquinelli 2013; Tran & Hutvagner 2013; Ha & Kim 2014).

Another source of pre-miRNAs is derived via the RNA splicing machinery of the cell. These pre-miRNAs are directly spliced out of small introns, bypassing the microprocessor complex (Drosha/DGCR8) (Berezikov et al. 2007; Westholm & Lai 2010; Yang & Lai 2011). These Drosha-independent miRNAs, known as 'mirtrons', constitute a class of non-canonical miRNAs, as opposed to the canonical miRNA class (most of the miRNAs), which are Drosha and Dicer-dependent (Abdelfattah et al. 2015).

The released hairpin-shaped pre-miRNAs, are being transported through the nuclear pore complexes and into the cytoplasm by Exportin 5 (EXP5), a Ran-guanosine-5'-triphosphate (GTP)-dependent nuclear transport receptor (Lund et al. 2004; Yi et al. 2003; Bohnsack et al. 2004). Hydrolysis of GTP is necessary to free the pre-miRNAs.

In the cytoplasm, pre-miRNA hairpins are cleaved by the endonuclease cytoplasmic RNase III enzyme, Dicer, which form a complex with its co-factor, the transactivation response cytoplasmic RNA-binding protein (TRBP) (Bernstein et al. 2001; Chendrimada et al. 2005; Lund & Dahlberg 2006). This endonuclease interacts with the two-nucleotide overhang at the 3' end of the hairpins, generated after Drosha cleavage, and cuts about 22 nucleotides away the loop joining 3' and 5' arms (Lee et al. 2003).

The imperfect double-stranded RNA duplexes generated, miRNA/miRNA\*, are 20-25 nucleotides in length with 3' two- nucleotide overhangs at both ends (Hutvagner et al. 2001; Zhang et al. 2002; H. Zhang et al. 2004); they are composed of a passenger strand/miRNA\*, and a guide strand/miRNA. Although mature miRNAs can reside on either strand of the hairpin stem, strand selection is dictated mainly by the relative thermodynamic stability of the two ends of the duplex: the strand whose 5' terminal nucleotides are less stable is most likely to be selected as mature miRNA (Schwarz et al. 2003; Khvorova et al. 2003). MiRNA originating from the 5' and 3' strands of pre-miRNA are referred to as 5p and 3p miRNAs, respectively.

Regulation of the precursor to mature step can occur in a variety of ways. Specific precursor miRNAs can be detained in the nucleus to prevent maturation in certain cell types. Precursor maturation can also be regulated by affecting Dicer levels and activity. For example, Ma *et al.* shown that human Dicer protein is able to negatively regulate its own catalytic activity through its helicase domain and that this auto-inhibitory effect could be modulated by binding of Dicer cofactors to the helicase domain (Ma et al. 2008).

Dicer activity can also be altered through many protein interactions. One Dicer interactor that increases cleavage efficiency is the TRBP. Interaction with TRBP modulates the processing efficiency of some pre-miRNAs and tunes the length of mature miRNAs. TRBP can be phosphorylated by the mitogen-activated protein kinase extracellular signal-regulated kinase (MAPK ERK) leading to upregulation of growth-promoting miRNAs; the reduction of the TRBP protein leads to the destabilisation of the Dicer protein and to the decrease of miRNA levels (Finnegan & Pasquinelli 2013; Ha & Kim 2014).

The relative instability of the guide strand, mature miRNA, facilitates its preferential incorporation into the RNA-Induced Silencing Complex (RISC). In order to be further processed, the two stands of the duplexes will have to be unwound; however, it is not clear whether this unwinding is happening before or after incorporation into RISC. Two models have been proposed. The first one, called ‘helicase model’, suggests that the unwinding occurs before incorporation and implicates direct activity of a putative RNA helicase; however, to date, this ‘unwindase’ remains unidentified and the model has not been validated. The second model, called ‘duplex-loading model’, suggests the incorporation of miRNA duplexes, as double-stranded, into RISC. Recent works tend to support this latest model by having demonstrated the presence of Ago proteins into RISC able to separate the two strands and degrade the passenger one, before further processing (Diederichs & Haber 2007; Kawamata & Tomari 2010).

RISC contains Argonaute (AGO), a multi-functional catalytic protein, Dicer, and TRBP, responsible for recruiting Dicer to AGO (Hammond 2001; Chendrimada et al. 2005). AGO contains two conserved RNA binding domains, a PAZ domain and a PIWI domain that can bind the single-stranded 3’ end and 5’ end, respectively (Pratt & MacRae 2009). AGO is needed for miRNA-induced silencing; it binds mature miRNAs and orients them for interaction with their target mRNAs.

AGO proteins can be modulated by numerous modifications. For example, hydroxylation of human AGO2 increases its stability or localisation within processing bodies; these processing bodies are cytoplasmic loci involved in mRNA turnover and RNA silencing, but also thought to be sites for translational suppression and/or mRNA decay. Phosphorylation of AGO2 has also been reported to be mediated by MAPK-activated protein kinase 2 (MAPKAPK2) or RAC $\gamma$  Serine/Threonine protein kinase (AKT3), resulting in its localisation to processing bodies or translational repression, respectively (Finnegan & Pasquinelli 2013; Ha & Kim 2014).

### **1.2.3. RNA silencing**

The mature miRNAs that have been loaded into RISC, target complementary sequences in mRNA 3'UTR using classic Watson-Crick base-pairing but also rely on the thermodynamic favourability of such interaction (minimal folding free energy) and the site accessibility (Bartel 2009; Peterson et al. 2014; Akhtar et al. 2016). The mRNA downregulation takes place through two main functional pathways, either mRNA translational repression or the mRNA cleavage (Doench & Sharp 2004; Filipowicz et al. 2008).

Mature miRNAs have a specific sequence at position 2-7 from the 5' end, called 'seed' sequence (Lewis et al. 2005; Grimson et al. 2007). When the same seed sequence is found in more than one miRNA they form a 'family'. Often miRNA/target interaction algorithms will predict the same targets for these miRNAs, however as seed sequence complementarity is only one of the criteria important for miRNA targeting, being members of the same family does not always mean that they will have the same real target(s).

Each miRNA can target several hundred mRNAs. In addition, because of their short seed sequences, multiple miRNAs can simultaneously regulate the expression of a specific mRNA, by targeting different sites on its 3'UTR (Selbach et al. 2008; Bartel 2009; Friedman et al. 2009). It is estimated that miRNAs may regulate over 60% of transcripts in humans (Friedman et al. 2009).

The miRNA/RISC complex (miRISC) uses the seed sequence to find target sequences, called 'miRNA response elements' (MREs), usually localised in the 3'UTRs of mRNAs (Bartel 2009).

The miRISC binds to mRNA(s) depending on the degree of complementarity between seed and target sequences. In plants, miRNAs often have targets with perfect, or near-perfect complementarity causing, in most of the cases, target mRNA degradation (Llave et al. 2002; Rhoades et al. 2002). Although perfect complementarity can happen in animals, it is more common to encounter miRNAs which bind to their targets with partial complementarity, leading to inhibition of protein synthesis by translation repression and/or mRNA instability (Wahid et al. 2010; Axtell et al. 2011).

Although the pairing to the seed region is often sufficient for functional binding specificity, it has been suggested that some sites in the remainder of the miRNA sequence might also be involved, contributing to enhance binding specificity and affinity (Doench & Sharp 2004), however, this process is not fully understood and will need to be further investigated.

Recent investigations have also provided evidence that miRNAs can act at different sites, including 5'UTRs, promoters and coding regions (Lee et al. 2009; Place et al. 2008; Qin et al. 2010; Tay et al. 2008; Forman et al. 2008; Lee 2014).

For example, Place *et al.* found a putative miR-373 target site in the promoter of E-cadherin, and showed that miR-373 induces E-cadherin expression (Place et al. 2008). These findings reveal a new mode by which miRNAs, in some cases, can up-regulate gene expression. This new miRNA function, called 'RNA activation' (RNAa), is a relatively poorly characterised phenomenon, compared to RNA inhibition; however, it seems likely conserved across mammals including rat, mouse, and some primates (Huang et al. 2010; Lee 2014).

#### **1.2.4. MicroRNAs and myogenesis**

##### **a. Dicer-dependent myogenesis**

Since their discovery in *C. elegans*, emerging evidences have highlighted key roles miRNAs play during development.

This was first demonstrated by Bernstein *et al.* who knocked-out Dicer in mice in order to prevent the processing of pre-miRNAs into functional mature miRNAs. They found that a complete loss of Dicer in mice results in early embryonic lethality (E7.5), with development likely halting in gastrulation (Bernstein et al. 2003).

Furthermore, a more targeted approach has demonstrated that miRNAs are not all ubiquitous, like let-7; many miRNAs have been shown to be expressed in a tissue-specific manner (Lee & Ambros 2001; Lagos-Quintana et al. 2002). Further investigation using tissue-specific Dicer deletion revealed that miRNAs are required for skeletal (O'Rourke et al. 2007) and cardiac (Chen et al. 2006; Zhao et al. 2007) muscle development. Tissue-specific Dicer knock-out mice were generated and showed that Dicer knock-out led to a significant decrease in muscle mass, a lower number of myofibres and abnormal myofibre morphology, as well as increased apoptosis of myogenic cells and enhanced cell death in myoblasts (O'Rourke et al. 2007). Similar results have been observed in zebrafish with loss-of-function mutation in Dicer (Mishima et al. 2009). These studies provide convincing genetic evidence for the essential role of miRNAs in muscle development and function.

#### **b. Muscle-specific microRNAs: myomiRs**

With respect to skeletal muscle, miRNAs can be divided into two categories: miRNAs that are exclusively or preferentially expressed in muscle, the myomiRs (McCarthy 2008); and miRNAs expressed exclusively in non-muscle tissue or broadly expressed across many cell types. Both categories have significant impacts on muscle proliferation and differentiation (Wang 2013).

The myomiR group, which initially was composed of miR-1, miR-133a and 133b, and miR-206, has recently expanded to include miR-208a and 208b, miR-486 and miR-499 (Lagos-Quintana et al. 2002; McCarthy & Esser 2007; McCarthy 2008; van Rooij et al. 2007; van Rooij et al. 2009; Small et al. 2010). With exception of miR-206 and 208a, most myomiRs are expressed in both cardiac and skeletal muscles.

MiR-206 is expressed specifically in skeletal muscle, in somites (Sempere et al. 2004; Sweetman et al. 2008), while miR-208a is reported to be expressed predominantly in cardiac muscles (van Rooij et al. 2007).

Some studies have reported evidence that not all myomiRs are solely expressed in a muscle-specific manner but may be detected in low levels in other tissues; however, their main function is still confined to muscle. For example, miR-486 is sometimes considered ‘muscle-enriched’ rather than ‘muscle-specific’ as it is also expressed in other tissues (Small et al. 2010; Lee et al. 2008).

### **MiR-1, miR-133 and miR-206:**

MiR-1, -133, and -206 were the first miRNAs described as myomiRs. They have been studied extensively and their roles in the regulation of the myogenic programme have been well established (reviewed in Horak et al. 2016).

These miRNAs are members of the miR-1/206 and the miR-133 families. They are organised in bicistronic clusters on the same chromosome (miR-1-1/miR-133a-2 on chromosome 20 in human and chicken, miR-1-2/miR-133a-1 on chromosome 18 in human and 2 in chicken, and miR-206/miR-133b on chromosome 6 in human and 3 in chicken), and are generally transcribed together (Nohata et al. 2012). They produce very similar mature miRNAs.

In chicken, the sequence of mature miR-1a-1 is identical to miR-1a-2, while the sequence of miR-133a-1 is identical to miR-133a-2; Gga-miR-133b differs from these by only a single nucleotide at the 3’ end. They share the same seed sequence, respectively.

MiR-1b differs from miR-1a-1/1a-2 by only 1 nucleotide; miR-1a and miR-1b differ from miR-206 in 3 and 4 nucleotides, respectively, in the 3’ region while sharing the same seed sequence.

In skeletal muscle, miR-1/206 and miR-133, play important roles in proliferation, differentiation, and cell fate specification (van Rooij et al. 2008). They are up-regulated during the early stages of muscle differentiation in both cell culture models (Chen et al. 2006; Kim et al. 2006) and in developing embryos (Wienholds et al. 2005; Sweetman et al. 2006; Darnell et al. 2006).

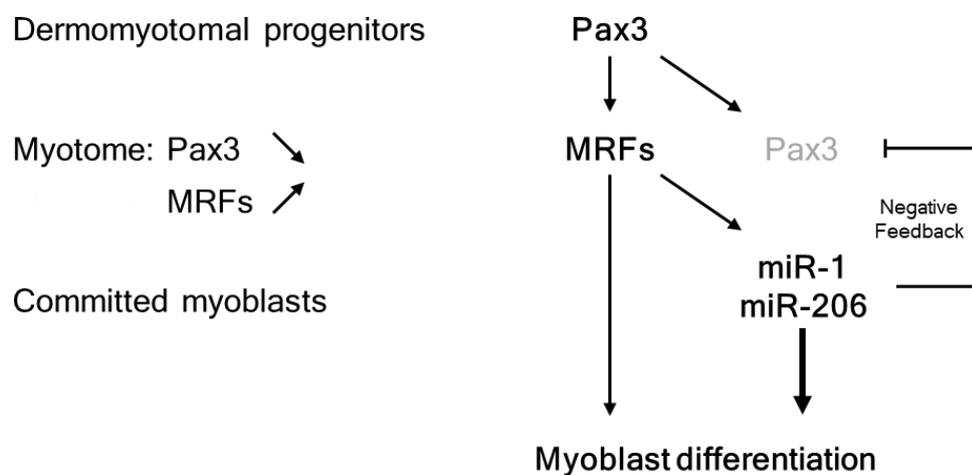
In zebrafish embryos, loss of miR-1 and miR-133 leads to the disorganisation of muscle segments (sarcomeres) and muscle gene expression (sarcomeric actin) (Mishima et al. 2009).

Although deletion of miR-1 in mice leads to heart defects (ventricular septal defects), skeletal muscles appear to form normally (Zhao et al. 2007). Similarly, mice lacking miR-206 do not display an overt muscle phenotype (Williams et al. 2009). One possible explanation could be the overlap in target genes among myomiR family members. This idea is supported by the double knock-out of miR-133a-1 and miR-133a-2 in which mice showed heart defects (~50% of lethality) and skeletal muscle myopathy that was not present in the single miR-133a knock-out mice (Liu et al. 2008; Liu et al. 2011). In chicken embryos, myogenesis is delayed after inhibition of miR-206 with absence of complete downregulation of Pax3, which is an important requirement for muscle differentiation (Goljanek-Whysall et al. 2011).

At the onset of myogenic differentiation, muscle gene expression is regulated by serum response factor (SRF), myocyte enhancer factor 2 (MEF2) and the MRFs, including MYOD1, MYF5, MRF4 and MYOG (Chen et al. 2006; Liu et al. 2007; Ge & Chen 2011). In skeletal muscle, SRF and MEF2 cooperate with MYOD1 and MYOG to transcriptionally activate the expression of the three pairs of muscle-specific miRNAs: miR-1-1/miR-133a-2, miR-1-2/miR-133a-1 and miR-206/133b (Rao et al. 2006; Sweetman et al. 2006; Sweetman et al. 2008).

As a consequence, several major changes in miRNA levels have been observed, especially a significant increase in miR-1 and miR-206 levels during C2C12 myoblast differentiation *in vitro* (Chen et al. 2006; Kim et al. 2006). The elevated expression of miR-1a and miR-206 results in promoted differentiation and in the blockade of proliferation due to the anti-proliferative effects of both miRNAs.

Along with miR-1/206, the expression of miR-133a and miR-133b is massively induced during myogenesis (Koutsoulidou et al. 2011). Their expression levels, like the levels of miR-1 and miR-206, correspond to the capacity of myoblasts to form myotubes. Chen *et al.* provided the evidence that miR-133 plays an opposing role, compared to miR-1/206, in skeletal myogenesis (Chen et al. 2006). They showed that an overexpression of miR-133 was able to repress the expression of MyoG and Myosin heavy chain (MHC) and promote myoblast proliferation; potentially by repressing SRF. However, a contradictory study reported on miR-133 participation in the suppression of myoblast proliferation and promotion of differentiation (Kim et al. 2006). These data suggest that miR-133 might control both cell proliferation and differentiation in a context-dependent manner.



**Fig. 1.5: MiR-1 and miR-206 contribute to the downregulation of Pax3 during myoblast differentiation.** Pax3 is strongly expressed in the dermomyotomal progenitors. As the somite develops and differentiates, Pax3 is downregulated and MRFs are upregulated. In committed myoblasts MRFs activate miR-1/miR-206 which are then able to target residual Pax3 transcripts (in grey; negative feedback). The complete silencing of Pax3 confers robustness to developmental timing of differentiation. Adapted from Goljanek-Whysall et al. 2011.

Recent studies provided new evidence that miR-1 and miR-206 play a major role in myoblast differentiation by regulation of multiple target genes.

Notably, inhibition of endogenous miR-1 and miR-206 was shown to block the downregulation of most targets in differentiation cells (Gagan et al. 2012), thus indicating that miRNA activity and target interaction is required for muscle differentiation (Goljanek-Whysall et al. 2012). Goljanek-Whysall *et al.* showed that miR-1 and miR-206 play a major role during myogenesis; they confer robustness to timing of myogenesis by regulating the transition from myogenic progenitor cells to committed myoblast. MiR-1 and miR-206 contribute to the downregulation of Pax3 as committed myoblasts start to differentiate; this step being essential to initiate the myogenic programme (Fig. 1.5) (Goljanek-Whysall et al. 2011; Goljanek-Whysall et al. 2012).

### **MiR-208a/b, miR-486, and miR-499:**

Newly identified and recently classified as myomiRs, miR-208a/b, miR-486 and miR-499, are less characterised than miR-1/206 and miR-133 (reviewed in Kirby et al. 2016; Horak et al. 2016). These new members of the myomiR family are monocistronic and located in protein coding genes.

Northern blot analyses showed that they are either strictly striated muscle-specific (miR-208a/b, miR-499), being derived from the intron of muscle-specific myosin heavy chain MYH6, MYH7 and MYH7B genes, respectively; or highly enriched in muscle (miR-486) and encoded in the intronic region of ANK1 (ankyrin 1) gene (van Rooij et al. 2007; van Rooij et al. 2009; Small et al. 2010). van Rooij *et al.* discovered that heart-specific miR-208a, co-expressed from MYH6 gene, encoding fast myosin, is essential for the upregulation of slow myosin MYH7 and miR-208b in the adult heart. They also described that miR-208a regulates the expression of MYH7B, another slow myosin, and its intronic miR-499 (van Rooij et al. 2009). Although miR-486 does not have a muscle-specific expression, it is involved in important skeletal muscle development processes. It has been reported that miR-486, highly upregulated during muscle differentiation, directly targets PAX7 and subsequently accelerate myoblast differentiation (Dey et al. 2011).

Generally, some miRNAs are seen as playing key roles during myogenesis, e.g. miR-1/206 or miR-133, while others are playing more subtle roles, including miR-208a/b which influences muscle performance by myosin switching.

The dysregulation of myomiRs has been reported to be associated with various skeletal muscle pathologies, in several types of cancers, muscle atrophy, myopathies and also in sarcopenia (age-related muscle wasting) (reviewed in Kirby et al. 2016).

Rhabdomyosarcoma (RMS), a type of soft tissue sarcoma, is derived from skeletal muscle progenitor cells that maintain a proliferative capacity by poorly differentiate. A dramatic decrease in miR-1, miR-206 and miR-133a/b expression was reported in RMS cell line and human RMS samples, resulting in the upregulation of the oncogene c-Met, a validated target of miR-1/206 (Yan et al. 2009; Rao et al. 2010). Several non-myomiRs have also been reported to regulate tumour formation in RMS cell lines, such as miR-29, which promotes myogenic differentiation (H. Wang et al. 2008).

### **1.2.5. Identification of novel microRNAs**

Although some miRNAs have been characterised in detail, in most cases there is only limited information about their function(s). This is in part due to the incomplete complementarity of miRNAs with their targets, which makes the identification of biologically relevant targets more challenging (Bartel 2009; Friedman et al. 2009).

MiRNA identification is complicated and requires an interdisciplinary strategy (reviewed in Gomes et al. 2013; Akhtar et al. 2016). Different approaches have been adapted to identify putative miRNAs over the years. The traditional experimental method used to discover miRNAs consisted of cloning size-fractionated RNA followed by Next Generation sequencing and experimental validation, then bioinformatics tools were used to locate their origin in the genome and to assess structural requirements for miRNA biogenesis. Other experimental approaches have been used to investigate and validate new miRNAs, such as Northern blot and *in situ* hybridisation.

In recent years, biological and bioinformatics approaches have enabled discovery of thousands of miRNAs in plant, animals, unicellular eukaryotes, and viruses. To date, about 30,000 mature miRNAs have been discovered and collated in miRBase, the main online repository of miRNA sequences and annotation (release 21 – June 2014) (Kozomara & Griffiths-Jones 2014); a large number of these predicted miRNAs have not been investigated yet and still need to be validated by experiments as real miRNAs.

The current release has catalogued 434 mature miRNAs in *C. elegans*, 466 in *Drosophila*, and 994 in chicken. These figures, significantly lower than those for mouse and human, 1915 and 2588, respectively, have considerably increased in just a few years, suggesting that more miRNAs, especially chicken miRNAs, have still to be discovered.

Recent technological advances like high-throughput sequencing have made possible the determination of miRNA tissue distribution (Rathjen et al. 2009; Milagro et al. 2013).

It is particularly of interest for the identification of new miRNAs that are enriched in specific tissues, like in skeletal muscle. For example, Rathjen *et al.* used high-throughput Solexa sequencing of short RNA libraries generated from chicken developing somites and identified new variants of known miRNAs (isomiRs) (Morin et al. 2008), but also potential novel skeletal muscle-specific miRNAs (Rathjen et al. 2009).

### **1.3. Research aims and objectives**

Multiple miRNAs have been identified by combining biological and bioinformatics approaches, with some of them shown to be involved in skeletal muscle development and differentiation. However, the precise roles of most of these miRNAs and how they act to regulate these processes remain to be identified.

This project was developed in order to get a better understanding of the mechanisms underlying interactions between miRNAs and their mRNA targets.

We focused our investigation on miR-128, and looked at the interaction with one of its candidate targets, *Eya4*. We also assessed the potential impact of this interaction on skeletal myogenesis in the chicken embryo.

The hypothesis and specific aims of the project were:

**Hypothesis:** MiR-128 regulates the early myogenic differentiation by targeting *Eya4*, a member of the PAX-SIX-EYA-DACH (PSED) network.

**Specific aims:**

- 1- Characterise miRNAs (specifically) expressed in skeletal muscle by LNA *in situ* hybridisation (chapter 3.1)
- 2- Identify candidate target mRNAs by using bioinformatics tools (chapter 3.2)
- 3- Characterise miRNA targets by molecular cloning and RNA *in situ* hybridisation (chapter 4.1)
- 4- Investigate miR/mRNA interactions *in vitro* by luciferase reporter assays (chapter 4.2)
- 5- Investigate miR/mRNA interactions *in vivo* by functional experiments (chapter 5)

## **CHAPTER 2: MATERIALS AND METHODS**

### **2.1. Culture of chicken eggs**

White leghorn chicken eggs (Henry Stewart & Co Ltd, UK) were stored at 17°C prior to incubation at 38-39°C for the required period of time.

### **2.2. Embryo staging, harvesting and fixation**

Embryos were incubated in 38-39°C humidified incubator until they reached the desired stages, according to the Hamburger and Hamilton table (Hamburger & Hamilton 1992).

Embryos were harvested from eggs by cutting away connecting tissues with fine scissors and removing the embryos using forceps or a spoon. The embryos were placed in a Petri dish, containing 1X DEPC-PBS (Phosphate-buffered saline), and dissected free of membranes and other attached tissues.

The embryos were then fixed overnight at 4°C in 4% (w/v) paraformaldehyde (PFA)/PBS.

### **2.3. Embryo dehydration**

Embryos were dehydrated by washing twice in 1X PBT (DEPC-PBS with 0.1% Tween-20), 50% (v/v) methanol, and 100% methanol, at least 5 minutes each wash. Once dehydrated, embryos were stored in fresh methanol at -20°C.

### **2.4. Embryo rehydration**

Embryos were rehydrated by washing them in 100% methanol, then 75%, 50% and 25% (v/v) methanol/PBT, at least 5 minutes each wash, followed by two additional washes in 1X PBT.

## **2.5. Whole-mount *in situ* hybridisation (WMISH)**

### **2.5.1. LNA probe pre-absorption**

In order to increase efficiency and reduce non-specific binding, LNA probes were pre-absorbed on late stage embryos (HH27-HH30).

Following rehydration, embryos were treated with 30-35 µg/mL Proteinase K for 30 minutes. Embryos were then washed twice in 1X PBT and fixed in 4% (w/v) PFA with 0.1% (w/v) glutaraldehyde for 20 minutes and rinsed in 1X PBT for 5 minutes.

After removing the 1X PBT, Hybridisation Buffer, pre-heated at 48°C, was mixed 1:1 with 1X PBT and added to the embryos; they were then allowed to settle and transferred to fresh pre-heated Hybridisation Buffer. The Hybridisation Buffer was replaced and embryos were incubated for at least 2.5 hours at 48°C. Embryos were incubated overnight in Hybridisation Buffer with probe (Exiqon) at 48°C using a rocking tray; LNA probes were used at 20 nM. Next day, the probe-containing solution was collected and stored at -20°C for the next use.

With LNA probes (miRCURY LNA™, Exiqon), the temperature of hybridisation is usually 20°C lower than their Melting Temperature (T<sub>m</sub>). For all the LNA probes used in this project, the temperature of incubation was 48°C.

This pre-absorption step needs to be repeated 4 to 6 times prior to first use.

### **2.5.2. *In situ* hybridisation protocol**

To establish the pattern of expression of genes and miRNAs of interest in chicken embryos, RNA *in situ* hybridisations (ISH) and LNA ISH were performed, respectively.

Following rehydration, embryos from stage HH10 to HH22 were treated with 5-35 µg/mL Proteinase K for 30 minutes, depending on the stage of development. Embryos were then washed twice in 1X PBT and fixed in 4% PFA with 0.1% (w/v) glutaraldehyde for 20 minutes and rinsed in 1X PBT for 5 minutes.

After removing the 1X PBT, Hybridisation Buffer, pre-heated at 48°C for LNA ISH or 65°C for RNA ISH, was mixed 1:1 with 1X PBT and added to the embryos; embryos were allowed to settle and transferred to fresh pre-heated Hybridisation Buffer.

The Hybridisation Buffer was then replaced and embryos were incubated for at least 2.5 hours at 48°C or 65°C. After pre-hybridisation step, the buffer was removed and replaced with fresh Hybridisation Buffer containing either LNA probe (20 nM; Exiqon) or RNA probe (0.2-0.4 ng/μL), and incubated at 48°C or 65°C, respectively, overnight using a rocking tray.

The next day, the probe-containing Hybridisation Buffer was recovered and stored, and any unwound probe was removed from the sample by doing two quick washes in Hybridisation Buffer and followed by a 10-minute wash in fresh Hybridisation Buffer. The embryos were washed four times in Wash Buffer for 30 minutes each wash, and then in Wash Buffer mixed 1:1 with 1X MABT for 10 minutes. Embryos were rinsed three times in 1X MABT, and then twice in 1X MABT, 30 minutes each at room temperature. Non-specific protein interactions were blocked by incubating embryos in Blocking solution (2% (w/v) BBR, 1X MABT) for 1 hour and in Blocking solution with 20% (v/v) goat serum for two hours at room temperature. Goat serum was previously heat inactivated at 55°C for 30 minutes. Antibody incubation was done overnight with anti-Digoxigenin (DIG) (1:2000), or anti-fluorescein isothiocyanate (FITC) (1:5000), in the Antibody solution (2% (v/v) BBR-MABT-20% (v/v) goat serum).

Excess antibody was removed by doing six washes of 1X MABT, at least 30 minutes each wash, at room temperature. Colour reaction was performed after washing embryos in freshly prepared NTMT Buffer twice for 10 minutes at room temperature. Detection was performed by using NBT/BCIP substrates in Alkaline Phosphatase/NTMT Buffer for probes conjugated to DIG, or in Fast Red/0.1M Tris (pH 8.2) for probes conjugated to FITC. As soon as the background appeared, embryos were washed in fresh 5X TBST (Tris-Buffered Saline, Tween-20). The colour reaction was resumed if necessary. Embryos were fixed in 4% (w/v) PFA. The microscopic analysis was carried out using an upright microscope (Zeiss) and images were captured using QCapture software.

### **2.5.3. *In situ* hybridisation – Buffers and solutions**

**Hybridisation Buffer:** The solution was made up with 50% (v/v) Formamide, 1.3X SSC (pH 5), 0.5% (w/v) CHAPS, 0.2% (v/v) Tween-20, 5 mM EDTA (pH 8) NaOH, 0.1 mg/mL Heparin, 50 μg/mL tRNA (yeast), in H<sub>2</sub>O.

**Wash Buffer:** The solution was made up with 50% Formamide, 1X SSC (pH 5), 0.1% (v/v) Tween-20, and H<sub>2</sub>O.

**10% Boehringer Mannheim Blocking Reagent (BBR):** The solution was made up with 10 g of Blocking reagent and 1X MAB in a total volume 100 mL. The solution was autoclaved, aliquoted and stored at -20°C.

**Blocking solution:** The solution was made up with 2% (w/v) BBR in 1X MAB.

**Antibody solution:** The solution was made up with 2% (w/v) BBR, 20% (v/v) goat serum, anti-DIG Fab fragment (1:2000) or anti- fluorescein isothiocyanate Fab fragment (FITC) in 1X MAB.

**Alkaline phosphatase buffer (NTMT):** The solution was made up with 0.1 M NaCl, 0.1 M Tris (pH 9.5) HCl, 50 mM MgCl, and 1% (v/v) Tween-20, in H<sub>2</sub>O.

**Colour development substrate:**

**5-bromo-4-chloro-3-indolyl-phosphate (BCIP):** 50 mg/mL, in 100% dimethylformamide (DMF).

**4-Nitro-Blue Tetrazolium chloride (NBT):** 75 mg/mL, in 70% (v/v) DMF.

**Fast Red solution:** The solution was made by dissolving 1 tablet of SIGMAFAST™ Fast Red TR/Naphthol AS-MX in 10 mL 0.1M Tris (pH 8.2) buffer.

**5X TBST:** The solution was made up with 80 g NaCl, 0.25M Tris (pH 7.5) HCl, 2 g KCl, and 10% (v/v) Tween-20, in a total volume of 1000 mL.

**1X PBT:** Phosphate-buffered saline (PBS) with 0.1% (v/v) Tween-20.

**1X Maleic acid buffer (MAB):** The solution was made up with 100 mM Maleic acid, 150 mM NaCl, adjusted to pH 7.5, in DEPC H<sub>2</sub>O.

**1X MAB with Tween-20 (MABT):** The solution was made up with 100 mM Maleic acid, 150 mM NaCl, and 0.1% (v/v) Tween-20, adjusted to pH 7.5, in DEPC H<sub>2</sub>O.

**20X Saline sodium citrate buffer (SSC) (pH 5):** 20X SSC buffer was made up with 175.3 g NaCl, and 88.2 g tri-sodium citrate, adjusted with citric acid to pH 5, in a total volume of 1000 mL.

**0.5M EDTA (pH 8) NaOH:** The solution was made up with 186.1 g of EDTA, adjusted to pH 8, in a total volume of 1000 mL.

**5M NaCl:** The solution was made up with 282.2 g of NaCl in a total volume of 1000 mL.

**2M Tris (pH 9.5) HCl:** The solution was made up with 242.2 g of Tris base, adjusted to pH 9.5, in a total volume of 1000 mL.

**1M Tris (pH 7.5) HCl:** The solution was made up with 121.1 g of Tris base, adjusted to pH 7.5, in a total volume of 1000 mL.

**2M MgCl:** The solution was made up with 406.6 g of Magnesium Chloride hexahydrate in a total volume of 1000 mL.

#### **2.5.4. Embedding, sectioning and imaging**

After WMISH, embryos were rinsed in 1X PBS five times, 5 minutes each wash. Then, embryos were incubated in 20% (w/v) sucrose overnight at room temperature, with gentle rotation. Embryos were placed in tubes containing O.C.T. embedding medium (Miles Inc.) and allowed to settle. They were manually positioned with tweezers, or a needle, according to the desired plane of sectioning. Tubes were rapidly placed in cold isopentane on dry ice, and then stored at -20°C.

20-30 micron sections were cut at -23°C on a Leica Cryostat. Sections were transferred to positively charged-slides (SuperFrost-Plus, Thermo Fisher Scientific) and dried at room temperature overnight. Then, the slides were washed twice in 1X PBS, 5 minutes each wash, and mounted in Hydromount (National Diagnostics). The slides were dried overnight at 4°C and then at room temperature.

Subsequently, microscopic analysis was performed using an upright microscope (Zeiss). Images were captured and analysed using AxioVision software, available in the Henry Wellcome Laboratory for Cell Imaging.

## **2.6. Locked-Nucleic Acid (LNA) probes for WMISH**

Locked-Nucleic Acids (LNA) are a class of high-affinity RNA analogues in which the ribose is 'locked' by a methylene bridge connecting the 2'-O atom and the 4'-C atom.

LNA oligonucleotides consist of a mixture of LNA and DNA or RNA, containing the common nucleotide bases (A, T, U, C, and G) and are able to form base pairs according to standard Watson-Crick binding. They exhibit thermal stability when hybridised to a complementary DNA or RNA strand and can be made shorter (20-25 nucleotides in length) than traditional DNA or RNA oligonucleotide probes (250-1500 nucleotides in length). They can be used to discriminate between highly similar sequences, and also to detect low abundance nucleic acids.

Due to their hybridisation properties, LNA probes are the ideal choice for the detection of short non-coding RNA such as microRNAs (miRNAs) (Darnell et al. 2006; Sweetman et al. 2008).

LNA modified DNA oligonucleotide probes labelled with DIG, at both 5' and 3' ends, were supplied by Exiqon. As part of the XenmiR project (Wheeler laboratory), the positions of the LNAs in their sequences were optimised such that these LNA probes had similar melting temperatures (TM), therefore allowing to test them at the same temperature of hybridisation in *in situ* experiments. LNA probes were designed using the Primer3 primer design programme (Untergasser et al. 2007) and checked using the LNA Oligo Optimizer Tool on the Exiqon website. Each probe sequence was then screened against all known chicken sequences using Basic Local Alignment Search Tool (BLAST).

## **2.7. Labelled RNA probe synthesis by polymerase chain reaction (PCR)**

### **2.7.1. DNA template preparation**

Fragments of coding sequence of genes of interest were cloned into pGEMT-Easy vector (Promega) (Appendix I Table 1). This contains M13 forward and reverse primer binding sites, as well as SP6 and T7 RNA polymerase promoters, flanking the insertion site. To make a linear DNA template for probe synthesis, with the correct orientation (5'-3'), appropriate sets of primers were used to amplify by PCR the inserted gene fragments.

PCR reaction mix (BioMix Red Kit (Bioline)):

- 1  $\mu$ L Plasmid DNA (1-5 ng/ $\mu$ L)
- 5  $\mu$ L 2X BioMix Red buffer (Bioline)
- 1  $\mu$ L M13 Forward primer (10  $\mu$ M)
- 1  $\mu$ L M13 Reverse primer (10  $\mu$ M)
- 2  $\mu$ L H<sub>2</sub>O (Sigma)

Reaction conditions were:

95°C 3 minutes

Next 25 cycles of:

95°C 1 minute

55°C 1 minute

72°C 1 minute

Followed by:

72°C 10 minutes

Then 1/10<sup>th</sup> of the PCR reactions were run on a 1% (w/v) agarose TAE gel to ensure that single linear products had been amplified.

### **2.7.2. Probe synthesis (Transcription Reaction)**

Labelled RNA probes were generated from the amplified DNA template using the appropriate RNA polymerase.

For polymerases from Promega:

- 1  $\mu$ L PCR product
- 10  $\mu$ L 5X Transcription buffer (Promega)
- 5  $\mu$ L DTT (100 mM; Invitrogen)
- 2  $\mu$ L T7 or SP6 RNA polymerase (20 units/ $\mu$ L; Promega)
- 5  $\mu$ L 10X DIG-UTP Labelling Mix (10 mM; Roche)
- 1  $\mu$ L RNasin (40 units/ $\mu$ L; Promega)
- 26  $\mu$ L H<sub>2</sub>O (Sigma)

Transcription reactions were incubated for 2 hours at 37°C for T7 polymerases, or 40°C for SP6 polymerase.

Following the transcription reactions, DNA templates were degraded by addition of 1  $\mu\text{L}$  DNase (2 units/ $\mu\text{L}$ ; Ambion) per reaction and incubated for 30 minutes at 37°C. Then, 1/20<sup>th</sup> of each transcription reaction was run on a 1% (w/v) agarose TAE gel to check the probe integrity.

### **2.7.3. Probe purification**

Probes were purified using Illustra ProbeQuant G-50 Micro Columns (GE Healthcare) following manufacturer's instructions. Finally, 5  $\mu\text{L}$  of probe was added to 1 mL of Hybridisation Buffer, and then stored at -20°C until use for ISH.

## **2.8. RNA extraction from embryos and dissected somites**

After 2, 3, and 4 days of incubation, chicken embryos were harvested and placed in 1X DEPC-PBS on ice. Samples were centrifuged at 5,000 rpm for 2 minutes at 4°C, 1X DEPC-PBS was removed and 1 mL of TRIzol (Ambion), which deactivates RNases, was added to 2-5 embryos depending on the stage.

Embryos were vortexed until homogenisation and complete dissociation of nucleoprotein complexes for 5 minutes at room temperature. Chloroform was added (1/5<sup>th</sup> of the TRIzol volume; Sigma) and mixed, and after 5 minutes at room temperature, the samples were centrifuged at 12,000 rpm for 10 minutes at 4°C. The aqueous upper phases were transferred into new tubes, isopropanol was added (1/2 of the TRIzol volume; Fisher) and the samples were centrifuged at 12,000 rpm for 10 minutes at 4°C. The supernatant was removed and the pellets gently washed with 70% (v/v) ethanol before centrifugation for 5 minutes at 4°C.

The pellets were then air-dried for no more than 10 minutes and re-suspended in 20  $\mu\text{L}$  DEPC-H<sub>2</sub>O warmed up to 55°C. The concentration and quality of RNA samples were verified using a NanoDrop spectrophotometer with an expected ratio of over 1.8 for readings at 260/280 and 260/230 nm. RNA integrity was analysed by agarose gel electrophoresis.

The somites from day-2, -3, and -4 chicken embryos were carefully dissected in cold 1X DEPC-PBS, using sharp forceps and needles, and processed as described above, in order to get somite-enriched RNAs.

## 2.9. Reverse transcription

RNA extracted from chicken embryos or dissected somites was used to synthesise cDNA.

Reverse transcription reaction mix:

- 2  $\mu\text{g}$  RNA
- 1  $\mu\text{L}$  Random Hexamer primers (50  $\mu\text{M}$ ; Invitrogen)
- x  $\mu\text{L}$  H<sub>2</sub>O (Sigma) (Q.S. 11  $\mu\text{L}$ )

The reaction mix was incubated at 70°C for 10 minutes and then transferred immediately on ice.

The following were added:

- 4  $\mu\text{L}$  5X Reverse Transcriptase Buffer (Invitrogen)
- 2  $\mu\text{L}$  DTT (100 mM; Invitrogen)
- 1  $\mu\text{L}$  dNTP (10 mM; Promega)
- 1  $\mu\text{L}$  RNasin Ribonuclease inhibitor (40 units/ $\mu\text{L}$ ; Promega)
- 1  $\mu\text{L}$  SuperScript II Reverse Transcriptase (200 units/ $\mu\text{L}$ ; Invitrogen)

The samples were then incubated at 42°C for 1 hour, and stored at -20°C.

Following a similar protocol, SuperScript III Reverse Transcriptase Kit (Invitrogen) was also used. Due to an increased thermal stability, it was possible to synthesise more stable and longer cDNAs. This was particularly useful for the cloning of low expression level genes, or full-length genes, longer than 1.5 kb.

## 2.10. Amplification of cDNA by Polymerase Chain Reaction (PCR)

Polymerase Chain Reaction (PCR) is used to amplify DNA and produce millions of copies of a specific DNA sequence based on the repetition of cycles involving three different steps at three different temperatures: denaturation, annealing and elongation.

The success of the PCR reaction was established by electrophoresis using the appropriate percentage agarose gel.

Molecular weight markers (1 kb and 100 bp DNA ladder; NEB) were used to establish product size. Water was used as a negative control in all polymerase chain reactions.

Two different kits were used, BioMix Red and Velocity Kit (Bioline). BioMix Red Kit contains an ultra-stable *Taq* DNA polymerase, while the Velocity Kit contains a high-fidelity proofreading DNA polymerase from archaeal origin (5'-3' DNA polymerase and 3'-5' proofreading exonuclease activities).

PCR reaction mix (BioMix Red Kit):

- 1  $\mu$ L cDNA
- 5  $\mu$ L 2X BioMix Red Buffer (Bioline)
- 1  $\mu$ L Forward primer (10  $\mu$ M)
- 1  $\mu$ L Reverse primer (10  $\mu$ M)
- 2  $\mu$ L H<sub>2</sub>O (Sigma)

PCR reaction mix (Velocity Kit):

- 1  $\mu$ L cDNA
- 5  $\mu$ L 5X Hi-Fi Reaction Buffer (Bioline)
- 1  $\mu$ L dNTP (100 mM; Promega)
- 1  $\mu$ L Forward primer (10  $\mu$ M)
- 1  $\mu$ L Reverse primer (10  $\mu$ M)
- 0.5  $\mu$ L Enzyme
- 16.5  $\mu$ L H<sub>2</sub>O (Sigma)

Reaction conditions were specific for each primer set and sample used. In order to determine the optimum annealing temperature, gradient PCRs were performed prior to the main experiments. The samples were then loaded onto a 1% (w/v) agarose TAE gel.

Primers used for the amplification of cDNA are listed in Appendix I.

### **2.11. Agarose gel electrophoresis of DNA and RNA**

Agarose (Sigma) gels were made at 1% (w/v) concentration by pouring a warm solution of melted agarose-TAE (Tris base-Acetic acid-EDTA) in 1X TAE buffer. Ethidium bromide was added to the gels at 0.5 µg/mL to visualise DNA or RNA using an UV trans-illuminator following electrophoresis. Prior to loading the gel, each sample was mixed with 1/10<sup>th</sup> volume of 10X Loading Buffer. Electrophoresis was carried out at 80 volts for 30-40 minutes.

### **2.12. Restriction digestion**

All restriction digestions were carried out at 37°C for 2 hours using the recommended amount of restriction enzyme for the amount of DNA used, and the appropriate supplied buffer. The digestion products were run on a 1% (w/v) agarose TAE gel.

Restriction enzymes were purchased from Promega (Promega UK) or Roche (Roche Biochemical Reagents, Sigma UK).

### **2.13. Purification of PCR products and digested plasmid DNAs**

PCR products and digested plasmid DNAs were purified by gel electrophoresis. The appropriate band(s), visualised on a UV transilluminator, were excised from agarose gel and transferred into 1.5 mL Eppendorf tubes. Each tube was weighted before and after the addition of the gel slice. Purification of DNA fragments was performed using Thermo Scientific Gene JET Gel extraction Kit (Thermo Fisher Scientific), following the manufacturer's instructions.

Binding Buffer was added to the gel slice (100µL for every 100 mg of agarose gel for 1% (w/v) gel) with its approximate volume determined by the weight. After incubation at 55°C for 10 minutes, the gel mixture was briefly vortexed and loaded onto a provided column, and centrifuged at full-speed for 1 minute, the flow-through was then discarded. The column was washed with 700 µL of Wash Buffer, centrifuged at full-speed for 1 minute, and again the flow-through was discarded. An additional centrifugation was done to completely remove residual Wash Buffer.

The column was then transferred to a new tube and the purified DNA was eluted from the column with 20-30  $\mu\text{L}$   $\text{H}_2\text{O}$  (Sigma). Sample was quantified using a NanoDrop spectrophotometer before storage at  $-20^\circ\text{C}$ .

## **2.14. Ligation of DNA into a plasmid vector**

The PCR products were cloned into the pGEM-T Easy vector (Promega). Once validated by sequencing, these constructs could then be used as template to generate WMISH probes. The PCR products were then sub-cloned into pGL3-modified vector (from Dalmay lab, UEA; Tuddenham et al. 2006) or into the pCAB vector (from Dietrich lab, University of Portsmouth), after BglII/NheI or NotI/EcoRI restriction digestions, respectively; depending on the experiments they were needed for.

### **2.14.1. Ligation efficiency**

For all the ligations, the optimal amount of insert (ng) to use, taking into account the size (kb) and the amount of vector (ng), and the ratio insert:vector (2:1 or 3:1 for ligation into pGEM-T Easy vector (~3 kb) or pCAB vector (~5.8 kb); 3:1 or 6:1 for pGL3-modified vector (~5.3 kb)), was determined using the following equation:

$$\text{insert (ng)} = [\text{vector (ng)} \times \text{insert (kb)} / \text{vector (kb)}] \times \text{insert:vector ratio}$$

insert (ng): amount of insert

vector (ng) amount of vector

insert (kb): size of the insert

vector (kb): size of the vector

insert:vector ratio: molar ratio (for example, if 3:1, 3 times more insert than vector)

Negative control ligation reactions were performed with all vectors used to test the efficiency of restriction digestion prior ligation.

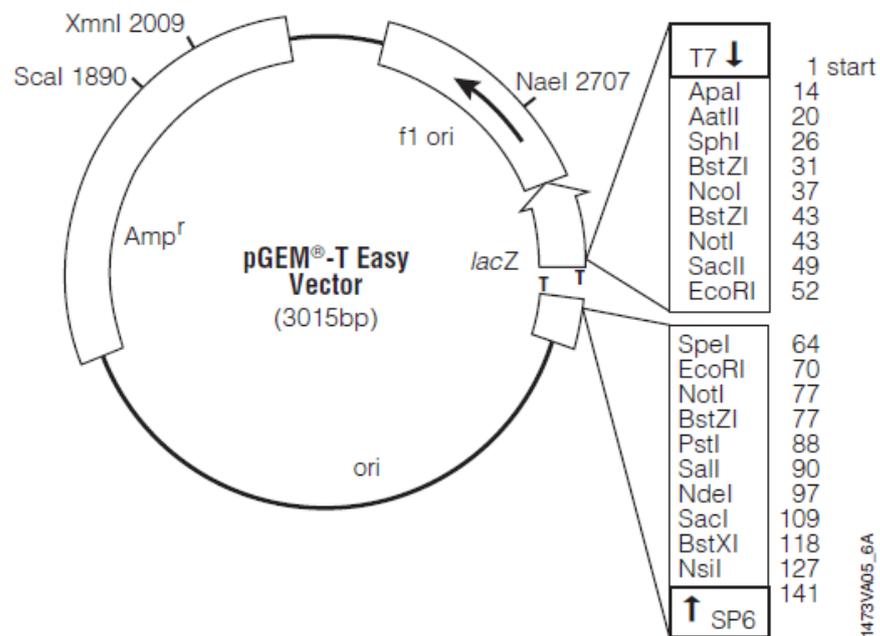
### 2.14.2. Ligation into pGEM-T Easy vector

Purified PCR fragments were cloned into pGEM-T Easy vector (3015 bp; Promega). This vector is a linearised vector with a single 3'-terminal thymidine at both ends which prevents re-circularisation and provides compatible overhangs improving the efficiency of ligation of PCR products.

The ligation reaction mix was as follows:

- 0.5  $\mu$ L pGEM-T Easy vector (25 ng)
- 5  $\mu$ L 2X Ligase Buffer (Promega)
- 0.5  $\mu$ L T4 DNA Ligase (3 units/ $\mu$ L; Promega)
- 4  $\mu$ L DNA (depending on the concentration and the insert:vector ratio) and H<sub>2</sub>O (Sigma)

The ligation was performed at room temperature for 2 hours, or overnight at 17°C.



**Fig. 2.1: pGEM-T Easy vector map.** The 3'-T overhangs prevent self-ligation and enable efficient ligation of PCR fragments. The plasmid contains an Ampicillin resistance cassette (Amp<sup>R</sup>) enabling for selection in *Escherichia coli* (*E. coli*).

### 2.14.3. Ligation into pGL3-modified vector and pCAB vector

The pGL3-modified vector is a pGL3 control vector (Promega) which was modified by deleting the region between SacI and BglII upstream of the SV40 promoter, and inserting a multiple cloning site (MCS) into the XbaI site downstream of the Luciferase stop codon (Tuddenham et al. 2006).

For example, 3'UTR fragments of predicted target genes, which had first been cloned into pGEM-T Easy vector, were digested out, using BglII and NheI restriction enzymes, and then sub-cloned into this vector into the MCS.

The pCAB vector is an expression vector; containing an internal ribosome entry site (IRES) upstream of GFP gene, this vector can be used for expressing a gene together with GFP.

For example, full-length Gga-Eya4 cDNA, which had been first cloned into pGEM-T Easy vector (Appendix I Table 4), was sub-cloned into the pCAB vector using NotI and EcoRI restriction enzyme. This construct could then be used to perform overexpression experiments.

The ligation reaction mix was as follows:

- x  $\mu$ L pGL3 or pCAB vector (50 ng)
- 5  $\mu$ L 2X Ligase Buffer (Promega)
- 0.5  $\mu$ L T4 DNA Ligase (3 units/ $\mu$ L; Promega)
- 4  $\mu$ L DNA (depending on the concentration and the insert:vector ratio) and H<sub>2</sub>O (Sigma)

### 2.15. Preparation of DH5 $\alpha$ *Escherichia coli* (*E. coli*) competent cells

Competent cells were spread onto a LB (Lysogeny Broth) plate using a metal hoop, and incubated overnight at 37°C.

A single colony was selected, placed into 5 mL of LB medium and incubated overnight at 37°C.

The 5 mL of culture were added to 200 mL of LB medium in a 2L sterile Erlenmeyer flask and grown at 37°C with constant shaking until the optical density (OD) at 600 nm reached 0.3-0.4 (approximately 1h30).

Then, the culture was split into four 50 mL tubes and kept on ice for 15 minutes. After a 15-minute centrifugation at 4,500 rpm at 4°C, the supernatant was removed and the pellets resuspended into 16 mL of Buffer TBI per tube. After 15 minutes of incubation on ice, and a 10-minute centrifugation at 4,500 rpm at 4°C, supernatant was removed and the pellets were resuspended into 4 mL of Buffer TBII per tube. Aliquots of 100 to 200 µL were prepared, snapped freeze using dry ice and stored at -80°C.

### 2.15.1. Solutions and Buffers

**Buffer TBI:** The solution was made up with 15% (v/v) Glycerol, 3 g RbCl<sub>2</sub>, 2.48 g MnCl<sub>2</sub>·4H<sub>2</sub>O, 0.38 g CaCl<sub>2</sub>·2H<sub>2</sub>O, and 30 mM KAc (pH 7.5), in a total volume of 250 mL (H<sub>2</sub>O). 0.2M glacial Acetic acid was used to adjust the pH to 5.8.

**Buffer TBII:** The solution was made up with 15% (v/v) Glycerol, 3 g RbCl<sub>2</sub>, 2.75 g CaCl<sub>2</sub>·2H<sub>2</sub>O, and 10 mM MOPS (pH 6.8), in a total volume of 250 mL (H<sub>2</sub>O).

**1M Potassium Acetate (KAc) (pH 7.5):** The solution was made up with 29.44 g of KAc in a total volume of 300 mL.

**0.5M 3-(N-morpholino) propane sulfonic acid (MOPS) (pH 6.8):** The solution was made up with 31.40 g of MOPS in a total volume of 300 mL.

**0.2M Acetic acid:** The solution was made up with 0.58 mL glacial Acetic acid in a total volume of 50 mL.

### 2.15.2. Transfection efficiency test

To test the efficiency of the newly made competent cells, 1 µL of a known plasmid at the concentration of 100 ng/µL was used to transform 100 µL of DH5α cells. After a 30-minute incubation on ice, and a heat shock at 37°C for 5 minutes, the cells were placed 5 minutes on ice. 900 mL of LB was added to the cells and after 1-hour incubation at 37°C, the transformation mix was spread onto LB plates:

Plate 1: 10 µL of the transformation volume

Plate 2: 100 µL of the transformation volume

After overnight incubation at 37°C, colonies were counted on each plate based on the following calculation:

$$\text{Transformation Efficiency (TE; cfu/}\mu\text{g)} = \text{Colonies} / \mu\text{g plasmid DNA} / \text{Dilution}$$

Colonies: number of colonies counted on a plate

$\mu\text{g}$  plasmid DNA: amount of DNA transformed expressed in  $\mu\text{g}$

Dilution: total dilution of the DNA before plating

TE: Colony-forming unit (cfu) per  $\mu\text{g}$  of plasmid DNA transformed

## **2.16. Plasmid transformation into DH5 $\alpha$ competent cells**

In order to increase the concentration of fragments of interest, competent cells were transformed using heat shock method.

5  $\mu\text{L}$  of the ligation mixture were added to 100  $\mu\text{L}$  DH5 $\alpha$  competent cells for 30 minutes on ice. After a heat shock at 37°C for 5 minutes, samples were immediately placed back on ice for 5 minutes. 1 mL of LB medium was added and the culture was incubated at 37°C for one hour with constant shaking. The bacterial culture was then centrifuged at 6,000 rpm for 5 minutes and 80% of the supernatant were removed. The pellet was re-suspended in the remaining supernatant. When doing a transformation with pGEM-T Easy vector, 5-10  $\mu\text{L}$  of X-gal (5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside; 20 mg/mL) were added. The cultures were spread onto LB plates with carbenicillin, an antibiotic, and incubated overnight at 37°C. An additional condition with the vector only was done for each experiment and used as negative control.

### **2.16.1. Selection of transformants**

Colonies that grew were either white or blue. In pGEM-T Easy vector, the multiple cloning site (MCS), where a gene of interest may be ligated, is located within the lacZ gene. Successful ligation disrupts the lacZ gene leading to the absence of active  $\beta$ -galactosidase resulting in white colonies.

At least three white colonies were selected from each plate, transferred into culture tubes containing 5 mL of LB medium and 5  $\mu$ L of carbenicillin (0.1 mg/mL), and incubated at 37°C on a shaker overnight.

### **2.16.2. Colony PCR**

Colony PCR reactions were performed to identify positive colonies (with an insert) for verification prior bacterial culture, or plasmid isolation.

A sterile tip was used to pick a colony which was first dipped into a PCR tube containing 10  $\mu$ L of the following PCR mix; the same tip was then used for overnight bacterial culture:

- 5  $\mu$ L 2X BioMix Red Buffer (Bioline)
- 1  $\mu$ L Forward primer (20  $\mu$ M)
- 1  $\mu$ L Reverse primer (20  $\mu$ M)
- 3  $\mu$ L H<sub>2</sub>O (Sigma)

The reaction conditions were:

94°C 4 minutes

Next 25-35 cycles of:

94°C 30 seconds

55°C 30 seconds

72°C 1-2 minute(s)

Followed by:

72°C 7 minutes

For pGEM-T and pCAB constructs, M13 Forward and Reverse primers were used. For pGL3 constructs, pGL3 Forward and Reverse primers were used, as well as primer sets specific to the inserts (see Appendix I Table 2 and 5).

### **2.17. Plasmid isolation from bacterial cultures**

Plasmid DNA extractions were done by using reSource Plasmid Mini Kit (LifeScience) and manufacturer's instructions were followed.

For mini preparation of DNA, 1 to 5 mL of overnight bacterial culture was transferred into an Eppendorf tube. After centrifugation at 8,000 rpm for 3 minutes, the supernatant was removed and the pellet was resuspended in 250  $\mu$ L of Resuspension Buffer (Buffer 1), then 250  $\mu$ L of Lysis Buffer (Buffer 2) were added and samples were incubated. After a 5-minute incubation at room temperature, 350  $\mu$ L of Neutralisation Buffer (Buffer 3) were added and immediately mixed to stop the lysis reaction. After centrifugation at 13,000 rpm for 10 minutes, a compact pellet was observed. The supernatant was loaded onto a provided column. After 1 minute, the tube was centrifuged for 1 minute, and the column was then washed with 750  $\mu$ L of Washing Buffer (Buffer E). After centrifugation for 1 minute, and an additional centrifugation of 2 minutes, the column was placed in a new tube and 30  $\mu$ L H<sub>2</sub>O (Sigma) were added. After 1 minute and centrifugation of 1-2 minutes, DNA was eluted, quantified and stored at -20°C, before sequence validation and then used.

Midi preparation were done using the NucleoBond Xtra Midi Plus Kit (Macherey-Nagel) and manufacturer's instructions were followed.

50 mL of overnight bacterial culture were transferred into a 50 mL Falcon tube and centrifuged at 4,500 rpm for 15 minutes. After removing the supernatant, the pellet was resuspended in 8 mL of Resuspension Buffer (Buffer RES), then 8 mL of Lysis Buffer (Buffer LYS) were added and the sample was incubated for 5 minutes at room temperature. 8 mL of Neutralisation Buffer (Buffer NEU) were added to stop the reaction. Then, 12 mL of Equilibration Buffer (Buffer EQU) were applied onto the rim of a provided column filter, and allowed to empty by gravity. The equilibrated column was loaded with the homogenised suspension and allowed to empty by gravity. The column was washed with 5 mL of Buffer EQU, the filter was removed, and the column was then washed with 8 mL of Wash Buffer. 5 mL of Elution Buffer (Buffer ELU) were added to the column and the eluted plasmid DNA was collected in a Falcon tube. Then, 3.5 mL of isopropanol were added to precipitate the eluted plasmid DNA. The mix was split into 6 Eppendorf tubes and centrifuged at 13,000 rpm for 15 minutes. The supernatant was removed, the pellets were washed with 70% (v/v) ethanol. After centrifugation at full-speed for 5 minutes, the supernatant was removed. An additional wash with 100% ethanol was done. The pellets were then air-dried for 5-10 minutes and resuspended in 50  $\mu$ L, final volume. DNA was quantified and stored at -20°C.

## **2.18. Bioinformatics: databases, software and algorithms**

MiRNA sequences were collected from XenmiR, GEISHA (Gallus Expression *In Situ* Hybridisation Analysis) and miRBase databases. XenmiR is a database of *Xenopus* miRNA expression patterns developed by the Wheeler laboratory (Ahmed et al. 2015). GEISHA is an *in situ* hybridisation gene expression resource for the chicken embryo (Bell et al. 2004). miRBase is a miRNA database of published miRNA sequences and annotation (Kozomara & Griffiths-Jones 2014).

Potential miRNA targets were identified using TargetScan, online software for miRNA target prediction and analysis (Lewis et al. 2005; Agarwal et al. 2015). TargetScan predicts biological targets of miRNAs by searching for the presence of conserved 7-mer and 8-mer sites that match the seed sequence of miRNAs of interest.

Identification of potential miRNAs targeting mRNAs of interest was done using the miRanda algorithm (John et al. 2004; Betel et al. 2008). The work was done with the help of Simon Moxon (Earlham Institute, Norwich UK).

GO term analysis (Gene Ontology) was assessed using DAVID bioinformatics resources (Database for Annotation, Visualisation and Integrated Discovery) (Huang, Sherman, et al. 2009; Huang, Lempicki, et al. 2009), and g:Profiler, a web server for functional profiling and interpretation of gene lists (Reimand et al. 2007; Reimand et al. 2016) These powerful tools provide a significant amount of information to understand biological meaning behind large lists of genes.

## **2.19. MiRNA target validation – *In vitro* interaction**

Putative miRNA targets were identified using several databases. 3'UTR fragments of these genes, containing putative miRNA binding site(s) were amplified by PCR from somite-enriched cDNA (Appendix I Table 2). They were cloned into pGEM-T Easy vector, verified by sequencing, and then sub-cloned into pGL3-modified vector and verified by sequencing. Cloning was performed using standard procedures as previously described.

### 2.19.1. Mutagenesis of inserts cloned into pGL3-modified vector

To generate pGL3 mutant constructs, the FastCloning method was used (Li et al. 2011). The predicted miRNA target site(s) were replaced by restriction enzyme sites, introducing point mutations (1-4 nucleotides modified).

Overlapping Forward and Reverse primers, with the mutation(s) to introduce, were designed (Appendix I Table 3), as well as overlapping primers in the Ampicillin resistance (AmpR) gene present in the pGL3-modified vector (Appendix I Table 5). The PCRs were done with Phusion High-fidelity DNA polymerase (NEB), a 5'-3' DNA polymerase with a 3'-5' proofreading exonuclease activity. The two products generated for each construct were blunt-ended.

Phusion PCR reaction:

Mix 1:

2.5  $\mu$ L Forward primer (10  $\mu$ M)  
2.5  $\mu$ L Reverse primer (10  $\mu$ M)  
2  $\mu$ L Plasmid DNA (10 ng/ $\mu$ L)  
7  $\mu$ L/tube

Where the following primer combinations were used to amplify the 2 fragments:

Forward primer gene – Reverse primer AmpR  
Reverse primer gene – Forward primer AmpR

Mix 2:

10  $\mu$ L 5X Buffer Phusion High Fidelity (HF) (NEB)  
1  $\mu$ L dNTPs (10 mM; Promega)  
0.5  $\mu$ L Phusion enzyme (NEB)  
31.5  $\mu$ L H<sub>2</sub>O (Sigma)  
43  $\mu$ L/tube

The reaction conditions were:

Depending on the size of the fragment to amplify:

<u>98°C</u> -----	30 sec	
98°C	10 sec	
59°C	30 sec	→ 22 cycles
<u>72°C</u> -----	15-30 sec / kb*	
72°C	10 min	

\*[2-2.5kb: 2 min; 2.5-3.5kb: 3 min; 3.5-4.5kb: 4 min]

To check if the PCRs had worked, 1µL of each PCR product was loaded onto a 1% (w/v) agarose TAE gel. Then, 5 µL of PCR product Fragment 1 and Fragment 2 were mixed together with 0.5 µL of DpnI restriction enzyme (Promega), in order to get rid of the methylated template. After two hours of incubation at 37°C, 1 µL of the digestion mix was loaded onto a 1% (w/v) agarose TAE gel for verification. 2 µL of the digestion mix was then added to 200 µL of DH5α competent cells and processed as described in chapter 2.16. The DH5α cells are able to recombine the 2 fragments. Then, and as described previously, cells were plated, transformants were selected, and plasmid DNA was purified from bacterial cultures. Control digestions – using the appropriate restriction enzymes for the mutations inserted – were performed, run on gel, and the samples were then sequence verified and stored at -20°C.

### 2.19.2. Cell culture

Chicken DF1 fibroblast cells were cultured in Dulbecco's modified Eagle medium containing 1g/L of D-glucose (DMEM + Glutamax; Gibco), 10% (v/v) heat-inactivated foetal bovine serum (FBS) (Gibco), 100 units/mL penicillin and 100 µg/mL streptomycin (Gibco) (DMEM complete medium).

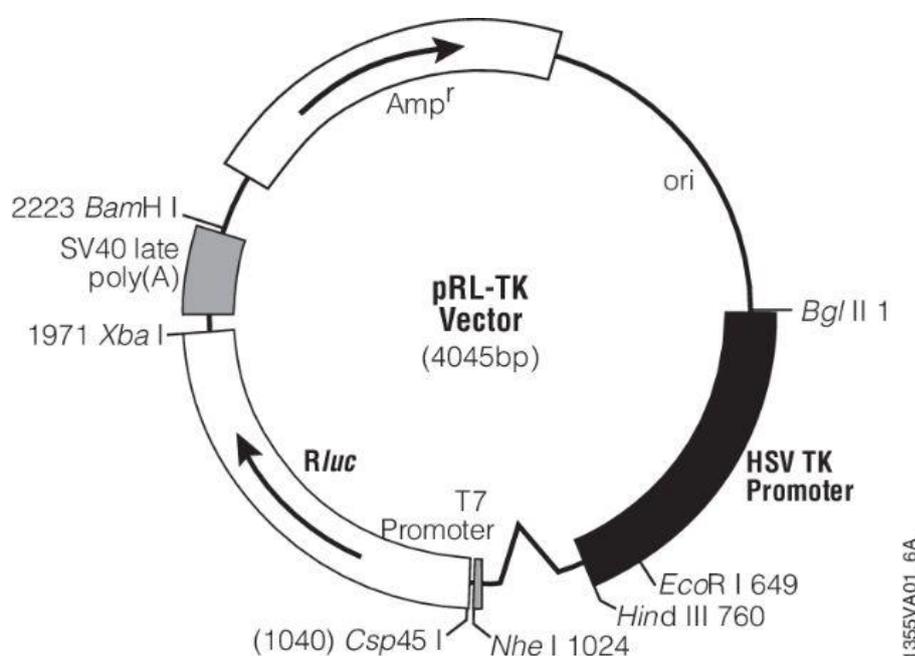
Cells were cultured in a humidified incubator and maintained at 37°C and 5% (v/v) CO<sub>2</sub>. The cells were passaged 1:4 every 2 to 3 days as follows: after removing culture medium, cells were rinsed in sterile 1X PBS (Gibco). Then, they were treated with 1 mL Trypsin/EDTA (0.25%) (Thermo Fisher Scientific) to cover the base of the flask for 5-15 seconds. The trypsin was removed from the cells and the flasks were returned to the incubator until cells had become dissociated from the base of the flask (~ 1-2 minutes). Cells were then resuspended in an appropriate volume of fresh culture medium and split.

### 2.19.3. Cell transfection

For luciferase reporter assays, DF1 cells were plated into 96-well plates. Cells were counted using a haemocytometer in order to calculate the volume needed to seed 7,000 cells/well. Then, DMEM complete medium was added up to 100  $\mu$ L, and the cells were returned to the incubator for 24 hours.

To mimic the miRNA action in the assays, siRNAs were used. The siRNAs (Sigma) were designed such that both strands of the siRNA represented the endogenous miRNA-3p/miRNA-5p duplexes with all the mismatches present in their sequences.

Transfections were done in serum-free medium (DMEM + Glutamax only). The complete medium was removed from the cells, and replaced by serum-free medium (50  $\mu$ L/well). The cells were transfected with Lipofectamine 2000 (0.2  $\mu$ L/well; Invitrogen) with either pGL3 wild-type (WT) or mutant constructs (100 ng), with or without siRNA (50 nM).



**Fig. 2.2: pRL-TK vector map.** Vector containing the cDNA encoding *Renilla* luciferase (*Rluc*) cloned from the anthozoan coelenterate *Renilla reniformis*. The plasmid contains a HSV-thymidine kinase promoter providing neutral constitutive expression of *Renilla* luciferase control reporter, a T7 promoter upstream of *Rluc* allowing *in vitro* synthesis of *Renilla* luciferase, SV40 late poly(A) signal sequence and an Ampicillin resistance cassette (AmpR).

The pRL-TK vector containing the cDNA encoding *Renilla* luciferase (Renilla vector; 25 ng) was co-transfected at the same time to check the transfection efficiency. The transfected cells were then returned to the incubator for 24 hours. After 24 hours, the Lipofectamine-containing medium was removed and luciferase reporter assays were performed.

Transfections with pGL3 vector containing a 3'UTR insert, Renilla vector, and no siRNA were done to check the impact of adding a siRNA on the transfection efficiency. The pGL3 vector containing a 3'UTR insert, Renilla vector and siC (universal negative control siRNA #1; Sigma) treated cells served as negative control.

#### **2.19.4. Luciferase Reporter Assays**

The Dual-Luciferase Reporter Assay System (Promega) was used to perform luciferase reporter assays. The activities of Firefly (*Photinus pyralis*) and Renilla (*Renilla reniformis*) luciferases were measured sequentially from a single sample.

After 24-hour transfection, cells were rinsed twice with cold sterile 1X PBS. Then, using a multi-pipette, Passive Lysis Buffer (PLB) was added to the cells (50  $\mu$ L/well) and incubated for 15-20 minutes at room temperature on a rocker.

10  $\mu$ L of each cell lysate were transferred to a white 96-well plate. The samples were organised on the plate such that, using a multi-pipette, WT and mutant samples for the same conditions were tested at the same time. The enzymatic reaction was started by adding 50  $\mu$ L/well of Luciferase Assays Reagent II (LARII), and the Firefly luciferase activity was measured by using a luminometer (Perkin-Elmer EnVision Applied Biosystems ABI 7500). Then, the Renilla luciferase activity was measured by adding 50  $\mu$ L of Stop & Glo reagent to the same sample.

#### **2.19.5. Normalisation of luciferase assays data**

All raw Firefly luciferase assays values were normalised to those of the Renilla luciferase. At least 3 assays using triplicate samples in each were performed. The activity of the pGL3 constructs containing a 3'UTR, with or without mutation, and transfected with the siC were set as 100% as shown in chapter 4.

The luciferase activities from the sensor constructs, with or without mutation, co-transfected with siRNAs of interest were then compared to the activity of pGL3(3'UTR) + siC (100%).

To assess the effect of the siRNA on the target constructs, the ratio values obtained by co-transfection of the WT or mutant construct with the siRNA were compared. Two tailed Student's unpaired t-test was performed to compare difference between two groups.

## **2.20. MicroRNA target validation – *In vivo* interaction**

Chicken embryos were used as *in vivo* model system in order to further validate miRNA target genes and to investigate the effect of their inhibition on myogenesis.

### **2.20.1. Microinjection**

Microinjection needles were prepared by pulling glass capillaries (1.0 mm O.D. x 0.78 mm I.D.; Harvard apparatus, UK) on a micropipette puller (P-97, Sutter Instrument Co., CA). Needles were filled by gravity with 2-4  $\mu$ L of solution to inject.

A small hole was performed at the blunt end of the egg shell and the hole was enlarged after the embryo had been located. After removing all the membranes covering the embryo using fine forceps (Dumont n°5), a micromanipulator was used to direct the needle into the somites.

Fine forceps were used to break the tip of the capillary; the opening generated would have to be wide enough to allow the antagomiR solution to go through (viscous), but still sharp in order not to damage the targeted somites. Pressure was exerted to fill these somites with the injected solution.

AntagomiRs were designed to target specific miRNAs (Appendix Table 6). The reverse complementary sequence was synthesized by Dharmacon with all bases replaced by 2'O-methyl-bases and, for example for the antagomiR-128 (AM-128), the phosphodiester bonds were replaced by thiol bonds between bases 1-2, 2-3, 18-19, 19-20 and 20-21. The antagomiRs also included a 3' cholesterol moiety. Scrambled antagomiR (AM-scr) was used as control (miR-206 scrambled sequence). For injections the antagomiRs were resuspended in H<sub>2</sub>O (Sigma) at a concentration of 1 mM. For injections of two antagomiRs, both antagomiRs were mixed based on a 1:1 ratio.

Chicken embryos at stage HH14-15 were injected into the 6 most posterior somites on one side. For each injected embryo, the non-injected contralateral side was used as control. After injection, the egg shell was re-sealed with tape and returned to the incubator at 38-39°C for the desired time.

### **2.20.2. Harvesting of treated chicken embryos**

The embryos for which the somites have been injected with antagomiRs, were incubated for 6, 9, 12 or 24 hours. All the treated embryos were collected in 1X PBS and the attached membranes were removed. The antagomiRs are fused to Fluorescein which can be detected by using a microscope equipped with GFP filter. Only embryos which developed to the expected stage and showed GFP expression in the injected somites were selected.

For *in situ* hybridisation, harvested embryos were fixed in 4% (w/v) PFA overnight and processed as described in chapter 2.5.

For Real-time quantitative PCR (RT-qPCR), injected somites were localised by detecting the GFP, and dissected out. RNAs were extracted and cDNA synthesised. The non-injected contralateral somites were also dissected and used as control.

### **2.20.3. RNA extraction from dissected somites – qPCR**

After dissecting out injected and non-injected somites from antagomiR-treated embryos, RNAs were extracted by using the protocol described in chapter 2.8, but with additional steps to obtain RNAs of better quality to use for qPCR.

The dissected somites were put in Eppendorf tubes containing cold 1X DEPC-PBS and kept on ice. After centrifugation at 5,000 rpm for 2 minutes at 4°C, 1X DEPC-PBS was removed, and 1 mL of TRIzol (Ambion) was added. Samples were vortexed and incubated for 5 minutes at room temperature. 200 µL of chloroform (Sigma) was added and mixed, and after 5 minutes at room temperature, the samples were centrifuged at 12,800 rpm for 30 minutes at 4°C. The aqueous upper phases were transferred into new tubes, 50 µL 5M NaCl, 1 µL Glycoblue (Invitrogen), and then 500 µL of isopropanol were added to each tube and incubated for 10 minutes at room temperature. The samples were centrifuged for 30 minutes at 4°C. The supernatants were discarded, the pellets were gently washed with 500 µL 70% (v/v) ethanol and centrifuged for 10 minutes at 4°C.

After removing the supernatant, an additional wash with 500  $\mu$ L 100% ethanol was done. The pellets were air-dried for no more than 10 minutes and re-suspended in 15  $\mu$ L H<sub>2</sub>O (Sigma), preliminary warmed up to 55°C. Then the RNAs were DNase treated by adding the following mix to each sample:

- 2 $\mu$ L 10X DNase Buffer (Ambion)
- 1  $\mu$ L DNase I (2 units/ $\mu$ L; Ambion)
- 2  $\mu$ L H<sub>2</sub>O (Sigma)

The samples were incubated at 37°C for 1 hour, and then a second RNA precipitation was performed using acid phenol-chloroform (Ambion). The steps previously described were resumed. Two washes with 70% (v/v) ethanol were done, followed by one with 100% ethanol, before to air-dry the pellets and re-suspend them in 20  $\mu$ L H<sub>2</sub>O (Sigma). The concentration and the quality of RNA samples were checked using a NanoDrop spectrophotometer (good quality: 260/280 and 260/230 ratios over 1.8).

cDNAs were synthesised as described in chapter 2.9, and 100-600 ng of RNA were used. Samples with RNA but no RNasin or SuperScript II Reverse Transcriptase were also generated and used as negative controls (RT- samples).

#### **2.20.4. Real-Time quantitative PCR (RT-qPCR)**

RT-qPCRs were performed in order to quantify potential changes in expression of genes of interest after antagomiR injection (see Appendix I Table 7 for qPCR primer sequences). The cDNAs were diluted 100 times before being used.

The reaction mix contained:

- 7.5  $\mu$ L 2X SYBR Green Master mix (Applied Biosystems)
- 5  $\mu$ L cDNA
- 0.5  $\mu$ L primer mix (Forward + Reverse; 10  $\mu$ M)
- 2  $\mu$ L H<sub>2</sub>O (Sigma)

The reaction conditions were as follows:

- 50°C 2 minutes
- 95°C 10 minutes

Followed by 40 cycles:

95°C 15 seconds

60°C 1 minute

Reactions were performed in 96-well plates in ABI Prism 7500 (Applied Biosystems).

#### **2.20.5. Normalisation of RT-qPCR data**

Results were analysed based on the Relative Standard Curve method (Larionov et al. 2005). By using this method, the quantity of each experimental samples was first determined by using a standard curve.

Five 5-fold serial dilutions of cDNA template, known to express the genes of interest (somite-enriched cDNAs from 3- or 4-day chicken embryos), was used to generate standard curves. A calibration curve was generated for each gene of interest and each housekeeping gene ( $\beta$ -actin and glyceraldehyde 3-phosphate dehydrogenase (GAPDH)), and used to extrapolate the relative expression for these same genes in unknown experimental samples.

Based on their respective calibration curve, the relative quantification results for each gene of interest were determined and normalised to the averaged relative quantification of  $\beta$ -actin and GAPDH housekeeping genes. Then the normalised values were compared between injected and non-injected samples to identify potential fold change in expression.

## CHAPTER 3: MicroRNA CHARACTERISATION AND TARGET IDENTIFICATION

### 3.1. Introduction:

In this chapter, two essential aspects for a better understanding of microRNA (miRNA) functions are going to be addressed: miRNA identification and characterisation, and target identification.

Since their discovery in 1993, in *C. elegans*, miRNAs have been increasingly studied due to the fundamental role(s) they have in regulating biological processes, through reshaping the cellular transcriptome and proteome (Lee et al. 1993); but also because of their extensive conservation across species, from nematode to human (He & Hannon 2004; Lagos-Quintana et al. 2001).

The first miRNAs – in particular muscle-specific miRNAs, such as the myomiRs miR-1/206 and miR-133 – were identified by using conventional techniques including cloning (Bentwich et al. 2005), Northern blotting (Sempere et al. 2004; Várallyay et al. 2007), and *in situ* hybridisation (Thomsen et al. 2005; Nelson et al. 2006). Although this allowed identification of a substantial number of miRNAs, these techniques are time-consuming and not cost effective (Mendes et al. 2009); they also do not provide much information regarding miRNA function(s).

With the idea that knowing the entire repertoire of these small molecules would help to gain a better understanding of their function(s), several new methods for DNA sequencing were developed leading to the identification of a large number of new miRNAs. Next-generation sequencing (NGS), also known as high-throughput sequencing, is a reliable and sensitive method to quantify miRNAs and detect less abundant miRNAs, which can either be ubiquitously represented or tissue-specific (Bar et al. 2008; Rathjen et al. 2009).

For efficient miRNA identification from NGS data, good prediction algorithms are necessary.

These tools consider some major miRNA characteristics, such as sequence conservation, and structural features, like hairpin and minimal folding free energy (Akhtar et al. 2016). There are two types of computational identification; in both of these methods the main signal used is the hairpin secondary structure of precursor miRNAs (pre-miRNAs).

The first method – comparative identification – is based on sequence conservation across different species and has been used for most of the known miRNAs (Lindow & Gorodkin 2007). Comparative genomics were used to filter out hairpins that are not evolutionary conserved in related species and then, based on sequence similarity, aligned unknown RNA sequences to known pre-miRNAs with a BLAST-like algorithm. The second method – non-comparative identification – does not rely on phylogenetic conservation and thus can be used to find non-conserved and/or species-specific miRNAs (Batuwita & Palade 2009); this approach mainly relies on the effective identification of pre-miRNAs among the predicted hairpin secondary structures. The first step consists in an initial screen identifying millions of hairpin structures from the genome, then, by combining bioinformatics predictions with microarray analysis, novel miRNAs were detected.

Thousands of miRNAs have now been discovered in several species and tissues by combining biological and bioinformatics approaches. All known information has been collected in online repositories, like miRBase for precursor sequences, mature miRNA sequences, and genomic location (Kozomara & Griffiths-Jones 2014); or species-specific databases, such as GEISHA for chicken, ZFIN for zebrafish, MGI for mouse and Xenbase for *Xenopus*, which include LNA probe sequences and expression patterns (Bell et al. 2004; Howe et al. 2013; Smith et al. 2014; Bowes et al. 2009).

However, most of these new miRNAs have been identified by ‘prediction tools’ using diverse algorithms, and it cannot be guaranteed that all of them are real. Each miRNA will have to be validated experimentally. In GEISHA, in addition to next-generation sequencing (NGS), miRNA microarray and PCR data, published expression patterns for most of the known chicken miRNAs are referenced (Bell et al. 2004; Darnell et al. 2006); however, except for the most studied miRNAs, expression data are often limited to whole-mount *in situ* hybridisations (WMISH) of a few stages, with hardly any sections. Work still need to be done to precisely determine where and when miRNAs are expressed in order to fully understand their function(s) in general and in skeletal muscle in particular.

The development of computational tools was not only important for the discovery of new miRNAs, but also for the identification of their targets. However, in animals, the complementarity between miRNAs and mRNA target sites is partial, therefore it is more difficult to determine potential targets with a high specificity (Axtell et al. 2011).

Numbers of studies have been focused on predicting accurate target mRNAs for miRNAs (Ekimler & Sahin 2014).

They were using various algorithms taking into account Watson-Crick complementarity in the seed sequence between miRNA and target mRNA, sequence comparison between species, and thermodynamic favourability of the miRNA-mRNA duplex (free energy calculations and site accessibility) (Peterson et al. 2014). Although seed regions are considered crucial for mRNA targeting, algorithms depending on simple base-pairing result in high false positive rates (Bartel 2009).

Taking into account most of these criteria, algorithms like TargetScan, miRBase, and miRanda provide long lists of potential target mRNAs for each miRNA (Agarwal et al. 2015; Kozomara & Griffiths-Jones 2014; Betel et al. 2008). Experiments will then have to be carried out in order to validate these putative miRNA/mRNA interactions in a given biological context.

The first part of the project was dedicated to (1) characterising the expression of interesting miRNAs in the chicken embryo during somitogenesis, and to (2) the identification of their potential targets. MiRNAs with a potential expression in skeletal muscle in various species were selected, with a particular focus on miR-128.

### **MiR-128:**

MiR-128 is an intronic miRNA, encoded by two distinct genes, miR-128-1 and miR-128-2, which are embedded in the introns of R3HDM1 (R3H domain containing 1) and ARPP21 (cyclic adenosine monophosphate (AMP)-regulated phosphoprotein, 21 kDa) (Bruno et al. 2011). Recently, miR-128 and ARPP21 have been associated with neuropsychiatric phenotypes (fear response, anxiety, movement disorders). Highly expressed in mammalian central nervous system, ARPP21, a calmodulin (CaM) signalling regulator, was shown to be a direct target of miR-128. Calmodulin plays important roles at the synapse by regulating the release of neurotransmitters from the presynaptic terminal (Ching & Ahmad-Annuar 2015).

In chicken, these genes are located on chromosome 7 and 2, respectively. Both miR-128-1 and miR-128-2 are processed to generate the same mature miRNA with identical sequence, miR-128.

MiR-128 is a ‘brain-enriched’ miRNA first identified in mouse, where its expression level increases during brain development and is maintained in adult brain tissues (Lagos-Quintana et al. 2002; Smirnova et al. 2005). Similar results were observed in chicken and zebrafish (Xu et al. 2006; Kapsimali et al. 2007). In addition, miR-128 has been shown to be involved in the repression of the RNA surveillance pathway, called nonsense-mediated decay (NMD); inactivation of NMD being necessary for the differentiation of neuronal cells during brain development (Bruno et al. 2011; Karam & Wilkinson 2012).

MiR-128 is found in cardiac tissue. A recent study examining newt cardiac regeneration by Witman *et al*, demonstrated that miR-128 regulates the expression of the transcription factor Islet1, a transcription factor expressed in cardiac progenitor cells. By targeting Islet1, miR-128 could be acting as a negative regulator of progenitor cell activity, emphasising a need for differentiation into cardiac cell lineages necessary during the process of regeneration (Witman et al. 2013).

In chicken, miR-128 is found in the developing heart, however its expression in this tissue appears to be limited to a short time-window as it is only seen in stage HH13 embryos (Darnell et al. 2006).

As well as being involved in neuronal and cardiac development, miR-128 expression was also detected in skeletal muscle. MiR-128 is found in adult mouse muscle (Sempere et al. 2004; Lee et al. 2008); adult and embryo porcine skeletal muscle (Zhou et al. 2010); and adult and embryo (somites) chicken skeletal muscle (Darnell et al. 2006; Lin et al. 2012; Abu-Elmagd et al. 2015). In mouse, the inhibition of insulin receptor substrate 1, Irs1, by miR-128, leads to the inhibition of myoblast proliferation and induction of myotube formation (Motohashi et al. 2013). These results are consistent with the increase in miR-128 expression observed during myoblast differentiation in differentiating mouse C2C12 myoblast cells (Sun et al. 2010). The role of miR-128 in the inhibition of proliferation and promotion of myoblast cell differentiation was also demonstrated in a recent work from Shi *et al*. done in mouse.

They showed that miR-128 promoted myotube formation by targeting myostatin, a negative regulator of myogenesis and muscle growth (Shi et al. 2015). They also showed that ectopic miR-128 is able to induce the expression of Pax3/7 and MRFs, like Myf5 and MyoG. Although miR-128 appears to be expressed in both adult and developing skeletal muscle, only its functions in adult muscle have been studied in these works, mostly in mouse and in *ex vivo* or *in vitro* experiments.

Although hundreds of miRNAs have been identified and deposited into miRBase repository, only limited data are available on their expression patterns during key developmental stages. In order to begin to understand the function(s) of miRNAs during embryo development, it is important to characterise their spatiotemporal expression patterns throughout development, and identify their targets.

A group of miRNAs, predicted to be expressed in skeletal muscle, has been studied; the process of their selection, and characterisation of their expression profiles will be presented in the first part of this chapter.

Some of these miRNAs, and in particular miR-128, which have interesting muscle expression, were further investigated in order to identify potential targets; results will be presented in a second part.

## **3.2. Results and discussion**

### **3.2.1. Characterisation of microRNA expression patterns**

Based on information available from miRBase, GEISHA, and Xenbase database (Kozomara & Griffiths-Jones 2014; Bell et al. 2004; Bowes et al. 2009), as well as high-throughput sequencing data (Rathjen et al. 2009), and information found in the literature, a short list of miRNAs potentially expressed in skeletal muscle was established.

To determine miRNA expression patterns, WMISH were performed in chicken embryos at different stages of development (see chapter 2.5.). Due to the small size of the miRNAs (20-25 nucleotides in length), performing classic WMISH with conventional RNA probes had been technically challenging. The technology developed by Exiqon (<http://www.exiqon.com/>), using modified oligos containing miRCURY™ Locked-Nucleic Acid (LNA) nucleotides as probes, helped to solve this problem (Nielsen et al. 1999; Kubota et al. 2006) (details in chapter 2.6).

To date, these probes are one of the most efficient ways to determine miRNA expression profiles in tissues and embryos, however their cost is an important limitation to any study of a large number of miRNAs.

#### The XenmiR project – miRNA identification in the model *Xenopus*:

The Wheeler laboratory (UEA, Norwich UK) was involved in the development of a database, similar to GEISHA for the chicken, entirely dedicated to miRNAs expressed in *Xenopus* (*Xenopus tropicalis* and *laevis*), called XenmiR. The aim of the XenmiR project was to determine the expression patterns of miRNAs expressed during *Xenopus* development using LNA oligonucleotides (Ahmed et al. 2015). To do so, LNA probes were developed by Exiqon. The positioning of the LNAs in the sequences was done such that all the LNA probes had a similar melting temperature ( $T_m$ ) and could be used at the same hybridisation temperature (48°C). They determined the expression patterns of 180 miRNAs in *Xenopus laevis* and found a large number being expressed in neural tissue and in the somites.

##### **a. MicroRNAs: conservation across species**

With the XenmiR project, 56 miRNAs were detected in *Xenopus laevis* (Xla) and *tropicalis* (Xtr), in somites. Based on the fact that most of the miRNAs are predicted to be conserved across species, it was tempting to think that these miRNAs could also be present in chicken, and their sequences conserved.

The first step was to determine if *Xenopus* and chicken miRNA sequences were conserved. *Xenopus* miRNA sequences, extracted from the XenmiR project data, and the corresponding chicken miRNA sequences, from miRBase, were collected and used for comparison. Sequence alignments showed that 42 of the 56 *Xenopus* miRNAs were sequence conserved in chicken, with 23 expressed in somites (based on whole-mount analysis). These 23 miRNAs are listed in Table 3.1.

Within these 23 miRNAs, 14 were completely conserved between the two species (100% identity), and 8 were 1-2 nucleotide(s) shorter in chicken at their 3' end, but the 5' ends, containing the seed sequence, were identical. The miRNA Xla-miR-1306 was the least conserved with nucleotides missing at both 5' and 3' ends.

XenmiR miRNAs	X/C	chicken miRNAs	RNAseq AEM	GEISHA	Whole-mount in situ hybridisation (WMISH)													Sectioning												
					HH10	HH11	HH12	HH13	HH14	HH15	HH16	HH17	HH18	HH19	HH20	HH21	HH22	HH23	HH10	HH11	HH12	HH13	HH14	HH15	HH16	HH17	HH18	HH19	HH20	HH21
Xla-miR-1306	(-3) x (-2)	Gga-miR-1306-3p					x	x			x			x	x					x					x					
Xtr-let-7a	x	Gga-let-7a-3p	YES	x		x	x				x	x		x	x					x							x			
Xtr-miR-1a	x	Gga-miR-1a-3p	YES	x [S]	x	x		x			x			x						x						x				
Xtr-miR-1b	x	Gga-miR-1b-3p	NO	x [S]	x	x				x		x		x						x				x		x				
Xtr-miR-10a	x (-1)	Gga-miR-10b-5p*		x		x				x		x		x						x						x				
Xtr-miR-15a	x (-1)	Gga-miR-15a	NO	x [S]		x	x				x				x	x					x					x				
Xtr-miR-15b		Gga-miR-15b-5p	YES	x [S]	x				x	x		x	x		x	x				x							x			
Xtr-miR-15c	x (-2)	Gga-miR-15c-5p				x			x	x	x		x	x	x					x					x					
Xtr-miR-17-5p	x	Gga-miR-17-5p	NO	x [S]		x	x			x		x		x						x						x				
Xtr-miR-18b	x (-1)	Gga-miR-18b-5p	NO	x			x			x		x	x	x		x				x						x				
Xtr-miR-23b	x (-1)	Gga-miR-23b-3p	yes			x	x				x	x	x		x					x				x				x		
Xtr-miR-24a	x	Gga-miR-24-3p	NO			x					x	x	x							x				x				x		
Xtr-miR-30a-5p	x	Gga-miR-30a-5p	YES	x [S]	x			x			x	x		x		x					x					x				
Xtr-miR-31	x	Gga-miR-31-5p	NO			x	x			x		x		x	x						x							x		
Xtr-miR-128	x (-1)	Gga-miR-128-3p	YES	x [S]	x	x	x	x	x	x	x	x	x	x	x	x					x				x		x			
Xtr-miR-130a	x (-2)	Gga-miR-130a-3p*		x [S]			x				x	x	x	x							x				x					
xtr-miR-133a	x	Gga-miR-133a-3p	YES	x [S]		x	x				x	x	x	x		x					x					x				
Xtr-miR-133b	x	Gga-miR-133b	YES			x				x		x	x	x	x							x						x		
Xtr-miR-133c	x	Gga-miR-133c-3p	NO	x		x	x	x			x	x			x	x						x						x		
Xtr-miR-194	x	Gga-miR-194	NO	x		x				x	x	x	x		x						x					x				
Xtr-miR-203	x	Gga-miR-203a	NO	x			x			x	x	x	x	x		x					x				x			x		
Xtr-miR-206	x	Gga-miR-206	YES	x [S]	x		x				x			x	x						x					x				
Xtr-miR-223	x	Gga-miR-223	YES	x [S]		x	x				x	x	x			x									x			x		

**Table 3.1: Summary – Characterisation of miRNAs during chicken development.** XenmiR miRNAs: list of 23 Xenopus miRNAs with an expression in somites. X/C: Xenopus vs chicken miRNA sequence comparison. x indicates 100% identity; number of missing nucleotides between sequences is indicated in brackets. Chicken miRNAs: Corresponding Xenopus miRNAs in chicken. RNAseq AEM: presence, or not, of the miRNAs in chicken somite dataset generated by NGS technology (Rathjen et al. 2009). GEISHA: Available information for a specific miRNA in GEISHA database (indicated by x). [S]: miRNA presents in somites. WMISH and Sectioning parts: x indicates the stages used to perform WMISH and for which sectioning was done.

With the help from Simon Moxon (Earlham Institute, Norwich UK), these 23 miRNAs were checked against a chicken somite dataset (RNAseq AEM), previously generated by our laboratory using NGS technology (Rathjen et al. 2009). This analysis revealed that from these 23 microRNAs found expressed in *Xenopus* somite and having conserved sequence between *Xenopus* and chicken, 10 were identified by RNA sequencing performed on chicken somite samples; amongst them were found the myomiRs miR-1a, miR-133a/b and miR-206. In addition, some information was available in GEISHA (chicken database) for 17 of these miRNAs, with 11 expressed in somites (Bell et al. 2004; Darnell et al. 2006).

### **b. MicroRNAs: expression patterns**

*Xenopus laevis* and *tropicalis* LNA oligonucleotides, and chicken LNA probes when available, designed for the 23 selected miRNAs, were used to perform LNA WMISH in chicken embryos at different stages of development, from HH10 (2 days of incubation) to HH23 (4 days of incubation), according to the Hamburger and Hamilton table (Hamburger & Hamilton 1992). After LNA WMISH, miRNA expression patterns were analysed in whole-mount embryos first, and then on sections (see Table 3.1 for details) (Ahmed et al. 2015).

#### **i) New insights into the myomiRs**

The myomiRs family, initially composed of miR-1, miR-133a/b and miR-206, has recently been expanded to include miR-208a/b, miR-486, and miR-499 (Lagos-Quintana et al. 2002; McCarthy & Esser 2007; McCarthy 2008; van Rooij et al. 2007; van Rooij et al. 2009; Small et al. 2010). Compared to the ‘new myomiRs’, miR-1/206 and miR-133a/b have been extensively studied since their discovery. Conserved across species, with well characterised skeletal muscle-specific expression patterns, these miRNAs were used as a starting point in the process of learning how to perform LNA WMISH in chicken embryos (McCarthy & Esser 2007; McCarthy 2008; Sweetman et al. 2008). Results are presented in Fig. 3.1-3.

Consistent with profiles already published in chicken (GEISHA) and in *Xenopus* (XenmiR database; Ahmed et al. 2015), miR-1, miR-206, and miR-133 are strongly expressed in the somites.

In addition, these series of WMISH provided some new elements to what was already known.

According to the expression patterns presented in GEISHA database, miR-1a is not detected by ISH before stage HH11 where it is found in the heart; and its somitic expression is observed later at about stage HH14.

Our results showed that miR-1a (Fig. 3.1a) was already expressed and detectable in HH9 embryos (i), in the heart, while its expression in somites, although very weak, was detected in the most anterior somites of HH12-13 embryos (ii). Its heart expression became stronger as the embryos developed (i-iii), as well as in the differentiating somites (ii, iii, iii', iv), and particularly in the myotome, as observed in transverse sections (ii', iii'', iv'). Interestingly, with only one nucleotide difference in the middle of its sequence compare to miR-1a, the expression of miR-1b (Fig. 3.1b) was less specific with a lot of background in whole-mount (v, v', vi). Sectioning showed that miR-1b was expressed in somites, in the myotome, at HH14-15 and onwards (v'', vi').

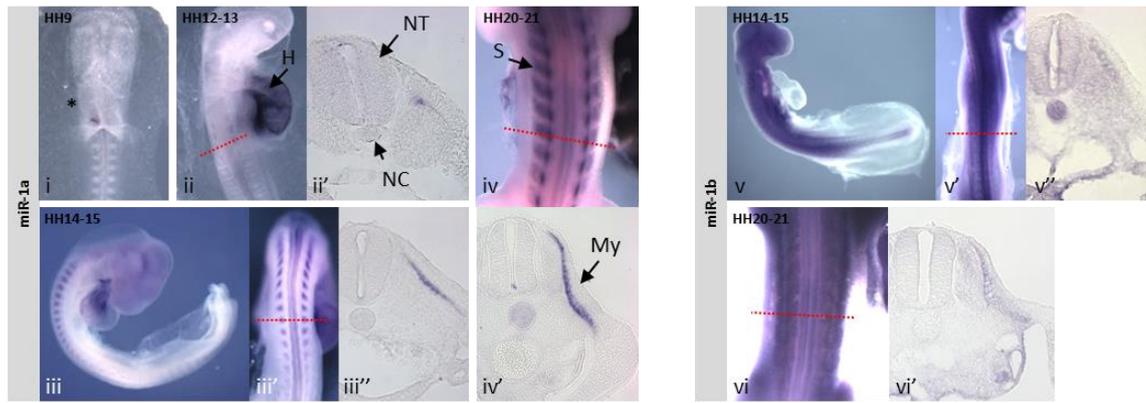
A similar observation was made for miR-206 (Fig. 3.2). In GEISHA database, there is no information on the expression of miR-206 before stage HH14.

Like for miR-1a, we observed that miR-206 is already expressed at HH12-13 in the most anterior somites (i); this suggests that miR-1a and miR-206 could already play important roles in the undifferentiated epithelial somites. MiR-206 was also found in the neural tube.

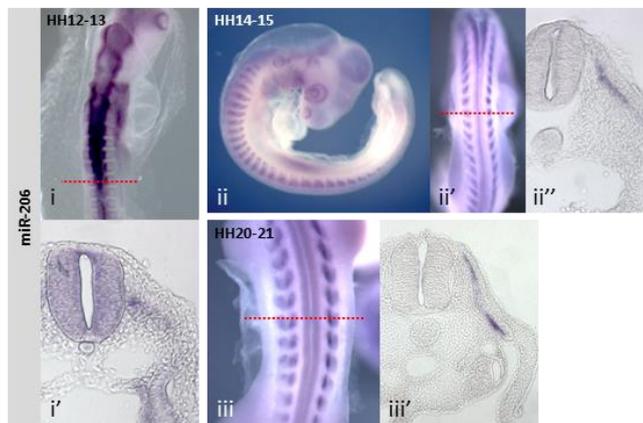
As the embryo developed, its expression in the neural tube disappeared; miR-206 was detected in most somites along the anterior-posterior axis by HH14-15 (ii, ii', iii). In transverse sections, miR-206 was strongly expressed in the myotome (ii'', iii'). MiR-206 was not detectable in early stage HH10-11 embryos, indicating that it probably starts to be expressed between stage HH10-11 and HH12-13.

While miR-1a and miR-206 are both expressed in the somites, only miR-1a is found in the heart. This could be explained by the fact that their sequences, extremely similar, differ by 3 nucleotides (see Appendix II Table 1).

In chicken embryos, miR-133 family is composed of 3 members: miR-133a, and miR-133b and miR-133c (Fig. 3.3).



**Fig. 3.1: Expression patterns of miR-1a (a) and miR-1b (b).** LNA WMISH were performed on chicken embryos using *Xenopus* LNA probes. WMISH and transverse sections (20x magnification) for miR-1a (a), and miR-1b (b). (a) MiR-1a expression pattern at HH9 (i), HH12-13 (ii), HH14-15 (iii), and HH20-21 (iv). MiR-1a was first observed in the heart at HH9 (i; \*), and later in somites from HH12-13 (ii, iii, iii' and iv (interlimb portion)). Transverse sections showed the expression of miR-1a in the developing somite (ii') and in the myotome (iii'', iv'). (b) MiR1b expression pattern at HH14-15 (v-v'') and HH20-21 (vi, vi'). MiR-1b was expressed ubiquitously in whole-mount (v, v', vi). Transverse sections showed its expression in the myotome and the notochord at stage HH14-15 (v''); notochord expression disappearing in later stage HH20-21 (vi'). Red dotted line indicates the location of the transverse sections. H: heart; My: myotome; NC: notochord; NT: neural tube; S: somite.



**Fig. 3.2: Expression pattern of miR-206.** LNA WMISH were performed on chicken embryos using *Xenopus* LNA probes. WMISH (i- iii) and transverse sections (i'-iii'; 20x magnification) for miR-206. MiR-206 expression pattern at HH12-13 (i), HH14-15 (ii), and HH20-21 (iii). MiR-206 was already expressed in the most anterior somites at HH12-13 (i). As the embryos developed, this expression was also found in most posterior somites (ii, iii). Transverse sections showed the expression of miR-206 in the developing somite (i') and in the myotome (ii'', iii'). Red dotted line indicates the location of the transverse sections. My: myotome; NC: notochord; NT: neural tube; S: somite.

MiR-133a is already expressed at HH10, in the heart tube; whereas its expression in somites is only observed from HH15 (GEISHA database).

With no embryo younger than HH11-12, we could not confirm this observation or determine a more precise expression start point, however, we observed that miR-133a was already strongly expressed in the heart at this stage (HH11-12); it was also found in the most anterior somites, in the neural tube, and notochord (i, i'); the expression in neural tube and notochord has also been reported in *Xenopus* WMISH (Ahmed et al. 2015). In older embryos, miR-133a was no longer expressed in the neural tube and the notochord, and its expression in somites became restricted to the myotome (i-iii').

There is no profile of expression available in GEISHA database for miR-133b, and miR-133c is described as ubiquitously expressed.

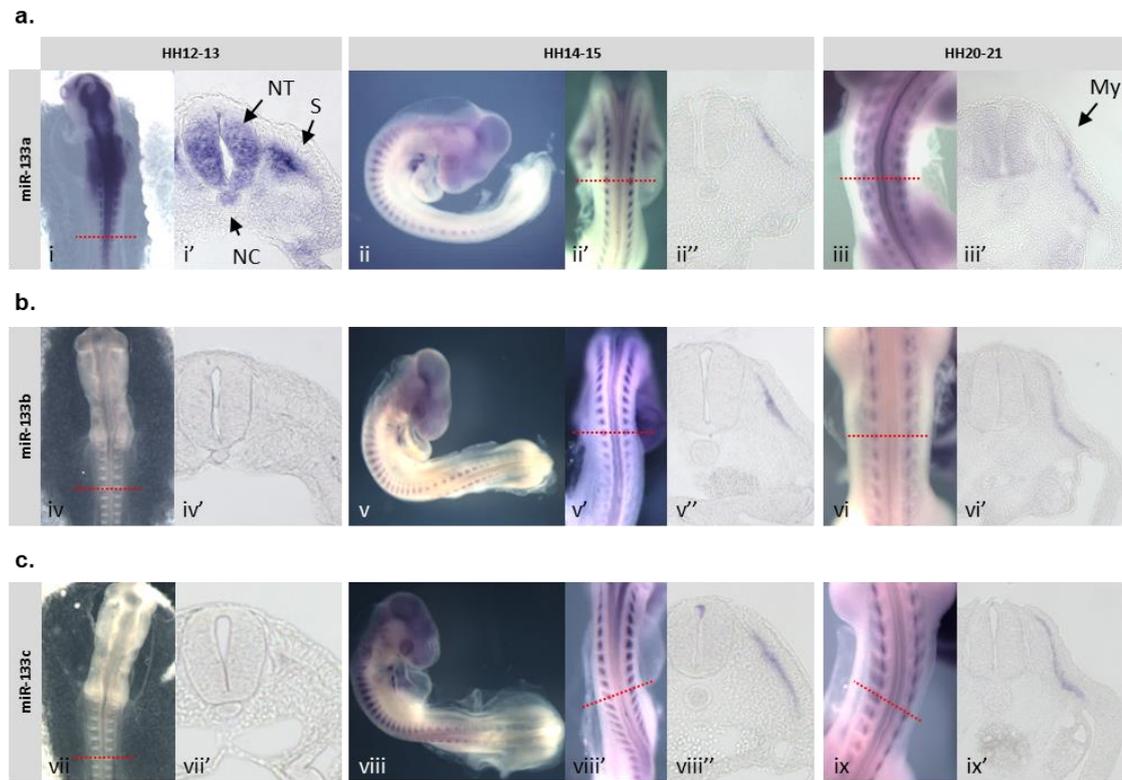
WMISH for these miRNAs indicated that miR-133b and miR-133c have expression profiles similar to miR-133a. This could be explained by the fact that their sequences are extremely similar and only differ by 1 or 2 nucleotides at the 3' end (see Appendix II Table 1). However, their expression profiles were not identical indicating that even 1 nucleotide of difference is enough to yield a specific signal when using LNA probes.

Interestingly, miR-133b and miR-133c seemed to start to be expressed with a slight delay compared to miR-133a (i', iv', vii'). While miR-133a was already strongly represented in the somites at HH11-12, miR-133b was not expressed yet, and miR-133c was only just becoming detectable. MiR-133c could be observed from HH12-13 in somites, miR-133b appeared later at stage HH14-15. From HH14-15, no difference was observed between the 3 members of miR-133 family, with a strong expression in somites and in particular in the myotome (ii-iii'; v-vi'; viii-ix').

## **ii) Characterisation of a group of 16 somitic miRNAs**

The same procedure was used to determine the expression patterns of 16 miRNAs which were previously shown to have partially conserved (seed sequence not affected) or conserved sequences between *Xenopus* and chicken (Table 3.1).

*Xenopus* LNA probes were used to perform WMISH in chicken embryos at different stages of development. Results are presented in Fig. 3.4 (see also Ahmed et al. 2015).



**Fig. 3.3: Expression patterns of miR-133 family members: miR-133a (a), mir-133b (b), and miR-133c (c).** LNA WMISH were performed on chicken embryos using Xenopus LNA probes. WMISH and transverse sections (20x magnification) for miR-133a (a), miR-133b (b), and miR-133c (c). Expression patterns of miR-133a/b/c at HH12-13 (i, iv, vii), HH14-15 (ii, v, viii), and HH20-21 (iii, vi, ix). MiR-133a was the first member to be detected. It was already strongly represented in the somites, and the neural tube at HH12-13 (i'), while miR-133b was not yet expressed (iv') and miR-133c was just starting to be detected (vii'). From HH14-15, they were all expressed in somites (ii, ii', iii; v, v', vi; viii, viii', ix) and transverse sections indicated their presence in the myotome (ii''-iii''; v''-vi''; viii''-ix'). Red dotted line indicates the location of the transverse sections. My: myotome; NC: notochord; NT: neural tube; S: somite.

All the *Xenopus* probes used worked in chicken embryos and a specific expression pattern was observed for each miRNA. Although sometimes a strong background interfered with the determination of miRNA patterns in whole-mount embryos, their specific expressions were revealed in transverse sections.

For most of these miRNAs there was no information in GEISHA database about their expression patterns. When WMISH and, sometimes, transverse section data were available, miRNA expressions were often not clear, very weak, with ubiquitous or widespread patterns.

#### MiR-1306:

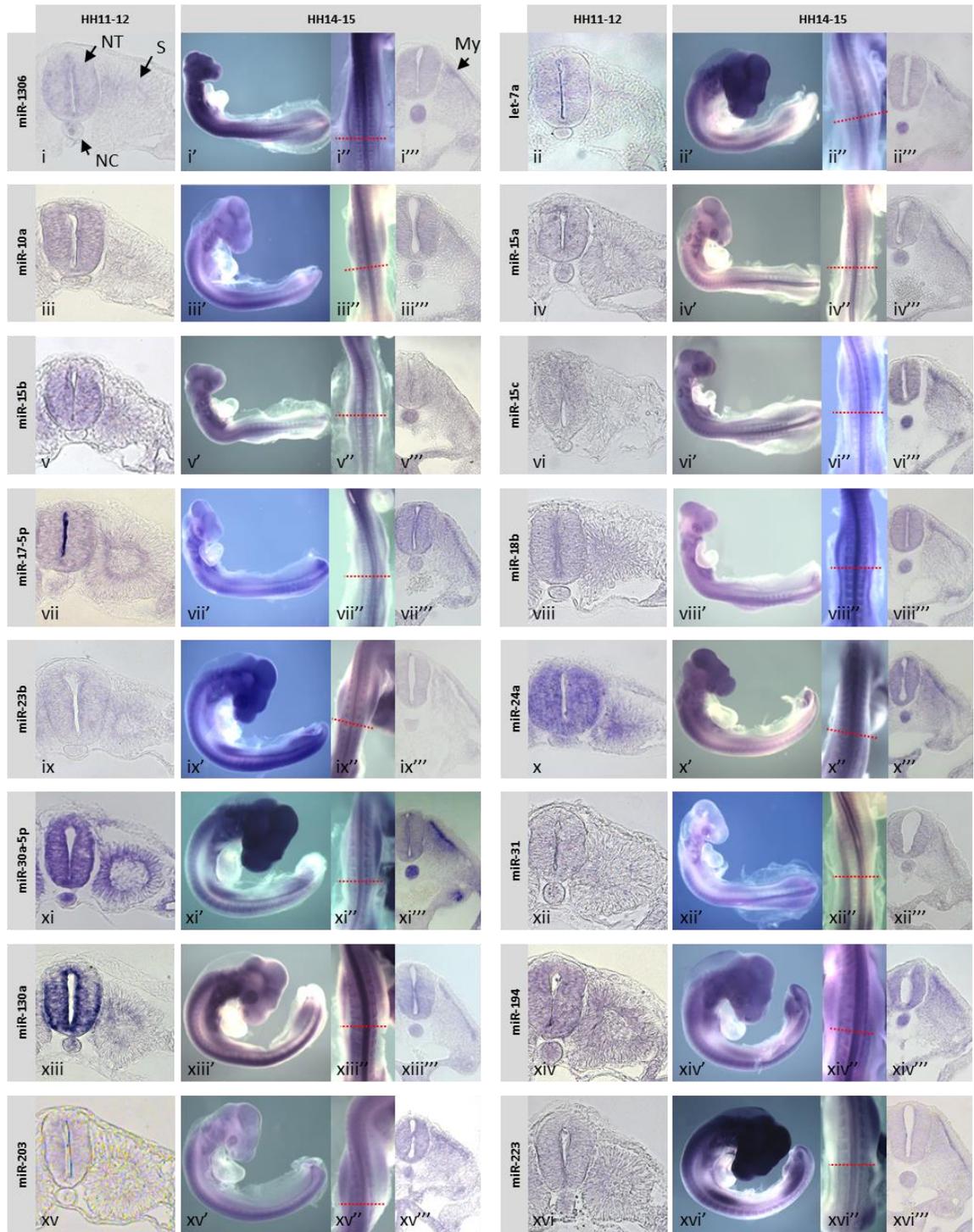
This miRNA was described in the literature as being the product resulting from the regulation of the microprocessor, which is involved in miRNA synthesis and is composed of Drosha and DGCR8 (Ha & Kim 2014). In order to maintain an optimal level of miRNA produced, DGCR8 was shown to stabilise Drosha, whereas Drosha was destabilising DGCR8 mRNA by cleaving it at a hairpin in its second exon (Han et al. 2009; Ha & Kim 2014). This cross-regulatory loop, reported to happen in at least certain cell types (dog peripheral blood (Friedländer et al. 2008); human embryonic stem cells (Morin et al. 2008)), could enable the homeostatic maintenance of the microprocessor activity; this process is deeply conserved throughout the animal kingdom. The mRNA fragment resulting from this Drosha-mediated cleavage, miR-1306, has not been validated as a functional miRNA yet (Ha & Kim 2014).

Chicken WMISH showed that miR-1306 was very weakly expressed at HH11-12 in the neural tube, the notochord and the developing somite (Fig. 3.4i).

In HH15-16 embryos, its expression was widespread in whole-mount (i', i'') (potentially linked to the widespread Drosha-mediated regulation); however on transverse section, miR-1306 was clearly still found in the neural tube and the notochord, as well as in the myotome of the differentiated somite (i'''). This could be explained by a potential stabilisation of miR-1306 in certain tissue.

#### Let-7a:

Let-7 was one of the first miRNAs discovered in *C. elegans* (Reinhart et al. 2000). In mammals, several isoforms for this miRNA exist (isomiRs), and amongst them let-7a. No information was found for this miRNA in databases.



**Fig. 3.4: Expression patterns of 16 selected miRNAs.** LNA WMISH were performed on chicken embryos using *Xenopus* LNA probes. MiRNA expression patterns at HH11-12 in transverse section (i-xvi; 20x magnification), and HH14-15: whole-mount (i'-xvi'), detail of somite expression in dorsal view (i''-xv''), and transverse section (i'''-xvi'''); 20x magnification). Red dotted line indicates the location of the transverse sections. My: myotome; NC: notochord; NT: neural tube; S: somite.

In ISH at HH11-12, let-7a was found weakly expressed in the neural tube (Fig. 3.4ii); expression maintained as the embryos developed. At HH15-16, let-7a was also expressed in the notochord and in the somites, in the myotome (ii'-ii''').

#### MiR-15 family:

MiR-15 family is composed of 3 members: miR-15a, miR-15b, and miR-15c. Some information was available in GEISHA database for miR-15a, indicating an expression in the limb buds; miR-15b was described with a widespread expression; and there was no information for miR-15c.

Although their sequences are very close, with just few differences, they showed slightly different expression patterns (Fig. 3.4iv-vi'''). While miR-15b was already expressed at HH11-12 in the neural tube and developing somite (v), miR-15a was only weakly detected in these tissues (iv), and miR-15c was not even detectable (vi). At HH15-16, they displayed similar expression patterns (iv'-vi'). They were all expressed in the neural tube, the notochord and the myotome (iv''-vi''); miR-15c was the most strongly expressed of the three (vi'''). At later stages they were also all expressed in the limb buds.

#### MiR-24a, miR-30a-5p, and miR-130a:

No information was found in GEISHA database for miR-24a; miR-30a-5p is apparently expressed in somites and limbs (whole-mount data); and widespread expression for miR-130a.

Of the 16 miRNAs studied here, these three miRNAs were the most strongly expressed in the early stage HH11-12. They were found in the neural tube, the notochord and the developing somite (Fig. 3.4x, xi, xiii).

MiR-130a did not appear to be expressed in the most dorsal part of the neural tube as shown in transverse section (xiii). At HH15-16, they were still expressed in the neural tube, the notochord and the differentiated somite, in the myotome (x'-x'''; xi'-xi'''; xiii'-xiii'''). At later stages, miR-24a, miR-30a-5p, and miR-130a were expressed in the limb buds.

#### MiR-10a, miR-17-5p, miR-18b, and miR-194:

There is no information for miR-10a in GEISHA database. MiR-17-5p is found in the surface ectoderm in early stages, in the neural tube and the somite (section), and has a widespread expression at later stages.

MiR-18b and miR-194 are weakly expressed in the somites; despite a widespread expression, they are also found in the limb buds at later stages.

These miRNAs were found weakly expressed in the neural tube and the somite at HH11-12 (Fig. 3.4iii; vii; viii; xiv); expression maintained as the embryo developed. At HH15-16, they were expressed in the myotome (iii''''; vii''''; viii''''; xiv''''); miR-10a (weakly) (iii''') and miR-18b (strongly) (viii''') were expressed in the notochord.

#### MiR-23b, miR-31, miR-203, and miR-223:

From GEISHA database, no specific expression was reported for miR-23b and miR-31; miR-203 and miR-223 have ubiquitous expression.

MiR-31, miR-203 and miR-223 were not expressed at HH11-12 (Fig. 3.4xii; xv; xvi). At HH15-16, miR-203 was found in the neural tube, the notochord and the myotome (xv'''), whereas miR-31 (xii''') and miR-223 (xvi''') were only expressed in the myotome. MiR-23b was found very weakly expressed in the neural tube and the somite (ix), in the myotome (ix'''), at all stages tested. These miRNAs were later expressed in the limb buds.

All 16 miRNAs were found expressed in chicken embryos, in the somites. Although the expression patterns were similar, depending on the stage and the tissue considered, they were not identical. For example, while most of these miRNAs were not detected or very weakly detected at HH11-12, miR-24a, miR-30a-5p and miR-130a were already strongly expressed, indicating potential important role(s) for these 3 miRNAs in neural tube, notochord and somite early in the development. Some of these miRNAs were also expressed in other tissues (see Appendix II Table 2).

### **iii) Characterisation of miR-128**

For this miRNA, both *Xenopus* and chicken probes were available in the laboratory; and miR-128 was shown to be conserved between these two species. However, because the 2 probes were not designed at the same time, the position of the LNA in their sequences (not communicated by Exiqon) might not be the same; this could have possible consequences on their TM and the optimal temperature of hybridisation to use, leading to potential differences in affinity and detection of expression.

A series of *in situ* hybridisations, using both probes, were performed at different temperatures (42°C, 48°C = optimal temperature for the *Xenopus* probes, and 54°C) in order to determine the optimal temperature to use for the chicken probe (good signal vs background); the range of temperature was determined based on the TM of the chicken probe (LNA hybridisation temperature ~ TM - 20°C). It was also a good opportunity to confirm that 48°C was the optimal temperature for the *Xenopus* probe (Fig. 3.5).

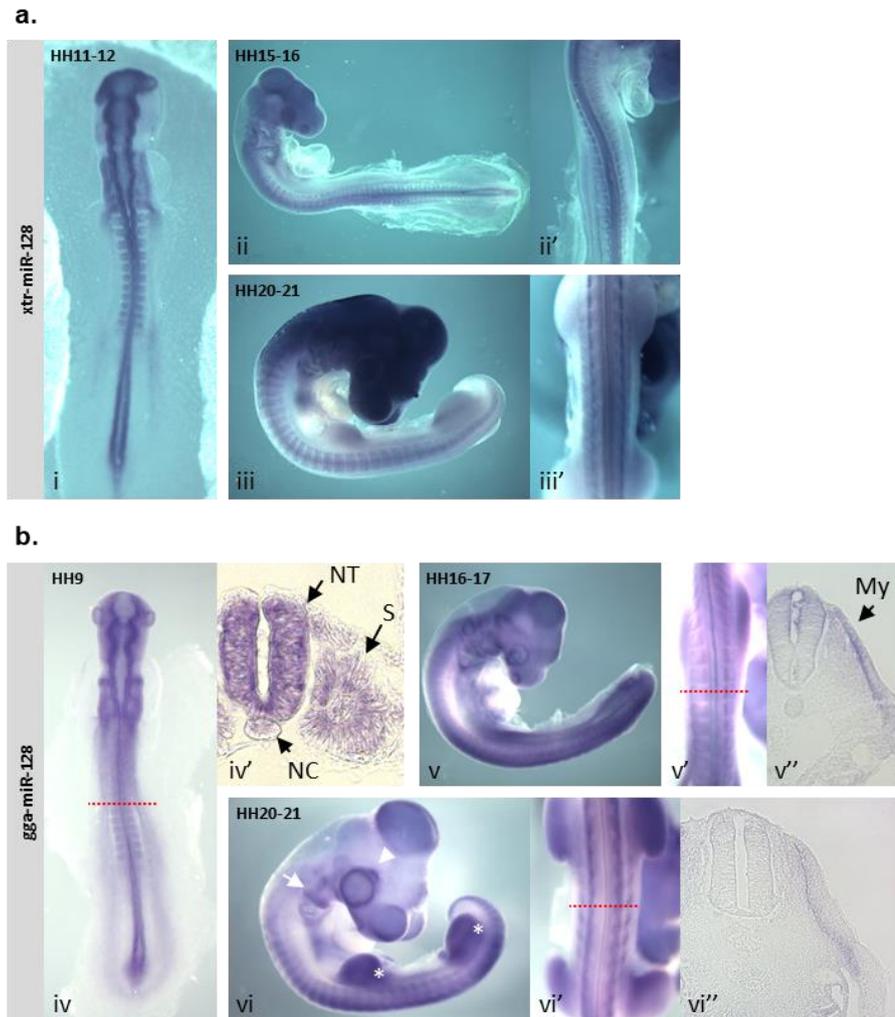
The 2 probes had the best signal (against background), at 48°C, as expected; embryos treated with the chicken probe were cleaner, with less background and a stronger and more specific pattern.

According to the information available in GEISHA database, miR-128 is strongly expressed in the heart at HH13 (only stage showing heart expression), and as the embryos developed it is found in the somites and in the limb buds.

*In situ* hybridisations performed for miR-128 (*Xenopus* and chicken probes), showed no expression in the heart in any of the tested stages (HH10-23; Fig. 3.5i-iii'; iv-vi). This expression of miR-128 in the heart, might happen at a very specific time-point during the development. At HH11-12, miR-128 was found in the neural tube, the developing somite, and weakly in the notochord (iv'). At HH16-17, it was strongly expressed in the myotome, with no expression detected in the notochord and a weak expression in the dorsal part of the neural tube (v''). At HH20-21 (vi), miR-128 was also found in the branchial arches (white arrow), around the eye (white arrowhead) and in the limbs (white asterisk). Its expression in the neural tube was consistent with the fact that miR-128 has been described in other species as a brain-enriched miRNA. However, according to our ISH results, it appears that, in the chicken, miR-128 is only necessary in this tissue in early stages, as it was not found in the neural tube at later stages.

In this first part, 23 miRNAs were studied. Their specific expression patterns in the chicken at different stages of development were confirmed and completed with additional information, like for the myomiRs miR-1a, miR-133a and miR-206; or determined for most of the other miRNAs, like for miR-128, as no, or few expression patterns could be found in the databases.

All the miRNAs tested were expressed in the myotome in somites. Their expression patterns were similar but not identical. This indicates that even with short, and sometimes closely related sequences, LNA technology permits specific detection of miRNAs and their spatiotemporal expression patterns can be determined.



**Fig. 3.5: Expression pattern of miR-128.** (a) MiR-128 expression pattern, using Xenopus LNA probe, at HH11-12 (i), HH15-16 (ii; dorsal view (ii')), and HH20-21 (iii; dorsal view (iii')). (b) MiR-128 expression pattern, using chicken LNA probe, at HH11-12 (iv; transverse section (iv')), HH16-17 (v; dorsal view (v'); interlimb transverse section (v'')), and HH20-21 (vi; dorsal view (vi'')); interlimb transverse section (vi'')). MiR-128 was expressed in the neural tube, the notochord at HH11-12, in the developing somites and in the myotome. At HH20-21, it was also found in the branchial arches (white arrow), around the eye (white arrowhead), and both fore and hind limbs (white asterisk) (vi). Red dotted line indicates the location of the transverse sections. Transverse sections: 20x magnification. My: myotome; NC: notochord; NT: neural tube; S: somite.

The following part of this chapter aimed to identify lists of potential targets for these miRNAs.

### **3.2.2. Identification of microRNA targets**

After having characterised the expression of these 23 miRNAs by LNA ISH, different computational strategies (see chapter 2.18) were used to identify their targets in order to get a better understanding of their functions.

For reasons of clarity, only the analysis done for miR-128 target identification will be presented here. A summarised analysis for the other miRNAs can be found in Appendix II Table 3.

By using bioinformatics tools (DAVID Bioinformatics resources and g:Profiler), this analysis allowed identifying interesting targets, and a particular interest was given to predicted targets with an expression in skeletal muscle, such as for example *Eya4*.

*Eya4* is a member of the EYA family, part of the PSED network. Together with the other members of this network, from the PAX, SIX, and DACH superfamilies, they have been implicated in the regulation of myogenesis (Heanue et al. 1999; Bajard et al. 2006; Buckingham & Relaix 2015). In addition, experiments performed in chicken embryos have shown that the expression of *Pax3* was regulated by miR-1a and miR-206 (Goljanek-Whysall et al. 2011) suggesting that more miRNAs could potentially play a role in the regulation of this network. However, the regulation of the PSED network by miRNAs has not been really investigated.

In chicken, the 3'untranslated region (UTR) sequences of the PSED members – important for miR targeting – have not been completely annotated. So by using the miRanda algorithm, predicted miRNA sites were identified. This will be presented in a second part.

#### **a. *Eya4*: a candidate target for miR-128**

To identify potential target genes of miR-128, TargetScan (release 7.1; June 2016), an online computational program for miRNA target identification and analysis (<http://www.targetscan.org/>), was used and a list of 507 genes was generated.

The molecular functions of miR-128 predicted targets was determined by comparing the results from two analyses:

- Gene Ontology (GO) term and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway annotation analysis in the Database for Annotation, Visualisation and Integrated Discovery tool (DAVID Bioinformatics resources: <https://david.ncifcrf.gov/>) (Huang, Lempicki, et al. 2009; Huang, Sherman, et al. 2009);

- and g:GOSSt (GO Statistics), for pathway enrichment analysis, in g:Profiler, a web server for functional profiling and interpretation of gene lists (<http://biit.cs.ut.ee/gprofiler/>) (Reimand et al. 2007; Reimand et al. 2016). These powerful tools provide a significant amount of information to understand biological meaning behind large lists of genes.

GO term annotation analysis showed that nearly 70% of the 507 miR-128-predicted targets were classified directly as ‘cellular process’, while other GO terms, like ‘biological regulation’ and ‘developmental process’ were also enriched, including 55.3% and 28.8% of the targets, respectively [GOTERM\_BP\_1] (Table 3.2). About a quarter of miR-128 targets were indicated as implicated in the regulation of transcription (125), transcription (102), and regulation of RNA metabolic process (83) [GOTERM\_BP\_FAT] (Table 3.3). KEGG pathway annotation analysis also revealed that some of its predicted targets were involved in the mitogen-activated protein kinase (MAPK) signalling pathway (17), in Insulin signalling pathway (14) and in mTOR signalling pathway (7) [KEGG\_PATHWAY] (Table 3.4).

For example, recent work done by Motohashi *et al.* showed that the miR-128 regulation of IRS1 (insulin receptor substrate 1), involved in IRS1/Akt insulin signalling, had an effect on myogenesis; they demonstrated that miR-128 is able to regulate myoblast proliferation and myotube hypertrophy through IRS1-dependent insulin signalling (Motohashi et al. 2013).

The g:GOSSt analysis performed using g:Profiler tool showed similar results, with more than 70% of miR-128 targets involved in biological regulation (73.9%), regulation of biological process (72.3%) and cellular process (69.4%); nearly 50% of miR-128 targets were classified as involved in ‘developmental process’ (Table 3.5).

CATEGORY	TERM	COUNT	%	P-value
<b>GOTERM_BP_1</b>	<b>Cellular process</b>	<b>322</b>	<b>69.2</b>	<b>2.5E-6</b>
<b>GOTERM_BP_1</b>	<b>Biological regulation</b>	<b>257</b>	<b>55.3</b>	<b>7.2E-9</b>
GOTERM_BP_1	Metabolic process	234	50.3	2.9E-3
GOTERM_BP_1	Multicellular organismal process	144	31.0	1.3E-3
<b>GOTERM_BP_1</b>	<b>Developmental process</b>	<b>134</b>	<b>28.8</b>	<b>7.9E-9</b>
GOTERM_BP_1	Cellular component organisation	97	20.9	1.2E-4
GOTERM_BP_1	Cellular component biogenesis	37	8.0	4.8E-2
GOTERM_BP_1	Locomotion	21	4.5	1.3E-2
GOTERM_BP_1	Growth	10	2.2	6.0E-2

**Table 3.2: GOTERM\_BP\_1 results (465 out of 507 targets were included in the analysis).**

CATEGORY	TERM	COUNT	%	P-value
<b>GOTERM_BP_FAT</b>	<b>Regulation of transcription</b>	<b>125</b>	<b>26.9</b>	<b>8.9E-12</b>
<b>GOTERM_BP_FAT</b>	<b>Transcription</b>	<b>102</b>	<b>21.9</b>	<b>1.2E-9</b>
<b>GOTERM_BP_FAT</b>	<b>Regulation of RNA metabolic</b>	<b>83</b>	<b>17.8</b>	<b>1.5E-4</b>
GOTERM_BP_FAT	Regulation of transcription, DNA-dependent	80	17.2	3.0E-6
GOTERM_BP_FAT	Intracellular signalling cascade	54	11.6	6.4E-4
GOTERM_BP_FAT	Phosphate metabolic process	52	11.2	3.1E-6
GOTERM_BP_FAT	Positive regulation of nitrogen	43	9.2	9.4E-8
GOTERM_BP_FAT	Positive regulation of macromolecule biosynthetic process	42	9.0	3.9E-7
GOTERM_BP_FAT	Positive regulation of transcription	39	8.4	1.8E-7

**Table 3.3: GOTERM\_BT\_FAT results (252 out of 507 targets were included in the analysis).**

CATEGORY	TERM	COUNT	%	P-value
KEGG_PATHWAY	Focal adhesion	17	3.7	3.6E-5
<b>KEGG_PATHWAY</b>	<b>MAPK signalling pathway</b>	<b>17</b>	<b>3.7</b>	<b>9.7E-4</b>
<b>KEGG_PATHWAY</b>	<b>Insulin signalling pathway</b>	<b>14</b>	<b>3.0</b>	<b>2.8E-5</b>
KEGG_PATHWAY	Pathways in cancer	14	3.0	6.8E-2
KEGG_PATHWAY	Neurotrophin signalling pathway	11	2.4	1.0E-3
KEGG_PATHWAY	Regulation of actin cytoskeleton	11	2.4	4.3E-2
KEGG_PATHWAY	Endocytosis	9	1.9	9.1E-2
<b>KEGG_PATHWAY</b>	<b>mTOR signalling pathway</b>	<b>7</b>	<b>1.5</b>	<b>1.8E-3</b>
KEGG_PATHWAY	ErbB signalling pathway	7	1.5	2.2E-2

**Table 3.4: KEGG\_PATHWAY results (465 out of 507 targets were included in the analysis).**

Based on different algorithms, g:Profiler was including more targets in its analysis compare to the DAVID one; a larger number of targets were classified as playing a role in the development in the g:GOSSt analysis (256 (g:GOSSt) vs 134 (GOTERM\_BP\_1)).

The miR-128 targets listed in the category ‘Developmental process’ from the two analysis, GOTERM\_BP\_1 (134) and g:GOSSt (236) were compared; 126 targets were common between the two analysis.

These 126 targets were then used to perform a new GOTERM analysis. More than 50% of these miR-128 targets were found in the brain (65), interesting because miR-128 was described as a brain-enriched miRNA, about 10% in the eye (12), 8% in muscle (10) and less than 2% in the heart (2) [UP\_TISSUE] (Table 3.6).

A closer look was given at the ‘muscle’ category; the 10 predicted miR-128 ‘muscle’ targets, which includes EYA4, are listed in Table 6. For each of these targets, information found in the literature about potential role(s) in muscle are summarised below. No information was found for NUS1.

#### MSTN:

Myostatin, member of the TGF $\beta$  protein family, is the only validated miR-128 target of this list. Shi *et al.* recently showed, in mouse, that by targeting Myostatin, miR-128 was involved in the inhibition of proliferation and the promotion of myoblast cell differentiation (Shi et al. 2015). In chicken, Myostatin is found in somites, in the dermomyotome, and during limb bud development (Amthor et al. 2002). In addition, ectopic expression of Myostatin in the developing limb bud results in a downregulation of Pax3 and Myf5, both associated with proliferation of myogenic cells.

#### BMI1:

In mouse, Bmi1 was found expressed in postnatal myogenic satellite cells, where it has been involved in their maintenance and plays an essential role in repeated muscle regeneration (Robson et al. 2011).

In Bmi1<sup>-/-</sup> mice, a depletion in Pax7<sup>+</sup>/Myf5<sup>-</sup> stem cell population was observed, with a reciprocal increase in Pax7<sup>+</sup>/Myf5<sup>+</sup> committed myogenic progenitor population, leading to a reduction in postnatal muscle fibre size and impaired regeneration upon injury.

CATEGORY	TERM	COUNT	%	P-value
<b>g:GOST</b>	<b>Biological regulation</b>	<b>357</b>	<b>73.9</b>	<b>2.0E-9</b>
<b>g:GOST</b>	<b>Regulation of biological process</b>	<b>349</b>	<b>72.3</b>	<b>1.4E-10</b>
<b>g:GOST</b>	<b>Regulation of cellular process</b>	<b>335</b>	<b>69.4</b>	<b>3.7E-10</b>
g:GOST	Regulation of metabolic process	256	53.0	7.9E-14
<b>g:GOST</b>	<b>Developmental process</b>	<b>236</b>	<b>48.9</b>	<b>3.6E-15</b>
g:GOST	Positive regulation of biological process	206	42.7	6.4E-9
g:GOST	Gene expression	198	41.0	2.8E-7
g:GOST	Regulation of gene expression	187	38.7	6.3E-14
g:GOST	Regulation of developmental process	109	22.6	4.4E-10
g:GOST	Tissue development	83	17.2	1.2E-5

**Table 3.5: g:GOST results (483 out of 507 targets were included in the analysis).**

a.

CATEGORY	TERM	COUNT	%	P-value
<b>UP_TISSUE</b>	<b>Brain</b>	<b>65</b>	<b>51.6</b>	<b>2.0E-2</b>
UP_TISSUE	Placenta	31	24.6	4.1E-2
UP_TISSUE	Epithelium	28	22.2	1.1E-2
UP_TISSUE	Foetal brain	12	9.5	1.4E-2
UP_TISSUE	Eye	12	9.5	7.6E-2
UP_TISSUE	Amygdala	11	8.7	7.5E-3
<b>UP_TISSUE</b>	<b>Muscle</b>	<b>10</b>	<b>7.9</b>	<b>8.9E-2</b>
UP_TISSUE	Frontal cortex	3	2.4	1.5E-2
UP_TISSUE	Foetal lung	3	2.4	7.1E-2
UP_TISSUE	Embryonic heart	2	1.6	5.9E-2
UP_TISSUE	Thyroid carcinoma	2	1.6	7.8E-2

b.

GENE SYMBOL	GENE NAME
BMI1	BMI1 polycomb ring finger oncogene
<b>EYA4</b>	<b>eyes absent homolog 4 (Drosophila)</b>
HOXA10	homeobox A10
MEIS2	Meis homeobox 2
MSTN	myostatin
MYH10	myosin, heavy chain 10, non-muscle
NUS1	nuclear undecaprenyl pyrophosphate synthase 1 homolog (S. cerevisiae)
RORA	RAR-related orphan receptor A
RYBP	RING1 and YY1 binding protein
SPRY2	sprouty homolog 2 (Drosophila)

**Table 3.6: UP\_TISSUE results (a) and list of the 10 ‘muscle’ targets (b).**

In chicken, *Bmi1* is found early in development in the primitive streak; and then in the heart, neural tube, and in somites, in the dermomyotome (Fraser & Sauka-Spengler 2004).

#### HOXA10:

*HoxA10* was shown to play a role in the regulation of the timing of cardiac differentiation by cooperating with *Nkx2-5*, involved in mesodermal patterning (Behrens et al. 2013). It has also been implicated as a regulator for hematopoietic stem cells and erythroid/megakaryocyte development (Magnusson et al. 2007).

In chicken, *HoxA10* is found in somites and limb buds (Alvares et al. 2003).

#### MEIS2:

*MEIS2* plays important role in formation of embryonic brain, eye, heart, cartilage and hematopoiesis. Mouse embryos lacking *Meis2* display defects in tissues derived from neural crests, such as abnormal heart outflow tract, and cranial nerves (Machon et al. 2015). Perturbations in craniofacial skeleton development were also observed, with anomalies in cranial bones and cartilages. Cecconi *et al.* also showed that *Meis2* plays a role in the cascade of induction leading to somitic mesoderm differentiation as well as in brain regionalisation (Cecconi et al. 1997).

In chicken, *Meis2* is found in the somites and limb buds (Sánchez-Guardado et al. 2011).

#### RYBP:

*Rybp* was implicated in transcriptional regulation, apoptotic signalling and, as a member of the polycomb repressive complex 1 (PRC1), in regulating pluripotency and differentiation of embryonic stem cells (ESCs). As well as playing an important role in mouse brain development, it has been identified as a critical regulator of heart development (Ujhelly et al. 2015).

In addition, work done by Zhou *et al.* showed that *Rybp*, as a negative regulator of skeletal myogenesis, is targeted by miR-29, a pro-myogenic miRNA, in order to downregulate its expression during myogenesis (Zhou et al. 2012).

#### RORA (or ROR $\alpha$ ):

The only information available for ROR $\alpha$  was found in mouse, where it has been involved in regulating the Akt2/adenosine monophosphate (AMP)-activated protein kinase (AMPK) signalling pathways in the context of lipid homeostasis in skeletal muscle (Raichur et al. 2010).

#### MYH10:

MYH10 encodes for non-muscle myosin heavy chain II B (NMHC-IIB); member of the non-muscle myosin IIs (NM IIs), a group of ubiquitously expressed proteins. NMHC-IIB is critical for cardiac and brain development (Ma & Adelstein 2014); it is expressed in cardiac myocytes, and enriched in neuronal tissue where it is thought to be important in neuronal migration.

In mouse, the knock-out of NMHC-IIB lead to embryonic death (E14.5) with severe cardiac defects and neurodevelopmental disorders. This phenotype was also observed in zebrafish (Huang et al. 2013; Gutzman et al. 2015) and in human (Tuzovic et al. 2013).

#### SPRY2:

In chicken, Spry2 is found in early stages in the primitive streak, in the neural tube, mesoderm of the branchial arches, retina, in somites in the myotome, and in later stages in the limb buds (Chambers & Mason 2000).

In addition, Abu-Elmagd *et al.* recently showed that overexpression of Spry2 results in reduction of somite myogenesis indicated by a loss of MyoD expression; suggesting that Spry2 could play a crucial role during chicken myogenesis by regulating myogenic cell proliferation (Abu-Elmagd et al. 2015).

#### EYA4:

In mouse, Eya4 is strongly expressed in skeletal muscle. It is found in somites, in the dermomyotome (Y. Zhang et al. 2004). Borsani *et al.* reported that at limb level, Eya4-positive cells appear to be migrating away from the dermomyotome into the limb structures in a pattern resembling that of migrating muscle precursor cells (Borsani et al. 1999); suggesting a potential role for Eya4 in limb muscle development.

In addition, *eya4* has also been reported as being important in the regulation of Na<sup>+</sup>/K<sup>+</sup>-ATPase, essential for zebrafish heart development. In *eya4* morphant fish, Na<sup>+</sup>/K<sup>+</sup>-ATPase level was decreased and amongst other phenotypes, heart failure was observed (L. Wang et al. 2008).

EYA4 is also known to be part of the PAX-SIX-EYA-DACH (PSED) network of transcriptional regulators acting early in the myogenesis, and upstream of the MRFs.

### **b. Identification of microRNAs targeting PSED members**

EYA4 is a member of the PSED network, and, as shown in the previous part, it appears that its expression might be regulated by miR-128. But what about the other members of this network?

The 3'UTR information and annotations found in databases, like TargetScan, are sometimes incomplete and often based on human sequences, which is most likely not highly conserved with the chicken (miRNA response elements (MREs) might be conserved, but not the surrounding sequence). In chicken, the 3'UTR sequences of the PSED members – important for miR targeting – have not been completely annotated. So by using the miRanda algorithm, predicted MREs were identified.

The miRanda algorithm, developed by Enright *et al.*, is an interesting computational tool for the prediction of miRNA targets (Enright et al. 2003; Betel et al. 2008).

For each miRNA, targets are selected on the basis of three properties: (1) sequence complementarity using a position-weighted local alignment algorithm taking into account moderate mismatches and complementarity at the 5' end (seed sequence location); (2) free energies of miRNA-mRNA duplexes with calculation of optimal interaction; and (3) conservation of target sites in related genomes including UTR matches between species, nucleotide identity, and equivalent target site positions according to a cross-species UTR alignment.

In order to identify potential MREs in the 3'UTRs of PSED members, the miRanda algorithm was used; it was run by Simon Moxon (Earlham Institute, UK).

a.

```

=====
Performing Scan: gga-mir-128 (21 nt) vs gga-Eya4-3UTR (3962 nt)
=====

Score: 130.000; Query: 2-19; Position: 493-513
Alignment: Length: 17; Identity: 58.82%; Similarity: 70.59%

      gga-mir-128:  3' uuuCUCUGGCCAAGUGACACu 5'
                    || | : |:|||||
      gga-Eya4-3'UTR: 5' gatGAAAATAACTTACTGTGa 3'

Energy: -11.480 kCal/Mol

```

b.

MiRNA	3'UTR						
	Eya1	Eya2	Eya3	Eya4	Six1	Six4	Dach1
gga-miR-1306-3p	NO	NO	NO	YES	NO	YES	YES
gga-let-7a-3p	YES	NO	NO	YES	YES	YES	YES
<b>gga-miR-1a</b>	<b>YES</b>	<b>YES</b>	<b>YES</b>	<b>YES #</b>	<b>YES</b>	<b>YES</b>	<b>YES</b>
gga-miR-1b	NO	YES	NO	YES #	YES	YES	YES
gga-miR-10b	YES #	YES	NO	NO	NO	YES	YES
gga-miR-15a	YES #	NO	YES	YES	YES	YES	YES
gga-miR-15b-5p	YES #	NO	YES	YES	YES	YES	YES
gga-miR-15c-5p	YES #	YES	YES	YES	YES	YES	YES
gga-miR-17-5p	YES	YES	NO	YES	YES	YES	YES
gga-miR-18b	YES	YES	YES	YES	YES	YES	YES
gga-miR-23b	YES #	YES	YES	YES	NO	YES	YES #
gga-miR-24a	YES	YES	YES	YES #*	NO	YES	YES
<b>gga-miR-27b-3p</b>	<b>YES #</b>	<b>YES</b>	<b>YES</b>	<b>YES #</b>	<b>YES</b>	<b>YES</b>	<b>YES</b>
gga-miR-30a-5p	NO	NO	NO	YES	NO	YES #	YES
gga-miR-31	YES	YES	NO	YES	NO	YES	YES #
<b>gga-miR-128</b>	<b>YES</b>	<b>YES</b>	<b>NO</b>	<b>YES #</b>	<b>NO</b>	<b>YES</b>	<b>NO</b>
gga-miR-130a	YES	NO	NO	YES	NO	YES	YES
<b>gga-miR-133a</b>	<b>NO</b>	<b>YES</b>	<b>YES</b>	<b>YES #</b>	<b>NO</b>	<b>YES</b>	<b>YES</b>
gga-miR-133b	NO	YES	YES	YES #	NO	YES	YES
gga-miR-133c	NO	YES	YES	YES #	NO	YES	YES
gga-miR-194	YES	YES	YES	YES	YES	YES	YES
gga-miR-203	YES	YES	YES	YES	YES	YES	YES #
<b>gga-miR-206</b>	<b>NO</b>	<b>NO</b>	<b>YES</b>	<b>YES #</b>	<b>YES</b>	<b>YES</b>	<b>YES</b>
gga-miR-223	YES	NO	YES	YES	NO	YES	YES
<b>gga-miR-499-5p</b>	<b>YES</b>	<b>NO</b>	<b>NO</b>	<b>YES #</b>	<b>NO</b>	<b>YES</b>	<b>YES</b>

**Table 3.7: MiRanda analysis.** (a) Example of raw miRanda data (miR-128 against Eya4 3'UTR). Score: alignment score based on complementarity between the sequences. The alignment score threshold used was 125. Only alignments with score greater or equal to this value have been considered for further analysis. Position of the interaction, conservation between sequences (identity and similarity percentages), and free energy score (the smaller, the better) were also indicated. (b) Table summarising the results generated from the miRanda analysis. Each miRNA was used to scan the 3'UTR sequences of chicken Eya1 [ENSGALT00000025181.4], Eya2 [ENSGALT00000007180.4], Eya3 [ENSGALT00000001127.4], Eya4 [ENSGALT00000022662.4], Six1 [NM\_001044685.1], predicted Six4 [XM\_003641442.2], and Dach1 [ENSGALT00000027373.3]. #: MRE annotated in human sequence (TargetScan 'human'), and conserved in chicken (TargetScan 'chicken'). MiRNAs of particular interest for this project are indicated in bold.

The 3'UTR sequences of Gga-Eya1, Gga-Eya2, Gga-Eya3, Gga-Eya4, Gga-Six1, Gga-Six4, and Gga-Dach1 were collected from Ensembl and NCBI, and the miRNA sequences, from miRBase.

The miRNAs previously selected for their expression in skeletal muscle (somite), and presented in section 3.2.1, were used for this analysis; two other miRNAs were added to the initial selection of 23: miR-27b (it has the same seed sequence as miR-128), and miR-499 (a cardiac myomiR). The complete list of miRNAs can be found in Appendix II Table 1.

The generated data from the algorithm run were then analysed, and the results are presented in Table 3.7.

These results were then compared to the information already available in TargetScan. In TargetScan, human sequences are the best annotated and therefore the most detailed; alignment with 3'UTRs of other species are also provided showing the position of the MREs in human sequences and the corresponding MRE positions in other species, when conserved.

However, 3'UTR sequences are usually not highly conserved between human and chicken: sometimes sequences are different, MRE positions are different, and sometimes only a part of the MRE sequence is conserved. Adding to the complexity of miRNA/target interaction identification, TargetScan 'chicken' does not always use the same sequence as the one used for alignment in TargetScan 'human'. Moreover, in some cases, 'chicken' MRE identified in TargetScan 'human' alignments, cannot be found in TargetScan 'chicken'.

In Table 3.7b, 'chicken' MRE predicted by miRanda, annotated in human sequence and conserved in chicken (TargetScan 'human' and TargetScan 'chicken') are indicated by #.

Most of the interesting sites predicted by TargetScan were also found by miRanda; it was the case, for example, for Gga-Eya4 3'UTR with miR-1/206, miR-27b/128, miR-133 and miR-499 sites. However, due to differences in their algorithms and in alignment score threshold used (threshold set at 125 for this analysis), the miRanda algorithm was able to predict additional miRNA sites.

Specific miRNAs were focused on, such as the myomiRs miR-1a, miR-133a, miR-206, and miR-499, but also miR-128 and miR-27b as they have the same seed sequence; they are indicated in bold in Table 3.7b.

MiR-27b/miR-128 sites were identified in the *Eya1* 3'UTR sequence; a miR-133 site was also found, but only predicted by TargetScan.

No 3'UTR sequence was available in TargetScan chicken for *Gga-Eya2*. An Ensembl sequence was found and used for miRanda analysis, and predicted sites for miR-1, miR-27b/128, and miR-133a were identified. With no annotation of its 3'UTR sequence in TargetScan, *Gga-Eya3* was predicted by miRanda analysis to be targeted by miR-133a. *Gga-Six1*, *Gga-Six4*, and *Gga-Dach1* were predicted to be the target of miR-1a and miR-206; as well as miR-133a and miR-499 for *Gga-Six4*. Interestingly, while *Gga-Dach1* was predicted to be a target of miR-27b, it did not seem to be a miR-128 predicted target even though these two miRNAs have the same seed sequence.

These results show that several miRNAs, such as the myomiRs and miR-128, might be able to target several members of the PSED network.

### **3.3. Conclusion**

In this chapter, 23 selected miRNAs have been studied. Their expression patterns, characterised by performing LNA ISH, have been established at different stages during chicken embryonic development. All expressed in skeletal muscle, they displayed similar but also specific patterns in the somites, as well as other tissues for some of them. One of these miRNAs, miR-128, gave interesting results. At HH11-12, miR-128 was expressed in the neural tube and developing somites. At HH15-16 and onwards, its expression in the neural tube disappeared; miR-128 was mostly expressed in skeletal muscle tissues, such as differentiating somites, myotome, and the limb buds.

In order to better understand 'muscle' miRNA function(s), and miR-128 function(s) in particular, predicted miR-128 targets were collected from TargetScan (507). By using a combination of different computational tools (DAVID and g:Profiler), 'muscle' targets were identified (n=10); amongst them, *Eya4*, member of the PSED network, was of particular interest.

*Eya4*, as well as members of the PSED network (*Six1/4*, *Eya1/2/3*, and *Dach1*), were investigated further, due to the interesting role(s) they seem to play with *Pax3/7* during skeletal myogenesis. The miRanda algorithm was used to scan selected miRNA sequences against their 3'UTR sequences; and locations of predicted miRNA sites were determined.

This analysis showed that most of the members of the PSED network could potentially be targeted by miRNAs; most importantly, miR-128 and myomiR sites were found in the 3'UTR sequences of most of these PSED members.

Although computational approaches help narrow down the number of targets for a miRNA, they are only predictions. Interactions between miRNAs and mRNA targets still need to be validated by performing *in vitro* and *in vivo* experiments.

In chapter 4, some of the miRNA/mRNA target interactions identified in chapter 3 are going to be investigated further. By using molecular cloning strategies, constructs will be generated in order to (1) determine expression patterns of relevant targets, and (2) validate miRNA/mRNA interactions by performing *in vitro* experiments.

## CHAPTER 4: CHARACTERISATION OF MicroRNA TARGETS AND *in vitro* INTERACTION INVESTIGATION

### 4.1. Introduction:

In chapter 3, relevant microRNA (miRNA) targets have been identified by using bioinformatics tools. Although prediction algorithms have become more precise and efficient at identifying miRNA targets, it is essential to experimentally validate these miRNA/mRNA target interactions.

In this chapter, two important aspects are going to be developed: characterisation of miRNA targets, and investigation of miRNA/mRNA target interactions using an *in vitro* model.

MiRNAs like the myomiRs (miR-1a, miR-206, and miR-133a) and miR-128 were predicted to target members of the PAX-SIX-EYA-DACH (PSED) network, which has been shown to play key regulatory roles in skeletal muscle development (Relaix & Buckingham 1999).

Pax3 and Pax7 have been shown to be directly targeted by miR-1 and miR-206 leading to delayed myogenic differentiation in developing somites, indicated by a transient loss of MyoG (Myogenin) expression (Hirai et al. 2010; Goljanek-Whysall et al. 2011). In addition, miR-1 and miR-206 also play a role in facilitating the differentiation of satellite cells, adult muscle stem cells, through regulation of the transcription factor Pax7 (Chen et al. 2010); Goljanek-Whysall *et al.* showed that a sustained expression of miR-1 and/or miR-206 targets resulted in increased proliferation and inhibition of myogenesis in mouse myoblast C2C12 cells (Goljanek-Whysall, Pais, et al. 2012).

We focused our investigations on Eya4 and the other members of the PSED network, Six1/4, Eya1-3, and Dach1/2; together they constitute the Six-Eya-Dach cascade.

### **Eya4 and the other PSED members:**

Similar to Pax3/7, the transcription factors Six1/4, Eya1/2/4, and Dach1/2 have been shown to synergistically regulate myogenesis and play a key role in the migration of myogenic precursors (Heanue et al. 1999). Moreover, Heanue *et al.* showed that ectopic expression of Six1 and Eya in chicken embryo resulted in the activation of Pax3 and the myogenic regulatory genes, suggesting an upstream function in myogenesis for these factors (Heanue et al. 1999).

#### Six1/4:

In vertebrate embryos, Six1 and Six4 are found in neural placodes, dorsal root ganglia, limb bud mesenchyme, and in migrating myogenic precursors (Oliver et al. 1995; Esteve & Bovolenta 1999; Fougousse et al. 2002; Grifone et al. 2005). They are also co-expressed in the newly formed somites, developing dermomyotome, and the myotome (Grifone et al. 2005; Wu et al. 2014).

In the chicken embryo, at HH4, Six1 is expressed in the non-somitic head mesoderm and the pre-placodal ectoderm. At HH8-HH10, Six1 is also found in the developing somites, and in differentiating somites, in the dermomyotome. At later stages, Six1 is expressed in the dorsomedial and ventrolateral lips of the dermomyotome and the myotome (Heanue et al. 1999; Berti et al. 2015). Six4 is expressed at HH4 where it is found in the ectoderm surrounding the developing axial midline. At stage HH8-HH12, it is expressed in the pre-placodal ectoderm, the newly formed optic vesicle, the otic, olfactory and neural placodes, and in the eyes. From HH14, Six4 is found in the trigeminal ganglia and developing limb buds, as well as in the notochord. Six4 is also strongly expressed in the paraxial mesoderm and the developing somites, first in the dorsal portion, then the dermomyotome and become finally limited to the myotome (Esteve & Bovolenta 1999).

Six4 knock-out mice have no developmental defects (Ozaki et al. 2001), while Six1 knock-out mice die at birth and show multiple organ developmental defects, including kidney, thymus, ear and rib, craniofacial and muscle deficiencies (Ozaki et al. 2004; Laclef et al. 2003). However, Six1/Six4 double knock-out mice show an aggravation of the phenotype previously reported for the single Six1 knock-out (Grifone et al. 2005). Six1/Six4 double knock-out mice are characterised by severe craniofacial, rib and muscle defects.

In addition, at the limb bud level, in the absence of Six1/4 in the myotome, no muscle is detected. Myogenic progenitor cells are lost and the expression of Pax3 in the hypaxial dermomyotome is lacking; the expression of the MRFs MyoG and MyoD1 is impaired, and Mrf4 expression becomes undetectable (Grifone et al. 2005). These data suggest that Six1/4 is an upstream regulatory factor of Pax3 and is essential for the genesis of muscle progenitors. Six1 is also expressed at high levels in adult skeletal muscle where it participates, in synergy with Eya1, in the establishment of the fast/glycolytic phenotype of the myofibre (Grifone et al. 2004; Niro et al. 2010; Richard et al. 2011). More recently, in zebrafish, O'Brian *et al.* found a microRNA-mediated regulatory mechanism for Six1; miR-30a regulates myogenesis via direct targeting and inhibition of Six1a/b expression (O'Brien et al. 2014).

#### Dach1/2:

In vertebrates, DACH1 and DACH2 are expressed in similar tissues to those observed in *Drosophila* (Davis et al. 1999). DACH1 and DACH2 are detected in multiple adult human tissues including kidney and heart. In mouse, Dach1 is expressed in the developing kidneys, eyes, and ear; it is also found in the somites, the anterior and proximal mesenchyme and the apical ectodermal ridge (AER) of the limb buds, as well as in gut and heart (Ayres et al. 2001; Heanue et al. 2002). Mouse Dach2 displays a similar pattern to those of Dach1 suggesting potential redundant roles for these genes during development (Davis, Shen, Sandler, Heanue, et al. 2001).

In chicken, Dach1 is found expressed in the developing eye and ear, and in the neural tube (Heanue et al. 2002; Kida et al. 2004; Litsiou et al. 2005); in the limb buds, Dach1 is expressed in migrating myoblast precursors. Dach1 is not expressed in the early stages of limb development (Heanue et al. 1999); its expression starts to be detected at HH20 in the AER, suggesting a role in the maintenance of the AER rather than in its initiation (Kida et al. 2004). In early epithelial somites, Dach2 is expressed dorsally as well as in the dorsal neural tube and in the intermediate mesoderm. In the differentiating somites, Dach2 is detected throughout the dermomyotome; it is also found in the nephrogenic ducts. At the limb level, Dach2 is also found in the migrating hypaxial myoblast precursors (Heanue et al. 1999).

Dach2 mutant mice are viable and fertile, and they do not exhibit gross defects in eye development or brain function, and although Dach1 mutant die postnatally, these mutants seem to have a normal development (Davis et al. 2006).

The lack of significant phenotype in *Dach1* mutants may be due to functional redundancy with *Dach2* as their expression profiles overlap in many tissues. Interestingly, *Dach1/Dach2* double mutant mice die after birth with a similar phenotype to *Dach1* homozygotes. Unlike *Drosophila dachshund* mutants that lack eyes and exhibit leg truncations, the eyes and limbs of *Dach* double mutant mice are present, suggesting potential differences between *Dach* and *dachshund* gene function during embryonic eye and limb formation (Davis et al. 2006).

*Eya1/2/3*:

Widely expressed during development, in mouse, *Eya1* and *Eya2* are co-expressed in the dermomyotome, and later in the myotome; they are also found in developing limbs, migrating muscle precursors and tendons (Xu, Cheng, et al. 1997). Similar expression patterns have been reported in *Xenopus*, and zebrafish; however, in chicken, *Eya1* and *Eya2* display differential expression profiles (Ishihara et al. 2008).

In chicken, *Eya1* and *Eya2* are expressed early during development. *Eya1* is found in the primitive streak at HH4, and appears in the ectoderm, the mesenchyme and the somites at HH6-HH11; at HH11, *Eya1* is expressed in the dermomyotome. At HH17 and HH19-20, in the trunk region, *Eya1* is detected in the dorsomedial and ventrolateral lips of the dermomyotome and in the myotome (Berti et al. 2015). In contrast, *Eya2* is already expressed in the endoderm at HH4, and is found in the endoderm and cranial placodes at HH6-HH11, earlier than *Eya1* (Ishihara et al. 2008). At HH11, *Eya2* is weakly expressed in the somites. At HH15, *Eya2* is only expressed in a restricted region of the olfactory placode; at the trunk level, *Eya2* is expressed in the myotome, and throughout the entire newly formed somites. At HH17, *Eya2* is restricted to the myotome in the somites at the anterior region of the trunk, whereas the whole somite region is positive for *Eya2* in the posterior region (Heanue et al. 1999). At later stage, *Eya2* is also found in the limb buds and the eyes (Mishima & Tomarev 1998). *Eya1* and *Eya2* are expressed in distinct locations of the chicken embryo with little overlap, suggesting distinct and unique functions in chicken early development.

In human, *EYA1* gene mutations have been associated with the dominant inherited branchio-oto-renal (BOR) syndrome, which alter the formation of branchial derivatives, ear and kidney (Vincent et al. 1997; Abdelhak, Kalatzis et al. 1997a; Abdelhak, Kalatzis et al. 1997b). *Eya1*-deficient mice show similar phenotypes with ear, kidney and branchial organ defects (Xu et al. 1999), and delayed myogenesis (Grifone et al. 2007).

Eya1 mutants of zebrafish also display defective development of the inner ear and lateral line (Kozlowski et al. 2005).

In mouse and zebrafish, Eya3 showed an abundant and widespread expression throughout development in brain, eyes, heart, somites and limbs. This is in contrast to the restricted expression pattern observed in *Xenopus* embryos (Söker et al. 2008).

Eya3-deficient mice showed minor phenotypes with weak effects on respiratory, heart and muscle function, with a decrease in locomotion activity. *In vitro* experiments showed that Eya3 was important for cell-autonomous proliferation of murine C2C12 cells (Li et al. 2003). There is no information for Eya3 in chicken.

Eya4:

By northern blot performed on various mouse adult tissues, Eya4 was found expressed in skeletal muscle; data consistent with the fact that a large number of EYA4 cDNAs were isolated from human skeletal muscle cDNA library, suggesting a potential similar expression pattern in human (Borsani et al. 1999). In the developing mouse embryo, Eya4 is expressed in the craniofacial mesenchyme, the dermomyotome and at later stages in the limbs. Expressed in the nasal placode and the otic vesicle at E9.5; from E10.5, Eya4 is found in the branchial arches and the somites. At E11.5, Eya4 is strongly expressed in the region of the somites, in the dermomyotome, and in cells migrating away from the dermomyotome to populate the limbs. Like Eya1-3, Eya4 is widely expressed during development; however, there is no evidence for expression of Eya4 in the developing eye, contrary to what was observed for Eya1, Eya2, and Eya3 (Borsani et al. 1999; Xu, Woo, et al. 1997). In human and zebrafish, EYA4 has also been shown to be involved in cardiac processes (Schönberger et al. 2005).

Mutations of the EYA4 gene resulting in truncated EYA4 proteins have been associated with the human autosomal non-syndromic sensorineural hearing loss (SNHL), which may be associated with dilated cardiomyopathy (DCM) (Wayne et al. 2001; Y. Zhang et al. 2004; Schönberger et al. 2005; Hildebrand et al. 2007; Makishima et al. 2007). Depreux *et al.* produced Eya4-deficient mice and showed that these mice have severe hearing deficits, similar to those observed in human (Depreux et al. 2008).

Injection of morpholino oligonucleotides against *eya4* into zebrafish embryos indicated an abnormal morphological and physiological phenotype of the heart (Schönberger et al. 2005; L. Wang et al. 2008). Surprisingly, no *Eya4* transcripts were detected in mouse embryonic and adult heart and no heart defects have been reported in the *Eya4*<sup>-/-</sup> knock-out mice (Borsani et al. 1999; Depreux et al. 2008). These different phenotypes could suggest that *Eya4* functions may not have been conserved during evolution.

One conserved feature of *EYA4* is its expression in skeletal muscle tissue, although no muscle alteration has been yet associated to *EYA4* mutations (Borsani et al. 1999; Schönberger et al. 2005). Because the other three *EYA* genes are also expressed in this tissue (Xu, Woo, et al. 1997; Heanue et al. 1999; Söker et al. 2008; Berti et al. 2015), the absence of muscle phenotype may be due to functional redundancies. This is consistent with the results observed in *Eya1/Eya2* double mutant mice; *Eya1/Eya2*-deficient mice have no diaphragm and present severe limb muscle hypoplasia (Grifone et al. 2007). There is no information for *Eya4* in the chicken.

In vertebrates, *SIX1/4*, *EYA1-4*, and *DACH1/2* are, like *PAX3/7*, all expressed in cells prior to skeletal muscle differentiation, and their overlapping expression continues in skeletal muscle derivatives where the cells are maintained in an undifferentiated state. Acting together upstream of the myogenic regulatory factors, they regulate early phases of skeletal myogenesis.

The regulatory PSED network plays a very important role in the regulation of myogenesis. However, in the chicken, not all the members of this network have been characterised and their function(s) investigated. Moreover, little is known about the potential miRNA regulation of the PSED network.

Found in human, mouse, and zebrafish, *Eya4* has not been characterised in the chicken. By using molecular cloning strategies, fragments of the coding sequence and the 3' untranslated region (UTR) of chicken *Eya4* (*Gga-Eya4*) were cloned. They were used to (1) determine its expression pattern and (2) perform luciferase reporter assays in order to investigate the ability of miR-128, in particular, but also other miRNAs like miR-1a, miR-206, miR-133a, and miR-499, previously identified in chapter 3, to target *Gga-Eya4* and regulate its expression *in vitro*. Results will be presented in the first part of this chapter.

The same approach was used to identify miRNA/target interactions for the other PSED members, Gga-Six1/4, Gga-Eya1-3, and Gga-Dach1/2. Although most of these genes have been studied and their expression patterns determined in several species, such as mouse, chicken and zebrafish, RNA probes were generated and whole-mount *in situ* hybridisation (WMISH) performed in order to confirm, complete or determine their specific expression profiles. Results will be presented in a second part.

## **4.2. Results and discussion**

### **4.2.1. Validation of Eya4 as a target for miR-128**

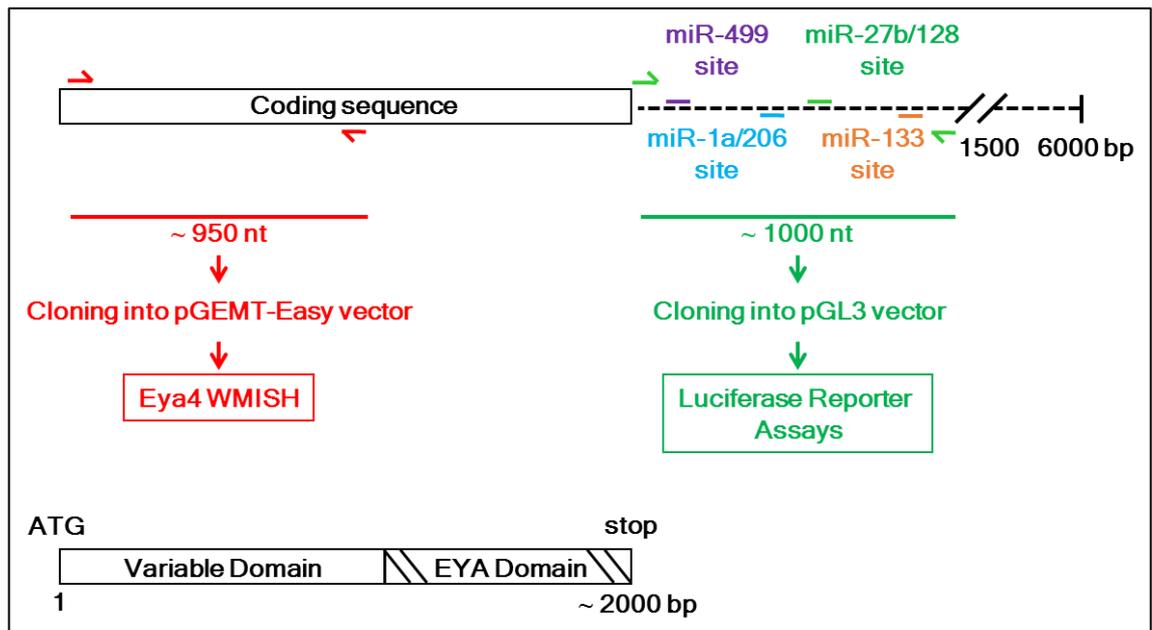
In order to study the interaction between miR-128 and Eya4, the first step was to clone Gga-Eya4. At the time this cloning was done, the chicken genome sequence was not completely annotated; Eya4 sequence was pulled out from the chicken Genebuild provided by Ensembl. This sequence, generated using the most reliable data available in the databases, contains the coding sequence, and the 5' and 3'UTR sequences; Gga-Eya4 sequence is now referenced on Ensembl as ENSGALT00000022662.4. The strategy used for the cloning of Gga-Eya4, coding sequence and 3'UTR fragments, is summarised in Fig. 4.1.

#### **a. Cloning of a fragment of the coding sequence of Gga-Eya4**

In order to characterise the expression profile of Gga-Eya4, it was first necessary to clone all or part of its coding sequence.

The sequences of human, mouse and Xenopus for EYA4 were collected from Ensembl and NCBI, and by performing sequence alignment, conserved portions between the different sequences were identified. A conserved portion in the EYA domain was used to scan against the chicken Genebuild (Ensembl) allowing extracting a sequence, being potentially Gga-Eya4.

A protein sequence alignment with the predicted Gga-Eya4, and those of human, mouse, and Xenopus, as well as an alignment with the three other members of the chicken Eya family, were performed.



**Fig. 4.1: Cloning of Gga-Eya4.** (Top panel) Schematic of Gga-Eya4 transcript (6000 base pairs (bp)) representing the coding (box) and 3'UTR (dotted line) sequences. Red and green arrows indicate the position of the pairs of primers used to clone the fragments of coding and 3'UTR sequences, respectively. The fragment of the coding region (~950 nucleotides (nt)) was used to perform whole-mount *in situ* hybridisation (WMISH). The fragment of the 3'UTR region (~1000 bp) was used to do luciferase reporter assays. The positions of miR-27b/128, miR-1a/206, miR-133a, and miR-499 sites are indicated on the 3'UTR sequence. (Bottom panel) Structure of Gga-Eya4 protein (~2000 bp) with variable domain and, at the C-terminal end, the conserved EYA domain.

The percentage of identity and similarity between protein sequences was determined using The European Molecular Biology Open Software Suite (EMBOSS) Needle software ([http://www.ebi.ac.uk/Tools/psa/emboss\\_needle/](http://www.ebi.ac.uk/Tools/psa/emboss_needle/)) (Hancock et al. 2004; Li et al. 2015); results are presented in Table 4.1.

It appeared that the predicted Gga-Eya4 protein shares a lower percentage of identity with its homologous Gga-Eya1, Gga-Eya2, and Gga-Eya3 (Table 4.1a), than with its human, mouse, and *Xenopus* orthologous (Hsa-EYA4, Mmu-Eya4, Xla-eya4) (Table 4.1b). With 72.1% identity, the predicted Gga-Eya4 is closer to Gga-Eya1, than to Gga-Eya2 and Gga-Eya3, which have 52.7% and 48.8% identity, respectively (Table 4.1a). The predicted Gga-Eya4 shows more than 85.8% identity with Xla-eya4, 91.3% with Mmu-Eya4, and up to 94.7% with Hsa-EYA4 (Table 4.1b).

EYA4 proteins have been highly conserved across species during evolution. This is particularly true in the EYA domain localised in the C-terminal region; this domain being common to all the EYA family members. Based on nucleotide alignment (not shown), a pair of primers was designed according to the following criteria: ~ 20 base pairs (bp) with about 50% of GC, generating a product of approximately 800-1000 bp in length, and localised in highly conserved sequence, but outside the EYA domain. These criteria have been defined in order to maximise the chances to specifically amplify Gga-Eya4 (Fig. 4.1).

Total RNAs, from whole chicken embryos (day 2, day 3, and day 4), dissected somites (day 3 and day 4), as well as dissected forelimbs and hindlimbs from day 7 embryos, were *in vitro* transcribed (see chapters 2.8 and 2.9; Appendix III). Generated cDNAs, with the specifically designed Gga-Eya4 pair of primers, were then used in a polymerase chain reaction (PCR) (see chapter 2.10); PCR products of about 950 bp were amplified from all the tissue samples tested. The product from the 'forelimbs day 7' cDNA sample was used to do the rest of the experiments as it was the one with the strongest expression (data not shown).

After excision from gel and purification (see chapters 2.11 and 2.13), the PCR product was ligated into pGEMT-Easy vector (Promega), a linearised vector with single 3' terminal thymidine at both ends providing a high efficiency of insertion (see chapter 2.14). DH5 $\alpha$  competent cells were transformed with the construct (see chapter 2.16).

a.

ID/SIM (%)	Gga-Eya1	Gga-Eya2	Gga-Eya3	Gga-Eya4
Gga-Eya1	100	54,5/62,8	46,9/59,9	72,1/82,8
Gga-Eya2		100	46,9/61	52,7/63,6
Gga-Eya3			100	48,8/63,3
Gga-Eya4				100

b.

ID/SIM (%)	Gga-Eya4	Hsa-EYA4	Mmu-Eya4	Xtr-eya4
Gga-Eya4	100	94.7/97.1	91.3/93.5	85.8/91.7
Hsa-EYA4		100	93.0/94.7	86.1/92.4
Mmu-Eya4			100	82.7/88.7
Xtr-eya4				100

**Table 4.1: Percentage identity and similarity between EYA protein sequences. (a)** Percentage identity (ID) and similarity (SIM) between predicted Gga-Eya4 [ENSGALT00000022662.4] and the three other chicken Eya members, predicted Gga-Eya1 [XP\_418290.3], Gga-Eya2 [NP\_990246.1], and predicted Gga-Eya3 [XP\_417715.2]. **(b)** Percentage ID and SIM between predicted Gga-Eya4 and human (Hsa) [NP\_004091.3], mouse (Mmu) [NP\_034297.2], and *Xenopus tropicalis* (Xtr) [ENSXETT00000000214.3] EYA4 protein sequences. Percentage identity (ID) and similarity (SIM) were determined using EMBOSS Needle software ([http://www.ebi.ac.uk/Tools/psa/emboss\\_needle/](http://www.ebi.ac.uk/Tools/psa/emboss_needle/)) (Hancock et al. 2004; Li et al. 2015).

After selection of 2 clones (or recombinants) by blue/white screening (white colonies contain an insert), they were purified, and a restriction digestion using EcoRI enzyme allowed to verify the presence of an insert of about 900-1000 bp in length into the vector (see chapter 2.12). The two plasmid DNAs cloned were validated by sequencing (Source BioScience, Cambridge UK).

The two identical sequences, which were not 947 bp long as expected, but 873 bp long, were identified by Basic Local Alignment Search Tool (BLAST) as EYA4 (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The cloned Gga-Eya4 sequence (Cloned\_seq) was aligned with the Eya4 sequence pulled out from the chicken Genebuild (Gga\_Eya4\_GB) showing a gap in its sequence (Fig. 4.2). A comparison with the human, mouse and Xenopus EYA4 sequences indicated that this gap corresponds to exon 5.

Interestingly, information found in the databases revealed the existence of different EYA4 transcripts with differences in exon 5. Alternative transcripts have been identified in human and mouse, displaying sequences with or without exon 5; other isoforms of EYA4 with truncated exon 6, 8 or 16, as well as substitution between exon 19 and 20, were also reported. For example, exon 5 is always absent in mouse transcript sequences (Borsani et al. 1999; Y. Zhang et al. 2004).

The Gga-Eya4 cloned from skeletal muscle-enriched sample corresponds to the isoform a, where exon 5 is absent; this is consistent with the fact that this isoform was also found in a large number of cDNA clones identified from a human skeletal muscle cDNA library.

The functional relevance for these alternative transcripts is not yet known, but their conservation across species and the fact that they appear to be tissue-specific, might indicate potential interesting roles that still remain to be discovered.

### **b. Characterisation of Gga-Eya4**

Sense and antisense RNA probes for Gga-Eya4 were synthesised by PCR using the plasmid containing the fragment of 947 bp, previously cloned, as a template (see chapter 2.7; Appendix III List 1).

Gga-Eya4_GB	MEDSQDLSEQSVKKTCTESDVSEPPNARSMEMQDLASPHNLVGSSDAFGSSKLDKPNLSS
Hsa-EYA4	MEDSQDLNEQSVKKTCTESDVSQSQNSRSMEMQDLASPHNLVGGGDTFGSSKLEKSNLSS
Mmu-Eya4	MEDTQDLNEQSVKKTCPADVSEPNQNSRSMEMQDLASPHALVGGSDTFGSSKLDKSGLSS
Xtr-eya4	MDDSQDLNEQPVKKPCTESDVTEPQNTRSMEMQDLASPHNRVSSGDTSGASKLDKSNINN
Cloned_seq	-----DSPNARSMEMQDLASPHNLVGSSDAFGSSKLDKPNLSS
Gga-Eya4_GB	TSVTTNGTGGDNMTVLNTADWLLSCSTPSSATMSLLAVKTEPMNSNETTTTGGDGLDTE
Hsa-EYA4	TSVTTNGTGGENMTVLNTADWLLSCNTPSSATMSLLAVKTEPLNSSETTATTGGDGLDTE
Mmu-Eya4	TSVTTNGTG-----VSLAVKTEPLHSESTTTTGGDGLDTE
Xtr-eya4	TSITTTNGTGGENMAVLNTADWLLSCSTPSSATVSLAVKTEPMNSNETTATTGGDGLDTE
Cloned_seq	TSVTTNGTG-----VSLAVKTEPMNSNETTTTGGDGLDTE
Gga-Eya4_GB	TGSVITSSGYSPRSAHQYSPQIYPSKPYPHILSTPAAQTMSAYAGQTQYSGMQQPAVYTA
Hsa-EYA4	TGSVITSSGYSPRSAHQYSPQLYPSKPYPHILSTPAAQTMSAYAGQTQYSGMQQPAVYTA
Mmu-Eya4	TGSVITSSGYSPRSAQQYSPQLYPSKPYPHILSTPAAQTMSAYAGQTQYSGMQQPAVYTA
Xtr-eya4	TGSVITSSGYSPRSAHQYSPQLYPSKPYPHILSTPAAQTMSAYGGQSQYSGMQQPVTYTA
Cloned_seq	TGNNK-----

**Fig. 4.2: Multiple protein sequence alignment of EYA4.** Alignment of the cloned Gga-Eya4 sequence (Cloned\_seq) with the chicken Genebuild (Gga-Eya4\_GB), human (Hsa-EYA4), mouse (Mmu-Eya4) and Xenopus (Xtr-eya4) sequences. Residues conserved in all sequences are indicated by dark background. Residues conserved in 4 or 3 orthologous are indicated in dark or light grey, respectively. The cloned Gga-Eya4 sequence is shorter than the others; the missing part of the sequence (indicated by a red frame) corresponds to the exon 5. For accession numbers see Fig. 4.1.

The plasmid was first linearised by PCR with M13 forward and reverse primers, then sense and antisense RNA probes were synthesised and Digoxigenin (DIG)-labelled by *in vitro* transcription with SP6 and T7 RNA polymerases, respectively. Sense probe was used as a negative control (Appendix III Table 1).

Eya4 WMISH was performed on chicken embryos fixed at different stages of development, from HH10 to HH23; expression pattern in whole-mount and on transverse sections at HH11-12, HH16, and HH21-22 are presented in Fig. 4.3a. Embryos treated with the sense RNA probe display no specific expression pattern; the trapping in the head being non-specific (data not shown). Expression pattern of miR-128, predicted to target Gga-Eya4, is also presented for comparison (Fig. 4.3b).

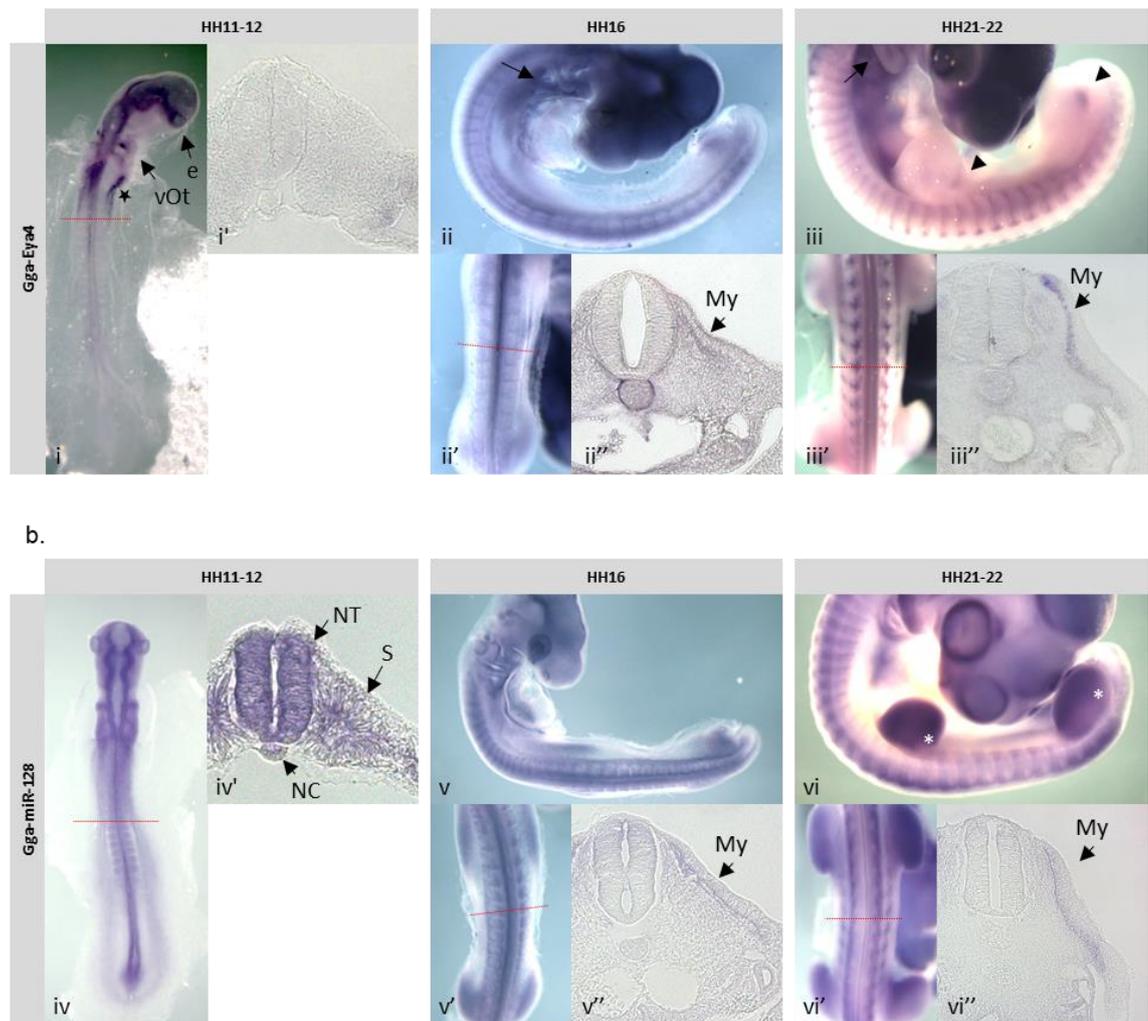
At HH11-12, Gga-Eya4 was mainly expressed in the head region in some cranial placodes, such as the nasal and otic vesicles (Fig. 4.3i). Eya4 was also expressed in a pool of cells, close to the heart region; non-identified, they seem to be migrating in an anterior-to-posterior fashion. No expression was detected in the somites. As the embryo developed, Gga-Eya4 was detected in the branchial arches and in the somites.

At HH16, Gga-Eya4 was expressed in the myotome; expression becoming even stronger at HH21-22.

At HH21-22, Gga-Eya4 was strongly expressed in the ventromedial lip of the dermomyotome and in the myotome; it was also found in the dorsal root ganglia, and in a restricted posterior region of the developing limbs. This expression was stronger in the hindlimbs (arrowhead), probably due to the fact that the hindlimbs start to develop before the forelimbs.

The expression of Gga-Eya4 in the otic vesicle, branchial arches and in the somites are consistent with the profile previously found in mouse. However, although Eya4 was not observed in the developing eye in mouse, Gga-Eya4 is expressed in this tissue and at all stages studied; this is similar to what was described for Gga-Eya1, Gga-Eya2 and Gga-Eya3 (Borsani et al. 1999; Xu, Woo, et al. 1997).

In addition, while Eya4 in mouse was found in the dermomyotome, in chicken, it is weakly present in the dorsomedial lip of the dermomyotome, and strongly expressed in the myotome. The expression in the otic vesicle is in line with the mouse pattern, and the hearing deficit associated to the mutation of EYA4 in human (Wayne et al. 2001; Zhang et al. 2004; Hildebrand et al. 2007; Makishima et al. 2007).



**Fig. 4.3: Expression profile of Gga-Eya4 (a), and comparison with miR-128 pattern (b).** (a) WMISH performed with antisense DIG-labelled RNA probe, and transverse sections at HH11-12 (i, i'), HH16 (ii-ii''), and HH21-22 (iii-iii''). At HH11-12, Gga-Eya4 is expressed in the eye (e), the otic vesicle (vOt), and in a pool of non-identified migrating cells close to the heart region (i; star). At HH16, Gga-Eya4 is expressed in the eye, the branchial arches (ii; arrow) and in the somites (s), in the myotome (ii''; My). At HH21-22, Gga-Eya4 is still expressed in the branchial arches (iii; arrow), and is strongly expressed in the myotome (iii''). Gga-Eya4 is also found in the dorsal part of the limb buds (iii; arrowhead). Embryos treated with the sense probe (negative control) did not show any expression (data not shown). (b) MiR-128 expression pattern determined by LNA ISH in whole-mount and transverse sections at HH11-12 (iv, iv'), HH16 (v-v''), and HH21-22 (vi-vi''). At HH11-12, miR-128 is expressed in the neural tube (NT) and the developing somites (iv'). At HH16 and HH21-22, miR-128 is in the branchial arches, in the myotome, and the developing limbs (v-vi''). At HH21-22, miR-128 is also expressed around the eye, and in the limbs (vi; white asterisk). Red dotted lines indicate the location of the transverse sections. Transverse sections: 20x magnification. e: eye; My: myotome; NC: notochord; NT: neural tube; S: somite; vOt: otic vesicle.

However, *Eya4* was not observed in the heart at any stages studied during chicken development, differing from the mouse and human phenotypes (Schönberger et al. 2005).

Because *Eya4* was predicted to be targeted by miR-128, their expression patterns were compared. While *Eya4* was only starting to be expressed at HH11-12, mainly in the head region (i, i'), miR-128 was already strongly expressed in the neural tube, notochord, and the somites (iv'). At HH16 and HH21-22, they were also both expressed in the branchial arches, and in the somites, in the myotome (ii-iii''; v-vi''). At HH21-22, they were both found in the developing limbs, with miR-128 being expressed in the entire limb buds, and *Eya4* restricted to the dorsal part of the limb buds (iii, vi).

MiRNAs are fundamental regulators that can silence gene expression at post-transcriptional level. Multiple modes of miRNA-mediated regulation have been described and include translational inhibition, increased mRNA de-adenylation and degradation, and/or mRNA sequestration (Nilsen 2007; Selbach et al. 2008); this is dependent on the target and its function, the stage of development, and the tissue considered.

Gga-*Eya4* and miR-128 display very similar profiles, especially in the somites from HH16. Therefore, because Gga-*Eya4* and miR-128 expression patterns are not exclusive, miR-128 is more likely to be a regulator/modulator of *Eya4* expression rather than an absolute repressor, if a direct interaction could be confirmed (see section 4.2.1e).

### **c. Cloning of a fragment of the 3' untranslated region (UTR) of Gga-*Eya4***

In order to identify a potential interaction between miR-128 and Gga-*Eya4* by luciferase reporter assays, it was necessary to clone the 3'UTR of Gga-*Eya4* (Fig. 4.1; Appendix III List 2). Gga-*Eya4* 3'UTR was used to generate reporter constructs, wild-type (WT) and mutants, and perform luciferase reporter assays (see chapter 2.19).

Localised in the 5' part of the 3'UTR sequence, a potential miR-128 site was predicted using different bioinformatics tools and algorithms (TargetScan, MiRanda); analysis was presented in chapter 3. Using a similar strategy, other miRNAs have also been predicted to target Gga-*Eya4* 3'UTR, including the myomiRs miR-1a, miR-206 and miR-133a, the cardiac miRNA miR-499, or miR-27b which has the same seed sequence as miR-128 and is predicted to target the same site. All these sites are located within the first 1,000 bp of Gga-*Eya4* 3'UTR sequence.

The 3'UTR of Gga-Eya4 is extremely long, 6,000 bp in length; too long to be cloned and used for the luciferase reporter assays. Taking into account that predicted miRNA sites located at the 5' and 3' extremities of a 3'UTR sequence are more likely to be functional (Long et al. 2007; Ekimler & Sahin 2014), a pair of primers was designed in order to clone a fragment of about 1,000 bp, in the 5' part of Gga-Eya4 3'UTR, containing miR-27b/128, miR-1a/206, miR-133a, and miR-499 sites (Fig. 4.1). BglII and NheI restriction sequences were added to the forward and reverse primers, respectively.

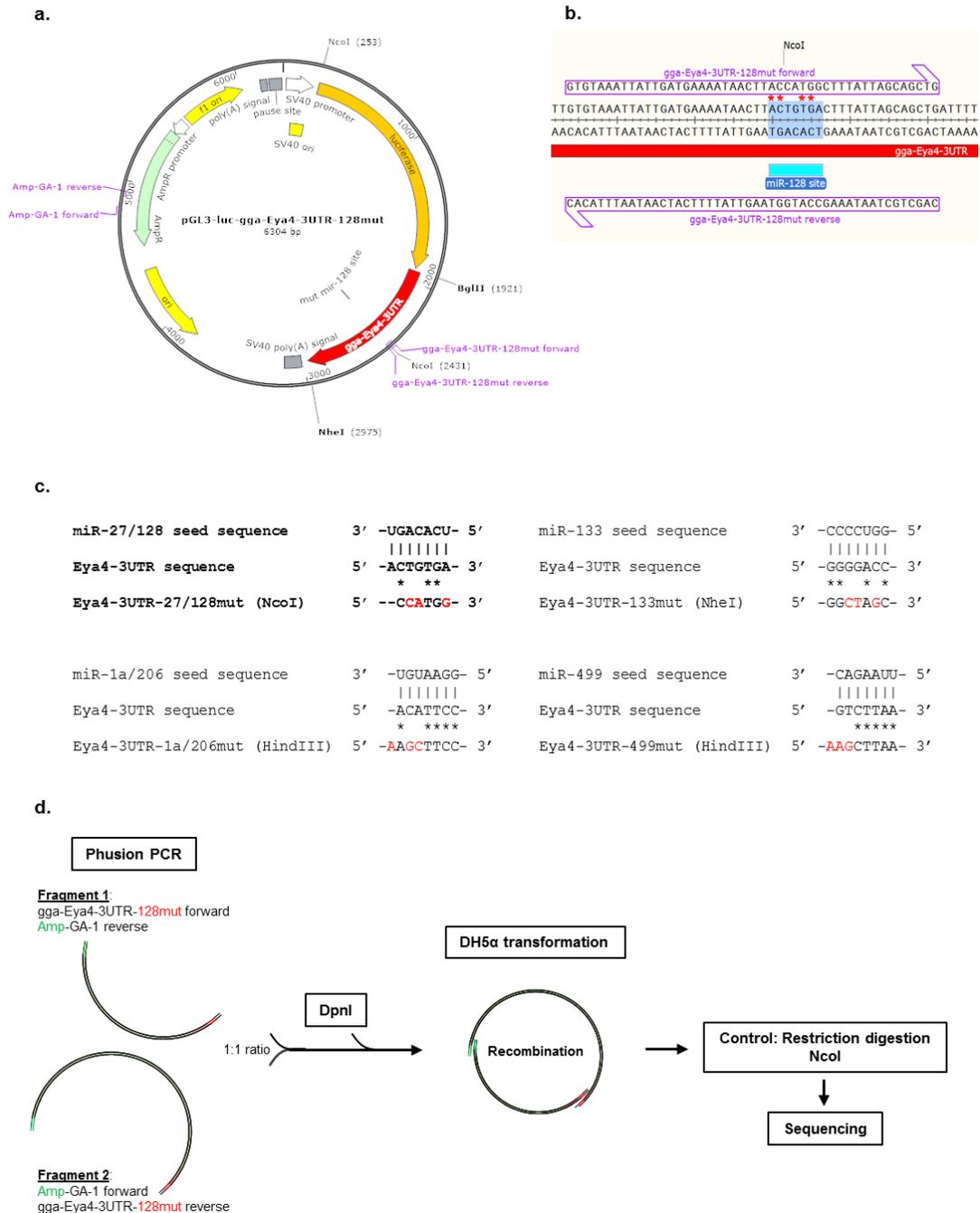
Following the same strategy previously described for the cloning of a fragment of Gga-Eya4 coding sequence, a fragment of Gga-Eya4 3'UTR was amplified by PCR. The PCR product was cloned into pGEMT-Easy vector; the construct was purified, quantified, and verified by sequencing. Then, Gga-Eya4 3'UTR fragment was excised from pGEMT-Easy vector by restriction digestion using BglII and NheI enzyme, and sub-cloned into linearised pGL3-Luciferase modified vector, which had also been previously BglII/NheI digested (see chapter 2.14) (Tuddenham et al. 2006). Gga-Eya4 3'UTR was inserted downstream of the Luciferase gene.

With this disposition, the pGL3(Gga-Eya4 3'UTR) construct is a good tool to study how a gene – in this case the Luciferase gene – can be regulated by action on its 3'UTR (Gga-Eya4 3'UTR), and how miRNAs – miR-27/128, miR-1a/206, miR-133a or/and miR-499 – which potentially interact with the Gga-Eya4 3'UTR can modulate its expression. This construct was then used to perform luciferase reporter assays.

#### **d. Mutagenesis: Gga-Eya4 3'UTR mutants**

In addition to pGL3(Gga-Eya4 3'UTR) construct, four mutant constructs were generated: 27b/128mut, 1a/206mut, 133amut, and 499mut. MiR-27b and miR-128, and miR-1a and miR-206, respectively, have the same seed sequence and are predicted to recognise the same target site. Mutagenesis was performed based on the FastCloning strategy (Li et al. 2011).

Developed by Li *et al.*, this PCR-based cloning technique can be used to insert any DNA fragment into a plasmid vector or into a gene in a vector at any desired position. Purification-free, sequence- and ligation-independent, the FastCloning technique is simple, fast, economic and as efficient as commercial assembly kits, like the Gibson assembly one (Gibson et al. 2009; Gibson 2011). Mutagenesis strategy is summarised in Fig. 4.4 (also see chapter 2.19).



**Fig. 4.4: Mutagenesis strategy.** (a) Schematic of pGL3(Gga-Eya4 3'UTR-128mut) construct. The position of the miR-128 site is indicated, as well as the position of the overlapping primers used to introduce to mutation. (b) MiR-128 site in the Gga-Eya4 3'UTR sequence. The overlapping primers (forward and reverse) containing the mutation (restriction enzyme site; in this case: NcoI) are indicated in purple. Red stars represent what is left of the target site after insertion of the mutation. (c) Alignments of Gga-Eya4 3'UTR sequence at the different miRNA sites with the seed sequence of their respective miRNA; mutated nucleotides are indicated in red. Vertical lines and stars indicate complementarity and identity between sequences, respectively. (d) Protocol used to generate mutant based on the FastCloning technique (Li et al. 2011).

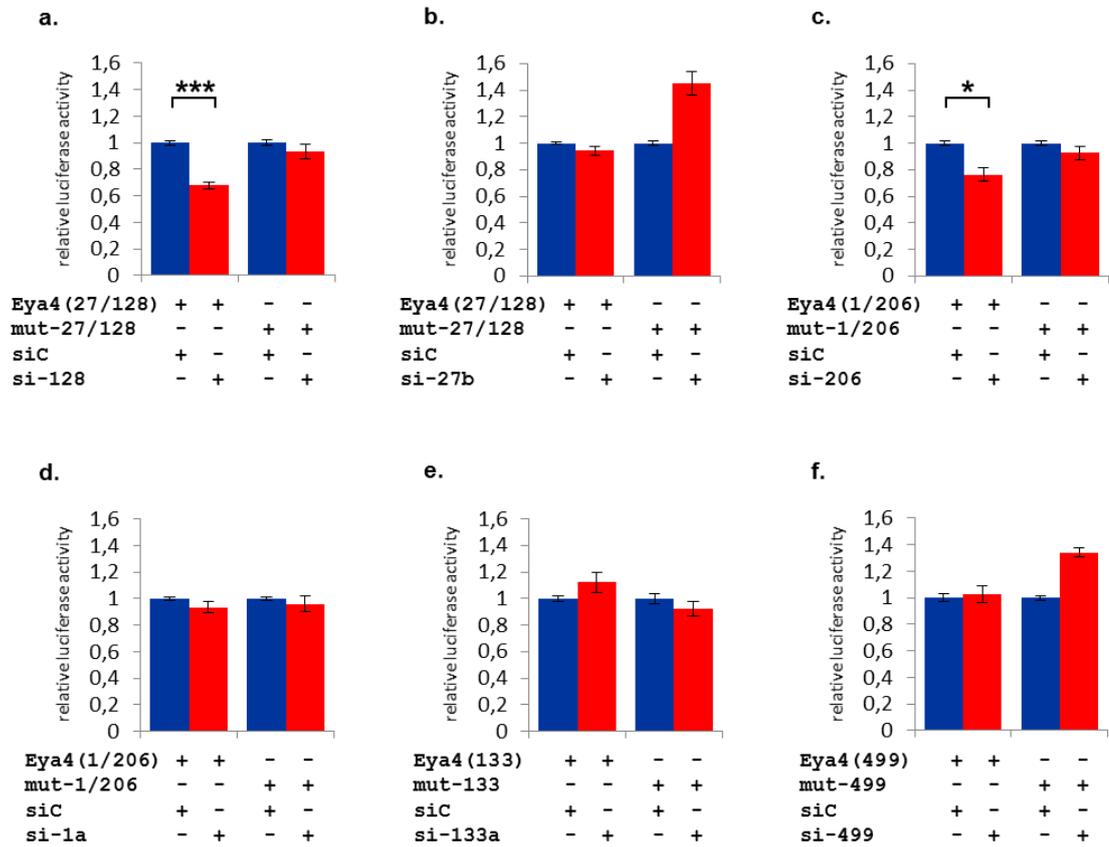
The pGL3(Gga-Eya4 3'UTR) construct was used as a template (Fig. 4.4a). Overlapping primers were designed to introduce a mutation in place of the target site (Fig. 4.4b); 1-3 nucleotides were replaced in order to create a restriction enzyme site preventing the miRNA from binding its target (Fig. 4.4c). The restriction sites to introduce were determined based on three main criteria: destruction of the target site with a minimum of nucleotides to change; no site for this enzyme in the 3'UTR fragment sequence, and a maximum of one site for this enzyme in the pGL3-Luciferase modified vector.

For each mutant construct, two PCR products were amplified using Phusion High-Fidelity polymerase (NEB) and a combination of primers (specific primers introducing the mutation (forward and reverse) + primers designed in the Ampicillin resistance (AmpR) gene (forward and reverse)). The PCR products were then mixed together (1:1 ratio), and DpnI digested to destroy methylated templates. After transformation into DH5 $\alpha$  competent cells, which are able to recombine the two fragments together, several colonies were tested. The presence of the mutation (introduction of a restriction site) in the recombined constructs was controlled by restriction digestions; samples were validated by sequencing. These mutant constructs were then used to perform 'rescue experiments' in luciferase reporter assays.

#### **e. Luciferase reporter assays: miRNA/target interaction *in vitro***

To identify potential interaction between Gga-Eya4 3'UTR and the six miRNAs, miR-128, miR-27b, miR-1a, miR-206, miR-133a and miR-499, luciferase reporter assays were performed (Promega; see chapter 2.19). This bioluminescence assay is a quantitative method based on sequential measurement of Firefly and Renilla luciferases activities in a single sample.

WT or mutant constructs (100 ng) were co-transfected, into chicken DF1 fibroblast cells, with Renilla vector (25 ng; used as an internal control of the transfection) and either without or with one of the following siRNAs (50 nM; Sigma): si-128; si-27b; si-1a; si-206; si-133a; si-499; siRNAs are used in these assays to mimic miRNA action. A universal negative control siRNA (siC; Sigma) was used as negative control; it also allowed to check the impact of adding siRNA on the transfection efficiency. Results are presented in Fig. 4.5 and Table 1.



**Fig. 4.5: Luciferase reporter assays – miRNA/Gga-Eya4 3'UTR interaction.** Luciferase activity for Gga-Eya4 3'UTR constructs, wild-type (WT) and mutants, co-transfected with control siRNA (siC) or one of these siRNAs: si-128 (a); si-27b (b); si-206 (c); si-1a (d); si-133a (e); si-499 (f). Normalised luciferase activity was plotted relative to the condition [WT or mutants construct + siC] (in blue in the graphics). Experiments were repeated 4 times independently with triplicate samples in each; 5 times for experiments with si-128, si-1a, and si-206. Error bars represent the standard error of the mean (SEM) (n=12 or 15). T-test: p<0.05: \*, p<0.001: \*\*\*.

Experiments were repeated 4 times with triplicate samples; 5 times for experiments with miR-128, miR-1a, and miR-206. The normalised luciferase activity (Firefly/Renilla) was plotted relative to the condition [WT or mutants construct + siC].

### **i) MiR-128 targets Gga-Eya4 3'UTR**

Luciferase reporter assays showed that miR-128 targets the Gga-Eya4 3'UTR, leading to a relative decrease in luciferase activity of 32% (68% activity; t-test:  $p < 0.001$ ). This was rescued significantly by mutating the miR-128 site; the luciferase activity going back to 93% of control (Fig. 4.5a).

Interestingly, with the same seed sequence and predicted target site as miR-128, miR-27b did not seem to be able to interact with Gga-Eya4 3'UTR (Fig. 4.5b). This indicates that miR/target interactions are not based only on a match between the seed nucleotides and target sequences, additional elements need to be taken into account, such as partial complementarity of the rest of the miRNA sequence with the target sequence.

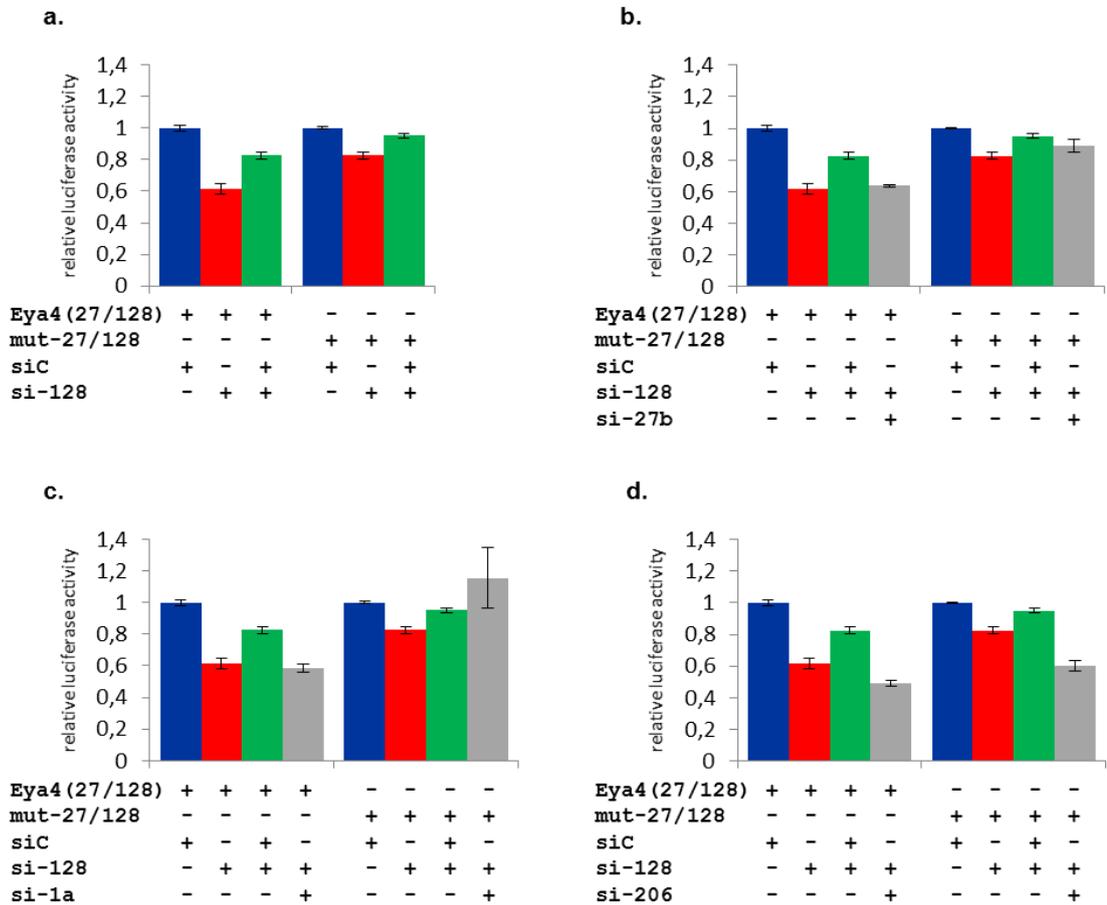
Furthermore, miR-206 could target Gga-Eya4 3'UTR. A decrease of 24% in luciferase activity was observed (76% activity; t-test:  $p < 0.05$ ); this was rescued significantly by mutating the miR-206 site (92.5% activity) (Fig. 4.5c).

Similar to what was observed for miR-27b and miR-128, miR-1a and miR-206, which have the same seed sequence and target the same site, showed different profiles; miR-1a did not seem to interact with Gga-Eya4 3'UTR (Fig. 4.5d).

No effect on luciferase activity was observed in the experiments performed with si-133 and si-499; thus Gga-Eya4 could not be validated as a target of miR-133 and miR-499 (Fig. 4.5e, f).

### **ii) Potential synergism between miRNAs in the regulation of Gga-Eya4 expression**

Due to their short sequences, miRNAs can interact and regulate several hundred targets. On the other hand, multiple miRNAs can simultaneously regulate the expression of a specific gene by targeting different sites on the 3'UTR of its mRNA (Selbach et al. 2008; Bartel 2009).



**Fig. 4.6: Luciferase reporter assays – miRNA combinations and interaction with Gga-Eya4 3'UTR.** Luciferase activity for Gga-Eya4 3'UTR constructs, WT and mutants, co-transfected with control siRNA (siC), si-128 alone or in combination with either siC, si-27b, si-1a or si-206. Normalised luciferase activity was plotted relative to the condition [WT or mutants construct + siC] (in blue in the graphs). Experiments were repeated two-three times with triplicate samples. Error bars represent the standard error of the mean (SEM) (n=6 or 9).

However, the concept of synergism between miRNAs has not been really investigated.

In order to test if Gga-Eya4 could be regulated by a combination of miRNAs, a pilot experiment was performed.

Previous experiments showed that miR-128, and miR-206, were able to interact with Gga-Eya4 3'UTR, resulting in a decrease in luciferase activity. Combinations of two siRNAs were prepared (1:1) such that the same final concentration of total siRNA was used in the transfection; therefore each siRNA was half-concentrated compared to the previous experiments.

WT and mutant construct mut27b/128 were co-transfected with Renilla, and a combination of si-128 with either siC, si-27b, si-1a, or si-206. Only 2-3 experiments were performed, each of them with triplicate samples. The normalised luciferase activities (Firefly/Renilla) were plotted relative to the condition [WT or mutants construct + siC]. Results are presented in Fig. 4.6.

First, luciferase reporter assays combining si-128 and siC were performed in order to determine the luciferase activity with half the amount of si-128; siC being added to maintain the final siRNA molarity of 50 nM. Transfections with only si-128 were also performed in parallel.

In the condition [WT + si-128], the luciferase activity was decreased by 38.5% (Fig. 4.6a); decrease similar to the one observed in the previous set of experiments (32%; Fig. 4.5a). However, contrary to the previous experiments showing rescue to 94% activity when mutating the miR-128 site, the luciferase activity was only partially rescued (82.5%). This may be due to the low number of repeats. As expected, the experiment with a combination of si-128 and siC showed a smaller decrease in luciferase activity of 17.5%, which was rescued by mutating the miR-128 site (95%); less si-128 leading to less impact on luciferase activity, with a better rescue.

Luciferase reporter assays were then performed with a combination of si-128 and si-206, both shown to interact, on their own, with Gga-Eya4 3'UTR. When these two siRNAs were transfected together with the WT Gga-Eya4 construct, the luciferase activity level was lower than the one observed after transfection with si-128 alone (51% in decrease compare to 32% with si-128 alone) (Fig. 4.6d).

The luciferase activity was not rescued after mutation of the miR-128 site (only 11% rescue; luciferase activity decrease: 40%).

The previous experiments showed that miR-128 and miR-206 alone were able to target Gga-Eya4 3'UTR leading to a decrease in luciferase activity by 32% and 24% respectively (Fig. 4.5a, c). The synergy experiments showed that when they are used in combination, with half the amount for each siRNA, the luciferase activity was decreased by 51% (Fig. 4.6d); rescue experiments showed that the luciferase activity was rescued by only 11%.

The weak rescue observed could be explained by the fact that the mutant construct used for these experiments was mutated for miR-128 site only; miR-206 site was intact. Moreover, if the results for this rescue experiment (Fig. 4.6d) are compared to the one observed in the experiment [WT construct + si-206 alone] (Fig. 4.5c), the level of luciferase activity is still much lower (51% decrease compared to 24% decrease (si-206 alone)); this result was observed with half the amount of si-206. This suggests that miR-128 and miR-206, which can target Gga-Eya4 3'UTR on their own (Fig. 4.5a, c) (38.5% and 24% decrease in luciferase activity, respectively (50 nM each)), might also be able to work in cooperation, in an additive manner, in order to have a stronger effect on the regulation of Gga-Eya4 expression (Fig. 4.6d); their combined effect being similar to the sum of their separate effect at the same doses (miR-128+miR-206: 51% decrease in luciferase activity (25 nM each)) (Ivanovska & Cleary 2008; Lu & Clark 2012).

From the previous experiments, although miR-128 and miR-27b, and miR-206 and miR-1a, respectively, were predicted to target Gga-Eya4 (same seed sequence, same site), miR-1a and miR-27b, did not seem to be able to interact with Gga-Eya4 3'UTR (Fig. 4.5b, d). Luciferase assays were performed with a combination of si-128 and si-27b (Fig. 4.6b), or si-1a (Fig. 4.6c). For each combination experiment (si-128/si-27b and si-128/si-1a), a luciferase activity level similar to the one observed in the condition [WT construct + si-128 + siC] (Fig. 4.6a), around 17.5%, was expected; with si-27b or si-1a having no effect, and half the amount of si-128.

Interestingly, when these two pairs of siRNAs were transfected together with the WT Gga-Eya4 construct (Fig. 4.6b, c), the luciferase activity level was similar to the one observed after transfection with si-128 alone (si-128/si-27b pair: 64% and 36% decrease; si-128/si-1a pair: 58% and 42% decrease; si-128 alone: 61.5% and 38.5% decrease); the luciferase activity was rescued (90-100%) after mutation of the miR-128 site.

These results suggest that miR-1a and miR-27b, when in combination with miR-128, respectively (Fig. 4.6b, c), could potentially improve the interaction between miR-128 and Gga-Eya4 3'UTR. The complete rescue observed with the combination si-128/si-1a seems to confirm that even with the same seed sequence than miR-206, miR-1a does not target Gga-Eya4 3'UTR.

These combination experiments suggest that miR-128 and miR-206, as well as targeting Gga-Eya4 3'UTR on their own, could work in cooperation, in an additive manner, to regulate the expression of Gga-Eya4; miR-27b and miR-1a could also be involved in this regulation.

Part of a pilot experiment aiming at investigating potential cooperative actions between miRNAs in order to regulate targets, these assays, although showing promising trend results, would have to be repeated to make the results statistically significant. Experiments using different constructs, such as a miR-128/206 double mutant construct with both target sites mutated, could provide interesting additional information.

In this first part, Gga-Eya4, a predicted miR-128 target, was validated as a target *in vitro*. The cloning of a fragment of its coding sequence allowed to generate a specific RNA probe and its expression pattern was determined by RNA ISH. A comparison between Gga-Eya4 and miR-128 profiles revealed similar expressions, in particular in the somites, in the myotome.

A fragment of Gga-Eya4 3'UTR, containing miR-27b/128 site, but also miR-1a/206, miR-133a and miR-499 sites, was cloned and used to perform luciferase reporter assays. This quantitative method allowed investigation of the predicted interaction between these miRNAs and Gga-Eya4 3'UTR. MiR-128 and miR-206 were validated as miRNAs able to target Gga-Eya4. In addition, a pilot experiment looking at potential miRNA synergism, gave an insight into the complex collaboration that could exist between these miRNAs (miR-128, miR-27b, miR-206, and miR-1a) for target regulation; preliminary data showing that miR-128 and miR-206 could potentially work in an additive manner to more efficiently regulate Gga-Eya4 expression.

The following part of this chapter will focus on the miRNA regulation of the other members of the PSED network, specifically Six, Eya and Dach, including new insights into their expression patterns.

#### **4.2.2. MicroRNA regulation of members of the PSED network, and their expression patterns**

In this part, the same approach used to validate miR-128 and miR-206/Gga-Eya4 3'UTR interactions was applied to investigate potential miRNA interactions with some other PSED members: Six1/4, Eya1-3, and Dach1/2.

##### **a. Cloning of 3'UTR fragments of PSED members and mutagenesis**

In order to study potential miRNA interactions with Six1/4, Eya1-3, and Dach1/2, the first step was to clone their 3'UTR.

Gga-Six1, Gga-Six4, Gga-Eya1, Gga-Eya2, Gga-Eya3 and Gga-Dach1 3'UTR sequences were found in Ensembl and NCBI databases; no 3'UTR sequence was found for Gga-Dach2, most likely due to the fact that Gga-Dach2 has not been annotated yet.

Based on the information available in TargetScan database and the analysis performed on data generated by running the miRanda algorithm (analysis presented in chapter 3), several miRNAs were identified and predicted to target the PSED members; relevant predicted miRNAs are listed in Table 4.2; they were selected for further investigation. For example, Eya1 is predicted to be targeted by miR-27b and miR-128, but also miR-133a; Six1, Six4 and Dach1 by miR-1a and miR-206.

Pairs of primers were designed for PCR-mediated cloning of these 3'UTRs. Similar to Gga-Eya4 3'UTR, several fragments were cloned for some of these 3'UTRs, which were very long (Appendix III List 2). For example, in the case of Gga-Six4 3'UTR (4518 bp in length), 3 overlapping fragments were cloned. However, the intermediate fragment was not investigated further. This was due to the presence of the most interesting miRNA sites for this study, in the 5' and the 3' regions of the 3'UTR, and because predicted miRNA sites located at the 5' and 3' extremities of a 3'UTR sequence are more likely to be functional (Long et al. 2007; Ekimler & Sahin 2014).

GENE	3'UTR FRAGMENT		miR	DECREASE		T.TEST (p value)	MUTANTS	miR	RESCUE		T.TEST (p value)
	Name	Length (bp)		N(exp)	%				N(exp)	%	
Eya1	Eya1.1	1291	miR-133a	4	19%	1.1E-3	Eya1.1(mut-133a)	miR-133a	4	91%	0.029
			miR-27b	3	3.4%	0.54	Eya1.1(mut-128(1))	miR-27b	4	100%	0,67
			miR-128	3	37%	5.8E-11	Eya1.1(mut-128(2))	miR-128	4	73%	9.9E-14
							Eya1.1(mut-128-DM)	miR-27b miR-128	4 4	100% 83%	0.044 7.4E-5
Eya3	Eya3	487	miR-133a	4	22%	2.7E-5	Eya3(mut-133a)	miR-133a	4	97%	0.65
Eya4	Eya4	1043	miR-27b	4	6%	0.36	Eya4(mut-27/128)	miR-27b	4	>100%	*
			miR-128	5	32%	1.8E-4		miR-128	5	93%	0.53
			miR-1a	5	6%	0.44	Eya4(mut-1a/206)	miR-1a	5	96%	0.73
			miR-206	5	24%	1.4E-2		miR-206	5	92.5%	0.40
			miR-133a	4	no	0.42	Eya4(mut-133a)	miR-133a	4	92%	0.41
			miR-499	4	no	0.83	Eya4(mut-499)	miR-499	4	>100%	*
Six1	Six1	1043	miR-1a	4	33.5%	2.1E-6	Six1(mut-1a/206)	miR-1a	4	73%	9.6E-5
			miR-206	4	54%	1.5E-13		miR-206	4	63%	1.6E-12
Six4	Six4.1	1566	miR-133a	3	no	0.47	Six4(mut-133a)	miR-133a	3	100%	0.70
	Six4.3	1448	miR-1a	3	33%	9.4E-7	Six4.3(mut-1a/206)	miR-1a	4	76%	0.021
			miR-206	3	48%	3.5E-7		miR-206	4	61%	8.6E-9
			miR-499	3	no	0.086	Six4.3(mut-499)	miR-499	3	100%	0.63
Dach1	Dach1.2	1073	miR-1a	3	35%	1.3E-5	Dach1.2(mut-1a/206)	miR-1a	6	74%	2.2E-7
			miR-206	3	55%	7.2E-9		miR-206	6	58%	5.7E-14
							Dach1.2(mut-1a-2)	miR-1a miR-206	1 1	67% 57%	- -
						Dach1.2(mut-1a/206-DM)	miR-1a miR-206	3 3	56% 54%	8.6E-7 1.5E-7	

**Table 4.2: Summary – Luciferase reporter assays.** GENE: List of the genes, members of the SED network, selected for investigation. 3'UTR FRAGMENTS and miR: 3'UTR fragments cloned and their respective length (bp), and lists of interesting miRNAs predicted, by bioinformatics tools (TargetScan, MiRanda), to target the 3'UTR of Eya1, Eya3 and Eya4, Six1 and Six4, and Dach1. DECREASE: Number of independent experiments performed (N(exp)) and percentage of decrease in luciferase activity (%). T.TEST (p value). MUTANTS and miR: List of mutants generated for each 3'UTR (DM: double mutant) and miRNAs which were predicted to target these 3'UTR before introduction of the mutations. RESCUE: Number of independent experiments performed (N(exp)) and percentage of luciferase activity (%). T.TEST (p value). \*: no t-test performed due to rescue values going over 100%; - indicates that no t-test was performed due to insufficient number of experiments.

All the 3'UTRs, except for Gga-Eya2, were cloned into pGEMT-Easy vector, and then sub-cloned into pGL3-Luciferase modified vector by BglII/NheI restriction digestion (Tuddenham et al. 2006). Gga-Eya2 could not be successfully cloned even after several attempts and optimisations; this is maybe due to a problem with the sequence used to design primers, as although Ensembl provides a sequence for Gga-Eya2 3'UTR, in TargetScan it is mentioned that Gga-Eya2 does not appear to have a 3'UTR, thus the annotation may be incorrect.

The pGL3 constructs were validated by sequencing, and then used to perform luciferase reporter assays.

### **Mutagenesis:**

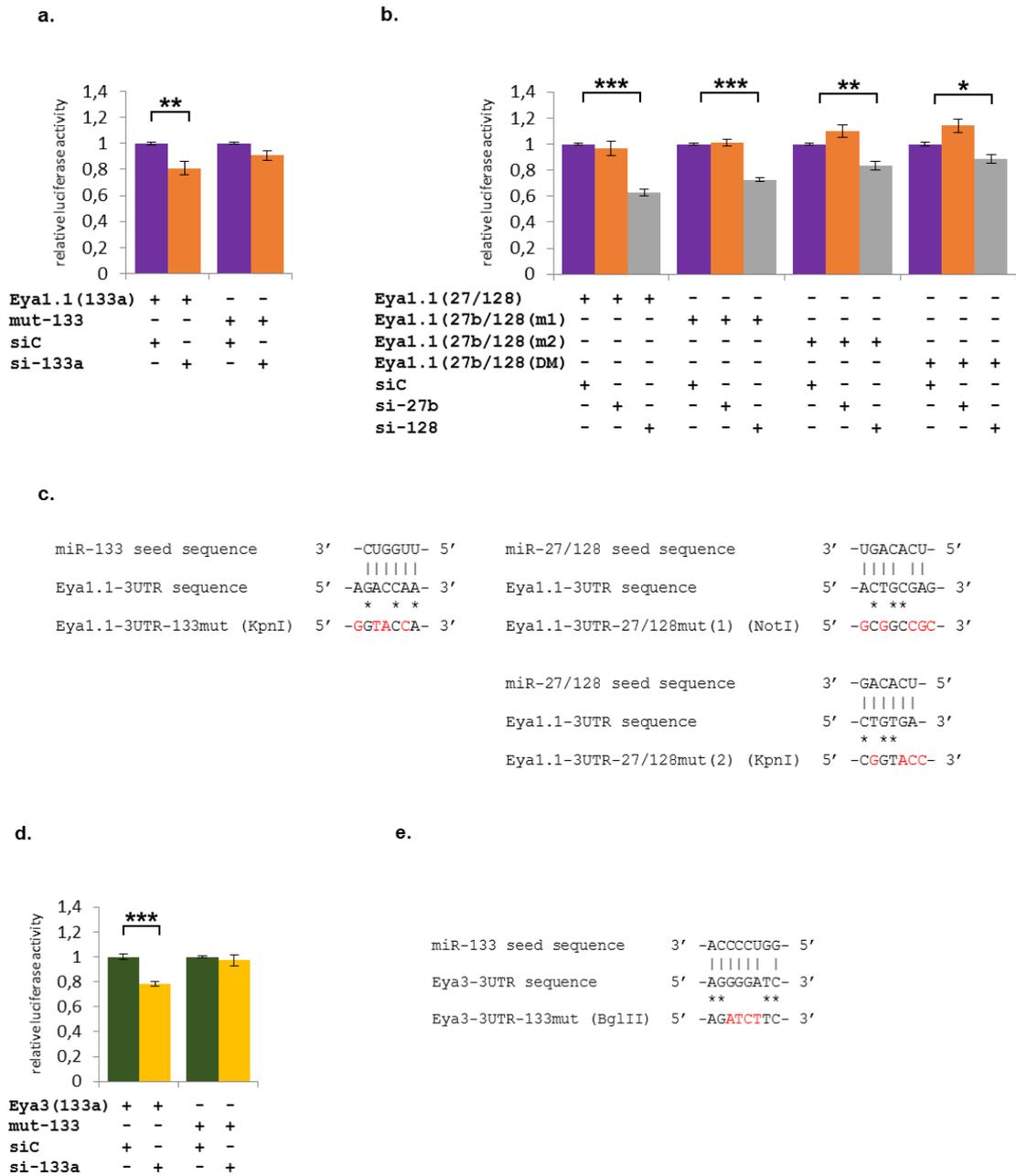
In addition to the WT pGL3 constructs, mutant constructs for Gga-Six1, Gga-Six4, Gga-Eya1, Gga-Eya3 and Gga-Dach1, and their respective miRNA sites were generated; this was done using the same strategy used to generate Gga-Eya4 3'UTR mutant constructs and presented in Fig. 4.4.

Overlapping primers were designed, containing a restriction enzyme site as mutation to introduce in place of the miRNA targeted site, and by using the FastCloning technique single or multiple miRNA sites were mutated.

In the case of Gga-Eya1, miR-128 and miR-27b are both predicted to target two sites in its 3'UTR. Single mutant constructs were generated for each site; a double mutant was also generated by using one of the single mutant constructs as a template and the appropriate set of primers. All the mutant constructs generated (listed in Table 4.2) were checked by restriction digestion and validated by sequencing, before being used to perform luciferase reporter assays.

### **b. Luciferase reporter assays: miRNA/target interaction in vitro**

For each predicted miRNA/target interaction mentioned in Table 4.2, luciferase reporter assays were performed, following the protocol described previously (see chapter 2.19).



**Fig. 4.7: Luciferase reporter assays – Gga-Eya1 and Gga-Eya3.** Luciferase activity for Gga-Eya1 3'UTR (**a-c**) and Gga-Eya3 3'UTR (**d, e**) constructs, WT and mutants, co-transfected with control siRNA (siC), or one of these siRNAs: si-133 (**a**; **d**); si-128 or si-27b (**b**). Normalised luciferase activity was plotted relative to the condition [WT or mutants construct + siC] (Eya1: in blue; Eya3: in green). Experiments were repeated 3-4 times with triplicate samples. Error bars represent the standard error of the mean (SEM) (n=9 or 12). T-test: p<0.05: \*, p<0.01: \*\*, p<0.001: \*\*\*. Alignments of Gga-Eya1 (**c**) and Gga-Eya3 3'UTR (**e**) sequences at the different miRNA sites with the seed sequence of their respective miRNA; mutated nucleotides are indicated in red. Vertical lines and stars indicate complementarity and identity between sequences, respectively.

### **i) MiR-128 and miR-133a target Gga-Eya1 3'UTR**

Luciferase reporter assays showed that miR-133 targets Gga-Eya1 3'UTR leading to a decrease in luciferase activity of 19% (81% activity; t-test:  $p < 0.01$ ), which was significantly rescued by mutating the miR-133 site; with relative luciferase activity going back to 91% (Fig. 4.7a).

Interestingly, while no effect was observed with miR-27b, miR-128 was able to target Gga-Eya1 3'UTR (Fig. 4.7b). A decrease of 37% of the luciferase activity was observed (63% activity), however after mutation of this site, the luciferase activity could only be partially rescued (73%; t-test:  $p < 0.001$ ). This was due to the presence of a second site. The mutation of this second site gave a similar or slightly better rescue (83%) and mutation of both sites brought the luciferase activity up to 89% (t-test:  $p < 0.01$ ). The significance of these promising results could be improved by performing an additional independent experiment; 4 independent experiments were performed with the WT and each single mutation constructs, while only 3 for the double mutant construct.

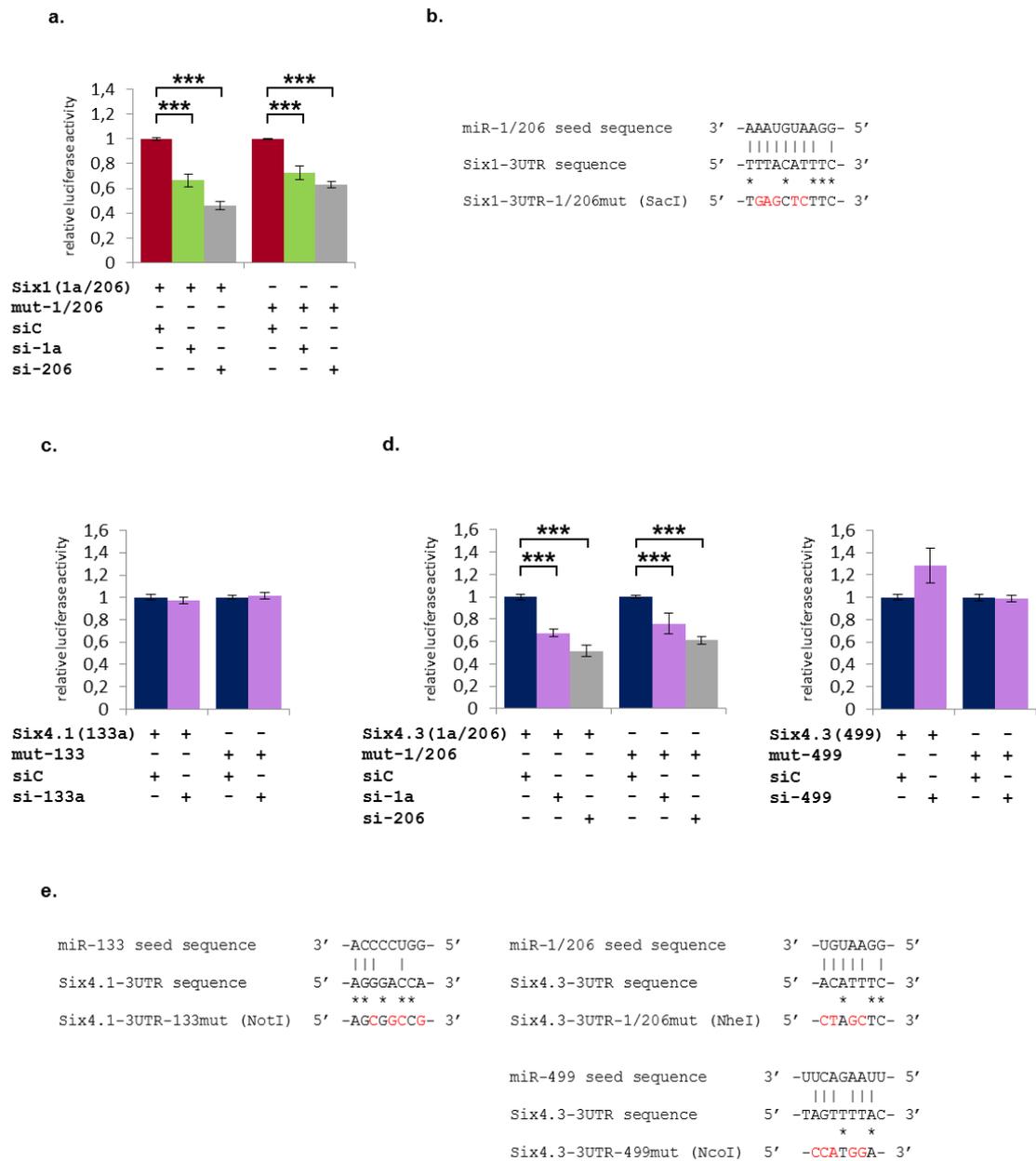
### **ii) MiR-133a targets Gga-Eya3 3'UTR**

Luciferase reporter assays showed that miR-133 targets Gga-Eya3 3UTR (Fig. 4.7d); a decrease in luciferase activity of 22% (78% activity; t-test:  $p < 0.001$ ) was observed, and 98% activity was restored after mutation of miR-133 site.

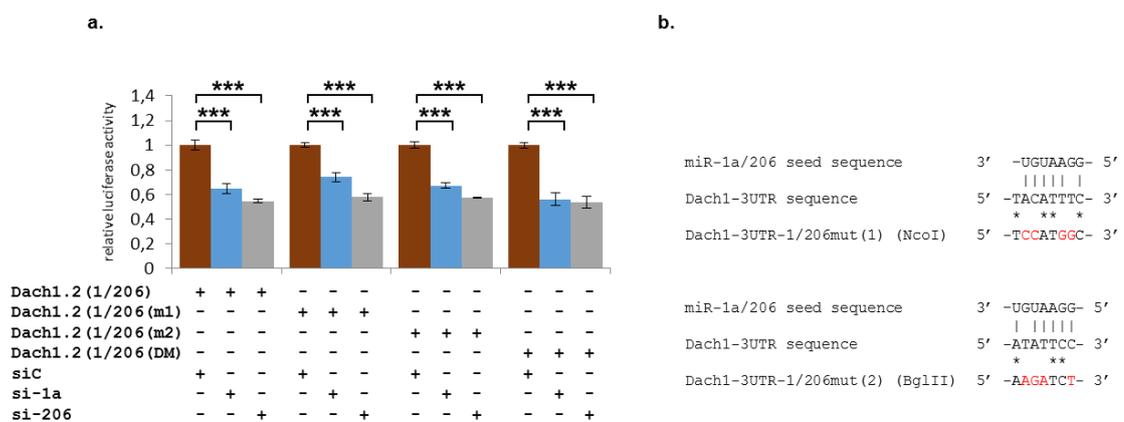
### **iii) Gga-Six1, Gga-Six4 and Gga-Dach1 3'UTRs, and miRNA regulation**

Luciferase reporter assays were performed and showed that although Gga-Six4 3'UTR is predicted to be a target of miR-133a and miR-499, no variation in luciferase activity was observed, suggesting that these miRNAs are not involved in the regulation of Gga-Six4 (Fig. 4.8c, d).

Luciferase reporter assays, aiming at validating Gga-Six1, Gga-Six4 and Gga-Dach1 as miR-1a and miR-206 targets were also performed (Fig. 4.8a, d and 4.9a).



**Fig. 4.8: Luciferase reporter assays – Gga-Six1 and Gga-Six4.** Luciferase activity for Gga-Six1 3'UTR (**a, b**) and Gga-Six4 3'UTR (**c-e**) constructs, WT and mutants, co-transfected with control siRNA (siC), or one of these siRNAs: si-1a or si-206 (**a; d**); si-133 (**c**); si-499 (**d**). Normalised luciferase activity was plotted relative to the condition [WT or mutants construct + siC] (Six1: in red; Six4: in blue). Experiments were repeated 3-4 times with triplicate samples. Error bars represent the standard error of the mean (SEM) (n=9 or 12). T-test: p<0.001: \*\*\*. Alignments of Gga-Six1 (**b**) and Gga-Six4 3'UTR (**e**) sequences at the different miRNA sites with the seed sequence of their respective miRNA; mutated nucleotides are indicated in red. Vertical lines and stars indicate complementarity and identity between sequences, respectively.



**Fig. 4.9: Luciferase reporter assays – Gga-Dach1.** (a) Luciferase activity for Gga-Dach1 3'UTR constructs, WT and mutants, co-transfected with control siRNA (siC), si-1a or si-206. Normalised luciferase activity was plotted relative to the condition [WT or mutants construct + siC] (in brown). Experiments were repeated 1-6 times with triplicate samples depending on the constructs tested. Error bars represent the standard error of the mean (SEM) (n=9 or 12). T-test: p<0.001: \*\*\*. (b) Alignments of Gga-Dach1 sequence at the different miRNA sites with the seed sequence of their respective miRNA; mutated nucleotides are indicated in red. Vertical lines and stars indicate complementarity and identity between sequences, respectively.

With all putative target 3'UTRs an important decrease in luciferase activity was observed when either si-1a or si-206 was used; the decrease being even more dramatic with si-206 (si-1a: 32-35% decrease; si-206: 45-54% decrease) (Table 4.2). However, in all cases no rescue or weak rescue was observed when performing the assays with the mutant constructs. This could be explained by the possibility that either the mutated sites were not functional, or by the presence of additional cryptic/non-canonical miR-1a/206 sites, which were not predicted by the bioinformatics tools. However, the first possibility appears unlikely given that the decrease in luciferase activity observed with the WT constructs was highly reproducible and based on a minimum of 3 independent experiments, with little variation between each experiment. Further investigations would have to be done in order to further confirm Gga-Six1, Gga-Six4 and Gga-Dach1 as targets of miR-1a and miR-206.

### **c. Chicken Six1/4, Eya1-3, Dach1 expression patterns**

In parallel to the cloning of the 3'UTR sequences of the PSED members, fragments of their coding sequences were also cloned in order to study their expression profiles (see Appendix III List 2).

Gga-Eya1, Gga-Eya2, Gga-Eya3, Gga-Six1, Gga-Six4 and Gga-Dach1 were collected from Ensembl and NCBI. Sequence alignments of members of the Eya and Six families, respectively, from several species (human, mouse and chicken), were performed in order to design specific pairs of primers for each gene (data not shown); criteria for choosing the position of the primers were that the sequence was conserved between species, but sufficiently different from the other members of their respective family to be able to discriminate each member.

Total RNAs, from dissected chicken somites (day 4), were *in vitro* transcribed (see chapters 2.8 and 2.9; Appendix III Table 1). The cDNAs generated were used to do PCRs (see chapter 2.10); PCR products were then ligated into pGEMT-Easy vector (Promega), and used to produce DIG-labelled antisense and sense RNA probes by PCR (see chapter 2.7).

For Gga-Eya1, Gga-Six1 and Gga-Dach1, the cloning of cDNAs containing open reading frame sequences failed; instead pGEMT-Easy constructs containing a fragment of their 3'UTRs were used as template to generate probes; the 3'UTR being often quite AT-rich, these probes could potentially be more challenging to use in RNA ISH, compared to classic probes made from coding sequences.

RNA ISH were performed on chicken embryos fixed at different stages of development, from HH10 to HH23. Expression patterns at HH11-12, HH16 and HH21-22, are presented in Fig. 4.10. Sense probes were used as negative controls.

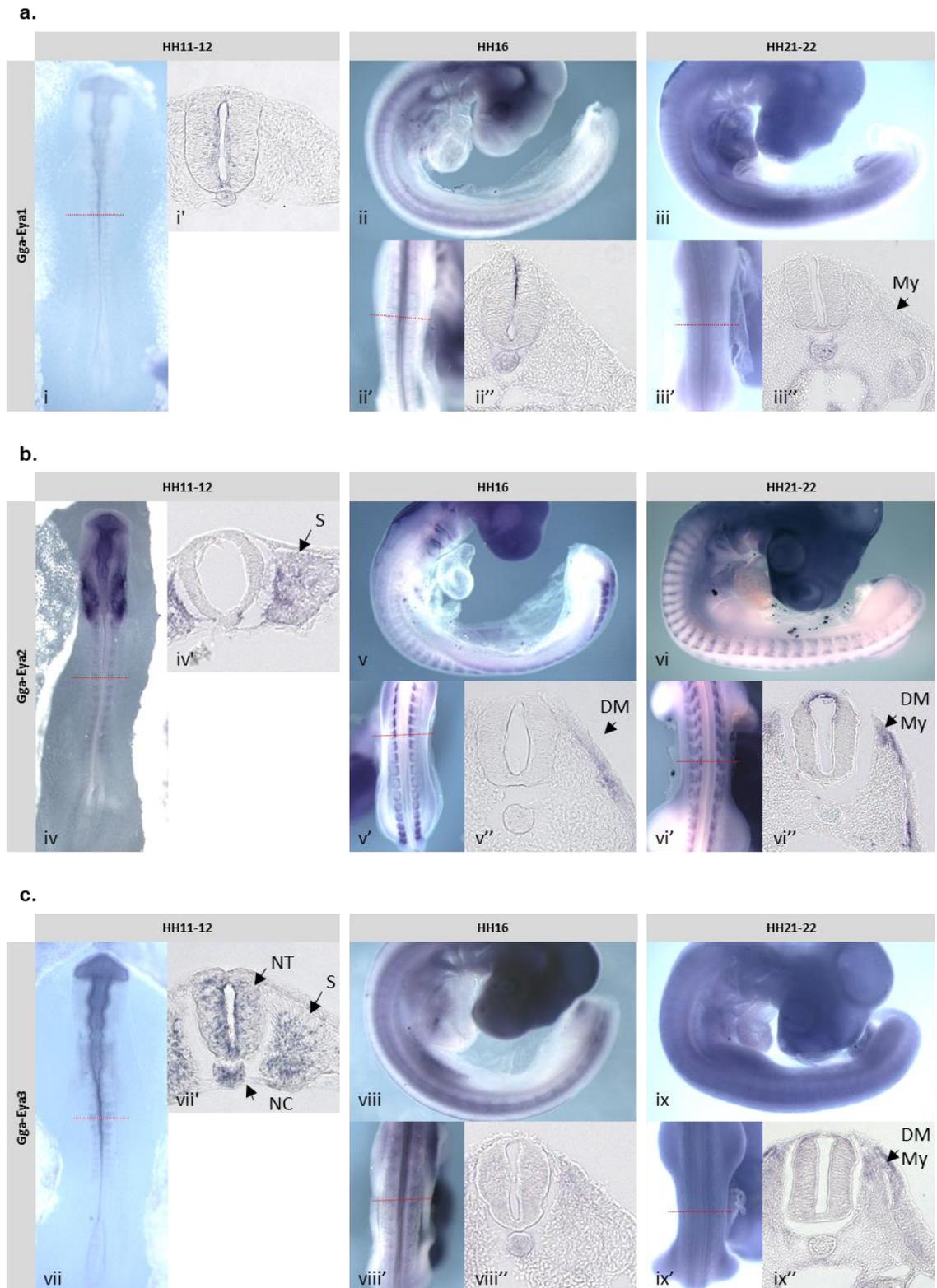
### **i) Gga-Eya1, Gga-Eya2 and Gga-Eya3**

At HH11-12, while Gga-Eya1 was not found expressed (Fig. 4.10a), and Gga-Eya4 was only detected in the head region (Fig. 4.3), Gga-Eya2 and Gga-Eya3 were already strongly expressed in the developing somites (Fig. 4.10b; c). At this stage, Gga-Eya3 was also found in the neural tube and the notochord (vii'); and Gga-Eya2 was expressed in the head region, in the presumptive cranial placodes (iv).

At HH16, Gga-Eya1 and Gga-Eya3 displayed a similar expression in the notochord, and a weak expression in the differentiating somites (ii''; viii''); Gga-Eya3 also appeared to be present in the developing limbs (viii). With a dynamic pattern, Gga-Eya2 was strongly expressed in the entire newly formed somites in the posterior region (v, v'), then as the somites differentiated, this expression became restricted to the dermomyotome (v''); Gga-Eya2 was also found in the branchial arches, but not in the limb buds (v).

At later stage (HH21-22), although Gga-Eya1 and Gga-Eya3 showed a widespread expression in whole-mount (iii; ix), on transverse sections Gga-Eya3 was found in the dermomyotome, in its dorsomedial lip, and in the myotome (ix''); Gga-Eya1 was still expressed in the notochord and displayed a weak expression in the myotome (iii''). At HH21-22, Gga-Eya2 was found strongly expressed in the ventromedial and dorsolateral lips of the dermomyotome, and in the myotome (vi'').

The differential expression profiles observed for Gga-Eya1 and Gga-Eya2 (Fig. 4.10a, b), at the different stages of development tested, are consistent with the published expression patterns described in chicken (Ishihara et al. 2008; Heanue et al. 1999; Mishima & Tomarev 1998).



**Fig. 4.10: Expression patterns of Gga-Eya1 (a), Gga-Eya2 (b) and Gga-Eya3 (c). (Part one).** WMISH performed with antisense DIG-labelled RNA probes, and transverse sections at HH11-12 (i, i'; iv, iv'; vii, vii'), HH16 (ii-ii''; v-v''; viii-viii''), and HH21-22 (iii-iii''; vi-vi''; ix-ix''). Embryos treated with the sense probes (negative control) did not show any expression (data not shown). Red dotted lines indicate the location of the transverse sections. Transverse section: 20x magnification. e: eye; DM: dermomyotome; My: myotome; NC: notochord; NT: neural tube; S: somite.

Gga-Eya2, expressed earlier than Gga-Eya1, and Gga-Eya4, displays a strong and dynamic expression in the developing somites, the dermomyotome and the myotome; however, Gga-Eya1 could not be detected in the dermomyotome, and neither Gga-Eya1 nor Gga-Eya2 were found in the limb buds. Although Gga-Eya3 displays a widespread expression in whole-mount, its expression at HH11-12 and HH21-22 is similar to those of Gga-Eya2 in the developing and differentiating somites. At HH16, only Gga-Eya3 appears to be expressed in the developing limb buds.

## ii) **Gga-Six1, Gga-Six4 and Gga-Dach1**

Gga-Six1 and Gga-Six4 displayed differential expression patterns at the different stages of development studied (Fig. 4.10d, e). At HH11-12, while Gga-Six1 was just becoming detectable (x, x'), Gga-Six4 was already well expressed in the neural tube, the notochord and the developing somites (xvi; xvi').

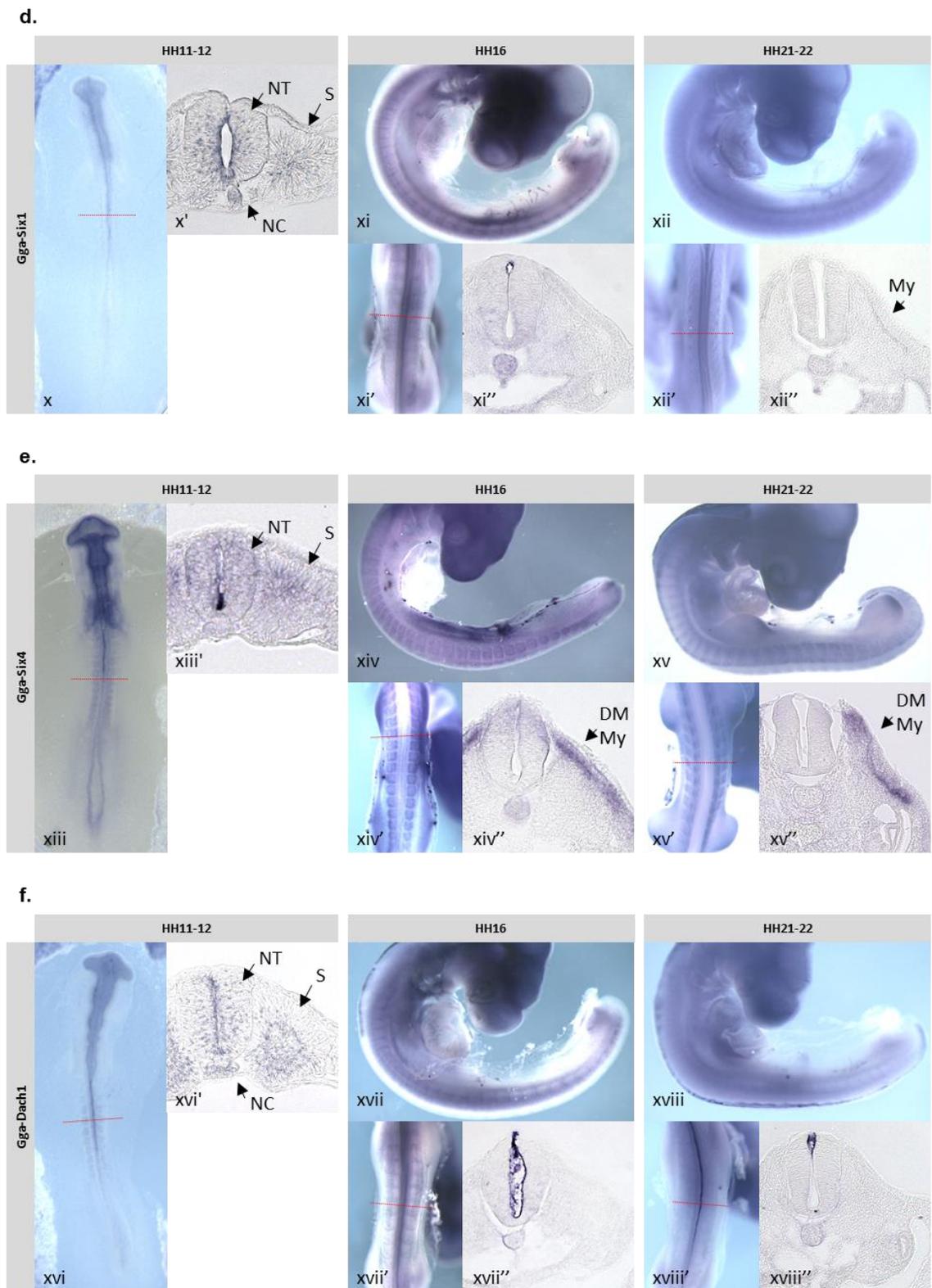
At HH16, they were both expressed in the notochord and in the neural tube (xi''; xiv-xiv''); Gga-Six1 being expressed in the most ventral part of the neural tube (xi''), and Gga-Six4 in the most dorsal part (xiv'').

From HH16, their respective expression in the notochord decreased (xi'', xii''; xiv'', xv''). Gga-Six4 was strongly expressed in the differentiating somites, in the dermomyotome and in the myotome (xiv''); Gga-Six1 was only weakly detected in these tissues (xi''). At this stage, Gga-Six4 was also expressed in the developing limbs (xiv); at HH21-22, this expression became restricted to the posterior part of the limbs (xv).

In these experiments, Gga-Dach1 was only detected at HH11-12 in the neural tube, the notochord and the developing somites (Fig. 4.10f; xvi, xvi'); no expression in the somites was observed at later stages (xvii-xviii').

The expression pattern observed for Gga-Six4 is consistent with the published expression pattern described in chicken (Esteve & Bovolenta 1999); Gga-Six4 is present in the head region, the notochord, in the somites, in the dermomyotome and the myotome. The expression profiles observed for Gga-Six1 and Gga-Dach1 are consistent with previously published patterns (Heanue et al. 1999; Heanue et al. 2002; Kida et al. 2004).

Gga-Six1 is expressed in the somites, in the myotome, however it is not found in the dermomyotome; and Gga-Dach1 is expressed in the neural tube at HH11-12, however it is not found in the somites or the developing limb buds as previously reported.



**Fig. 4.10: Expression patterns of Gga-Six1 (d), Gga-Six4 (e) and Gga-Dach1 (f). (Part two).** WMISH performed with antisense DIG-labelled RNA probes, and transverse sections at HH11-12 (x, x'; xiii, xiii'; xvi, xvi'), HH16 (xi-xi'; xiv-xiv'; xvii-xvii''), and HH21-22 (xii-xii'; xv-xv'; xviii-xviii''). Embryos treated with the sense probes (negative control) did not show any expression (data not shown). Red dotted lines indicate the location of the transverse sections. Transverse section: 20x magnification. e: eye; DM: dermomyotome; My: myotome; NC: notochord; NT: neural tube; S: somite.

### 4.3. Conclusion

In this chapter, *Eya4*, predicted target of miR-128, has been examined further and fragments of its coding region and 3'UTR sequences were cloned.

*Gga-Eya4* expression pattern was characterised for the first time in the model chicken. Found in the somites, in the dorsomedial lip of the dermomyotome and the myotome; its expression is similar to those of miR-128 in these tissues.

The predicted interaction between *Gga-Eya4* and miR-128 was investigated. *Gga-Eya4* 3'UTR constructs, WT and mutants, generated by FastCloning technique (Li et al. 2011), were used to perform luciferase reporter assays. These assays showed that miR-128 can target *Gga-Eya4* 3'UTR and reduce its expression activity; activity which can be rescued after mutation of miR-128 site. Other miRNAs were also tested, and amongst them miR-206 was also shown to be able to target *Gga-Eya4* 3'UTR, with slightly weaker effect than observed for miR-128. However, this effect was stronger when si-128 and si-206 were used in combination. This could suggest that miR-128 and miR-206, which can target *Gga-Eya4* 3'UTR on their own, might also be able to work in cooperation, in an additive manner, in order to have a stronger effect on the regulation of *Gga-Eya4* expression. This will need to be investigated further following on from the pilot experiment described here.

Using the same strategy, fragments of coding and 3'UTR sequences were cloned for *Gga-Six1*, *Gga-Six4*, *Gga-Eya1*, *Gga-Eya3* and *Gga-Dach1*.

Luciferase reporter assays showed that miR-128 and miR-133a can target *Gga-Eya1* 3'UTR, while miR-27b cannot, although it has the same seed sequence as miR-128 and was predicted to target the same site; miR-133 is also able to target *Gga-Eya3* 3'UTR. However, miR-133a and miR-499 did not seem to be able to target *Gga-Six4* 3'UTR. Other experiments were performed to determine if miR-1a and miR-206 could target *Gga-Six1*, *Gga-Six4* and *Gga-Dach1*. Although a decrease in luciferase activity was observed; this decrease could not be rescued after mutation of the different miRNA sites in their respective 3'UTR sequences.

RNA probes were also generated for *Gga-Six1*, *Gga-Six4*, *Gga-Eya1*, *Gga-Eya2*, *Gga-Eya3*, and *Gga-Dach1*, and RNA ISH performed on chicken embryos at different stages of development. Similar expression to those already published were observed.

Gga-Eya3, which had not been characterised in chicken before, seems to have a similar expression than what was described in mouse and zebrafish; widespread expression in whole-mount, neural tube and developing somites at HH11-12, dermomyotome and myotome at HH21-22.

Predicted by bioinformatics tools (chapter 3) and validated by *in vitro* luciferase reporter assays (this chapter), the interaction between Gga-Eya4 and miR-128 is going to be investigated further in chapter 5. MiRNA loss-of-function experiments, using antagomiRs (miRNA inhibitors), will be performed in order to (1) validate Gga-Eya4/miR-128 interaction *in vivo*, and (2) determine the impact of this interaction on myogenesis.

## **CHAPTER 5: MicroRNA/mRNA TARGET INTERACTIONS *in vivo* AND THEIR ROLE(S) DURING MYOGENESIS**

### **5.1. Introduction**

In chapter 4, miRNA/mRNA target interactions were investigated by luciferase reporter assays. Gga-Eya4 was identified as a target of miR-128 and miR-206; and miR-133a was shown to target Gga-Eya1 and Gga-Eya3. Although these experiments showed pertinent results, they were generated by using a cell-based approach, and needed to be confirmed *in vivo*.

In this chapter, miRNA/mRNA target interactions are going to be investigated using an *in vivo* model, the chicken embryo. The potential impact(s) of such interactions on myogenesis is also going to be assessed.

#### **MicroRNAs, the PSED network and myogenesis:**

PSED network and myogenesis:

The myogenic determination factors control entry into the myogenic program, which leads to the formation of skeletal muscle. Upstream of this obligatory step, other transcription factors, part of the regulatory PSED network, direct cells toward myogenesis; PAX, SIX, EYA and DACH are referred to as pre-myogenic factors.

In vertebrates, SIX1/4, EYA1/2/4, and DACH1/2, with overlapping expression patterns, are co-expressed in the myogenic precursor cells in the somites, in the dermomyotome and the myotome; at limb level, with PAX3/7, they play a crucial role in ensuring that the migrating myogenic precursor cells remain committed to their fate until they reach their final destination (Christ & Ordahl 1995; Xu, Cheng, et al. 1997; Oliver et al. 1995).

Since Heanue *et al.* showed in the chicken embryos that ectopic expression of Six and Eya resulted in activation of Pax3 and myogenic regulatory genes (Heanue et al. 1999), mouse mutants for the different PSED members have provided insight into their complex upstream roles in the activation and regulation of the myogenic programme.

As previously reported, Pax3 and Pax7 play an important role in the activation and control of the myogenic regulatory factors (MRFs) expression (Myf5 and MyoD1) (Williams & Ordahl 1994; Maroto et al. 1997; Bajard et al. 2006). While in Pax7<sup>-/-</sup> mutant mice skeletal muscle forms normally (Mansouri et al. 1996), Pax3<sup>-/-</sup> mutants have abnormal myotome formation, trunk muscle defects and absence of limb muscle (Bober et al. 1994; Goulding et al. 1994). Moreover, Pax3/Pax7 double mutants have major defects in myogenesis (Relaix et al. 2005).

Similarly, while in Six4 and Eya2 knock-out mice no developmental defects were observed (Ozaki et al. 2001; Grifone et al. 2007), Six1 and Eya1 mutants mice have important muscle deficiencies (Ozaki et al. 2004; Laclef et al. 2003; Grifone et al. 2007). In Six1/4 and Eya1/2 double mutants, defects are even more important with absence of all muscles derived from hypaxial dermomyotome (trunk muscles and limbs). In Six1/4 double mutant embryos, no Pax3-positive cells are detectable in the forelimb, and just a few in the hindlimb; Myf5 and MyoD-positive cells are absent (Tremblay et al. 1998; Giordani et al. 2007; Relaix et al. 2013). In addition, Six1/4/Myf5 mutants, with no expression of MyoD and no skeletal muscle formed, display a similar phenotype to what was observed in Pax3/Myf5 mutants (Relaix et al. 2013; Tajbakhsh et al. 1997). These results suggest that SIX and EYA are upstream of PAX.

Ohto *et al.* showed that transcription regulation of certain target genes by Six proteins requires cooperative interaction with Eya proteins (Ohto et al. 1999). In mouse, Six1/4 bind to Eya1/2 in the cytoplasm and translocate Eya into the nucleus. SIX, often associated with DACH, has been described as a repressor or weak activator, however, when interacting with EYA, the complex formed becomes a strong activator; this complex is then able to activate SIX target genes, such as Pax3 and MyoD1 (Grifone et al. 2005). Heanue *et al.* reported similar results in the chicken model where Eya2 interacts with Six1 and Dach2 in order to regulate Pax3 and therefore influence myogenic differentiation (Heanue et al. 1999).

In addition, it has also been shown that SIX/EYA complex can directly up-regulate MyoD and MyoG expression by targeting enhancer elements on their respective promoters (Tapscott 2005; Giordani et al. 2007; Spitz et al. 1998); activation being independent of Pax3. These results are consistent with the severe decrease of Myf5 and MyoD1 expression, in the myotome, observed in the Six1/Six4 double mutant mice (Tajbakhsh et al. 1997).

These effects on the activation of downstream transcription factors and muscle genes distinguish SIX/EYA from PAX regulation of myogenesis. While PAX3/7 control upstream events leading to myogenesis, with Pax3 active in the somite prior to SIX/EYA intervention; SIX/EYA complex plays a major role in the onset of hypaxial myogenesis, both directly through activation of MRFs (MyoD1 and MyoG), and indirectly through control of PAX3 (Buckingham & Rigby 2014).

MiRNAs and myogenesis:

Myogenesis starts in the dermomyotome and requires the commitment of a pool of cells into the skeletal muscle lineage. This process is associated with the activation of the MRFs (MyoD1, Myf5, Mrf4 and MyoG), and the downregulation of Pax3 (Williams & Ordahl 1994; Goulding et al. 1994; Gros et al. 2004).

Downregulation of Pax3 is essential to ignite the myogenic programme. Goljanek-Whysall *et al.* showed in the chicken that miR-1 and miR-206 directly target Pax3 and inhibit its expression (Goljanek-Whysall, Pais, et al. 2012). Consistent with this, the inhibition of miR-1 and miR-206 leads to de-repression of Pax3 and delayed myogenic differentiation in the somites (Goljanek-Whysall et al. 2011).

In skeletal muscle stem cells, miR-27b has also been identified to target the transcription factor Pax3. *In vivo* overexpression of miR-27b, leads to Pax3 downregulation in the myotome and premature differentiation, whereas in cell culture the inhibition of miR-27b permits Pax3 to induce increased proliferation and delay the differentiation process (Crist et al. 2009). Moreover, inhibiting miR-27b in regenerating muscle leads to an increase in Pax3 expression and fibres with smaller diameter.

Other miRNAs have also been shown to play a role in myogenesis. An increase in miR-128 expression was reported during myoblast differentiation (Sun et al. 2010). In addition, Shi *et al.* recently showed in mouse that ectopic expression of miR-128 induces the expression of skeletal muscle proliferation and differentiation related genes, such as Pax3/7 and the MRFs (Myf5 and MyoD1) (Shi et al. 2015). Conversely, these genes were down-regulated when miR-128 was repressed.

O'Brian *et al.* identified, in zebrafish, an upstream miRNA regulatory mechanism where miR-30a directly regulated myogenesis via inhibiting Six1a/b expression (O'Brien et al. 2014).

### **Functional experiments:**

To study miRNA/mRNA target interactions and investigate their potential role(s) in skeletal development, the chicken embryo, relatively large and accessible, is an ideal tool. Somites can be targeted with great precision by microinjection or/and electroporation, in order to enable the delivery of oligonucleotides or gene expression constructs.

Different approaches, involving loss-of-function and gain-of-function experiments, can be used to study the function(s) of miRNAs, and any target genes.

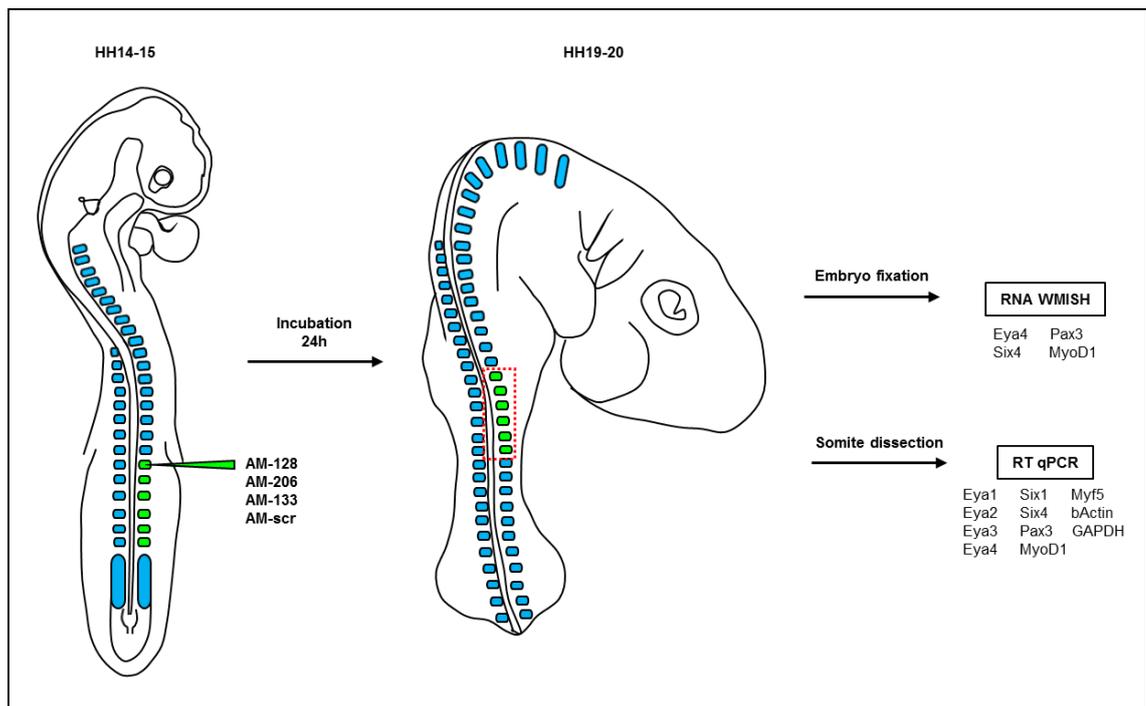
The silencing of miRNAs *in vivo* has been made possible by the development of chemically engineered oligonucleotides, called ‘antagomiRs’ (AM; Dharmacon) (Krützfeldt et al. 2005; McGlenn et al. 2009). These antagomiRs bind to their specific miRNAs and prevent them from interacting with their targets. MiRNAs can also be overexpressed by using mature miRNA mimics, or expression constructs containing miRNA-hairpin precursors; in addition to the endogenous miRNAs they will bind to their targets and prevent their expression.

The expression of target genes can be ‘knocked-down’ by vector-based approach using siRNA, or antisense morpholinos (MO) specifically designed to bind their respective mRNA 5’UTR regions, and block their translation (Voiculescu et al. 2008; Mende et al. 2008). Overexpression of target genes is also possible by using expression constructs containing cDNAs.

Followed by *in situ* analysis, or somite dissection and biochemical/molecular tests, such as western blot, RT-qPCR, or high-throughput sequencing, the results generated provide important information on miRNA function(s).

In order to validate *in vivo* the interaction between miR-128 and Eya4 (described in chapter 4), loss-of-function experiments were performed. Embryos, injected with an antagomiR (miRNA inhibitor) directed against miR-128 was then used to perform RNA *in situ* hybridisation (ISH) and identify potential phenotypes.

In parallel, injected somites, as well as non-injected contralateral somites (used as control) were dissected out and RT-qPCRs were performed, providing an insight into the potential role of miR-128 in the regulation of the PSED network. RNA ISH and RT-qPCR results will be presented in the first part of this chapter.



**Fig. 5.1: Functional experiments: miRNA loss-of-function – AntagomiR strategy.** Injection of antagomiR into the 6 most posterior somites (newly formed; indicated in green), on one side (right side), of HH14-15 chicken embryos. Embryos were incubated for 24 hours and harvested at HH19-20. Embryos were either fixed in paraformaldehyde (PFA) in order to perform RNA whole-mount *in situ* hybridisation (WMISH), or the injected somites were dissected out and used for real time quantitative PCR (RT-qPCR). The contralateral somites, from the non-injected side, were also dissected out and used as control.

Other miRNAs, miR-206 and miR-133, and their respective interactions with Eya4, and Eya1 and Eya3, were also investigated (described in chapter 4). RT-qPCR results will be presented in a second part.

## **5.2. Results and discussion**

In order to investigate the function(s) of miRNAs in skeletal muscle development, loss-of-function experiments were performed using antagomiRs. HH14-15 chicken embryos were injected in the 6 most posterior and newly formed somites, on one side, with one of the following antagomiRs: AM-128, AM-206, AM-133 or AM-scr. Scrambled antagomiR (AM-scr) was used as a control (scrambled miR-206 sequence). After 24-hour incubation, HH19-20 embryos were collected, and either fixed in paraformaldehyde (PFA) in order to perform RNA ISH, or the injected somites were dissected out and used for real-time quantitative polymerase chain reaction (RT-qPCR) (summarised in Fig. 5.1) (see chapter 2.20).

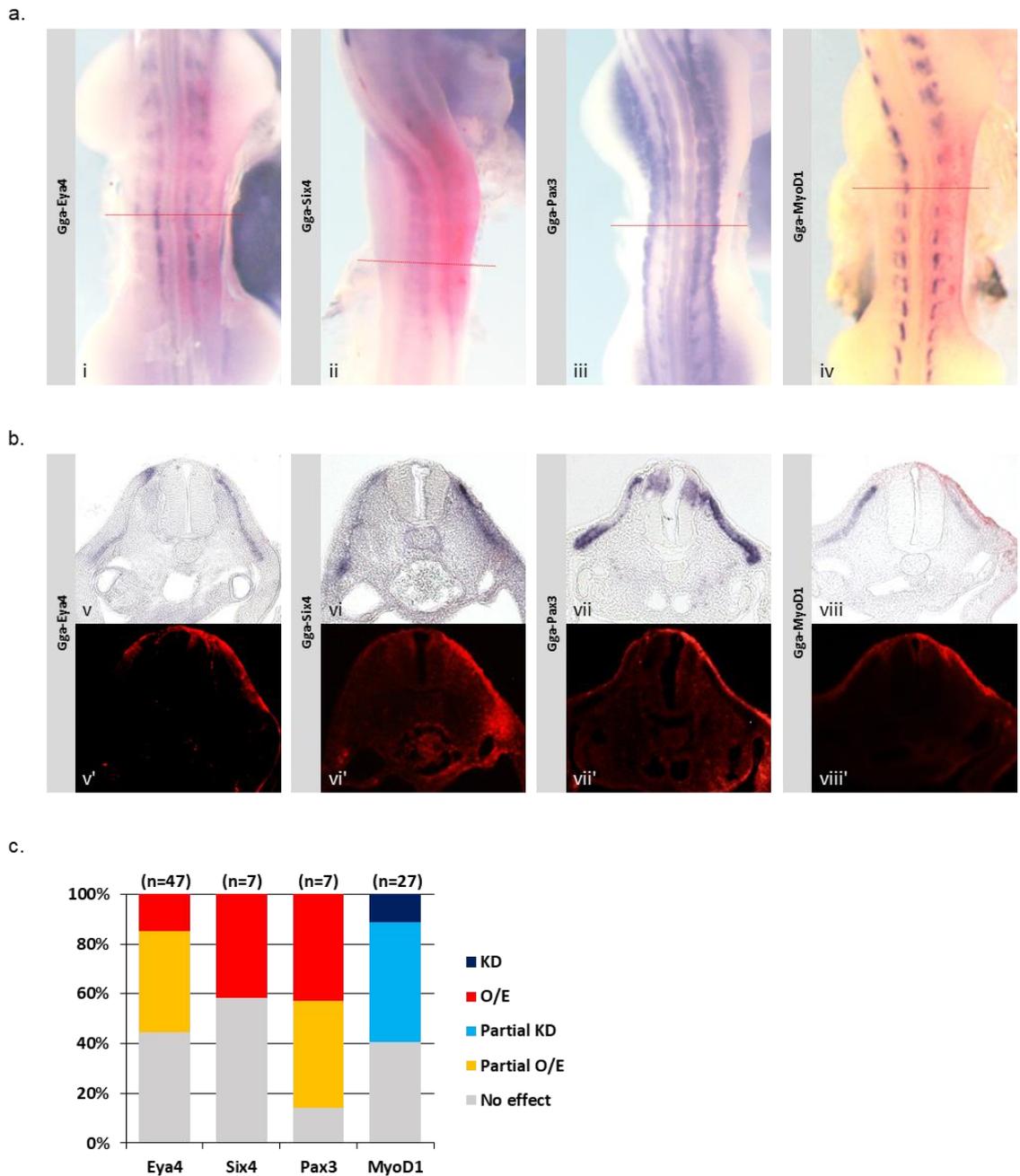
The optimal stage for injection was determined based on the respective expression of the different PSED genes we wanted to look at. As previously described in the literature and according to the profiles observed in chapter 4, Pax3, Six1/4 and Eya1-4 are all expressed at HH14-15. The antagomiR injections were performed in developing somites, before they start to differentiate.

### **5.2.1. MiR-128 regulates myogenesis via its interaction with Gga-Eya4**

#### **a. MiR-128 targets Gga-Eya4 *in vivo***

RNA ISH:

A first experiment was performed to determine for how long the embryos would have to be incubated in order to see a phenotype. After AM-128 injection, 4 groups of 10 embryos were incubated for 6h, 9h, 12h, and 24h, respectively. Embryos from each group were collected, and using a microscope with GFP filter, the somites injected with FITC-labelled AM-128 were localised.



**Fig. 5.2: Expression patterns of Gga-Eya4, Gga-Six4, Gga-Pax3 and Gga-MyoD1 after antagomiR-128 injection.** AntagomiR-128 (1 mM) was injected into the 6 most posterior somites, on one side (right side), of HH13-14 embryos. After 24h incubation, HH19-20 embryos were collected and fixed. Double RNA WMISH were performed using Digoxigenin (DIG)-labelled probes for Gga-Eya4 (a), Gga-Six4 (b), Gga-Pax3 (c), and Gga-MyoD1 (d) (purple). FITC labelled-antagomiR was detected by alkaline phosphatase coupled anti-FITC antibody developed with Fast Red. This localised the injected somites (red). Expression patterns on dorsal view. Transverse sections (red dotted lines in top panel) at the interlimb level showing Gga-Eya4 (e), Gga-Six4 (f), Gga-Pax3 (g) and Gga-MyoD1 (h) expression; antagomiR location was detected by fluorescent filter Alexa-Fluor-568 on the same sections. The contralateral non-injected side (left side), was used as control. (i) Quantification of phenotype observed. KD: knock down; O/E: overexpression. Transverse sections: 20x magnification.

Embryos which had reached the expected stage and were successfully injected (GFP-positive somites), were then used to perform a double ISH for Eya4 (Digoxigenin (DIG)-labelled antisense probe; revealed in Alkaline phosphatase/NTMT buffer) and for AM-128 (FITC-labelled; revealed in Fast Red buffer) (see chapter 2.5). The expression pattern for each embryo was analysed in whole-mount, and after sectioning, in the injected somites.

As antagomiRs are repressing miRNA activity, with a repression of miR-128 by AM-128, and given the fact that Gga-Eya4 is a potential target of miR-128, a de-repression of Gga-Eya4 expression was expected.

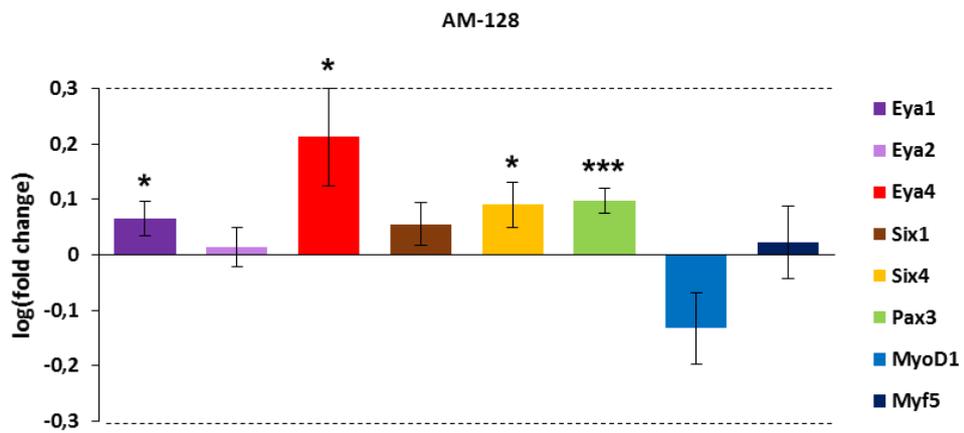
Embryos incubated for 24h after injection displayed the strongest phenotype with a stronger expression signal detected for Gga-Eya4 transcript in the myotome. After 24h, the injected somites were localised in the region facing the dorsal part of the developing forelimb and the anterior part of the interlimb region (Fig. 5.1).

AM-128 injection experiments with 24-hour incubation were repeated to increase the number of embryos (n=47) and an example of the strongest encountered phenotype is presented in Fig. 5.2a, e. The expression of Gga-Eya4 did not seem to be strongly different when compared to the contralateral non-injected side (left side) in whole-mount (Fig. 5.2a); however, on transverse section, Gga-Eya4 expression signal was increased, especially in the central part of the myotome. The phenotype was observed in 55.3% of the embryos: 7 embryos out of 47 (15%) with an overexpression similar to the one presented in Fig. 5.2a, e; 19 embryos with partial overexpression (40.4%). However, 44.7% of the embryos showed no change in expression on the injected side compared to the non-injected side (Fig.5.2i).

These results indicated that when miR-128 was inhibited by AM-128, Gga-Eya4 expression was de-repressed in the differentiating somite, in the myotome; suggesting that miR-128 could regulate the expression of Gga-Eya4 *in vivo*.

#### RT-qPCR:

In order to quantify potential changes in expression of Gga-Eya4, and consolidate the ISH results, RT-qPCRs were performed.



**Fig. 5.3: RT-qPCR results - PSED and MRF (MyoD1 and Myf5) transcript levels in somites injected with antagomiR (AM)-128.** Results are expressed in log(fold change). For each gene and each experiment, the injected somite data have been normalised to (GAPDH +  $\beta$ -actin) housekeeping gene, and then compared to the contralateral non-injected somite data of the same embryo. Number of independent experiments for each gene tested after AM-128 injection: Eya1, Eya2, Six1 (n=10); Eya4, Pax3 (n=8); Six4, MyoD1 (n=7); Myf5 (n=12). T-test: p<0.05: \*, p<0.001: \*\*\*. Dotted lines indicate a 2-fold change.

Injected somites from 3-4 HH19-20 embryos of 10 independent experiments were dissected out. The contralateral non-injected somites were also dissected out and used as control. Their total RNA were extracted and cDNAs synthesised (see chapter 2.20).

RT-qPCRs were performed using SYBR Green Master mix and specific designed primers for Gga-Eya4. Results were analysed based on the Relative Standard Curve method (Larionov et al. 2005); method allowing to quantify differences in the expression level of a specific target gene (for example Gga-Eya4) between different samples (in this study: injected vs non-injected somites). Two housekeeping genes were used for these experiments: Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), an important glycolytic pathway enzyme, and  $\beta$ -actin, essential for the structure and kinetics of the cytoskeleton (Choi et al. 1991; Kozera & Rapacz 2013).

Although housekeeping genes should not vary in the tissue under investigation, or in response to experimental treatment, many studies showed that their expressions can vary considerably (Vandesompele et al. 2002). In order to minimise these unwanted variations, data obtained for GAPDH and  $\beta$ -actin were averaged. For all samples, levels of both target and housekeeping genes (GAPDH,  $\beta$ -actin) were assessed. Gga-Eya4 data were normalised to the averaged relative quantification of  $\beta$ -actin and GAPDH housekeeping genes. Results for injected somite samples were then compared to their respective contralateral non-injected somite samples. Results, expressed in log(fold change) were then plotted on a linear scale where the x-axis corresponds to the non-injected condition set at 0 ( $\log(1)=0$ ).

RT-qPCR results, presented in Fig. 5.3, showed a 1.5-fold higher expression level of Gga-Eya4 in somites after AM-128 injection (8 independent experiments; t-test:  $p<0.05$ ); this also suggests that miR-128 may play a role in the regulation of Gga-Eya4 *in vivo*.

The results of these functional experiments (RNA ISHs and RT-qPCRs) are consistent with the luciferase reporter assays results presented in chapter 4.

In the reporter assays, the luciferase activity was reduced when Gga-Eya4 3'UTR-containing construct was co-transfected with si-128, and rescued after mutation of the miR-128 site in Gga-3'UTR; these results suggested that miR-128 was able to target Gga-Eya4 3'UTR and decrease its expression *in vitro*, and that an intact seed site was required for this interaction

The RNA ISHs and RT-qPCRs performed after AM-128 injections in somites of chicken embryos, showed that inhibiting miR-128 resulted in a significant de-repression of Gga-Eya4 expression *in vivo*.

Overall, these experiments validated miR-128/Gga-Eya4 interaction *in vitro*, and confirmed Gga-Eya4 as a target of miR-128 *in vivo*.

### **b. MiR-128, Gga-Eya4 and myogenesis**

In order to investigate the potential effect of miR-128/Gga-Eya4 interaction on the PSED network and myogenesis, functional experiments combining RNA ISH and RT-qPCRs were performed on the same AM-128 injected somites samples, and examined the expression of Gga-Eya1 and Gga-Eya2, Gga-Six1 and Gga-Six4, Gga-Pax3, Gga-MyoD1 and Gga-Myf5 (Fig. 5.3).

Eya1/2:

RT-qPCR showed that while no variation in expression was observed for Gga-Eya2, Gga-Eya1 expression level was 1.2-fold higher than the control (10 independent experiments; t-test:  $p < 0.05$ ) after AM-128 injection (Fig. 5.3).

This increase of Gga-Eya1 expression is consistent with the luciferase reporter assays results presented in chapter 4 (Fig. 4.7b). However, there was only a partial rescue after mutating the seed sites. Two miR-128 sites were identified in Gga-Eya1 3'UTR. The co-transfection of Gga-Eya1 construct with si-128 showed a decrease in luciferase activity of about 40%, and after mutation of these 2 miR-128 sites the luciferase activity was still reduced (73%) compared to control. It is possible that sequences outside the seed mediate interaction between miR-128 and the Gga-Eya1 3'UTR.

Taken together, these results indicate that miR-128 may interact with Gga-Eya1 *in vitro* and *in vivo*.

Six1/4:

Although Gga-Six1 and Gga-Six4 were not predicted to be targeted by miR-128, a de-repression of their expression levels was observed after antagomiR injection; Gga-Six4 expression level was increased by 1.25-fold (7 independent experiments; t-test:  $p < 0.05$ ) (Fig. 5.3). In addition, RNA ISH performed on AM-128 injected embryos also showed an increase in Gga-Six4 expression in the somites, in the dorsomedial lip of the dermomyotome and the myotome (Fig. 5.2b, f).

This phenotype was observed in 5 of 7 embryos (71.4%); no phenotype was observed for the 2 other embryos (Fig. 5.2i).

These results are consistent with a potential indirect effect and cross-regulation between SIX and EYA co-factors, which have been shown to form a strong activator complex that activates SIX target genes (Heanue et al. 1999; Ohto et al. 1999; Grifone et al. 2005). For example, with more Gga-Eya4 available, more SIX/Gga-Eya4 complex could potentially be formed, thus promoting Gga-Six4 expression. Gga-Six4 could be a potential partner for Gga-Eya4.

#### Pax3:

RT-qPCR results showed a relative increase in Gga-Pax3 expression level after AM-128 injection in the somites; this was statistically significant (8 independent experiments; 1.3-fold change; t-test:  $p < 0.001$ ). Moreover, RNA ISH also showed an increase of Gga-Pax3 in the dermomyotome, especially in its central part where its expression is usually weak (contralateral non-injected side) (Fig. 5.2c, g). An increase in Gga-Pax3 was observed in the majority of the embryos (85.8%): 3 of 7 embryos with an overexpression similar to the one presented in Fig. 5.2c, g; and another 3 embryos with partial overexpression (42.9%). One of the embryos showed no change in expression on the injected side compared to the non-injected side (Fig.5.2i).

This increase in Gga-Pax3 expression is consistent with the fact that PAX3 is a known target of SIX (Grifone et al. 2005). Moreover, it is also possible that Gga-Pax3 was a target of miR-128; a miR-128 site having been predicted, by TargetScan, in the human PAX3 3'UTR.

#### MRFs (MyoD1 and Myf5):

RT-qPCR results showed that while Gga-Myf5 expression did not seem to be affected by AM-128 injection, Gga-MyoD1 expression level was decreased by 0.78-fold compared to its expression in the contralateral non-injected somites (7 independent experiments; t-test:  $p = 0.07$ ). RNA ISH also showed a decrease in Gga-MyoD1 expression in the somite, in the myotome (Fig. 5.2d, h). A decrease in Gga-MyoD1 was observed in the majority of embryos (59.2%): 3 of 27 embryos with a phenotype similar to the one presented in Fig. 5.2d, h; and 13 of 27 embryos with partial loss of MyoD1 expression (Fig.5.2i).

These results are consistent with the fact that Gga-Pax3 remained expressed in the injected somites after antagomiR injection.

Thus, because Gga-Pax3 was not efficiently downregulated, the myogenic programme could not be activated, and therefore myogenic differentiation markers, including MyoD1 could not be expressed (Williams & Ordahl 1994; Goulding et al. 1994; Gros et al. 2004).

### **5.2.2. MiR-206/Gga-Eya4, miR-133/Gga-Eya1 and miR-133/Gga-Eya3 interactions *in vivo***

#### **a. MiR-206/Gga-Eya4 interaction and myogenesis**

In chapter 4, a potential interaction between miR-206 and Gga-Eya4 was identified by luciferase reporter assays (Fig. 4.5). In order to assess whether this interaction plays a role *in vivo*, a loss-of-function experiment, using AM-206 was performed, following the same strategy previously described.

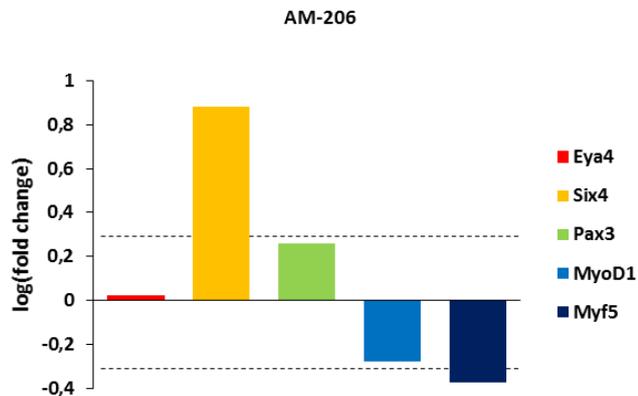
The 6 most posterior and newly formed somites, on one side, of HH14-15 embryos were injected with AM-206. After 24h incubation, HH19-20 embryos were collected, analysed, and injected somites were dissected in order to perform RT-qPCR.

Although these results are only based on 1 experiment, they provided pertinent information (Fig. 5.4a). No modification in Gga-Eya4 expression was observed (maybe due to the n=1), however, Gga-Six4 and Gga-Pax3 expression levels were de-repressed, with a strong 8-fold change increase for Gga-Six4 and nearly 2-fold change increase for Gga-Pax3. Importantly, a relative decrease in Gga-MyoD1 and Gga-Myf5 expression (about 2-fold change) was also observed.

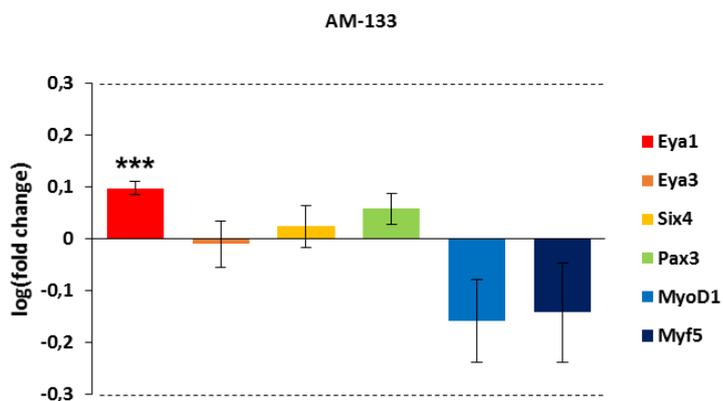
Gga-Pax3 increase, and Gga-MyoD1 and Gga-Myf5 decrease, are consistent with previous work showing, in chicken, that miR-206 directly targets Pax3 and inhibits its expression leading to a delayed myogenesis (Goljanek-Whysall et al. 2011). The de-repressed Gga-Pax3 expression could also be explained by the strong expression of Gga-Six4; PAX3 being a target of SIX.

More experiments will have to be performed in order to confirm these observations.

a.



b.



**Fig. 5.4: RT-qPCR results - PSED and MRF (MyoD1 and Myf5) transcript levels in in somites injected with antagomiR (AM)-206 (a) and AM-133 (b).** Results are expressed in log(fold change). For each gene and each experiment, the injected somite data have been normalised to (GAPDH +  $\beta$ -actin) housekeeping gene, and then compared to the contralateral non-injected somite data of the same embryo. (a) Number of independent experiments for each gene tested after AM-206 injection: n=1. (b) Number of independent experiments after AM-133 injection was n=5 for Eya1, Eya3, Six4, Pax3 and MyoD1. T-test:  $p < 0.001$ : \*\*\*. Dotted lines indicate a 2-fold change.

## **b. MiR-133/Gga-Eya1 and miR-133/Gga-Eya3 interactions *in vivo***

In chapter 4, Gga-Eya1 and Gga-Eya3 were identified as potential targets of miR-133 by luciferase reporter assays (Fig. 4.5). In order to assess whether this interaction plays a role *in vivo*, loss-of-function experiments, using AM-133 were performed, following the same strategy previously described; the 6 most posterior and newly formed somites, on one side, of HH14-15 embryos were injected with AM-133.

RT-qPCR results showed a significant 1.25-fold increase in the expression of Gga-Eya1 after antagomiR injection (5 independent experiments; t-test:  $p < 0.001$ ) (Fig. 5.4b). However, no change in Gga-Eya3 expression level was observed. Although not statistically significant, results also showed weak increase of Gga-Pax3 expression level, and a 1.5-fold decrease in Gga-MyoD1 and Gga-Myf5 expression levels. More independent experiments will have to be done in order to confirm these observations.

These results are consistent with study done in mouse, zebrafish and *Xenopus* reporting potential role(s) for miR-133 in promoting skeletal muscle differentiation (Kim et al. 2006; Chen et al. 2006; Mishima et al. 2009; Liu et al. 2008; Liu et al. 2011).

### **5.2.3. Control experiments – Scrambled antagomiR (AM-scr) injections**

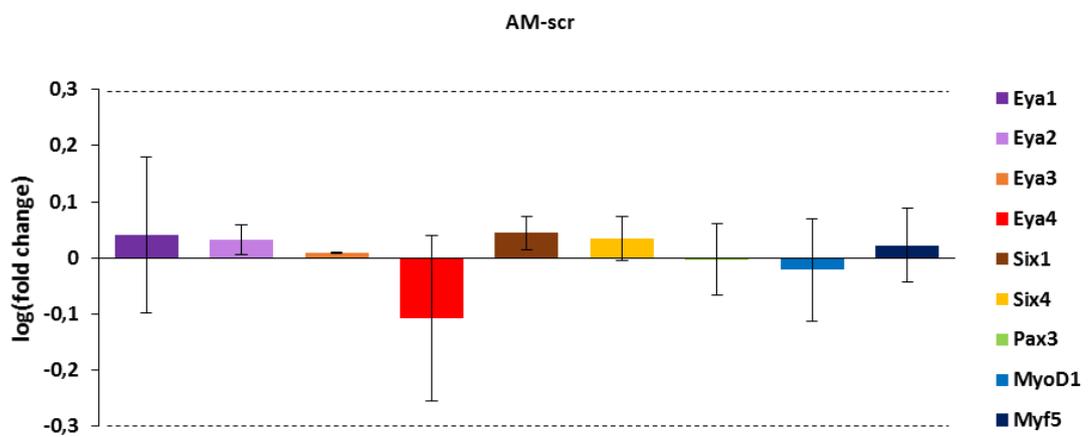
Although the contralateral non-injected side of each tested embryo was used as a control, additional experiments using a scrambled antagomiR (AM-scr; scrambled AM-206 sequence) were performed in order to determine the potential impact of the procedure and of injecting an antagomiR into somites.

Following the same procedure previously described, the 6 most posterior somites, on one side, of HH14-15 embryos were injected with AM-scr.

However, due to limited availability of the AM-scr, it was not possible to generate enough material in order to perform RNA ISH for all the genes of this study.

Instead, RT-qPCRs were performed on AM-scr injected somite samples. Since this is more sensitive and uses less material, this method allowed to look at expression level changes for several genes. Results, presented in Fig. 5.5, showed no significant change in expression, for all the PSED and MRF genes tested, between injected and non-injected somites.

RNA ISH experiments after AM-scr injection will be done in order to complete the study and validate the results observed in the different functional experiments.



**Fig. 5.5: RT-qPCR results - PSED and MRF (MyoD1 and Myf5) transcript levels in in somites injected with antagomiR scrambled (AM-scr).** Results are expressed in log(fold change). For each gene and each experiment, the injected somite data have been normalised to (GAPDH +  $\beta$ -actin) housekeeping gene, and then compared to the contralateral non-injected somite data of the same embryo. Number of independent experiments for each gene tested after AM-scr injection: Eya1, Six1 (n=3); Eya2, Eya3 (n=2); Eya4, Six4, Pax3, MyoD1 (n=5). Dotted lines indicate a 2-fold change.

### 5.3. Conclusion

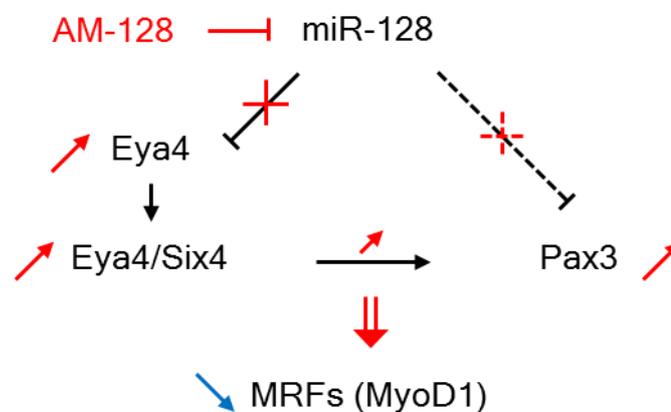
In this chapter, miRNA/mRNA target interactions were investigated by performing miRNA loss-of-function experiments, using antagomiRs, followed by RNA ISH and/or RT-qPCR.

MiR-128/Gga-Eya4 interaction, previously validated by luciferase reporter assays in chapter 4 (*in vitro*), was validated in this chapter by RNA ISH and RT-qPCR after AM-128 injection (*in vivo*). After AM-128 injection, a significant increase in Gga-Eya4 was observed (1.5-fold increase; 8 independent experiments; t-test:  $p < 0.05$ ). In addition, a significant increase in Gga-Six4 and Gga-Pax3 expression levels was observed (Gga-Six4: 1.25-fold increase; 7 independent experiments; t-test:  $p < 0.05$ ; Gga-Pax3: 1.3-fold increase; 8 independent experiments; t-test:  $p < 0.001$ ), whereas Gga-MyoD1 expression was decreased.

MiR-128/Gga-Eya1, miR-206/Gga-Eya4, miR-133/Gga-Eya1 and miR-133/Gga-Eya3 interactions were also investigated by RT-qPCR. After AM-128 and AM-133 injections, a significant increase in Gga-Eya1 expression level was observed (after AM-128 injection: 1.2-fold increase; 10 independent experiments; t-test:  $p < 0.05$ ; after AM-133 injection: 1.25-fold increase; 5 independent experiments; t-test:  $p < 0.001$ ), but not for Gga-Eya3. No change in Gga-Eya4 expression was observed after AM-206 injection; however this result was only based on one experiment. After AM-128, AM-133 and AM-206 injection, the expression levels of Gga-Six4 and Gga-Pax3 were de-repressed, and Gga-MyoD1 and Gga-Myf5 were relatively decreased; no change in Gga-Myf5 was observed after AM-128 injection.

Although Gga-Eya4 appears to be a direct target of miR-128 *in vitro* and *in vivo*, the observed effects after AM-128 injection on the expression of the other members of the PSED network and MRFs could only be indirect; to the exception of Gga-Pax3 (to be validated) any other PSED members or MRFs were predicted to be targeted by miR-128.

Based on elements from the literature and on the previous observations, a model was proposed to explain the impact that miR-128/Gga-Eya4 interaction could have on the PSED network, the MRFs, and on myogenesis (Fig. 5.6).



**Fig. 5.6: Proposed model – MiR-128/Eya4 interaction and potential effect on the PSED network and the MRFs.** AM-128 inhibits miR-128. In the absence of miR-128-mediated negative regulation, Eya4 expression is de-repressed. More Eya4 transcripts become available to interact with Six4 and both proteins can form a strong transcriptional activator complex. This complex can then activate SIX target genes, like Pax3. Pax3 expression is maintained or up-regulated by SIX action, and potentially also through the lack of direct repression by miR-128. In addition, the MRF MyoD1 is negatively regulated, possibly due to the maintained expression of Pax3.

From the literature it is known that: SIX and EYA can interact together to form a transcriptional activator complex, and one of the targets of SIX is PAX3 (Heanue et al. 1999; Ohto et al. 1999; Grifone et al. 2005); PAX3 downregulation is essential to ignite the myogenic programme, and myogenic differentiation is characterised by the activation of the MRFs (Goulding et al. 1994; Williams & Ordahl 1994; Maroto et al. 1997; Gros et al. 2004; Bajard et al. 2006).

After injection of AM-128 in developing somites of HH14-15 chicken embryos, and 24h incubation, the embryos are at HH19-20. At HH19-20, the injected somites are located in the region facing the dorsal part of the developing limb and the anterior part of the interlimb region. The RNA ISH and RT-qPCR performed on these AM-128 injected somites showed that they express more Gga-Eya4, Gga-Six4, and Gga-Pax3, but less Gga-MyoD1, compared to the contralateral non-injected somites.

We propose that after knock-down of miR-128 (inhibited by AM-128), Gga-Eya4 could be de-repressed (fails to be downregulated) (miR-128 directly targets Gga-Eya4). More Gga-Eya4 transcripts would become available, and the chances for interacting with one of its SIX co-factors (SIX1 and/or SIX4), probably Gga-Six4 (based on RNA ISH and RT-qPCR results), would be increased (overexpression of Gga-Six4 in injected somites). Considering that Gga-Eya4 and Gga-Six4 could interact together, they would form a strong transcriptional complex able to potentially activate Gga-Pax3 (overexpression of Gga-Pax3 in injected somites). This increase in Gga-Pax3 expression would prevent Gga-MyoD1 from being expressed (downregulation of Gga-MyoD1 in injected somites).

What would happen in the context of a normal myogenesis when the different tissues composing a somite and the location of the different actors potentially involved in the proposed model are taken into account?

MiR-128 would be necessary to downregulate Gga-Eya4 expression in the myotome. Gga-Eya4 being repressed, it could no longer interact with Gga-Six1/4 and form the SIX/EYA complex. Expressed in the dermomyotomal lips (dml and vll) and in the myotome, Gga-Six1/4, either repressor or weak activator on its own, would no longer be able to regulate the expression of its target Gga-Pax3 (de-repression of Gga-Pax3 in the dermomyotome after AM-128 injection (Fig. 5.2g)).

Pax3 being expressed in the dermomyotome it could suggest potential indirect mechanisms (FGF signalling) implying myotome/dermomyotome crosstalk/interaction. It is also possible that, similarly to what Goljanek-Whysall *et al.* showed with the regulation of Gga-Pax3 by miR-1/miR-206 (Goljanek-Whysall *et al.* 2011), miR-128 could also regulate Gga-Pax3, either directly or not, in the dermomyotomal progenitor cells migrating to populate the central part of the myotome. With a decrease in Gga-Pax3, the initiation of the myogenic programme would become possible, with activation of the MRFs, such as MyoD1, in the myotome.

These functional experiments provided the first elements in order to better understand miR-128 functions in skeletal muscle development, in the chicken.

Additional functional experiments will have to be done in order to consolidate the results presented in this chapter, and try to obtain more support for the proposed model. For example, a Gga-Eya4 expression construct has been generated (pCAB-Gga-Eya4 full length-HA tag) and could be used in overexpression experiments in order to see if it gives the same phenotype observed in the miR-128 loss-of-function experiments. An Eya4 morpholino could also be used to try to rescue the phenotype observed after AM-128 treatment.

## CHAPTER 6: DISCUSSION AND CONCLUSION

### 6.1. Summary and discussion

The aim of this project was to better understand the mechanisms underlying interactions between miRNAs and their mRNA targets, during skeletal muscle development. The investigation focused on miR-128, the interaction with one of its candidate targets, *Eya4*, and the potential impact of this interaction on myogenesis in the chicken embryo.

#### 6.1.1. MicroRNA characterisation

The expression patterns of 23 miRNAs, predicted to be expressed in skeletal muscle, were characterised by locked-nucleic acid *in situ* hybridisation (LNA ISH).

The expression patterns observed for the muscle-specific miRNAs, or myomiRs, miR-1a, miR-133a/b, and miR-206 were consistent with profiles already published in chicken (Goljanek-Whysall et al. 2011; GEISHA database) and in *Xenopus* (Ahmed et al. 2015; XenmiR database). Moreover, these LNA ISH also provided new insights into their respective expressions. No expression information had been reported for miR-1a and miR-206 at early stages, and no somitic expression had been detected for miR-1a, miR-206 or miR-133a in the somites before HH14-15 (Geisha database).

We observed that miR-1a was already expressed at HH9, in the heart. At HH11-12, miR-133a was already strongly expressed in the heart and the most anterior somites, while miR-1a and miR-206 were just starting to be detectable in the most anterior somites at HH12-13 (Fig. 3.1; Fig. 3.2).

Interestingly, with only one nucleotide of difference in the middle of its sequence compared to miR-1a, the expression of miR-1b was less specific with a lot of background in whole-mount. Sectioning showed that miR-1b was expressed in somites, in the myotome, at HH14-15 and onwards (Fig. 3.1).

MiR-133b and miR-133c had expression profiles similar to miR-133a. This could be explained by the fact that their sequences are extremely similar and only differ by 1 or 2 nucleotides. However, their expression profiles were not identical indicating that even 1 nucleotide of difference is enough to make a difference. This is also likely to be dependent on the positioning of the LNA within the probe, however the supplier (Exiqon) does not reveal this information.

Interestingly, miR-133b and miR-133c seemed to start to be expressed with a slight delay compared to miR-133a. While miR-133a was already strongly represented in the somites at HH11-12, miR-133b was not expressed yet, and miR-133c was only just becoming detectable (Fig. 3.3).

16 somitic miRNAs were also characterised, providing for most of them the first description of their expression patterns in the chicken embryo. All expressed in skeletal muscle, they displayed similar but also specific patterns in the somites, as well as other tissues for some of them (Fig. 3.4) (Ahmed et al. 2015).

MiR-128, expressed in the neural tube and the developing somites at HH11-12, and from HH15, mostly in skeletal muscle tissues (differentiating somites, myotome, and limb buds) (Fig. 3.5), was selected for further investigation.

### **MiR-128: myomiR?**

MiRNAs can be divided into two categories: miRNAs that are exclusively or preferentially expressed in muscle, the myomiRs (McCarthy 2008); and miRNAs expressed exclusively in non-muscle tissue or broadly expressed across many cell types. Both categories have significant impacts on muscle proliferation and differentiation (Wang 2013).

The first miRNAs classified as myomiRs, were miR-1a, miR-133a and miR-133b, and miR-206 (Lagos-Quintana et al. 2002; McCarthy & Esser 2007; McCarthy 2008) due to their specific expression in skeletal muscle tissue. MiR-206 is expressed specifically in somites; miR-1a and miR-133a/b, as well as being expressed in the somites, are also found in the heart (Sempere et al. 2004; Sweetman et al. 2008).

However, recently, the myomiR group has been extended to include miR-208a/b, miR-486 and miR-499 (van Rooij et al. 2007; van Rooij et al. 2009; Small et al. 2010).

To the exception of miR-208a, expressed predominantly in the heart (van Rooij et al. 2007), they are all expressed in both cardiac and skeletal muscles.

Some studies have reported evidence that not all myomiRs are solely expressed in a muscle-specific manner but may be detected in low levels in other tissues; however, their main function has to be confined to muscle. For example, miR-486 is sometimes considered ‘muscle-enriched’ rather than ‘muscle-specific’ as it is also expressed in other tissues (lung, bladder in adult mouse) (Small et al. 2010; Lee et al. 2008).

MiR-128 was first identified in mouse, where its expression level increases during brain development, and is maintained in adult brain tissues (Lagos-Quintana et al. 2002; Smirnova et al. 2005). Xu *et al.* and Kapsimali *et al.* reported similar observations in chicken (18-day embryos) and zebrafish (3-day embryos) (Xu et al. 2006; Kapsimali et al. 2007).

MiR-128 was found in cardiac tissue. A study done by Witman *et al.*, demonstrated that miR-128 regulates the expression of the transcription factor Islet1, a marker for cardiac progenitor cells, during newt cardiac regeneration (Witman et al. 2013).

MiR-128 was also found in adult mouse muscle (Sempere et al. 2004; Lee et al. 2008); adult and embryo porcine skeletal muscle (Zhou et al. 2010); and adult and embryo (somites) chicken skeletal muscle (Darnell et al. 2006; Lin et al. 2012; Abu-Elmagd et al. 2015). In mouse, the inhibition of insulin receptor substrate 1, Irs1, by miR-128, leads to the inhibition of myoblast proliferation and induction of myotube formation (Motohashi et al. 2013). These results are consistent with the increase in miR-128 expression observed during myoblast differentiation in differentiating mouse C2C12 myoblast cells (Sun et al. 2010). In addition, Shi *et al.* recently showed in mouse that ectopic expression of miR-128 affected the expression of skeletal muscle proliferation and differentiation related genes, such as Pax3/7 and the MRFs (Myf5 and MyoD1) (Shi et al. 2015).

With this present work, we showed by LNA ISH that miR-128 was expressed in skeletal muscle, but also in the neural tube at early stages during chicken development (Fig. 3.5). No expression in the heart was detected at any stages included in this study.

Based on these observations, miR-128 does not seem to match the requirements for being classified as a skeletal muscle-specific myomiR, like miR-1a, miR-133a/b, and miR-206. However, it would be reasonable to classify miR-128 as a brain and muscle-enriched miRNA.

### 6.1.2. MiR-128 target identification

In order to better understand miR-128 function(s), predicted miR-128 targets were collected from TargetScan database (n=507). By using a combination of different computational tools (DAVID and g:Profiler), 10 ‘muscle’ targets were identified (Table 3.6), and amongst them, Eya4, member of the PAX-SIX-EYA-DACH (PSED) network.

Eya4, as well as other members of the PSED network (Six1/4, Eya1/2/3, and Dach1), were investigated further, due to the interesting role(s) they seem to play together with Pax3/7 during skeletal myogenesis.

The miRanda algorithm was used to scan selected miRNA sequences against the 3’UTR sequences of PSED genes (Table 3.7; APPENDIX III); and predicted miRNA sites were identified. This analysis showed that most of the members of the PSED network could potentially be targeted by miRNAs; importantly, miR-128 and myomiR sites were found in the 3’UTR sequences of most of these PSED members (chapter 3).

Gga-Eya4 was predicted to be targeted by miR-1a/206, miR-27b/128, miR-133a and miR-499; Gga-Eya1 by miR-27b/128 and miR-133a; Gga-Eya3 by miR-133a; Gga-Six1 by miR-1a/206; Gga-Six4 by miR-1a/206 and miR-499.

In order to validate these miRNA/target interactions, predicted by bioinformatics tools, *in vitro* (chapter 4) and *in vivo* (chapter 5) experiments were undertaken.

#### **MicroRNAs and regulation of target expression:**

The effect of an individual miRNA on target expression level tends to be quite subtle. MiRNAs can use different strategies to accomplish significant regulation (Ebert & Sharp 2012).

For example, a miRNA can target multiple sites for a given target conferring a stronger repression. More often, different miRNAs can work together to co-target a given mRNA, leading to a combined repressive effect exceeding their individual contributions.

Another mechanism by which a miRNA can increase its impact is by targeting a set of genes that are in a shared pathway or protein complex. The reduction of concentration of several components in a signalling cascade induced by a miRNA can then potentially have a significant impact in the signal output over time.

The predicted involvement of several miRNAs, including miR-128 and the myomiRs, in the regulation of the expression of members of the PSED network (1) suggest that a strong regulation of this network is absolutely crucial; and (2) may indicate that a miRNA network, composed of miRNAs, that can work individually or in cooperation, would be able to act at different levels of the PSED cascade, which eventually could have an impact on myogenesis.

### **6.1.3. MiR-128/Gga-Eya4 interaction**

Gga-Eya4, candidate target of miR-128 was cloned; a fragment of its coding and 3'UTR sequences were cloned by PCR. Fragments of coding and 3'UTR sequences were also cloned for Gga-Six1, Gga-Six4, Gga-Eya1, Gga-Eya3 and Gga-Dach1.

The coding sequence fragment cloned was used to make a probe for Gga-Eya4 and RNA ISH was performed. Gga-Eya4 expression pattern was characterised for the first time in the model chicken. It was found in the somites at HH16; as the somites differentiated, Gga-Eya4 was expressed in the dorsomedial lip of the dermomyotome and the myotome. This pattern was similar to those of miR-128 (Fig. 4.3); miR-128 and Gga-Eya4 were both expressed in the somites.

The predicted interaction between Gga-Eya4 and miR-128 was investigated. Gga-Eya4 3'UTR constructs, WT and mutants, generated by FastCloning technique (Li et al. 2011), were used to perform luciferase reporter assays. These assays showed that miR-128 can target Gga-Eya4 3'UTR and reduce its expression activity (32% decrease in luciferase activity); activity being rescued after mutation of miR-128 site (93% of control) (Fig. 4.5). Other miRNAs were tested, and amongst them miR-206 was also shown to be able to target Gga-Eya4 3'UTR, with slightly weaker effect than observed for miR-128 (24% decrease in luciferase activity; rescue: 92.5% of control) (Fig. 4.5); no changes were observed with miR-1a, miR-133a, and miR-499. A stronger effect was observed when si-128 and si-206 were used in combination (Fig. 4.6).

Luciferase reporter assays also showed that miR-128 and miR-133a can target Gga-Eya1 3'UTR (37% and 19% decrease in luciferase activity, respectively), while miR-27b cannot, although it has the same seed sequence as miR-128 and was predicted to target the same site (Fig. 4.7a, b).

MiR-133a was also able to target Gga-Eya3 3'UTR (22% decrease in luciferase activity) (Fig. 4.7d). MiR-133a and miR-499 did not seem to be able to target Gga-Six4 3'UTR. Other experiments were performed to determine if miR-1a and miR-206 could target Gga-Six1, Gga-Six4 and Gga-Dach1. Although a decrease in luciferase activity was observed, it could not be rescued after mutation of the different miRNA sites in their respective 3'UTR sequences.

This could be explained by the presence of additional cryptic/non-canonical miR-1a/206 sites, which were not predicted by the bioinformatics tools. Further investigations would have to be done in order to confirm Gga-Six1, Gga-Six4 and Gga-Dach1 as targets of miR-1a and miR-206.

RNA probes were also generated for Gga-Six1, Gga-Six4, Gga-Eya1, Gga-Eya2, Gga-Eya3, and Gga-Dach1, and RNA ISH performed on chicken embryos at different stages of development (Fig. 4.10). Similar expression to those already published were observed for Gga-Six1, Gga-Six4, Gga-Eya1, Gga-Eya2, and Gga-Dach1 (Esteve & Bovolenta 1999; Ishihara et al. 2008; Heanue et al. 1999; Mishima & Tomarev 1998; Heanue et al. 2002; Kida et al. 2004).

Gga-Eya2, expressed earlier than Gga-Eya1, and Gga-Eya4, displayed a strong and dynamic expression in the developing somites, the dermomyotome and the myotome; however, Gga-Eya1 could not be detected in the dermomyotome, and neither Gga-Eya1 nor Gga-Eya2 were found in the limb buds. Gga-Six4 was present in the head region, the notochord, in somites, in the dermomyotome and the myotome. Gga-Six1 was expressed in the somites, in the myotome, however it was not found in the dermomyotome; and Gga-Dach1 was expressed in the neural tube at HH11-12, however it was not found in the somites or the developing limb buds as previously reported.

Gga-Eya3, which had not been characterised in chicken before, seemed to have a similar expression than what was described in mouse and zebrafish; with a widespread expression in whole-mount, Gga-Eya3 was expressed in the neural tube and developing somites at HH11-12, and in the dermomyotome and myotome at HH21-22.

### **MiR-128 and Gga-Eya4 – overlapping expression:**

Gga-Eya4 and miR-128 display very similar profiles, especially in the somites from HH16.

Therefore, because Gga-Eya4 and miR-128 expression patterns are not exclusive, miR-128 is more likely to be a regulator/modulator of Eya4 expression rather than an absolute repressor.

This is consistent with the 'miRNA buffer-like' model described by Hornstein and Shomron (Hornstein & Shomron 2006), where miRNA and target are co-expressed in intermediate levels. This model is to oppose to the 'anti-correlated' model where miRNA and target are expressed in mutually exclusive domains; miRNA expression level being typically higher than the target in order to efficiently prevent target expression.

This co-expression model suggests that the miRNA could play a role in buffering fluctuations in target expression at a post-transcriptional level. Reduction of this 'genetic noise' by the miRNA leading to the modulation of target expression level (bringing target gene expression back to its mean level) would be consistent with the fact that this target might serve different purposes from primary gene regulation (Hornstein & Shomron 2006; Shkumatava et al. 2009; Ebert & Sharp 2012).

In the case of the EYA family members, like Eya4, it has been shown that as well as being transcriptional cofactors, they are also protein tyrosine/threonine phosphatases (Rebay 2015); they belong to the phosphatase subgroup of the haloacid dehalogenase (HAD) superfamily. For example, Li *et al.* showed that EYA's tyrosine phosphatase activity could be necessary in order to switch Six1-Dach1 complexes from repressive to activating (Li et al. 2003). This suggests a potential role for EYA as a phosphatase in the process of myogenesis.

Three other protein substrates have been identified: the histone H2AX playing a role in the repair versus apoptosis response to DNA damage (Cook et al. 2009; Krishnan et al. 2009); atypical protein kinase C zeta (aPKC $\zeta$ ) important in the process of epithelial polarity and asymmetric cell division (El-Hashash et al. 2011; El-Hashash et al. 2012); estrogen receptor beta (ER $\beta$ ) involved in Six-independent modes of transcriptional regulation (Yuan et al. 2014). None of these three implicate EYA phosphatase activity in regulating EYA-SIX-mediated transcriptional events, suggesting new roles for EYA (Rebay 2015).

These elements and the overlapping expression of miR-128 and Gga-Eya4 in the somites observed by ISH suggest that a basal level of EYA4 expression might be necessary in this tissue; EYA4 being both transcription cofactor and protein phosphatase could also be involved in processes other than myogenesis.

### **Gga-Eya4: co-regulation by miR-128 and miR-206?**

Due to their short sequences, miRNAs can interact and regulate several hundred targets. For example, miR-128 is predicted to target 507 mRNAs. On the other hand, since the 3'UTR of an mRNA often displays multiple sites that can be targeted simultaneously, a cooperative effect of miRNA targeting can be expected (Selbach et al. 2008; Bartel 2009; Friedman et al. 2009).

MiRNAs can exert synergistic regulatory effects through 2 mechanisms: (1) a 3'UTR having target sites to multiple miRNAs; and (2) a 3'UTR with multiple target sites to the same miRNA. These miRNA sites have to be close to each other (around 100 nucleotides) (Lu & Clark 2012).

During this work looking at the interaction between miR-128 and Gga-Eya4, other miRNAs were also identified and predicted to target Gga-Eya4; amongst them, miR-206. MiR-128 and miR-206 sites are 162 bp apart.

The luciferase reporter assays performed showed that miR-128, and miR-206, could target Gga-Eya4 3'UTR and reduce its expression activity; activity which was rescued after mutation of miR-128 and miR-206 sites, respectively (Fig. 4.5). Interestingly, when si-128 and si-206 were used in combination, a stronger effect was observed.

This suggests that miR-128 and miR-206, which can target Gga-Eya4 3'UTR on their own (39% and 24% decrease in luciferase activity, respectively (50 nM each)), might also be able to work in cooperation in order to have a stronger effect on the regulation of Gga-Eya4 expression (miR-128+miR-206: 51% decrease in luciferase activity (25 nM each)). Their combined effect was similar to the sum of their separate effect at the same final dose; although each siRNA was used at half the concentration compared to the experiment where they were tested on their own.

With these results it is only possible to conclude on a potential additive effect between miR-128 and miR-206. More experiments, for example combining si-206 with siC, using miR-206 mutant or miR-128/miR-206 double mutant constructs, would have to be undertaken in order to determine a potential synergistic effect.

Either additive or synergistic, this effect observed *in vitro*, would have to be tested *in vivo*, where potential upstream transcription factor regulation might have an effect on this cooperativity between miR-128 and miR-206 and the downstream miRNA target expression.

#### **6.1.4. MiR-128/Gga-Eya4 interaction, PSED network and myogenesis**

Predicted by bioinformatics tools (chapter 3) and validated by *in vitro* luciferase reporter assays (chapter 4), the interaction between miR-128 and Gga-Eya4 in the chicken embryo, as well as its impact on skeletal muscle myogenesis was investigated.

MiR-128/Gga-Eya4 interaction was validated by RNA ISH and RT-qPCR after AM-128 injection (*in vivo*). After AM-128 injection, a significant increase in Gga-Eya4 was observed (1.5-fold increase; 8 independent experiments; t-test:  $p < 0.05$ ). In addition, a significant increase in Gga-Six4 and Gga-Pax3 expression levels was observed (Gga-Six4: 1.25-fold increase; 7 independent experiments; t-test:  $p < 0.05$ ; Gga-Pax3: 1.3-fold increase; 8 independent experiments; t-test:  $p < 0.001$ ), whereas Gga-MyoD1 expression was decreased (Fig. 5.3).

MiR-128/Gga-Eya1, miR-206/Gga-Eya4, miR-133/Gga-Eya1 and miR-133/Gga-Eya3 interactions were also investigated by RT-qPCR.

After AM-128 and AM-133 injections, a significant increase in Gga-Eya1 expression level was observed (after AM-128 injection: 19%; 10 independent experiments; t-test:  $p < 0.05$ ; after AM-133 injection: 1.25-fold increase; 5 independent experiments; t-test:  $p < 0.001$ ), but not for Gga-Eya3 (Fig. 5.3). After AM-128, AM-133 and AM-206 injections, the expression levels of Gga-Six4 and Gga-Pax3 were increased, and Gga-MyoD1 and Gga-Myf5 decreased (Fig. 5.4a, b).

#### **MiR-128/Gga-Eya4 interaction and myogenesis:**

The functional experiments showed that disrupting the interaction between miR-128 and Gga-Eya4, by blocking miR-128 with AM-128, not only had an impact on Gga-Eya4 expression (de-repression), but also on the PSED members Gga-Six1/4, Gga-Pax3 (de-repression), and on the MRF Gga-MyoD1 (downregulation). This suggest that Gga-Eya4 might be upstream of Gga-Six1/4, Gga-Pax3 and Gga-MyoD1.

Furthermore, it has been shown that SIX and EYA can interact together to form a transcriptional activator complex, and one of the known targets of SIX is PAX3 (Heanue et al. 1999; Ohto et al. 1999; Grifone et al. 2005).

PAX3 downregulation is essential to initiate the myogenic programme, and myogenic differentiation is characterised by the activation of the MRFs (Goulding et al. 1994; Williams & Ordahl 1994; Maroto et al. 1997; Gros et al. 2004; Bajard et al. 2006).

Taken together, it is tempting to hypothesise that in the absence of miR-128, Gga-Eya4 would interact with Gga-Six1/4, and form an activator complex able to regulate Gga-Pax3. Expression of Pax3, being maintained, would prevent the MRF Gga-MyoD1 from being activated (Fig. 5.6).

These are the potential effects that miR-128/Gga-Eya4 interaction could have in the context of the PSED network and Gga-Eya4 as a transcription factor during myogenesis. Another aspect to consider would be the fact that Gga-Eya4 is also a protein tyrosine/threonine phosphatase; therefore, the regulation of Gga-Eya4 by mir-128 could also have onward effects via alternative roles of the Eya4 protein during myogenesis (Li et al. 2003), or other processes (Rebay 2015).

## **6.2. Future work**

Despite achieving the overall aim of this project, there are a number of areas that can be developed further in the future, in order to consolidate the results presented and try to validate, or not, the proposed model.

### **Additional luciferase reporter assays:**

Other interesting interactions than the one between miR-128 and Gga-Eya4 were also identified during this work, constructs (WT and mutants) were generated and luciferase reporter assays were performed. Although these assays showed promising results, they were not completely successful; they did not give the complete rescue of luciferase activity expected after mutation of the miRNA sites. This was the case for the interactions between miR-128 and Gga-Eya1 (Fig. 4.7b), miR-1a/206 and Gga-Six1 (Fig. 4.8a), miR-1a/206 and Gga-Six4 (Fig. 4.8d), and miR-1a/206 and Gga-Dach2 (Fig. 4.9a).

Building on some interesting *in vivo* results (Fig. 5.3; Fig. 5.4a), these experiments will have to be repeated to make the results statistically significant.

For the assays involving miR-1a/206, if no improvement in the rescue experiments is observed, we could look for potential cryptic miRNA sites in Gga-Six1, Gga-Six4 and Gga-Dach2 3'UTR sequences; if such sites can be identified, they will then have to be mutated.

The pilot experiment looking at miRNA synergism could also be completed (Fig. 4.6). Different combinations of constructs and siRNAs would be used.

In order to investigate further the potential synergism identified between miR-128 and miR-206 in the regulation of Gga-Eya4 expression, the miR-128/miR-206 double mutant construct will need to be generated.

#### **Additional functional experiments:**

To completely validate the results observed in the RNA ISH and RT-qPCR performed after miRNA-128 loss-of-function (Fig. 5.2; Fig. 5.3), AM-128 injection experiments will be repeated in order to increase the number of embryos for each condition. The same experiments but using a control antagomiR (AM-scr) will also be done, and RNA ISH will be performed for the different PSED members and MRFs.

In order to investigate further miR-128/Gga-Eya4 interaction, gain-of-function and rescue experiments will also be done. A Gga-Eya4 expression construct has already been generated (pCAB(Gga-Eya4 full length-HA tag)) and will be used to perform overexpression experiments in order to see if it gives the same phenotype observed in the miR-128 loss-of-function experiments. In order to try to rescue the phenotype observed after AM-128 treatment it should be possible to use morpholinos directed against Gga-Eya4.

#### **New interesting miRNA/target interactions:**

In parallel to the main project focussing on miR-128 and Gga-Eya4, a group of miRNAs, with somitic expressions, were also characterised and bioinformatics tools allowed to identify potential relevant targets in skeletal muscle tissue. This analysis provided precious information that could be used in the future in order to investigate new potential miRNA/target interactions, and determine the function(s) these miRNAs could have in skeletal muscle tissue (see chapter 3; APPENDIX II Table 3).

### **6.3. Conclusion**

Over the past decade and a half, miRNAs have emerged as key component of gene regulation underlying the skeletal muscle development and function.

With this project, using a combination of cell-based experiments and whole-embryo studies, we showed that miR-128 could play an important role in the regulation of skeletal muscle myogenesis by targeting Gga-Eya4, a member of the PSED network.

The regulation of Eya4 by miR-128 could be one of the upstream regulatory steps contributing to the necessary downregulation of Pax3, in order to initiate the myogenic programme, characterised by activation of the MRFs.

## ABBREVIATIONS

<b>μL</b>	microliter
<b>AER</b>	Apical ectodermal ridge
<b>AGO</b>	Argonaute
<b>AM</b>	AntagomiR
<b>AMP</b>	Adenosine monophosphate
<b>ANK1</b>	Ankyrin 1
<b>aPKCζ</b>	atypical protein kinase C zeta
<b>BBR</b>	Boehringer Mannheim Blocking Reagent
<b>BCIP</b>	5-bromo-4-chloro-3-indolyl-phosphate
<b>bHLH</b>	Basic helix-loop-helix
<b>BLAST</b>	Basic Local Alignment Search Tool
<b>BMP</b>	Bone morphogenetic protein
<b>BOR</b>	Branchio-oto-renal
<b>bp</b>	base pair
<b>CaCl<sub>2</sub></b>	Calcium chloride
<b>CaM</b>	calmodulin
<b>CBP</b>	CREB-binding protein
<b>cDNA</b>	complementary DNA
<b>CHAPS</b>	3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate
<b>DACH</b>	dachshund-related homeobox
<b>DCM</b>	Dilated cardiomyopathy
<b>DEPC</b>	Diethylpyrocarbonate
<b>DIG</b>	Digoxigenin
<b>DMEM</b>	Dulbecco's modified eagle medium
<b>DMF</b>	Dimethylformamide
<b>DNA</b>	Deoxyribonucleic acid
<b>dNTP</b>	Deoxynucleotide
<b>DTT</b>	Dithiothreitol
<b>E</b>	Mouse embryonic day of development
<b>ECM</b>	Extracellular matrix
<b>ED</b>	EYA domain
<b>EDTA</b>	Ethylenediaminetetraacetic acid

<b>EMT</b>	Epithelial-to-mesenchymal transition
<b>ER<math>\beta</math></b>	Estrogen receptor beta
<b>ERK</b>	Extracellular signal-regulated kinase
<b>EXP5</b>	Exportin 5
<b>EYA</b>	eyes absent-related homeobox
<b>FGF</b>	Fibroblast growth factor
<b>Fig</b>	Figure
<b>FITC</b>	Fluorescein isothiocyanate
<b>GAPDH</b>	Glyceraldehyde 3-phosphate dehydrogenase
<b>GFP</b>	Green fluorescent protein
<b>Gga</b>	Gallus gallus
<b>GO</b>	Gene Ontology
<b>GTP</b>	Guanosine-5'-triphosphate
<b>h</b>	hour
<b>H<sub>2</sub>O</b>	Water
<b>HA</b>	Human influenza hemagglutinin
<b>HAD</b>	Haloacid dehalogenase
<b>HCl</b>	Hydrochloric acid
<b>HD</b>	Homeodomain
<b>HDAC</b>	Histone deacetylase
<b>HF</b>	High fidelity
<b>HH</b>	Hamburger-Hamilton stage
<b>Hsa</b>	Homo sapiens
<b>IRS1</b>	Insulin receptor substrate 1
<b>ISH</b>	<i>in situ</i> hybridisation
<b>KAc</b>	Potassium acetate
<b>kb</b>	kilo-base
<b>KCl</b>	Potassium chloride
<b>LB</b>	Lysogeny Broth
<b>Lfng</b>	Lunatic Fringe
<b>LNA</b>	Locked-nucleic acid
<b>Log</b>	Logarithm
<b>MAB</b>	Maleic acid buffer
<b>MABT</b>	Maleic acid buffer supplemented with Tween-20
<b>MAPK</b>	Mitogen-activated protein kinase

<b>MAPKAPK2</b>	Mitogen-activated protein kinase activated-protein kinase 2
<b>MCS</b>	Multiple cloning site
<b>MEF2</b>	Myocyte enhancer factor 2
<b>MET</b>	Mesenchymal-to-epithelial transition
<b>MgCl</b>	Magnesium chloride hexahydrate
<b>MHC</b>	Myosin heavy chain
<b>MiR</b>	MicroRNA
<b>MiRISC</b>	MiRNA/RISC
<b>MiRNA</b>	MicroRNA
<b>Mmu</b>	Mus musculus
<b>MnCl<sub>2</sub></b>	Manganese chloride
<b>MO</b>	Morpholino
<b>MOPS</b>	3-(N-morpholino) propane sulfonic acid
<b>MPC</b>	Myogenic progenitor cell
<b>MRE</b>	MicroRNA response element
<b>MRF</b>	Myogenic regulatory factor
<b>MRF4</b>	Myogenic regulatory factor 4
<b>mRNA</b>	messenger RNA
<b>mTOR</b>	Mammalian target of rapamycin
<b>mut</b>	mutant
<b>MYF5</b>	Myogenic factor 5
<b>MYOD1</b>	Myogenic differentiation 1
<b>MYOG</b>	Myogenin
<b>NaCl</b>	Sodium chloride
<b>NaOH</b>	Sodium hydroxide
<b>NBT</b>	4-nitro-blue tetrazolium chloride
<b>ng</b>	nanogram
<b>NGS</b>	Next generation sequencing
<b>NMD</b>	Nonsense-mediated decay
<b>nt</b>	nucleotide
<b>PAX</b>	Paired-box
<b>PBS</b>	Phosphate-buffered saline
<b>PBT</b>	Phosphate-buffered saline supplemented with Tween-20
<b>PCR</b>	Polymerase chain reaction
<b>PFA</b>	Paraformaldehyde

<b>Pre-miRNA</b>	Precursor microRNA
<b>Pri-miRNA</b>	Primary microRNA
<b>PSED</b>	PAX-SIX-EYA-DACH
<b>PSM</b>	Pre-somitic mesoderm
<b>RA</b>	Retinoic acid
<b>RbCl<sub>2</sub></b>	Rubidium chloride
<b>RISC</b>	RNA-induced silencing complex
<b>RMS</b>	Rhabdomyosarcoma
<b>RNAa</b>	RNA activation
<b>RNasin</b>	Ribonuclease inhibitor
<b>rpm</b>	revolutions per minute
<b>RT-qPCR</b>	Real-time quantitative PCR
<b>Scr</b>	Scrambled
<b>SD</b>	SIX domain
<b>SEM</b>	Standard error of the mean
<b>SHH</b>	Sonic hedgehog
<b>siC</b>	Control small interfering RNA
<b>SIX</b>	sine oculis-related homeobox
<b>SRF</b>	Serum response factor
<b>SSC</b>	Saline sodium citrate
<b>TAE</b>	Tris base-acetic acid-EDTA
<b>TM</b>	Melting temperature
<b>TRBP</b>	Transactivation response cytoplasmic RNA-binding protein
<b>tRNA</b>	torula RNA
<b>UTR</b>	Untranslated region
<b>UV</b>	Ultraviolet
<b>v/v</b>	volume (of solute) per volume (of solvent)
<b>WMISH</b>	Whole-mount <i>in situ</i> hybridisation
<b>WT</b>	Wild-type
<b>w/v</b>	weight (of solute) per volume (of solvent)
<b>Xtr</b>	Xenopus tropicalis

## PUBLICATION

- [1] A. Ahmed, N. Ward, S. Moxon, S. Lopez-Gomollon, **C. Viaut**, M. L. Tomlinson, I. Patrushev, M. J. Gilchrist, T. Dalmay, D. Dotlic, A. E. Münsterberg, and G. N. Wheeler, “A database of microRNA expression patterns in *Xenopus laevis*,” *PLoS One*, vol. 10, no. 10, pp. 1–10, 2015.

## REFERENCES

- Abdelfattah, A.M., Park, C. & Choi, M.Y., 2015. Update on non-canonical microRNAs. *Biomolecular Concepts*, 5(4), pp.275–287.
- Abdelhak, S., Kalatzis, V., Heilig, R., Compain, S., Samson, D., Vincent, C., Weil, D., et al., 1997. A human homologue of the *Drosophila* eyes absent gene underlies branchio-oto-renal (BOR) syndrome and identifies a novel gene family. *Nature genetics*, 15(2), pp.157–164.
- Abdelhak, S., Kalatzis, V., Heilig, R., Compain, S., Samson, D., Vincent, C., Levi-Acobas, F., et al., 1997. Clustering of mutations responsible for branchio-oto-renal (BOR) syndrome in the eyes absent homologous region (eyaHR) of EYA1. *Human molecular genetics*, 6(13), pp.2247–2255.
- Abrahante, J.E. et al., 2003. The *Caenorhabditis elegans* hunchback-like gene *lin-57/hbl-1* controls developmental time and is regulated by microRNAs. *Developmental Cell*, 4(5), pp.625–637.
- Abu-Elmagd, M. et al., 2015. Sprouty2 mediated tuning of signalling is essential for somite myogenesis. *BMC medical genomics*, 8 Suppl 1(1), p.S8.
- Agarwal, V. et al., 2015. Predicting effective microRNA target sites in mammalian mRNAs. *eLife*, 4(AUGUST2015), pp.1–38.
- Ahmed, A. et al., 2015. A database of microRNA expression patterns in *Xenopus laevis*. *PLoS ONE*, 10(10), pp.1–10.
- Akhtar, M.M. et al., 2016. Bioinformatic tools for microRNA dissection. *Nucleic Acids Research*, 44(1), pp.24–44.
- Alvares, L.E. et al., 2003. Intrinsic, Hox-dependent cues determine the fate of skeletal muscle precursors. *Developmental Cell*, 5(3), pp.379–390.
- Amthor, H. et al., 1998. The importance of timing differentiation during limb muscle development. *Current biology*, 8(11), pp.642–652.
- Amthor, H. et al., 2002. The Regulation and Action of Myostatin as a Negative Regulator of Muscle Development during Avian Embryogenesis. *Developmental Biology*, 251(2), pp.241–257.
- Aulehla, A. & Pourquié, O., 2008. Oscillating signaling pathways during embryonic development. *Current Opinion in Cell Biology*, 20(6), pp.632–637.
- Aulehla, A. & Pourquié, O., 2010. Signaling gradients during paraxial mesoderm development. *Cold Spring Harbor perspectives in biology*, 2(2).

- Axtell, M.J., Westholm, J.O. & Lai, E.C., 2011. Vive la différence: biogenesis and evolution of microRNAs in plants and animals. *Genome Biology*, 12(221), pp.1–13.
- Ayres, J. a et al., 2001. DACH: genomic characterization, evaluation as a candidate for postaxial polydactyly type A2, and developmental expression pattern of the mouse homologue. *Genomics*, 77(September), pp.18–26.
- Bajard, L. et al., 2006. A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly activates Myf5 in muscle progenitor cells in the limb. *Genes and Development*, 20(17), pp.2450–2464.
- Bar, M. et al., 2008. MicroRNA discovery and profiling in human embryonic stem cells by deep sequencing of small RNA libraries. *Stem cells*, 26(10), pp.2496–505.
- Barnes, G.L. et al., 1996. Chicken Pax-1 gene: Structure and expression during embryonic somite development. *Differentiation*, 61(1), pp.13–23.
- Bartel, D.P., 2004. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell*, 116(2), pp.281–297.
- Bartel, D.P., 2009. MicroRNAs: Target Recognition and Regulatory Functions. *Cell*, 136(2), pp.215–233.
- Batuwita, R. & Palade, V., 2009. microPred: Effective classification of pre-miRNAs for human miRNA gene prediction. *Bioinformatics*, 25(8), pp.989–995.
- Baykal, B. & Korkusuz, P., 2016. Development of the Musculoskeletal System. In F. Korkusuz, ed. *Musculoskeletal Research and Basic Science*. Cham: Springer International Publishing, pp. 289–302.
- Behrens, A.N. et al., 2013. Nkx2-5 mediates differential cardiac differentiation through interaction with Hoxa10. *Stem cells and development*, 22(15), pp.2211–20.
- Bell, G.W., Yatskievych, T.A. & Antin, P.B., 2004. GEISHA, a Whole-Mount in Situ Hybridization Gene Expression Screen in Chicken Embryos. *Developmental Dynamics*, 229(3), pp.677–687.
- Bénazéraf, B. & Pourquié, O., 2013. Formation and segmentation of the vertebrate body axis. *Annual review of cell and developmental biology*, 29, pp.1–26.
- Bentwich, I. et al., 2005. Identification of hundreds of conserved and nonconserved human microRNAs. *Nature genetics*, 37(7), pp.766–770.
- Bentzinger, C., Wang, Y.X. & Rudnicki, M.A., 2012. Building muscle: Molecular regulation of myogenesis. *Cold Spring Harbor Perspectives in Biology*, 4(2).
- Berezikov, E. et al., 2007. Mammalian Mirtron Genes. *Molecular Cell*, 28(2), pp.328–336.

- Bergstrom, D.A. & Tapscott, S.J., 2001. Molecular Distinction between Specification and Differentiation in the Myogenic Basic Helix-Loop-Helix Transcription Factor Family Molecular Distinction between Specification and Differentiation in the Myogenic Basic Helix-Loop-Helix Transcription Factor. *Molecular and Cellular Biology*, 21(7), pp.2404–2412.
- Bernstein, E. et al., 2003. Dicer is essential for mouse development. *Nature genetics*, 35(3), pp.215–217.
- Bernstein, E. et al., 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, 409(6818), pp.363–366.
- Berti, F. et al., 2015. Time course and side-by-side analysis of mesodermal, pre-myogenic, myogenic and differentiated cell markers in the chicken model for skeletal muscle formation. *Journal of anatomy*, 227(3), pp.361–82.
- Betel, D. et al., 2008. The microRNA.org resource: Targets and expression. *Nucleic Acids Research*, 36(SUPPL. 1), pp.149–153.
- Bober, E. et al., 1994. Pax-3 is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. *Development*, 120(3), pp.603–612.
- Bohnsack, M.T., Czaplinski, K. & Gorlich, D., 2004. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA*, 10(2), pp.185–91.
- Borsani, G. et al., 1999. EYA4, a novel vertebrate gene related to Drosophila eyes absent. *Human Molecular Genetics*, 8(1), pp.11–23.
- Bowes, J.B. et al., 2009. Xenbase: Gene expression and improved integration. *Nucleic Acids Research*, 38(SUPPL.1), pp.607–612.
- Brand-Saberi, B. et al., 1996. The formation of somite compartments in the avian embryo. *International Journal of Developmental Biology*, 40(1), pp.411–420.
- Brand-Saberi, B. et al., 1993. The ventralizing effect of the notochord on somite differentiation in chick embryos. *Anatomy and embryology*, 188(3), pp.239–245.
- Brent, A.E., Schweitzer, R. & Tabin, C.J., 2003. A somitic compartment of tendon progenitors. *Cell*, 113(2), pp.235–248.
- Bruno, I.G. et al., 2011. Identification of a MicroRNA that Activates Gene Expression by Repressing Nonsense-Mediated RNA Decay. *Molecular Cell*, 42(4), pp.500–510.
- Buckingham, M. & Relaix, F., 2015. PAX3 and PAX7 as upstream regulators of myogenesis. *Seminars in Cell and Developmental Biology*, 44, pp.115–125.

- Buckingham, M. & Relaix, F., 2007. The role of Pax genes in the development of tissues and organs: Pax3 and Pax7 regulate muscle progenitor cell functions. *Annual review of cell and developmental biology*, 23, pp.645–73.
- Buckingham, M. & Rigby, P.W.J., 2014. Gene Regulatory Networks and Transcriptional Mechanisms that Control Myogenesis. *Developmental Cell*, 28(3), pp.225–238.
- Cai, X., Hagedorn, C.H. & Cullen, B.R., 2004. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA*, 10(12), pp.1957–1966.
- Cecconi, F. et al., 1997. Expression of Meis2, a knotted-related murine homeobox gene, indicates a role in the differentiation of the forebrain and the somitic mesoderm. *Developmental Dynamics*, 210(2), pp.184–190.
- Chambers, D. & Mason, I., 2000. Expression of sprouty2 during early development of the chick embryo is coincident with known sites of FGF signalling. *Mechanisms of Development*, 91(1–2), pp.361–364.
- Chen, J.-F. et al., 2006. The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nature genetics*, 38(2), pp.228–233.
- Chen, J.F. et al., 2010. microRNA-1 and microRNA-206 regulate skeletal muscle satellite cell proliferation and differentiation by repressing Pax7. *Journal of Cell Biology*, 190(5), pp.867–879.
- Chen, R. et al., 1997. Dachshund and eyes absent proteins form a complex and function synergistically to induce ectopic eye development in *Drosophila*. *Cell*, 91(7), pp.893–903.
- Chendrimada, T.P. et al., 2005. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature*, 436(7051), pp.740–744.
- Cheng, L. et al., 2004. The epaxial-hypaxial subdivision of the avian somite. *Developmental Biology*, 274(2), pp.348–369.
- Ching, A.-S. & Ahmad-Annuar, A., 2015. A Perspective on the Role of microRNA-128 Regulation in Mental and Behavioral Disorders. *Frontiers in cellular neuroscience*, 9(DEC), p.465.
- Choi, J.K. et al., 1991. Phorbol esters selectively and reversibly inhibit a subset of myofibrillar genes responsible for the ongoing differentiation program of chick skeletal myotubes. *Molecular and cellular biology*, 11(9), pp.4473–82.
- Christ, B., Huang, R. & Scaal, M., 2007. Amniote somite derivatives. *Developmental Dynamics*, 236(9), pp.2382–2396.

- Christ, B. & Ordahl, C.P., 1995. Early stages of chick somite development. *Anatomy and Embryology*, 191(5), pp.381–396.
- Cinnamon, Y., Kahane, N. & Kalcheim, C., 1999. Characterization of the early development of specific hypaxial muscles from the ventrolateral myotome. *Development*, 126, pp.4305–4315.
- Cook, P.J. et al., 2009. Tyrosine dephosphorylation of H2AX modulates apoptosis and survival decisions. *Nature*, 458(7238), pp.591–596.
- Cooke, J. & 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(2), pp.455–476.
- Del Corral, R.D. & Storey, K.G., 2004. Opposing FGF and retinoid pathways: A signalling switch that controls differentiation and patterning onset in the extending vertebrate body axis. *BioEssays*, 26(8), pp.857–869.
- Crist, C.G. et al., 2009. Muscle stem cell behavior is modified by microRNA-27 regulation of Pax3 expression. *Proceedings of the National Academy of Sciences of the United States of America*, 106(32), pp.13383–7.
- Dale, J.K. et al., 2003. Periodic notch inhibition by lunatic fringe underlies the chick segmentation clock. *Nature*, 421(6920), pp.275–278.
- Dale, K.J. & Pourquié, O., 2000. A clock-work somite. *BioEssays*, 22(1), pp.72–83.
- Darnell, D.K. et al., 2006. MicroRNA expression during chick embryo development. *Developmental Dynamics*, 235(11), pp.3156–3165.
- Davis, J.R. et al., 1999. Mouse Dach, a homologue of Drosophila dachshund, is expressed in the developing retina, brain and limbs. *Development Genes and Evolution*, 209(9), pp.526–536.
- Davis, R.J., Shen, W., Sandler, Y.I., Heanue, T.A., et al., 2001. Characterization of mouse Dach2, a homologue of Drosophila dachshund. *Mechanisms of Development*, 102(1–2), pp.169–179.
- Davis, R.J., Shen, W., Sandler, Y.I., Amoui, M., et al., 2001. Dach1 mutant mice bear no gross abnormalities in eye, limb, and brain development and exhibit postnatal lethality. *Molecular and cellular biology*, 21(5), pp.1484–90.
- Davis, R.J. et al., 2006. Mouse Dach2 mutants do not exhibit gross defects in eye development or brain function. *Genesis*, 44(2), pp.84–92.
- Denetclaw, W.F., Christ, B. & Ordahl, C.P., 1997. Location and growth of epaxial myotome precursor cells. *Development*, 124(8), pp.1601–10.

- Denli, A.M. et al., 2004. Processing of primary microRNAs by the Microprocessor complex. *Nature*, 432(7014), pp.231–5.
- Depreux, F.F.S. et al., 2008. Eya4 -deficient mice are a model for heritable otitis media. *The Journal of Clinical Investigation*, 118(February).
- Dey, B.K., Gagan, J. & Dutta, A., 2011. miR-206 and -486 induce myoblast differentiation by downregulating Pax7. *Molecular and cellular biology*, 31(1), pp.203–14.
- Dias, A.S. et al., 2014. Somites Without a Clock. *Science*, 343(6172), pp.791–795.
- Diederichs, S. & Haber, D.A., 2007. Dual Role for Argonautes in MicroRNA Processing and Posttranscriptional Regulation of MicroRNA Expression. *Cell*, 131(6), pp.1097–1108.
- Dietrich, S. et al., 1998. Specification of the hypaxial musculature. *Development*, 125(12), pp.2235–2249.
- Dietrich, S., Schubert, F.R. & Lumsden, A., 1997. Control of dorsoventral pattern in the chick paraxial mesoderm. *Development*, 124(19), pp.3895–3908.
- Doench, J.G. & Sharp, P.A., 2004. Specificity of microRNA target selection in translational repression. *Genes*, 504, pp.504–511.
- Dubrulle, J. & Pourquié, O., 2004. fgf8 mRNA decay establishes a gradient that couples axial elongation to patterning in the vertebrate embryo. *Nature*, 427(6973), pp.419–422.
- Ebert, M.S. & Sharp, P.A., 2012. Roles for MicroRNAs in conferring robustness to biological processes. *Cell*, 149(3), pp.505–524.
- Ekimler, S. & Sahin, K., 2014. Computational Methods for MicroRNA Target Prediction. *Genes*, 5(3), pp.671–683.
- El-Hashash, A.H.K. et al., 2011. Eya1 controls cell polarity, spindle orientation, cell fate and Notch signaling in distal embryonic lung epithelium. *Development*, 138(7), pp.1395–407.
- El-Hashash, A.H.K. et al., 2012. Eya1 protein phosphatase regulates tight junction formation in lung distal epithelium. *Journal of Cell Science*, 125(Pt 17), pp.4036–48.
- Enright, A.J. et al., 2003. MicroRNA targets in Drosophila. *Genome Biology*, 5(1), pp.1–14.
- Esteve, P. & Bovolenta, P., 1999. cSix4, a member of the six gene family of transcription factors, is expressed during placode and somite development. *Mechanisms of Development*, 85(1–2), pp.161–165.

- Fan, C., Lee, C. & Tessier-Lavigne, M., 1997. A role for WNT proteins in induction of dermomyotome. *Developmental biology*, 191(1), pp.160–5.
- Fan, C.M. & Tessier-Lavigne, M., 1994. Patterning of mammalian somites by surface ectoderm and notochord: Evidence for sclerotome induction by a hedgehog homolog. *Cell*, 79(7), pp.1175–1186.
- Filipowicz, W., Bhattacharyya, S.N. & Sonenberg, N., 2008. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nature reviews. Genetics*, 9(february), pp.102–114.
- Finnegan, E.F. & Pasquinelli, A.E., 2013. MicroRNA biogenesis: regulating the regulators. *Critical reviews in biochemistry and molecular biology*, 48(1), pp.51–68.
- Forman, J.J., Legesse-Miller, A. & Collier, H.A., 2008. A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence. *Proceedings of the National Academy of Sciences of the United States of America*, 105(39), pp.14879–84.
- Fougerousse, F. et al., 2002. Six and Eya expression during human somitogenesis and MyoD gene family activation. *Journal of Muscle Research and Cell Motility*, 23(3), pp.255–264.
- Fraser, P.E. & Sauka-Spengler, T., 2004. Expression of the polycomb group gene bmi-1 in the early chick embryo. *Gene Expression Patterns*, 5(1), pp.23–27.
- Friedländer, M.R. et al., 2008. Discovering microRNAs from deep sequencing data using miRDeep. *Nature biotechnology*, 26(4), pp.407–15.
- Friedman, R.C. et al., 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Research*, 19(1), pp.92–105.
- Gagan, J. et al., 2012. Notch3 and Mef2c proteins are mutually antagonistic via Mkp1 protein and miR-1/206 microRNAs in differentiating myoblasts. *Journal of Biological Chemistry*, 287(48), pp.40360–40370.
- Galli, L.M. et al., 2008. Identification and characterization of subpopulations of Pax3 and Pax7 expressing cells in developing chick somites and limb buds. *Developmental Dynamics*, 237(7), pp.1862–1874.
- Gargalionis, A.N. & Basdra, E.K., 2013. Insights in microRNAs biology. *Current topics in medicinal chemistry*, 13(13), pp.1493–502.
- Ge, Y. & Chen, J., 2011. MicroRNAs in skeletal myogenesis. *Cell Cycle*, 10(3), pp.441–448.
- Gibson, D.G. et al., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Meth*, 6(5), pp.343–345.

- Gibson, D.G., 2011. Enzymatic assembly of overlapping DNA fragments. *Methods in Enzymology*, 498, pp.349–361.
- Giordani, J. et al., 2007. Six proteins regulate the activation of Myf5 expression in embryonic mouse limbs. *Proceedings of the National Academy of Sciences*, 104(27), pp.11310–11315.
- Goljanek-Whysall, K. et al., 2011. MicroRNA regulation of the paired-box transcription factor Pax3 confers robustness to developmental timing of myogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 108(29), pp.11936–41.
- Goljanek-Whysall, K., Pais, H., et al., 2012. Regulation of multiple target genes by miR-1 and miR-206 is pivotal for C2C12 myoblast differentiation. *Journal of Cell Science*, 125(15), pp.3590–3600.
- Goljanek-Whysall, K., Sweetman, D. & Münsterberg, A.E., 2012. microRNAs in skeletal muscle differentiation and disease. *Clinical science*, 123(11), pp.611–25.
- Gomes, C.P.C. et al., 2013. A review of computational tools in microRNA discovery. *Frontiers in Genetics*, 4(MAY), pp.1–9.
- Goulding, M., Lumsden, A. & Paquette, a J., 1994. Regulation of Pax-3 expression in the dermomyotome and its role in muscle development. *Development*, 120(4), pp.957–971.
- Gregory, R.I. et al., 2004. The Microprocessor complex mediates the genesis of microRNAs. *Nature*, 432(7014), pp.235–240.
- Griffiths-Jones, S. et al., 2008. miRBase: Tools for microRNA genomics. *Nucleic Acids Research*, 36(SUPPL. 1), pp.154–158.
- Grifone, R. et al., 2007. Eya1 and Eya2 proteins are required for hypaxial somitic myogenesis in the mouse embryo. *Developmental Biology*, 302(2), pp.602–616.
- Grifone, R. et al., 2004. Six1 and Eya1 expression can reprogram adult muscle from the slow-twitch phenotype into the fast-twitch phenotype. *Molecular and cellular biology*, 24(14), pp.6253–6267.
- Grifone, R. et al., 2005. Six1 and Six4 homeoproteins are required for Pax3 and Mrf expression during myogenesis in the mouse embryo. *Development*, 132(9), pp.2235–49.
- Grimson, A. et al., 2007. MicroRNA Targeting Specificity in Mammals: Determinants Beyond Seed Pairing. *Molecular Cell*, 27(1), pp.91–105.
- Gros, J., Scaal, M. & Marcelle, C., 2004. A two-Step mechanism for myotome formation in chick. *Developmental Cell*, 6(6), pp.875–882.

- Gros, J., Serralbo, O. & Marcelle, C., 2009. WNT11 acts as a directional cue to organize the elongation of early muscle fibres. *Nature*, 457(7229), pp.589–593.
- Gutzman, J.H., Sahu, S.U. & Kwas, C., 2015. Non-muscle myosin IIA and IIB differentially regulate cell shape changes during zebrafish brain morphogenesis. *Developmental Biology*, 397(1), pp.103–115.
- Ha, M. & Kim, V.N., 2014. Regulation of microRNA biogenesis. *Nature reviews. Molecular cell biology*, 15(8), pp.509–524.
- Hamburger, V. & Hamilton, H.L., 1992. A series of normal stages in the development of the chick embryo. *Developmental Dynamics*, 195, pp.231–272.
- Hammond, S.M., 2001. Argonaute2, a Link Between Genetic and Biochemical Analyses of RNAi. *Science*, 293(5532), pp.1146–1150.
- Han, J. et al., 2006. Molecular Basis for the Recognition of Primary microRNAs by the Drosha-DGCR8 Complex. *Cell*, 125(5), pp.887–901.
- Han, J. et al., 2009. Posttranscriptional Crossregulation between Drosha and DGCR8. *Cell*, 136(1), pp.75–84.
- Han, J. et al., 2004. The Drosha-DGCR8 complex in primary microRNA processing. *Genes and Development*, 18(24), pp.3016–3027.
- Hancock, J.M. et al., 2004. EMBOSS (The European Molecular Biology Open Software Suite). In *Dictionary of Bioinformatics and Computational Biology*. John Wiley & Sons, Ltd.
- Hanson, I.M., 2001. Mammalian homologues of the Drosophila eye specification genes. *Seminars in cell & developmental biology*, 12(6), pp.475–484.
- He, L. & Hannon, G.J., 2004. MicroRNAs: small RNAs with a big role in gene regulation. *Nature reviews. Genetics*, 5(7), pp.522–531.
- Heanue, T.A. et al., 2002. Dach1, a vertebrate homologue of Drosophila dachshund, is expressed in the developing eye and ear of both chick and mouse and is regulated independently of Pax and Eya genes. *Mechanisms of Development*, 111(1–2), pp.75–87.
- Heanue, T.A. et al., 1999. Synergistic regulation of vertebrate muscle development by Dach2, Eya2, and Six1, homologs of genes required for Drosophila eye formation. *Genes and Development*, 13(24), pp.3231–3243.
- Hildebrand, M.S. et al., 2007. A Novel Splice Site Mutation in EYA4 Causes DFNA10 Hearing Loss. *American journal of medical genetics. Part A*, 143A, pp.1599–1604.

- Hillier, L.W. et al., 2004. Sequencing and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*, 432(7018), pp.695–716.
- Hinsch, G.W. & Hamilton, H.L., 1956. The developmental fate of the first somite of the chick. *The Anatomical Record*, 125(2), pp.225–245.
- Hirai, H. et al., 2010. MyoD regulates apoptosis of myoblasts through microRNA-mediated down-regulation of Pax3. *Journal of Cell Biology*, 191(2), pp.347–365.
- Hirst, C.E. & Marcelle, C., 2015. The Avian Embryo as a Model System for Skeletal Myogenesis. *Results and Problems in Cell Differentiation*, 56, pp.99–122.
- Horak, M., Novak, J. & Bienertova-Vasku, J., 2016. Muscle-specific microRNAs in skeletal muscle development. *Developmental Biology*, 410(1), pp.1–13.
- Hornstein, E. & Shomron, N., 2006. Canalization of development by microRNAs. *Nature genetics*, 38(JUNE), pp.520–524.
- Howe, D.G. et al., 2013. ZFIN, the Zebrafish Model Organism Database: Increased support for mutants and transgenics. *Nucleic Acids Research*, 41(D1), pp.854–860.
- Huang, D.W., Lempicki, R. a & Sherman, B.T., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4(1), pp.44–57.
- Huang, D.W., Sherman, B.T. & Lempicki, R.A., 2009. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, 37(1), pp.1–13.
- Huang, R. et al., 1997. The fate of the first avian somite. *Anatomy and Embryology*, 195(5), pp.435–449.
- Huang, V. et al., 2010. RNAi is conserved in mammalian cells. *PLoS ONE*, 5(1), pp.1–8.
- Huang, Y. et al., 2013. Nonmuscle myosin II-B (myh10) expression analysis during zebrafish embryonic development. *Gene Expression Patterns*, 13(7), pp.265–270.
- Hutvagner, G. et al., 2001. A Cellular Function for the RNA-Interference Enzyme Dicer in the Maturation of the let-7 Small Temporal RNA. *Science*, 293, pp.834–838.
- Ikeya, M. & Takada, S., 1998. Wnt signaling from the dorsal neural tube is required for the formation of the medial dermomyotome. *Development*, 125, pp.4969–4976.
- Ishihara, T. et al., 2008. Differential expression of Eya1 and Eya2 during chick early embryonic development. *Gene Expression Patterns*, 8(5), pp.357–367.
- Ivanovska, I. & Cleary, M.A., 2008. Combinatorial microRNAs: Working together to make a difference. *Cell Cycle*, 7(20), pp.3137–3142.

- Jacobson, a G., 1988. Somitomeres: mesodermal segments of vertebrate embryos. *Development*, 104 Suppl, pp.209–20.
- Jemc, J. & Rebay, I., 2007. The eyes absent family of phosphotyrosine phosphatases: properties and roles in developmental regulation of transcription. *Annual review of biochemistry*, 76, pp.513–538.
- John, B. et al., 2004. Human MicroRNA targets. *PLoS biology*, 2(11), p.e363.
- Johnson, R.L. et al., 1994. Ectopic expression of Sonic hedgehog alters dorsal-ventral patterning of somites. *Cell*, 79(7), pp.1165–1173.
- Jouve, C. et al., 2000. Notch signalling is required for cyclic expression of the hairy-like gene HES1 in the presomitic mesoderm. *Development*, 127, pp.1421–1429.
- Kaehn, K. et al., 1988. The onset of myotome formation in the chick. *Anatomy and Embryology*, 177(3), pp.191–201.
- Kapsimali, M. et al., 2007. MicroRNAs show a wide diversity of expression profiles in the developing and mature central nervous system. *Genome biology*, 8(8), p.R173.
- Karam, R. & Wilkinson, M., 2012. A conserved microRNA/NMD regulatory circuit controls gene expression. *RNA biology*, 9(1), pp.22–6.
- Kassar-Duchossoy, L. et al., 2005. Pax3/Pax7 mark a novel population of primitive myogenic cells during development. *Genes and Development*, 19(12), pp.1426–1431.
- Kawamata, T. & Tomari, Y., 2010. Making RISC. *Trends in Biochemical Sciences*, 35(7), pp.368–376.
- Khvorova, A., Reynolds, A. & Jayasena, S.D., 2003. Functional siRNAs and miRNAs exhibit strand bias. *Cell*, 115(2), pp.209–216.
- Kida, Y. et al., 2004. Chick Dach1 interacts with the Smad complex and Sin3a to control AER formation and limb development along the proximodistal axis. *Development*, 131(17), pp.4179–4187.
- Kim, H.K. et al., 2006. Muscle-specific microRNA miR-206 promotes muscle differentiation. *Journal of Cell Biology*, 174(5), pp.677–687.
- Kim, Y.-K., Kim, B. & Kim, V.N., 2016. Re-evaluation of the roles of DROSHA, Exportin 5, and DICER in microRNA biogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 113(13), pp.E1881-1889.
- Kirby, T.J., Chaillou, T. & McCarthy, J.J., 2016. The role of microRNAs in skeletal muscle health and disease. *Frontiers in Biosciences*, 20, pp.37–77.

- Koutsoulidou, A. et al., 2011. Expression of miR-1, miR-133a, miR-133b and miR-206 increases during development of human skeletal muscle. *BMC developmental biology*, 11(1), p.34.
- Kozera, B. & Rapacz, M., 2013. Reference genes in real-time PCR. *Journal of Applied Genetics*, 54(4), pp.391–406.
- Kozłowski, D.J. et al., 2005. The zebrafish dog-eared mutation disrupts *eya1*, a gene required for cell survival and differentiation in the inner ear and lateral line. *Developmental Biology*, 277(1), pp.27–41.
- Kozomara, A. & Griffiths-Jones, S., 2014. MiRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research*, 42(D1), pp.68–73.
- Krishnan, N. et al., 2009. Dephosphorylation of the C-terminal tyrosyl residue of the DNA damage-related histone H2A.X is mediated by the protein phosphatase eyes absent. *Journal of Biological Chemistry*, 284(24), pp.16066–16070.
- Krützfeldt, J. et al., 2005. Silencing of microRNAs in vivo with “antagomirs”. *Nature*, 438(7068), pp.685–689.
- Kubota, K. et al., 2006. Improved in situ hybridization efficiency with locked-nucleic-acid-incorporated DNA probes. *Applied and environmental microbiology*, 72(8), pp.5311–7.
- Kumar, J., 2009. The sine oculis homeobox (SIX) family of transcription factors as regulators of development and disease. *Cell Mol Life Sci.*, 66(4), pp.565–583.
- Laclef, C. et al., 2003. Altered myogenesis in Six1-deficient mice. *Development*, 130(10), pp.2239–2252.
- Lagos-Quintana, M. et al., 2001. Identification of novel genes coding for small expressed RNAs. *Science*, 294(5543), pp.853–8.
- Lagos-Quintana, M. et al., 2002. Identification of tissue-specific MicroRNAs from mouse. *Current Biology*, 12(9), pp.735–739.
- Larionov, A., Krause, A. & Miller, W., 2005. A standard curve based method for relative real time PCR data processing. *BMC bioinformatics*, 6, p.62.
- Lee, E.J. et al., 2008. Systematic evaluation of microRNA processing patterns in tissues, cell lines, and tumors. *RNA*, 14(1), pp.35–42.
- Lee, H.-J., 2014. Additional stories of microRNAs. *Experimental Biology and Medicine*, 239, pp.1275–1279.
- Lee, I. et al., 2009. New class of microRNA targets containing simultaneous 5'-UTR and 3'-UTR interaction sites. *Genome Research*, 19(7), pp.1175–1183.

- Lee, R.C. & Ambros, V., 2001. An Extensive Class of Small RNAs in *Caenorhabditis elegans*. *Science*, 294(5543), pp.862–864.
- Lee, R.C., Feinbaum, R. & Ambros, V., 1993. The *C. elegans* Heterochronic Gene *lin-4* Encodes Small RNAs with Antisense Complementarity to *lin-4*. *Cell*, 75, pp.843–854.
- Lee, Y. et al., 2004. MicroRNA genes are transcribed by RNA polymerase II. *The European Molecular Biology Organization Journal*, 23(20), pp.4051–4060.
- Lee, Y. et al., 2002. MicroRNA maturation: stepwise processing and subcellular localization. *The EMBO Journal*, 21(17), pp.4663–4670.
- Lee, Y. et al., 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature*, 425(6956), pp.415–419.
- Lewis, B.P., Burge, C.B. & Bartel, D.P., 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, 120(1), pp.15–20.
- Li, C. et al., 2011. FastCloning: a highly simplified, purification-free, sequence- and ligation-independent PCR cloning method. *BMC Biotechnology*, 11(1), p.92.
- Li, W. et al., 2015. The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic acids research*, 43(W1), pp.W580-4.
- Li, X. et al., 2003. Eya protein phosphatase activity regulates Six1-Dach-Eya transcriptional effects in mammalian organogenesis. *Nature*, 426(6964), pp.247–254.
- Lin, S. et al., 2012. Let-7b regulates the expression of the growth hormone receptor gene in deletion-type dwarf chickens. *BMC genomics*, 13(1), p.306.
- Lindow, M. & Gorodkin, J., 2007. Principles and limitations of computational microRNA gene and target finding. *DNA and cell biology*, 26(5), pp.339–51.
- Litsiou, A., Hanson, S. & Streit, A., 2005. A balance of FGF, BMP and WNT signalling positions the future placode territory in the head. *Development*, 132(18), pp.4051–62.
- Liu, N. et al., 2007. An intragenic MEF2-dependent enhancer directs muscle-specific expression of microRNAs 1 and 133. *Proceedings of the National Academy of Sciences of the United States of America*, 104(52), pp.20844–9.
- Liu, N. et al., 2011. Mice lacking microRNA 133a develop dynamin 2-dependent centronuclear myopathy. *The Journal of Clinical Investigation*, 121(8), pp.3258–3268.

- Liu, N. et al., 2008. microRNA-133a regulates cardiomyocyte proliferation and suppresses smooth muscle gene expression in the heart. *Genes and Development*, 22(23), pp.3242–3254.
- Llave, C. et al., 2002. Cleavage of Scarecrow-like mRNA Targets Directed by a Class of Arabidopsis miRNA. *Science*, 297(5589), pp.2053–2056.
- Long, D. et al., 2007. Potent effect of target structure on microRNA function. *Nature structural & molecular biology*, 14(APRIL), pp.287–294.
- Lu, J. & Clark, A.G., 2012. Impact of microRNA regulation on variation in human gene expression. *Genome Research*, 22, pp.1243–1254.
- Lund, E. et al., 2004. Nuclear Export of MicroRNA Precursors. *Science*, 303(5654), pp.95–8.
- Lund, E. & Dahlberg, J.E., 2006. Substrate selectivity of exportin 5 and Dicer in the biogenesis of microRNAs. *Cold Spring Harbor Symposia on Quantitative Biology*, 71, pp.59–66.
- Ma, E. et al., 2008. Auto-inhibition of Human Dicer by its Internal Helicase Domain. *Journal of Molecular Biology*, 380(1), pp.237–243.
- Ma, H. et al., 2013. Lower and upper stem-single-stranded RNA junctions together determine the Drosha cleavage site. *Proceedings of the National Academy of Sciences of the United States of America*, 110(51), pp.20687–92.
- Ma, X. & Adelstein, R.S., 2014. A Point Mutation in Myh10 Causes Major Defects in Heart Development and Body Wall Closure. *Circulation. Cardiovascular Genetics*, 7(3), pp.257–265.
- Machon, O. et al., 2015. Meis2 is essential for cranial and cardiac neural crest development. *BMC developmental biology*, 15(1), p.40.
- Magnusson, M. et al., 2007. HOXA10 is a critical regulator for hematopoietic stem cells and erythroid / megakaryocyte development. *Regulation*, 109(9), pp.3687–3696.
- Makishima, T. et al., 2007. Nonsyndromic Hearing Loss DFNA10 and a Novel Mutation of EYA4: Evidence for Correlation of Normal Cardiac Phenotype With Truncating Mutations of the Eya Domain. *American journal of medical genetics. Part A*, 143A, pp.1592–1598.
- Manceau, M. et al., 2008. Myostatin promotes the terminal differentiation of embryonic muscle progenitors. *Genes and Development*, 22(5), pp.668–681.
- Mansouri, A. et al., 1996. Dysgenesis of cephalic neural crest derivatives in Pax7<sup>-/-</sup> mutant mice. *Development*, 122(3), pp.831–8.

- Marcelle, C., Stark, M.R. & Bronner-Fraser, M., 1997. Coordinate actions of BMPs, Wnts, Shh and noggin mediate patterning of the dorsal somite. *Development*, 124(20), pp.3955–3963.
- Maroto, M. et al., 1997. Ectopic Pax-3 activates MyoD and Myf-5 expression in embryonic mesoderm and neural tissue. *Cell*, 89(1), pp.139–148.
- McCarthy, J.J., 2008. MicroRNA-206: The skeletal muscle-specific myomiR. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1779(11), pp.682–691.
- McCarthy, J.J. & Esser, K. a, 2007. MicroRNA-1 and microRNA-133a expression are decreased during skeletal muscle hypertrophy. *Journal of applied physiology*, 102(1), pp.306–313.
- McGlinn, E. et al., 2009. In ovo application of antagomiRs indicates a role for miR-196 in patterning the chick axial skeleton through Hox gene regulation. *Proceedings of the National Academy of Sciences of the United States of America*, 106(44), pp.18610–5.
- Meier, S.P., 1982. The development of segmentation in the cranial region of vertebrate embryos. *Scanning electron microscopy*, (Pt 3), p.1269—1282.
- Mende, M., Christophorou, N.A.D. & Streit, A., 2008. Specific and effective gene knock-down in early chick embryos using morpholinos but not pRFPRNAi vectors. *Mechanisms of Development*, 125(11–12), pp.947–962.
- Mendes, N.D., Freitas, A.T. & Sagot, M.F., 2009. Survey and Summary: Current tools for the identification of miRNA genes and their targets. *Nucleic Acids Research*, 37(8), pp.2419–2433.
- Menko, A.S. & Boettiger, D., 1987. Occupation of the extracellular matrix receptor, integrin, is a control point for myogenic differentiation. *Cell*, 51(1), pp.51–57.
- Milagro, F.I. et al., 2013. High-Throughput Sequencing of microRNAs in Peripheral Blood Mononuclear Cells: Identification of Potential Weight Loss Biomarkers. *PLoS ONE*, 8(1), pp.1–10.
- Mishima, N. & Tomarev, S., 1998. Chicken Eyes absent 2 gene: Isolation and expression pattern during development. *International Journal of Developmental Biology*, 42(8), pp.1109–1115.
- Mishima, Y. et al., 2009. Zebrafish miR-1 and miR-133 shape muscle gene expression and regulate sarcomeric actin organization. *Genes and Development*, 23(5), pp.619–632.

- Mok, G.F. & Sweetman, D., 2011. Many routes to the same destination: Lessons from skeletal muscle development. *Reproduction*, 141(3), pp.301–312.
- Moncaut, N., Rigby, P.W.J. & Carvajal, J.J., 2013. Dial M(RF) for myogenesis. *FEBS Journal*, 280(17), pp.3980–3990.
- Monteys, A.M. et al., 2010. Structure and activity of putative intronic miRNA promoters. *RNA*, 16(3), pp.495–505.
- Morin, R.D. et al., 2008. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Research*, pp.610–621.
- Motohashi, N. et al., 2013. Regulation of IRS1/Akt insulin signaling by microRNA-128a during myogenesis. *Journal of Cell Science*, 126, pp.2678–2691.
- Munsterberg, A.E. et al., 1995. Combinatorial signaling by Sonic hedgehog and Wnt family members induces myogenic bHLH gene expression in the somite. *Genes and Development*, 9(23), pp.2911–2922.
- Musumeci, G. et al., 2015. Somitogenesis: From somite to skeletal muscle. *Acta Histochemica*, 117(4–5), pp.313–328.
- Nabeshima, Y. et al., 1993. Myogenin gene disruption results in perinatal lethality because of severe muscle defect. *Nature Letters*, 364, pp.532–535.
- Nakamura, H. & Funahashi, J., 2012. Electroporation: Past, present and future. *Development, Growth & Differentiation*, 55(1), pp.15–19.
- Nakaya, Y. et al., 2004. Mesenchymal-epithelial transition during somitic segmentation is regulated by differential roles of Cdc42 and Rac1. *Developmental Cell*, 7(3), pp.425–438.
- Nakaya, Y. & Sheng, G., 2008. Epithelial to mesenchymal transition during gastrulation: An embryological view. *Development Growth and Differentiation*, 50(9), pp.755–766.
- Nelson, P.T. et al., 2006. RAKE and LNA-ISH reveal microRNA expression and localization in RAKE and LNA-ISH reveal microRNA expression and localization in archival human brain. *Rna*, 12, pp.187–191.
- Nielsen, C.B. et al., 1999. The solution structure of a locked nucleic acid (LNA) hybridized to DNA. *Journal of biomolecular structure & dynamics*, 17(2), pp.175–91.
- Nilsen, T.W., 2007. Mechanisms of microRNA-mediated gene regulation in animal cells. *Trends Genet*, 23(5), pp.243–249.

- Niro, C. et al., 2010. Six1 and Six4 gene expression is necessary to activate the fast-type muscle gene program in the mouse primary myotome. *Developmental Biology*, 338(2), pp.168–182.
- Nohata, N. et al., 2012. microRNA-1/133a and microRNA-206/133b clusters: dysregulation and functional roles in human cancers. *Oncotarget*, 3(1), pp.9–21.
- O'Brien, J.H. et al., 2014. MicroRNA-30a regulates zebrafish myogenesis through targeting the transcription factor Six1. *Journal of cell science*, 127(Pt 10), pp.2291–301.
- O'Rourke, J.R. et al., 2007. Essential role for Dicer during skeletal muscle development. *Developmental Biology*, 311(2), pp.359–368.
- Ohto, H. et al., 1999. Cooperation of Six and Eya in Activation of Their Target Genes through Nuclear Translocation of Eya. *Molecular and cellular biology*, 19(10), pp.6815–6824.
- Oliver, G. et al., 1995. Homeobox genes and connective tissue patterning. *Development*, 121(3), pp.693–705.
- Ordahl, C.P. et al., 2001. The dermomyotome dorsomedial lip drives growth and morphogenesis of both the primary myotome and dermomyotome epithelium. *Development*, 128(10), pp.1731–44.
- Ordahl, C.P. & Le Douarin, N.M., 1992. Two myogenic lineages within the developing somite. *Development*, 114(2), pp.339–53.
- Otto, A., Schmidt, C. & Patel, K., 2006. Pax3 and Pax7 expression and regulation in the avian embryo. *Anatomy and Embryology*, 211(4), pp.293–310.
- Ozaki, H. et al., 2004. Six1 controls patterning of the mouse otic vesicle. *Development (Camb)*, 131(3), pp.551–562.
- Ozaki, H. et al., 2001. Six4, a putative myogenin gene regulator, is not essential for mouse embryonal development. *Molecular and cellular biology*, 21(10), pp.3343–50.
- Ozbudak, E.M. & Pourquié, O., 2008. The vertebrate segmentation clock: the tip of the iceberg. *Current Opinion in Genetics and Development*, 18(4), pp.317–323.
- Palmeirim, I. et al., 1997. Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. *Cell*, 91(5), pp.639–648.
- Palmeirim, I. et al., 2008. Development on Time. In M. Maroto & N. A. M. Monk, eds. *Advances in Experimental Medicine and Biology*. New York, NY: Springer New York, pp. 62–71.
- Pasquinelli, a E. et al., 2000. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature*, 408(6808), pp.86–89.

- Peterson, S.M. et al., 2014. Common features of microRNA target prediction tools. *Frontiers in Genetics*, 5(FEB), pp.1–10.
- Pignoni, F. et al., 1997. The eye-specification proteins So and Eya form a complex and regulate multiple steps in Drosophila eye development. *Cell*, 91(7), pp.881–891.
- Place, R.F. et al., 2008. MicroRNA-373 induces expression of genes with complementary promoter sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 105(5), pp.1608–13.
- Pourquié, O. et al., 1995. Control of somite patterning by signals from the lateral plate. *Proceedings of the National Academy of Sciences of the United States of America*, 92(8), pp.3219–3223.
- Pourquié, O. et al., 1996. Lateral and axial signals involved in avian somite patterning: A role for BMP4. *Cell*, 84(3), pp.461–471.
- Pratt, A.J. & MacRae, I.J., 2009. The RNA-induced silencing complex: A versatile gene-silencing machine. *Journal of Biological Chemistry*, 284(27), pp.17897–17901.
- Qin, W. et al., 2010. miR-24 regulates apoptosis by targeting the Open Reading Frame (ORF) region of FAF1 in cancer cells. *PLoS ONE*, 5(2), pp.1–8.
- Raichur, S. et al., 2010. Identification and validation of the pathways and functions regulated by the orphan nuclear receptor, ROR alpha1, in skeletal muscle. *Nucleic Acids Research*, 38(13), pp.4296–4312.
- Rao, P.K. et al., 2010. Distinct roles for miR-1 and miR-133a in the proliferation and differentiation of rhabdomyosarcoma cells. *The FASEB Journal*, 24(9), pp.3427–3437.
- Rao, P.K. et al., 2006. Myogenic factors that regulate expression of muscle-specific microRNAs. *Proceedings of the National Academy of Sciences of the United States of America*, 103(23), pp.8721–6.
- Rathjen, T. et al., 2009. High throughput sequencing of microRNAs in chicken somites. *FEBS Letters*, 583(9), pp.1422–1426.
- Rayapureddi, J.P. et al., 2003. Eyes absent represents a class of protein tyrosine phosphatases. *Nature*, 426(6964), pp.295–8.
- Rebay, I., 2015. Multiple functions of the Eya phosphotyrosine phosphatase. *Molecular and Cellular Biology*, 36(December), p.MCB.00976-15.
- Reimand, J. et al., 2007. G:Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Research*, 35(SUPPL.2), pp.193–200.

- Reimand, J. et al., 2016. g:Profiler—a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Research*, pp.1–7.
- Reinhart, B.J. et al., 2000. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature*, 403(6772), pp.901–906.
- Relaix, F. et al., 2005. A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature*, 435(7044), pp.948–53.
- Relaix, F. et al., 2013. Six Homeoproteins Directly Activate Myod Expression in the Gene Regulatory Networks That Control Early Myogenesis. *PLoS Genetics*, 9(4).
- Relaix, F. & Buckingham, M., 1999. From insect eye to vertebrate muscle: redeployment of a regulatory network. *Genes and Development*, pp.3171–3178.
- Rhoades, M.W. et al., 2002. Prediction of plant microRNA targets. *Cell*, 110(4), pp.513–520.
- Richard, A.F. et al., 2011. Genesis of muscle fiber-type diversity during mouse embryogenesis relies on Six1 and Six4 gene expression. *Developmental Biology*, 359(2), pp.303–320.
- Robson, L.G. et al., 2011. Bmi1 is expressed in postnatal myogenic satellite cells, controls their maintenance and plays an essential role in repeated muscle regeneration. *PLoS ONE*, 6(11).
- Rodriguez, A., 2004. Identification of Mammalian microRNA Host Genes and Transcription Units. *Genome Research*, 14(10a), pp.1902–1910.
- van Rooij, E. et al., 2009. A Family of microRNAs Encoded by Myosin Genes Governs Myosin Expression and Muscle Performance. *Developmental Cell*, 17(5), pp.662–673.
- van Rooij, E. et al., 2007. Control of stress-dependent cardiac growth and gene expression by a microRNA. *Science*, 316(5824), pp.575–579.
- van Rooij, E., Liu, N. & Olson, E.N., 2008. MicroRNAs flex their muscles. *Trends in Genetics*, 24(4), pp.159–166.
- Rudnicki, M.A. et al., 1993. MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell*, 75(7), pp.1351–1359.
- Ruvkun, G. & Giusto, J., 1989. The *Caenorhabditis elegans* heterochronic gene *lin-14* encodes a nuclear protein that forms a temporal developmental switch. *Nature*, 338(6213), pp.313–319.
- Sánchez-Guardado, L.Ó. et al., 2011. Distinct and redundant expression and transcriptional diversity of MEIS gene paralogs during chicken development. *Developmental Dynamics*, 240(6), pp.1475–1492.

- Sano, T. & Nagata, S., 2011. Characterization of the threonine-phosphatase of mouse eyes absent 3. *FEBS Letters*, 585(17), pp.2714–2719.
- Scaal, M. & Christ, B., 2004. Formation and differentiation of the avian dermomyotome. *Anatomy and Embryology*, 208(6), pp.411–424.
- Schönberger, J. et al., 2005. Mutation in the transcriptional coactivator EYA4 causes dilated cardiomyopathy and sensorineural hearing loss. *Nature genetics*, 37(4), pp.418–422.
- Schubert, F.R. et al., 2002. Wnt6 marks sites of epithelial transformations in the chick embryo. *Mechanisms of Development*, 114(1–2), pp.143–148.
- Schwarz, D.S. et al., 2003. Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, 115(2), pp.199–208.
- Selbach, M. et al., 2008. Widespread changes in protein synthesis induced by microRNAs. *Nature*, 455(7209), pp.58–63.
- Sempere, L.F. et al., 2004. Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation. *Genome biology*, 5(3), p.R13.
- Shi, L. et al., 2015. MicroRNA-128 targets myostatin at coding domain sequence to regulate myoblasts in skeletal muscle development. *Cellular Signalling*, 27(9), pp.1895–1904.
- Shkumatava, A. et al., 2009. Coherent but overlapping expression of microRNAs and their targets during vertebrate development. *Genes and Development*, 23, pp.466–481.
- Small, E.M. et al., 2010. Regulation of PI3-kinase/Akt signaling by muscle-enriched microRNA-486. *Proceedings of the National Academy of Sciences*, 107(9), pp.4218–4223.
- Smirnova, L. et al., 2005. Regulation of miRNA expression during neural cell specification. *European Journal of Neuroscience*, 21(6), pp.1469–1477.
- Smith, C.M. et al., 2014. The mouse Gene Expression Database (GXD): 2014 update. *Nucleic Acids Research*, 42(D1), pp.818–824.
- Söker, T. et al., 2008. Pleiotropic effects in Eya3 knockout mice. *BMC developmental biology*, 8(1), p.118.
- Solnica-Krezel, L., 2005. Conserved patterns of cell movements during vertebrate gastrulation. *Current Biology*, 15(6), pp.213–228.
- Solnica-Krezel, L. & Sepich, D.S., 2012. Gastrulation: Making and Shaping Germ Layers. *Annual Review of Cell and Developmental Biology*, 28(1), pp.687–717.

- Sosic, D. et al., 1997. Regulation of paraxis Expression and Somite Formation by Ectoderm- and Neural Tube-Derived Signals. *Developmental Biology*, 185(2), pp.229–243.
- Spitz, F. et al., 1998. Expression of myogenin during embryogenesis is controlled by Six/sine oculis homeoproteins through a conserved MEF3 binding site. *Proceedings of the National Academy of Sciences of the United States of America*, 95(November), pp.14220–14225.
- Stern, C.D., 2005. The chick: A great model system becomes even greater. *Developmental Cell*, 8(1), pp.9–17.
- Stern, C.D. & Piatkowska, A.M., 2015. Multiple roles of timing in somite formation. *Seminars in Cell and Developmental Biology*, 42, pp.134–139.
- Sun, Y. et al., 2010. Mammalian target of rapamycin regulates miRNA-1 and follistatin in skeletal myogenesis. *Journal of Cell Biology*, 189(7), pp.1157–1169.
- Sweetman, D. et al., 2006. FGF-4 signaling is involved in mir-206 expression in developing somites of chicken embryos. *Developmental Dynamics*, 235(8), pp.2185–2191.
- Sweetman, D. et al., 2008. Specific requirements of MRFs for the expression of muscle specific microRNAs, miR-1, miR-206 and miR-133. *Developmental Biology*, 321(2), pp.491–499.
- Tajbakhsh, S. et al., 1997. Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. *Cell*, 89(1), pp.127–138.
- Tapscott, S.J., 2005. The circuitry of a master switch: Myod and the regulation of skeletal muscle gene transcription. *Development*, 132(12), pp.2685–2695.
- Tay, Y. et al., 2008. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature*, 455(7216), pp.1124–1128.
- Thomsen, R., Nielsen, P.S. & Jensen, T.H., 2005. Dramatically improved RNA in situ hybridization signals using LNA-modified probes. *RNA*, 11(11), pp.1745–8.
- Tootle, T.L. et al., 2003. The transcription factor Eyes absent is a protein tyrosine phosphatase. *Nature*, 426(6964), pp.299–302.
- Tran, N. & Hutvagner, G., 2013. Biogenesis and the regulation of the maturation of miRNAs. *Essays in biochemistry*, 54, pp.17–28.
- Tremblay, P. et al., 1998. A crucial role for Pax3 in the development of the hypaxial musculature and the long-range migration of muscle precursors. *Developmental biology*, 203(1), pp.49–61.

- Tuddenham, L. et al., 2006. The cartilage specific microRNA-140 targets histone deacetylase 4 in mouse cells. *FEBS Letters*, 580(17), pp.4214–4217.
- Tuzovic, L. et al., 2013. A human de novo mutation in MYH10 phenocopies the loss of function mutation in mice. *Rare diseases*, 1, p.e26144.
- Ujhelly, O. et al., 2015. Lack of Rybp in Mouse Embryonic Stem Cells Impairs Cardiac Differentiation. *Stem Cells and Development*, 24(18), pp.2193–2205.
- Untergasser, A. et al., 2007. Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research*, 35(SUPPL.2), pp.71–74.
- Vandesompele, J. et al., 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology*, 3(7), pp.1–12.
- Várallyay, E., Burgyán, J. & Havelda, Z., 2007. Detection of microRNAs by Northern blot analyses using LNA probes. *Methods*, 43(2), pp.140–145.
- Vella, M.C. et al., 2004. The *C. elegans* microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR. *Genes and Development*, 18(2), pp.132–137.
- Vincent, C. et al., 1997. BOR and BO syndromes are allelic defects of EYA1. *European journal of human genetics : EJHG*, 5(4), pp.242–246.
- Voiculescu, O., Papanayotou, C. & Stern, C.D., 2008. Spatially and temporally controlled electroporation of early chick embryos. *Nat Protoc*, 3(3), pp.419–426.
- Wahid, F. et al., 2010. MicroRNAs: Synthesis, mechanism, function, and recent clinical trials. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1803(11), pp.1231–1243.
- Wang, H. et al., 2008. NF- $\kappa$ B-YY1-miR-29 Regulatory Circuitry in Skeletal Myogenesis and Rhabdomyosarcoma. *Cancer Cell*, 14(5), pp.369–381.
- Wang, L. et al., 2008. Eya4 regulation of Na<sup>+</sup>/K<sup>+</sup>-ATPase is required for sensory system development in zebrafish. *Development*, 135(20), pp.3425–34.
- Wang, X.H., 2013. MicroRNA in myogenesis and muscle atrophy. *Curr Opin Clin Nutr Metab Care*, 16(3), pp.258–266.
- Wayne, S. et al., 2001. Mutations in the transcriptional activator EYA4 cause late-onset deafness at the DFNA10 locus. *Human molecular genetics*, 10(3), pp.195–200.
- Westholm, J. & Lai, E.C., 2010. Mirtrons: microRNA biogenesis via splicing. *Biochimie*, 48(Suppl 2), pp.1–6.
- Wienholds, E. et al., 2005. MicroRNA expression in zebrafish embryonic development. *Science*, 309(5732), pp.310–311.

- Wightman, B. et al., 1991. Negative regulatory sequences in the lin-14 3'-untranslated region are necessary to generate a temporal switch during *Caenorhabditis elegans* development. *Genes and Development*, 5(10), pp.1813–1824.
- Williams, A.H. et al., 2009. MicroRNA-206 Delays ALS Progression and Promotes Regeneration of Neuromuscular Synapses in Mice. *Science*, 326(5959), pp.1549–1554.
- Williams, B. a & Ordahl, C.P., 1994. Pax-3 expression in segmental mesoderm marks early stages in myogenic cell specification. *Development*, 120(4), pp.785–796.
- Witman, N. et al., 2013. MiR-128 regulates non-myocyte hyperplasia, deposition of extracellular matrix and Islet1 expression during newt cardiac regeneration. *Developmental Biology*, 383(2), pp.253–263.
- Wu, W. et al., 2014. The role of Six1 in the genesis of muscle cell and skeletal muscle development. *International Journal of Biological Sciences*, 10(9), pp.983–989.
- Xu, H. et al., 2006. Identification of microRNAs from different tissues of chicken embryo and adult chicken. *FEBS Letters*, 580(15), pp.3610–3616.
- Xu, P.X. et al., 1999. Eya1-deficient mice lack ears and kidneys and show abnormal apoptosis of organ primordia. *Nature genetics*, 23(1), pp.113–117.
- Xu, P.X., Cheng, J., et al., 1997. Mouse Eya genes are expressed during limb tendon development and encode a transcriptional activation function. *Proceedings of the National Academy of Sciences of the United States of America*, 94(22), pp.11974–9.
- Xu, P.X., Woo, I., et al., 1997. Mouse Eya homologues of the *Drosophila* eyes absent gene require Pax6 for expression in lens and nasal placode. *Development*, 124(1), pp.219–231.
- Yan, D. et al., 2009. MicroRNA-1/206 targets c-met and inhibits rhabdomyosarcoma development. *Journal of Biological Chemistry*, 284(43), pp.29596–29604.
- Yang, J.S. & Lai, E.C., 2011. Alternative miRNA Biogenesis Pathways and the Interpretation of Core miRNA Pathway Mutants. *Molecular Cell*, 43(6), pp.892–903.
- Yi, R. et al., 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes and Development*, 17(24), pp.3011–3016.
- Yuan, B. et al., 2014. A phosphotyrosine switch determines the antitumor activity of ER $\beta$ . *The Journal of Clinical Investigation*, 124(8), pp.3378–3390.
- Yusuf, F. & Brand-Saber, B., 2006. The eventful somite: Patterning, fate determination and cell division in the somite. *Anatomy and Embryology*, 211(SUPPL. 1), pp.21–30.

- Zeng, Y., Yi, R. & Cullen, B.R., 2005. Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *The EMBO journal*, 24(1), pp.138–48.
- Zhang, H. et al., 2002. Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *EMBO Journal*, 21(21), pp.5875–5885.
- Zhang, H. et al., 2004. Single processing center models for human Dicer and bacterial RNase III. *Cell*, 118(1), pp.57–68.
- Zhang, Y. et al., 2004. A comparative study of Eya1 and Eya4 protein function and its implication in branchio-oto-renal syndrome and DFNA10. *JARO - Journal of the Association for Research in Otolaryngology*, 5(3), pp.295–304.
- Zhao, Y. et al., 2007. Dysregulation of Cardiogenesis, Cardiac Conduction, and Cell Cycle in Mice Lacking miRNA-1-2. *Cell*, 129(2), pp.303–317.
- Zhou, B. et al., 2010. MicroRNA expression profiles of porcine skeletal muscle. *Animal Genetics*, 41(5), pp.499–508.
- Zhou, L. et al., 2012. A novel target of microRNA-29, Ring1 and YY1-binding protein (Rybp), negatively regulates skeletal myogenesis. *Journal of Biological Chemistry*, 287(30), pp.25255–25265.

## APPENDIX I

**Table 1: Primers to generate ISH probes.**

Genes	Primers	Sequences (5'-3')
Eya1	ggaEya1(ISH)-F	GCTTCGCCCATCTGGAAAAC
	ggaEya1(ISH)-R	GCTGGCACCGTAGCTTAGAA
Eya2	ggaEya2(ISH)-F	ACATTTATGCCACGGCTCCA
	ggaEya2(ISH)-R	GCACAGACGTTGTTGTGTCC
Eya3	ggaEya3(ISH)-F	CCCAGACCACCCAAACCTAC
	ggaEya3(ISH)-R	ATGGTCTCATCCAGGTCCCA
Eya4	ggaEya4(ISH)-F	CTCCGAACGCCAGGTCTATG
	ggaEya4(ISH)-R	GGAGGGGCTCTGTACTGTGT
Six1	ggaSix1(ISH)-F	ATGTGCGATGCTGCCGTCGTT
	ggaSix1(ISH)-R	TTAGGAGCCCAGGTCCACCA
Six4	ggaSix4(ISH)-F	TGGAGAGCCACAACCTCGAC
	ggaSix4(ISH)-R	ACACCAGATGAGCTCAAGGC
Dach1	ggaDach1(ISH)-F	GCCTCGGGGACAAACCTATT
	ggaDach1(ISH)-R	CCTGGGACAGAATGTGGCAT
Dach2	ggaDach2(ISH)-F	CTCACCAACAGCCTCGTCAA
	ggaDach2(ISH)-R	CTGCCCTGGAAAGAGGACTG

**Table 2: Primer sequences for PCR amplification of the 3'UTR region of PSED members.** For sub-cloning purpose, restriction sites (bases underlined) were added to the 5' end of the primers. BglII (AGATCT), NheI (GCTAGC).

Genes	Primers	Sequences (5'-3')
Eya1	ggaEya1(3UTR-BglII)-F1	GC <u>AGATCT</u> CAGCTCAGCAGCACTTTGAA
	ggaEya1(3UTR-NheI)-R1	AT <u>GCTAGC</u> CTGACTCCTGGTGGAAAGAGG
	ggaEya1(3UTR-BglII)-F2	AG <u>AGATCT</u> CAACTACCTGCAAAGCTGCG
	ggaEya1(3UTR-NheI)-R2	GC <u>GCTAGC</u> TCAATGTAGCAACAAAACCCAG
Eya2	ggaEya2(3UTR-BglII)-F	AT <u>AGATCT</u> CAGACCCCAACATCTTAGCA
	ggaEya2(3UTR-NheI)-R	AT <u>GCTAGC</u> TCCCACCCAAACTGAGATGG
Eya3	ggaEya3(3UTR-BglII)-F	AT <u>AGATCT</u> GTAGTCTCCAGAGGGAGGGG
	ggaEya3(3UTR-NheI)-R	AT <u>GCTAGC</u> CTGTGGCATCTGTGGTCTGA
Eya4	ggaEya4(3UTR-BglII)-F	AG <u>AGATCT</u> TGTTCTAAAGTTGGCGATCCT
	ggaEya4(3UTR-NheI)-R	CT <u>GCTAGC</u> CACTCACTGCATGGCTTTCA
Six1	ggaSix1(3UTR-BglII)-F	AT <u>AGATCT</u> TAGCCAAATGCAGAGAGCGG
	ggaSix1(3UTR-NheI)-R	AT <u>GCTAGC</u> CGGCTGTTCCGGTACAATAGA
Six4	ggaSix4(3UTR-BglII)-F1	AT <u>AGATCT</u> GGAGAAAGGAAACGCCAGGG
	ggaSix4(3UTR-NheI)-R1	AT <u>GCTAGC</u> AGGATGTGCTTTTCCACGCTA
	ggaSix4(3UTR-BglII)-F2	GC <u>AGATCT</u> ATGGGTGTCCTCTCCCTTCA
	ggaSix4(3UTR-NheI)-R2	AT <u>GCTAGC</u> TGTACCGTGCAACGAGTCTTTA
	ggaSix4(3UTR-BglII)-F3	AT <u>AGATCT</u> ACCTCCCGTTCTTTCGTGG

	ggaSix4(3UTR-NheI)-R3	AT <u>GCTAGC</u> GGACCCTGCATGTCTGTTTCA
Dach1	ggaDach1(3UTR-BglII)-F1	AT <u>AGATCT</u> CCTGCTGAAGATACCTGTGCT
	ggaDach1(3UTR-NheI)-R1	GC <u>GCTAGC</u> TGTGTACCAGTATTGCAAGGAAG
	ggaDach1(3UTR-BglII)-F2	GC <u>AGATCT</u> TGCCTCATTGTTTGGCTTGG
	ggaDach1(3UTR-NheI)-R2	AT <u>GCTAGC</u> AACAACTGGATTACCCTCTCTG

**Table 3: Mutagenesis primers.** Bases constituting miRNA sites are underlined. Mutated nucleotides are indicated in red.

Genes	Primers	Sequences (5'-3')
Eya1	ggaEya1-1 (3UTR-m133a)-F	TAGATAAATTTGTCA <u>GGTACC</u> AAAGCATGGATGT CAAGTGTCAATATG
	ggaEya1-1 (3UTR-m133a)-R	CATATTGACACTTGACATCCATGCTTT <u>GGTACC</u> T GACAAATTTATCTA
	ggaEya1-1 (3UTR-m128/27(1))-F	AGAACAGCTGTTGACTCTGGT <u>GCGGCCGCT</u> CCAA CAAAAATAAGCCA
	ggaEya1-1 (3UTR-m128/27(1))-R	TGGCTTATTTTTGTTGGAG <u>GCGGCCG</u> CACCAGAGT CAACAGCTGTTCT
	ggaEya1-1 (3UTR-m128/27(2))-F	AGGAAATAAAGGTTCTGACC <u>GGTACC</u> AAAAGGAC CTGCAAGTGCTTTG
	ggaEya1-1 (3UTR-m128/27(2))-R	CAAAGCACTTGCAGGTCCTTTT <u>GGTACC</u> GGTACG AACCTTTATTTCT
Eya3	ggaEya3 (3UTR-m133a)-F	CTGTTAATGAGCAG <u>ATCTT</u> CATTAGATTCCAGCT GTCCATGAC
	ggaEya3 (3UTR-m133a)-R	GTCATGGACAGCTGGAATCTAATGA <u>AGATCT</u> GCT CATTAAACAG
Eya4	ggaEya4 (3UTR-m128/27)-F	TTTGTGTAATTATTGATGAAAATAACTT <u>ACCA</u> T <u>GGCTTT</u> TATTAGCAGCTGATTTT
	ggaEya4 (3UTR-m128/27)-R	AAAATCAGCTGCTAATAAAG <u>CCA</u> TGGTAAGTTAT TTTCATCAATAATTTACACAAA
	ggaEya4 (3UTR-m1a/206)-F	CAAAGTGGTGTTC AAC <u>AAGCTT</u> CCTCAAATGGG ATATATTCTCAG
	ggaEya4 (3UTR-m1a/206)-R	CTGAGAATATATCCATTTT <u>GAGGAAGCTT</u> GTTG AACACCACTTTG
	ggaEya4 (3UTR-m133a)-F	GTATGTTGGGAGGTTGATAGT <u>GGCTAGCG</u> CTACC TTGAAAGCTAAAAAAGA
	ggaEya4 (3UTR-m133a)-R	TCTTTTTTAGCTTTCAAGGTAGCG <u>CTAG</u> CCACTA TCAACCTCCCAACATAC
	ggaEya4 (3UTR-m499)-F	TCTGGCTTTACACATATGAAT <u>AAGCTT</u> AAGAAGG GAAGAAATATTTGGAATTA AAA
	ggaEya4 (3UTR-m499)-R	TTTTAATTCCAAATATTTCTTCCCTTCTTAAG <u>CT</u> <u>T</u> ATTCATATGTGTAAGCCAGA
Six1	ggaSix1 (3UTR-m1a/206)-F	AGGGGAACTTTTT <u>CGTGAGCTCTT</u> CTTTTTTTTC ATATTTAGCTTC
	ggaSix1 (3UTR-m1a/206)-R	GAAGCTAAATATGAAAAAAGAA <u>GAGCTC</u> ACGA AAAAGTTCCCCT

Six4	ggaSix4-1 (3UTR-m133a)-F	AAGGATTGTGCCGAG <u>GCGGCCG</u> CGTGCAGAGCAGT GCAGAGCAGTGCAG
	ggaSix4-1 (3UTR-m133a)-R	CTGCACTGCTCTGCACTGCTCTGCACG <u>GCGGCCG</u> TCGGCACAATCCTT
	ggaSix4-3 (3UTR-m1a/206)-F	GTCGTACTTAATTGG <u>CTAGCT</u> CAGCCACATCAGT CGTGGACGCCTAT
	ggaSix4-3 (3UTR-m1a/206)-R	ATAGGCGTCCACGACTGATGTGGCT <u>GAGCTAGCC</u> AATTAAGTACGAC
	ggaSix4-3 (3UTR-m499)-F	AGATATTTAAATAGT <u>CCATGGACT</u> CCTACTGTAA ATTAAGGGTTGG
	ggaSix4-3 (3UTR-m499)-R	CCAACCCTTAATTTACAGTAGGAGT <u>CCATGGACT</u> ATTTAAATATCT
Dach1	ggaDach1-2 (3UTR-m1a/206)-F	CTACATGATTTATTTATGT <u>CCATGGCT</u> CAGTTTA TGAAGCTGTTAT
	ggaDach1-2 (3UTR-m1a/206)-R	ATAACAGCTTCATAAACTGAG <u>CCATGGAC</u> ATAAAA TAAATCATGTAG
	ggaDach1- 2(3UTR-m1a-2)-F	ACCTTTTTTTATATATTGTGA <u>AGATCT</u> CATGATT CTTATTTTCTAGA
	ggaDach1-2 (3UTR-m1a-2)-R	TCTGAAATAAGAATCATG <u>AGATCT</u> TCACAATATA TAAAAAAGGT

**Table 4: Gga-Eya4 Full-length primers.** For sub-cloning purpose, restriction sites (bases underlined) were added to the 5' end of the primers. NotI (GCGGCCGC), EcoRI (GAATTC). Start and stop codons are indicated in red. The HA-tag is indicated in italic.

Genes	Primers	Sequences (5'-3')
Eya4	ggaEya4 (NotI/ATG)-F	GC <u>GCGGCCGC</u> <b>ATG</b> GAAGACTCTCAGGACCTA
	ggaEya4 (EcoRI/stop/HA)-R	GC <u>GAATTC</u> <b>TTA</b> <i>AGCGTAATCTGGAACATCGTATGGGTA</i> CAAATACTCTAGTTCCAG

**Table 5: Other primers used for sequencing and mutagenesis.**

Genes	Primers	Sequences (5'-3')
pGL3	pGL3-F	CTCATCAATGTATCTTATCATGTC
	pGL3-R	CCTCATAAAGGCCAAGAA
AmpR	Amp_GA_1-F	CAACTTTATCCGCTCCATCCAGTCTAT TAATTGTTGCCGGGAAGCT
	Amp_GA_1-R	GGAGCTGAATGAAGCCATACCAAACGACGAGCGT GA
pCAB	pCAB-F	GGCAGGAAGGAAATGGGCGGGGA
	pCAB-R	GGCCCTCACATTGCCAAAAGACG

**Table 6: AntagomiR sequences.** All bases were replaced by 2’O-methyl-bases, and some phosphodiester bonds were replaced by thiol bonds, indicated in the sequences by ‘m’ and ‘\*’, respectively. The antagomiRs were FITC-labelled at their 5’ end (Fl), and included a 3’ cholesterol moiety at their 3’ end (Chl).

AntagomiR	Sequences (5’-3’)
AntagomiR-206 (AM-206)	(Fl) mC (*) mC (*) mAmCmAmCmAmCmUmUmCmCmUmUmAmCmAmUmU (*) mC (*) mC (*) mA (Chl)
AntagomiR-133 (AM-133)	(Fl) mA (*) mC (*) mAmGmCmUmGmGmUmUmGmAmAmGmGmGmGmAmC (*) mC (*) mA (*) mA (Chl)
AntagomiR-128 (AM-128)	(Fl) mA (*) mA (*) mAmGmAmGmAmCmCmGmGmUmUmCmAmCmUmG (*) mU (*) mG (*) mA (Chl)
AntagomiR-scrambled (AM-scr)	(Fl) mC (*) mA (*) mUmCmCmAmUmCmAmCmUmCmAmCmUmCmCmAmU (*) mC (*) mA (*) mU (Chl)

**Table 7: qPCR primers.**

Genes	Primers	Sequences (5’-3’)
Eya1	gga-Eya1q-F	CCGTATCCCTCGCATTACATG
	gga-Eya1q-R	CTGGTATGTTGCGTTTGTGG
Eya2	gga-Eya2q-F	CGGCTCCAGATGACATAGAAG
	gga-Eya2q-R	GGATTCCCTGCACTGTACTGAG
Eya3	gga-Eya3q-F	AGAGCCACAAGATTTACCCG
	gga-Eya3q-R	CAGTTGACATGGGAAGGTTTG
Eya4	gga-Eya4q-F	GAATCAGATGTGTCAGAGCCTC
	gga-Eya4q-R	TTTATCCAGTTTAGAGCTACCCG
Six1	gga-Six1q-F	TCCTCAAGGCCAAAGCG
	gga-Six1q-R	CTTCTCGGCTTCCACGTAG
Six4	gga-Six4q-F	ATTTATCCCCACATCCGCTC
	gga-Six4q-R	ACGAGGTTCCCATTC AACAG
Dach1	gga-Dach1q-F	GATTTGAGACCCCTCTACAACG
	gga-Dach1q-R	GATTCCAGGAGACATTAGGCC
Pax3	gga-Pax3q-F	CCAAGTATGGCTTTTAACCAC
	gga-Pax3q-R	CTATGGACTGTACTGCTTGATC
MyoD1	gga-MyoD1q-F	ACTTCCACCAACCCCAAC
	gga-MyoD1q-R	TCTGACTCCCCGCTGTAG
Myf5	gga-Myf5q-F	CAACCCCAACCAGAGACTCC
	gga-Myf5q-R	GAGTCCGCCATCACATCGGA
MyoG	gga-MyoGq-F	AGCTGGAGTTTGGCACC
	gga-MyoGq-R	GAGAGCGAGTGGAGGTTG

## APPENDIX II

**Table 1: MiRNA sequences used to run the miRanda algorithm.**

CHICKEN MiRNA	ACCESSION NUMBER	SEQUENCE (5'-3')
<b>gga-miR-1306-3p</b>	MIMAT0007329	ACGUUGGCUCUGGUGGUG
<b>gga-let-7a-5p</b>	MIMAT0001101	UGAGGUAGUAGGUUGUAUAGUU
<b>gga-miR-1a-3p</b>	MIMAT0001127	UGGAAUGUAAAGAAGUAUGUA
<b>gga-miR-1b-3p</b>	MIMAT0001175	UGGAAUGUUAAGAAGUAUGUA
<b>gga-miR-10b-5p</b>	MIMAT0001148	UACCCUGUAGAUCCGAAUUUGU
<b>gga-miR-15a</b>	MIMAT0001117	UAGCAGCACAUAAUGGUUUUGU
<b>gga-miR-15b-5p</b>	MIMAT0001154	UAGCAGCACAUCAUGGUUUUG
<b>gga-miR-15c-5p</b>	MIMAT0007737	UAGCAGCACAUCAUGGUUUUGUA
<b>gga-miR-17-5p</b>	MIMAT0001114	CAAAGUGCUUACAGUGCAGGUAGU
<b>gga-miR-18b-5p</b>	MIMAT0001141	UAAGGUGCAUCUAGUGCAGUUA
<b>gga-miR-23b-3p</b>	MIMAT0001186	AUCACAUUGCCAGGGAAUUACC
<b>gga-miR-24-3p</b>	MIMAT0001188	UGGCUCAGUUCAGCAGGAACAG
<b>gga-miR-27b-3p</b>	MIMAT0001187	UUCACAGUGGCUAAGUUCUGC
<b>gga-miR-30a-5p</b>	MIMAT0001135	UGUAAACAUCUCGACUGGAAG
<b>gga-miR-31-5p</b>	MIMAT0001189	AGGCAAGAUGUUGGCAUAGCUG
<b>gga-miR-128-3p</b>	MIMAT0001123	UCACAGUGAACCGGUCUCUUU
<b>gga-miR-130a-3p</b>	MIMAT0001167	CAGUGCAAUGUUAAGGGC
<b>gga-miR-133a-3p</b>	MIMAT0001126	UUGGUCCCCUUAACCAGCUGU
<b>gga-miR-133b</b>	MIMAT0001138	UUGGUCCCCUUAACCAGCUA
<b>gga-miR-133c-3p</b>	MIMAT0001176	UUGGUCCCCUUAACCAGCUGC
<b>gga-miR-194</b>	MIMAT0001133	UGUAACAGCAACUCCAUGUGGA
<b>gga-miR-203a</b>	MIMAT0001146	GUGAAAUGUUUAGGACCACUUG
<b>gga-miR-206</b>	MIMAT0001139	UGGAAUGUAAGGAAGUGUGGG
<b>gga-miR-223</b>	MIMAT0001140	UGUCAGUUUGUCAAUACCCC
<b>gga-miR-499-5p</b>	MIMAT0003367	UUAAGACUUGUAGUGAUGUUUAG

**Table 2: Summary of miRNA expression in different tissues in chicken embryo.**

miRNA	Somites	Neural Tube	Notochord	Heart	Mesonephros
gga-miR-1306-3p	✓	✓(HH21-22; ventral)	✓	✓	✓
gga-let-7a-3p	✓	✓(dorsal; ventral)	✓		
gga-miR-1a	✓			✓	
gga-miR-1b	✓	✓(early stage)	✓(early stage)	✓	
gga-miR-10b	✓	✓	✓(early stage)		✓
gga-miR-15a	✓	✓(early stage)	✓(early stage)		
gga-miR-15b-5p	✓	✓(HH21-22; dorsal)	✓(early stage)	✓	
gga-miR-15c-5p	✓	✓	✓	✓	✓
gga-miR-17-5p	✓	✓	✓(early stage)		✓
gga-miR-18b	✓	✓	✓(early stage)	✓	✓
gga-miR-23b	✓	✓	✓		✓
gga-miR-24a	✓	✓(early stage)	✓(early stage)		✓
gga-miR-30a-5p	✓	✓(early stage)	✓(early stage)		✓
gga-miR-31	✓	✓(early stage)	✓(early stage)		
gga-miR-128	✓	✓(early stage)			
gga-miR-130a	✓	✓(ventral)	✓(early stage)		
gga-miR-133a	✓	✓(early stage)	✓(early stage)	✓	
gga-miR-133b	✓			✓	
gga-miR-133c	✓			✓	
gga-miR-194	✓	✓	✓(early stage)		✓
gga-miR-203	✓	✓(early stage)	✓(early stage)		✓
gga-miR-206	✓	✓(early stage)			
gga-miR-223	✓		✓		

**Table 3: Somitic miRNAs – Bioinformatics analysis.** Gene ontology (GO) term and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway annotation analysis performed using the Database for Annotation, Visualisation and Integrated Discovery tool (DAVID Bioinformatics resources). For each miRNA, the number of predicted targets is indicated. Genes from categories of interest (for example: Developmental process, signalling pathways, muscle) have been listed. Genes from the PSED network have been underlined.

MiRNA	Targets	DAVID analysis	
miR-1a-3p	268	<b>GOTERM_BP_1 (253)</b>	
		191	Cellular process
		147	Biological regulation
		135	Metabolic process
		67	Developmental process
		<b>GOTERM_BP_FAT (253)</b>	
		74	Regulation of transcription
		62	Transcription
		49	Regulation of RNA metabolic process
		<b>KEGG_PATHWAY (253)</b>	
		10	MAPK signalling pathway
		9	Neurotrophin signalling pathway
		8	Adherens function
		8	Wnt signalling pathway
miR-10a-5p	86	<b>GOTERM_BP_1 (84)</b>	
		65	Cellular process
		49	Biological regulation
		31	Multicellular organismal process
		25	Developmental process
		<b>GOTERM_BP_FAT (84)</b>	
		34	Regulation of transcription
		32	Transcription
		29	Regulation of RNA metabolic process
		<b>KEGG_PATHWAY (84)</b>	
		4	Regulation of actin cytoskeleton
		2	Neurotrophin signalling pathway
		<b>Developmental process_list (67):</b> ALX1; CD164; CITED2; E2F5; GFRA1; GLIS2; KLF4; MNT; MXD1; MEIS1; NAB1; RYBP; SH3GL1; SOX6; WHSC1L1; ACTB; AP3D1; ASPH; BMPR1B; BDNF; CREM; CTNND2; COL4A3BP; COL19A1; CCND1; CCND2; DICER1; DONSON; MECOM; EFNB2; ERBB2IP; EYA4; FGFR3; FRS2; FNDC3A; FOXP1; GJA1; GNAQ; HSPD1; HOXA3; KALRN; KIF2A; LEF1; MAB21L1; MEOX2; MAPK1; MAP3K1; MYEF2; NRCAM; NRP1; NFATC3; NR4A3; PAX6; LPPR4; PDGFA; PHLDA2; PDCD4; RARb; RNF111; SNAI2; SLC5A3; SPRED1; SNAP25; TCF7L2; UTRN; VEZF1; ZFP36L1	
		<b>MAPK signalling pathway_list (10):</b> RAP1B; TAOK3; ATF2; BDNF; MECOM; FGFR3; MAPK1; MAP3K1; PDGFA; PRKACB	
<b>Wnt signalling pathway_list (8):</b> RAP1B; TAOK3; ATF2; BDNF; MECOM; FGFR3; MAPK1; MAP3K1; PDGFA; PRKACB			
<b>Developmental process_list (25):</b> BCL6; BCL2L11; DAZAP1; GATA6; MDGA2; RORA; SKIL; TIAM1; ACTG1; ANK3; BDNF; CREB1; EBF2; HOXA3; HOXD10; JARID2; MBNL3; MYT1L; NCOA6; NR2C2; NR5A2; ONECUT1; SOBP; SLIT2			

miR-15a	196	<b>GOTERM_BP_1 (182)</b>		<b>Developmental process_list (57):</b> FKBP1A; KLF7; NAA15; ARHGDI A; SIX5; ASPH; BHLHE41; CHD7; CCND1; DLL1; EFNB2; FOXP2; GHR; GNAQ; HGF; HOXA3; INSR; IRS1; JPH1; LRP6; MAP7; NAV1; NRP2; NFE2L1; NFATC3; OMG; PARD6B; PPAP2B; PIK3R1; PLXNA1; PBX3; PDCD4; PTPRJ; PCDHA11; PCDHAC2; PCDHA1; PCDHA2; PCDHA3; PCDHA5; PCDHA7; PCDHA6; RREB1; RARB; RECK; RPS6KA3; RUNX1; SEMA3A; SIAH1; SHOX; SCN8A; SLC5A3; SNAP25; TCF3; UBR3; WNT10B; WNT7A
		139	Cellular process	
		97	Biological regulation	
		61	Multicellular organismal process	
		57	Developmental process	
		16	Biological adhesion	
		<b>GOTERM_BP_FAT (182)</b>		
			Regulation of transcription	
			Intracellular signalling cascade	
			Regulation of RNA metabolic process	
		<b>KEGG_PATHWAY (182)</b>		
		8	Insulin signalling pathway	
		8	Wnt signalling pathway	
8	MAPK signalling pathway			
6	Neurotrophin signalling pathway			
6	JAK-STAT signalling pathway			
4	mTOR signalling pathway			
miR-15b-5p	85	<b>GOTERM_BP_1 (77)</b>		<b>Developmental process_list (25):</b> DGCR2; NRARP; PAPP A; SOX6; ACVR2B; BTF3; BPTF; DLL4; DAB1; EYA1; GSK3B; LAMC1; MAP2K1; NF1; NLGN1; ONECUT2; PAX2; PLEKHA1; PLXNA2; RFX3; SALL4; SLIT2; TMEM189-UBE2V1; SKI; ZNF423
		41	Biological regulation	
		25	Developmental process	
		20	Cellular component organisation	
		<b>GOTERM_BP_FAT (77)</b>		
		19	Regulation of transcription	
		14	Regulation of RNA metabolic process	
		9	Regulation of cell death	
		<b>KEGG_PATHWAY (77)</b>		
		3	Insulin signalling pathway	
		<b>Insulin signalling pathway_list (8):</b> PDPK1; FOXO1; INSR; <u>IRS1</u> ; PIK3R1; PRKAR2A; SOS2; RAF1		
		<b>Wnt signalling pathway_list (8):</b> BTRC; CCND1; LRP6; NKD1; NFATC3; SIAH1; WNT10B; WNT7A		
		<b>MAPK signalling pathway_list (8):</b> ELK4; RAPGEF2; TAOK1; DUSP7; PPM1A; RPS6KA3; SOS2; RAF1		
		<b>Insulin signalling pathway_list (3):</b> EIF4E; GSK3B; MAP2KA		
		<b>Skeletal muscle_list (6):</b> CHD12; EEA1; GRB10; PLXNA2; SLC24A3; SLC4A4		

		<b>UP_TISSUE (77)</b>	
		41   Brain	
		10   Kidney	
		9   Foetal brain	
		6   Skeletal muscle	
miR-17-5p	111	<b>GOTERM_BP_1 (105)</b>	<b>MAPK signalling pathway_list (7):</b> MKNK2; RASA2; RASGRF2; TAOK3; HSPA8; MAP3K5; RPS6KA5
		55   Biological regulation	
		<b>GOTERM_BP_FAT (105)</b>	<b>Embryo_list (6):</b> ATG2B; RASL11B; CBX8; CSRNP3; OXR1; TANC2
		19   Intracellular signalling cascade	
		12   Protein localisation	
		11   Protein transport	
		<b>KEGG_PATHWAY (105)</b>	
		7   MAPK signalling pathway	
		<b>UP_TISSUE (105)</b>	
		60   Brain	
6   Stomach			
6   Embryo			
miR-23b-3p	470	<b>GOTERM_BP_1 (449)</b>	<b>Developmental process_list (128):</b> ARID3B; AGFG1; CREBBP; CELF1; DIP2A; DICER1; EGLN1; ELK3; HHIP; ISL1; LDB2; MEIS1; MEIS2; NKX2-1; NAA15; PRDM1; RORA; SOX11; SOX5; SOX-; TGIF1; ULK2; ADAM10; ANKRD17; AMBRA1; CAPZA2; CAR2; CTNND2; CDC42; CXCL12; CHD7; COL4A3BP; CBFA2T2; CUL3; CCND1; DACH1; DOCK7; DLL4; DST; EBF1; EBF3; ENC1; EFNA5; ESRRG; ETV1; EYA1; FOXA1; FOXP2; FMR1; GGNBP2; GABRG2; GSK3B; GREM1; GAP43; HAND2; HOXA3; HOXD10; HIPK2; IRS2; JAG1; JARID2; JPH1; LGR4; LRP5; KDM5A; MET; MTSS1; MITF; MACF1; MAPK8; MAP3K1; NLGN1; NFIB; NR6A1; OTP; PTEN; PICALM; PDE3A; PDGFA; POGZ; PBRM1; HMGB2; WWP1; PSEN1; PPP3CA; PTPN11; PRTG; RET; RDH10; RNF2; ROBO2; RUNX1T1; SIK1; SEMA3A; SEMA6D; SRPK2; STK4; SHROOM2; TRIM71; MEF2A; ZIC1; FBN2; PURA; SIX4; SIM1; SIRT1; SCN2A1; SPAST; SATB1; SATB2; SPRY2; SS18L1; TSHZ3; TET2; THRB; TJP1; TOP1; TGFB2; YWHAG; ERBB4; VANGL1; VCAN; ZEB1; ZFR; ZFH3; ZFP423
		283   Cellular process	
		226   Biological regulation	
		206   Metabolic process	
		139   Multicellular organismal process	
		128   Developmental process	
		<b>GOTERM_BP_FAT (449)</b>	
		111   Regulation of transcription	
		94   Transcription	
71   Regulation of RNA metabolic process			

		<b>KEGG_PATHWAY (449)</b>	<b>MAPK signalling pathway_list (19):</b> BRAF; RAP1B; RAPGEF2; ATF2; CDC42; FGF14; MAPK8; MAP3K1; MAP3K5; TAB2; MAP4K4; NLK; PAK2; PDGF1; RAP1A; PPP3CA; STK4; AKT3; TGFB2
		19	MAPK signalling pathway
		13	Wnt signalling pathway
		<b>UP_TISSUE (449)</b>	<b>Wnt signalling pathway_list (13):</b> CREBBP; CSNL2A1; CSNK2A2; CCND1; GSK3B; LRP5; MAPK8; NLK; NFAT5; PSEN1; PPP2R5E; PPP3CA; VANGL1
		246	Brain
		111	Liver
		68	Embryo
		57	Eye
		28	Heart
		16	Foetal brain
13	Limb	<b>Limb_list (13):</b> KDM6A; RAB39B; RBMS3; APPL1; CCNT2; <u>DACH1</u> ; HMGN2; HOXD10; KPNA4; HMGA2; SET; TLK1; SPRY1	
9	Skeletal muscle	<b>Skeletal muscle_list (9):</b> CFL2; FOXK1; JPH1; MYH1; MYH4; NFIB; NFAT5; PDE7A; <u>SIX4</u>	
miR-24a-3p	152	<b>GOTERM_BP_1 (141)</b>	<b>Developmental process_list (38):</b> AGPAT6; BCL2L11; DAZAP1; FREM2; HNF1B; MKL2; MEIS2; PRDM1; RASA1; B3GNT5; ANK3; BHLHE22; CDKN1B; DLL1; EBF3; ELL; <u>EYA4</u> ; INSIG1; KIF2A; LMTK2; MFM1; MKX; NRP1; NRP2; PDE3A; PLAG1; PROX1; PTPRF; PTPRQ; SEMA6A; MEF2A; ARID5B; TSHZ1; TLL1; TOP1; TNFRSF19; WNT4; ZFP217
		86	Cellular process
		63	Biological regulation
		38	Developmental process
		<b>GOTERM_BP_FAT (141)</b>	<b>MAPK signalling pathway_list (7):</b> RASA1; RAP1B; TAOK1; DUSP16; DUSP8; NLK; RAP1A
		34	Regulation of transcription
		28	Transcription
		<b>KEGG_PATHWAY (141)</b>	<b>Embryo_list (23):</b> DNAJC16; ERC2; HNF1B; MKL2; ERG; CDKN1B; DLL1; EBF3; <u>EYA4</u> ; HDGFRP3; IGF2BP1; MKX; PLAG1; REEP1; SESN1; BNIP3L; RAP2C; ARID5B; TCERG1; TMEM161B; TNFRSF19; VCPIP1; ZFP654
		7	MAPK signalling pathway
		<b>UP_TISSUE (141)</b>	<b>Limb_list (6):</b> UGCG; WHSC1; CDV3; KPNA4; RAP2C; SHOC2
71	Brain		
23	Embryo		
6	Limb		
miR-27b-3p	596	<b>GOTERM_BP_1 (559)</b>	
		401	Cellular process

		320	Biological regulation	<b>MAPK signalling pathway_list (27):</b> MKNK2; RAP1B; RAPGEF2; RASGRF1; TAOK1; ATF2; CACNA2D3; CACNB2; DUSP16; DUSP5; FGF14; GRB2; MAPK14; MAPK8IP3; MAP2K4; MAP3K4; TAB2; MEF2C; NLK; NF1; PDGFRA; PRKCB; PRKX; PPP3R1; PPP3R2; SOS1; KRAS	
		282	Metabolic process		
		184	Multicellular organismal process		
		170	Developmental process		
		<b>GOTERM_BP_FAT (559)</b>			<b>Insulin signalling pathway_list (19):</b> PDPK1; CBLB; MKNK2; FOXO1; GRB2; INSR; IRS1; PDE3A; PDE3B; PIK3CA; PIK3R3; PRKAA2; PRKX; PPP1CC; RHOQ; RPS6KB1; SOS1; TSC1; KRAS
		156	Regulation of transcription		
		128	Transcription		
		102	Regulation of RNA metabolic process	<b>KEGG_PATHWAY (559)</b>	<b>Wnt signalling pathway_list (12):</b> APC; CAMK2D; DAAM1; FZD4; NLK; PLCB4; PRKCB; PRKX; PPP3R1; PPP3R2; TBL1Y; WNT8B
		27	MAPK signalling pathway		
		19	Insulin signalling pathway		
		12	Wnt signalling pathway		
		<b>UP_TISSUE (141)</b>			<b>Muscle_list (37):</b> AKAP2; BMI1; EYA4; H3F3B; HMGXB3; INO80D; MEIS2; PALM2-AKAP2; PNISR; RORA; RASAL2; RYBP; TMEM189-UBE2V1; ATF3; ANKRD17; COLQ; DC1A; FOXN3; HOXA10; KPNA3; MAP1B; MEF2C; MYH10; MSTN; MARCKS; NCOA7; NAP1L4; OSBPL11; PALM2; PPP6R3; RFX3; RUNX1T1; SLC25A25; SPRY2; SDC2; TMEM189; UBE2V1
		327	Brain		
		37	Muscle		
26	Heart				
miR-30a-5p	191	<b>GOTERM_BP_1 (183)</b>		<b>Developmental process_list (46):</b> BCL11B; FYN; KLF10; SMAD2; MEIS2; MSI2; NKX2-1; SKIL; ACTC1; CSNK1A1; CHD7; COL9A3; DLL4; DAG1; ESRRG; FOXD1; FOXF2; FOXO3; FRZB; HOXA11; HOXA3; IRS1; IRS2; IGF1R; IRF4; JARID2; MBNL3; MYH10; MYH11; NF1; NFIB; NR6A1; OTP; PRRX1; PGP; PRKAR1A; RPS6KA2; SALL4; NR2F2; LRP6; ZFAND5; SNAIL1; SATB1; SATB2; TIMP3; ZEB2	
		110	Cellular process		
		86	Biological regulation		
		51	Multicellular organismal process		
		46	Developmental process		
		<b>GOTERM_BP_FAT (183)</b>			
		43	Regulation of transcription		
		35	Transcription		
		34	Regulation of RNA metabolic process		
<b>UP_TISSUE (183)</b>					
		100	Brain		

		44	Liver	
		34	Thymus	
miR-31-5p	121	<b>GOTERM_BP_1 (112)</b>		<b>Developmental process_list (39):</b> BD11B; CLASP2; DICER1; EGLN1; ISL1; POU2F1; PTK2; RASA1; SOX11; TAF4A; ADAM10; ACVR1; APBB2; BMPR1A; CTNND2; DMD; EBF3; EFNA5; FOXP1; FXR1; HOMER1; ICMT; JARID2; MAP3K1; NUMB, PP1R9A; PCDH18; PCDH8; RET; SEMA6D; QK; SS18; SPRR2A1; SCN2A1; SATB2; TACC2; YWHAE; UBN1; WNT11
		82	Cellular process	
		58	Biological regulation	
		41	Multicellular organismal process	
		39	Developmental process	<b>TGF-beta signalling pathway_list (3):</b> ACVR1; BMPR1A; LTBP1
		<b>GOTERM_BP_FAT (112)</b>		
		25	Regulation of transcription	<b>Embryo_list (23):</b> AHS2; EGLN1; POU2F1; PTK2; SOX11; BICD2; BMPR1A; EBF3; FOXP1; JARID2; NUFIP2; NCOA2; NUMB; PPP1R9A; PPP2R2A; RERB1; QK; SPRR2A1; SGMS1; TACC2; UBN1; VAMP4; ZFP618
		20	Transcription	
		17	Regulation of RNA metabolic process	
		<b>KEGG_PATHWAY (112)</b>		<b>Skeletal muscle_list (4):</b> KDEL2; DMD; NFAT5; TACC2
		3	TGF-beta signalling pathway	
		<b>UP_TISSUE (112)</b>		
		67	Brain	
		23	Embryo	
19	Heart			
4	Skeletal muscle			
miR-130a-3p	167	<b>GOTERM_BP_1 (151)</b>		<b>Developmental process_list (32):</b> BCL2L11; FYN; HECTD1; ARHGAP24; TAF4A; UGT8A; BMPR2; COL4A3BP; DMRT3; EBF3; ENAH; HOXA5; HPRT; IGF1; LMTK2; MEOX2; MBNL3; NHLH2; NFIB; NCOA3; OTX2; PTEN; PAFAH1B1; WTAP; PTPRF; ROBO2; ARID5B; S1PR1; TGFB2; TSC1; UHRF2; MAFB
		93	Cellular process	
		69	Biological regulation	
		69	Metabolic process	
		32	Developmental process	<b>Insulin signalling pathway_list (5):</b> PPARGC1A; CALM1; PRKACB; PTPRF; TSC1
		<b>GOTERM_BP_FAT (151)</b>		
		27	Regulation of transcription	
		23	Transcription	
20	Regulation of RNA metabolic process			

		<b>KEGG_PATHWAY (151)</b>	
		5   Insulin signalling pathway	<b>Skeletal muscle_list (5):</b> CFL2; IGF1; MSTN; NFIB; STIM2
		<b>UP_TISSUE (151)</b>	
		83   Brain	
		5   Skeletal muscle	
miR-133a-3p	151	<b>GOTERM_BP_1 (143)</b>	<b>Developmental process_list (39):</b> CTBP2; CELF1; EPHA7; GLI3; KLF7; MYCBP2; MEIS1; MEIS2; PRDM16; RB1CC1; AMD1; SOX11; SOX4; SOX8; TBPL1; BICC1; CREB1; COL8A2; CAND1; MECOM; FOXC1; FOXL2; FOXP2; GDNF; MLLT3; MYH9; PITPNB; ODC1; PPP2CA; PTPR21; RBPJ; RTN4RL1; RARB; RUNX1T1; SOBP; SP3; TGFBR1; ZBTB16; ZIC3
		95   Cellular process	
		69   Biological regulation	
		66   Metabolic process	
		43   Multicellular organismal process	
		39   Developmental process	
		<b>GOTERM_BP_FAT (143)</b>	<b>Notch signalling pathway_list (3):</b> CTBP2; MAML3; RBPJ
		39   Regulation of transcription	
		34   Transcription	
		32   Regulation of RNA metabolic process	
<b>KEGG_PATHWAY (143)</b>			
3   Notch signalling pathway			
miR-194	207	<b>GOTERM_BP_1 (198)</b>	<b>Developmental process_list (67) :</b> AGPAT6; ADAM17; BRSK2; HNF1B; KLF7; LHX6; MEIS2; PRDM16; RYBP; ARHGAP24; SMURF1; SOX11; SOX5; SOX6; SP3; TAF4; TEAD1; ZIC1; ACVR2B; AP3D1; APP; CDH11; CDH2; CADM1; CHD6; COL4ABP; CLASP1; DYRK1A; DMD; EFNB2; EVX2; FLNB; FOXP2; FXR1; IGF1R; LRRC4C; MITF; MAP2; MEF2C; NRP1; NTRK3; NRIP1; NR2F2; ONECUT2; OPCML; OTP; PAX5; PPARGC1A; PAFAH1B1; PDGFA; QKI; RFX3; SALL1; SALL4; SGCE; SEMA6A; ZFAND5; SOBP; SHH; SRI; SPRED1; SS18; TLL1; TCF7L2; TRPS1; ERBB4; ZEB1
		144   Cellular process	
		116   Biological regulation	
		106   Metabolic process	
		71   Multicellular organismal process	
		67   Developmental process	
		<b>GOTERM_BP_FAT (198)</b>	<b>TGF-beta signalling pathway_list (5):</b> ROCK2; SMURF1; SMURF2; ACVR2B; ZFYVE9
		63   Regulation of transcription	
		50   Regulation of RNA metabolic process	
46   Transcription			

		<b>KEGG_PATHWAY (198)</b>	
		5   TGF-beta signalling pathway	
		<b>UP_TISSUE (198)</b>	<b>Embryo_list (9)</b> PHF21A; SLTM; CADM1; ENOX1; JMJD1C; QKI; RHOQ; TRPS1; UBE2W
		112   Brain	
		23   Foetal brain	
		9   Embryo	
miR-203a	341	<b>GOTERM_BP_1 (324)</b>	<b>Developmental process_list (94) :</b> AGPAT6; ADAM23; AFF4; ATP5G3; BCL11B; BMI1; CTBP2; CITED2; FAT1; FAT3; FYN; GLI3; LMO4; LIMD1. NDRG3; RAN; RPGRIP1L; RASGRF1; SIX3; TIAM1; UGCG; ACVR1; ACVR2A; ACVR2B; APC; APP; BMPR2; CDH2; CDH6; CDK13; COL12A1; CNTN4; CUL1 CYR61; DICER1; DLX5; DYRK1A; EBF1; EBF3; EN2; EXT1; FGF10; FGF16; FOXP1; FOXP2; GPHN; GRHL3; HAND2; HOXA13; HIF1A; INHBA; ID4; LAMC1; MET; MAP3K1; MBNL1; MBNL3; MSH6; MEF2C; NRG2; NR5A2; NCL; PTCH1; PTEN; PPAP2B; PDGFRA; PROX1; PTP4A1; ROR2; RARB; ROBO2; RUNX1T1; RUNX2; SIK1; SGCD; SEMA3A; SEMA5A; SEMA6A; SOBP; SIM1; SNAI2; STRBP; SS18L1; TOP1; TCF12; TCF7L2; TGFB2; TRPS1; ULK2; MAFB; MAFK; SKI; VCAN ZMIZ1
		238   Cellular process	
		189   Biological regulation	
		177   Metabolic process	
		104   Multicellular organismal process	
		94   Developmental process	
		<b>GOTERM_BP_FAT (324)</b>	
		103   Regulation of transcription	
		79   Transcription	<b>MAPK signalling pathway_list (11) :</b> RAP1A; RASGRF1; FGF10; FGF16; MAPK10; MAP3K1; MEF2C; NLK; PDGFRA; PPM1A; TGFB2
		71   Regulation of RNA metabolic process	
<b>KEGG_PATHWAY (324)</b>			
11   MAPK signalling pathway	<b>TGF-beta signalling pathway_list (8):</b> ACVR1; ACVR2A; ACVR2B; BMPR2; CUL1; INHBA; ID4; TGFB2		
8   TGF-beta signalling pathway			
8   Wnt signalling pathway	<b>Wnt signalling pathway_list (8):</b> CTBP2; APC; CAMK2D; CU1; MAPK10; NLK; CACYBP; TCF7L2		
miR-206	72	<b>GOTERM_BP_1 (69)</b>	<b>Developmental process_list (23):</b> CTBP2; RORA; RASA1; SPEG; TIMP3; BICD1; CRIM1; DLG1; FZD7; HDAC4; IGF1; JARID2; NR4A2; <u>PAX3</u> ; <u>PAX7</u> ; PLEKHA1; SEMA6D; SCML2; SRI; THBS1; TGFB3; TRPS1; VEGFA
		53   Cellular process	
		42   Biological regulation	
		23   Developmental process	
		<b>GOTERM_BP_FAT (69)</b>	
		19   Regulation of transcription	<b>mTOR signalling pathway_list (3):</b> EIF4E; IGF1; VEGFA
16   Regulation of RNA metabolic process			

		15	Transcription	
		<b>KEGG_PATHWAY (69)</b>		
			mTOR signalling pathway	
		<b>UP_TISSUE (69)</b>		
		38	Brain	
		16	Epithelium	
		7	Foetal brain	
miR-223	138	<b>GOTERM_BP_1 (130)</b>		<b>Developmental process_list (41):</b> ERC1; FAT1; LMO2; LHX8; POU2F1; POU6F2; RAB3IP; RORB; RASA1; SRPK2; SOX11; SOX6; SP3; ACO1; ACVF2A; CRIM1; EBF3; EGLN1; FOXO3; FOXP1; GPM6B; HLF; ITPKB; MAP1B; MSI2; MBNL1; MYH10; NFASC; NFIB; <u>PAX6</u> ; RBPJ; RPS6KB1; SIAH1; SHOX2; TSH23; TOP2B; TSC1; TWIST1; ULK2; MAFB; ZEB1
		100	Cellular process	
		84	Biological regulation	
		76	Metabolic process	
		43	Multicellular organismal process	
		41	Developmental process	
		<b>GOTERM_BP_FAT (130)</b>		<b>MAPK signalling pathway_list (6):</b> ELK4; RASA1; TAOK3; FGFR2; PRKACB; RRAS2
		40	Regulation of transcription	
		36	Transcription	
		32	Regulation of RNA metabolic process	<b>Insulin signalling pathway_list (4):</b> CBLB; PRKACB; RPS6KB1; TSC1
		<b>KEGG_PATHWAY (130)</b>		
		6	MAPK signalling pathway	
		4	Insulin signalling pathway	<b>mTOR signalling pathway_list (3):</b> RPS6KB1; TSC1; ULK2
		3	mTOR signalling pathway	
<b>UP_TISSUE (130)</b>				
		78	Brain	
		15	Eye	
miR-499-5p	173	<b>GOTERM_BP_1 (166)</b>		
		111	Cellular process	
		82	Biological regulation	

	79	Metabolic process	<b>Developmental process_list (45):</b> BCL11B; FAT1; FYN; H2AFZ; KLF7; MEIS1; NHS; NOTCH1; FOXG1; SOX5; SOX6; ABI1; ATG7; CDH4; CTNND2; EBF1; ENAH; ESRRG; EYA4; FGF9; FNDC3A; FOXP2; FZD8; HOXD3; JPH1; MYH10; NTRK3; NRIP1; POGZ; WTAP; YBX1; PPP3CB; EVL; LRP6; ARID5B; SORL1; ZFAND5; SATB1; STRBP; SPRED2; TCF12; TAC2; ZEB2; ZIC3
	50	Multicellular organismal process	
	45	Developmental process	
	<b>GOTERM_BP_FAT (166)</b>		
	46	Regulation of transcription	<b>Wnt signalling pathway_list (5):</b> CSNK2A1; FZD8; NLK; PPP3CB; LRP6
	37	Transcription	
	34	Regulation of RNA metabolic process	
	<b>KEGG_PATHWAY (166)</b>		
	5	Wnt signalling pathway	<b>MAPK signalling pathway_list (5):</b> FGF9; MAP3K4; NLK; PPP3CB; SOS2
	5	MAPK signalling pathway	

## APPENDIX III

**List 1: Sequences of coding and 3'UTR fragments cloned and used to generate RNA oligonucleotides in order to perform RNA *in situ* hybridisation.** Primer sequences have been underlined in the sequences cloned from coding regions.

**>Gga-Eya1-ISH [970 bp; ENSGALT00000025181.4] (from pGEMT(Eya1.2-3UTR) construct)**

CAACTACCTGCAAAGCTGCGAAGCACC AATGTAGCTGCGAAGCAACACCATTCTTTATTG  
GAGGAATCATGTTGCTGCTAGAATTAATGTTTGTGGATTGCTTTATTACCGACAAAGAG  
GTTCTTATCCTGCCAAAGACTTTTCACATGTGCTTTACTTTACCCAGTATGAGTACTGGAAT  
TCACACAGTACTCACATACATACACGAAGCACATTGTGCATATCTTTTAAAAACCTAGACT  
CTAGCCCTTTCTTTCTCAAGCAAAGTAGGCAGAAATGGAGGTTGGTTGTCTTTTTTTTTT  
TTCCCCATCTCTTTTCTATCTTTAGTCACTTGCAGACTGAAAAAACAGTTTATCTGGGC  
CTTATGTACAAAAAGCGTGTGTGTCCACAATTGTGTACAGAATTTTCTTCATTAATT  
TTGTTTTAAATTAATAAAAATTGATTTGTGAACATATTAAGCAACTCTCTTGTATGCAAGT  
CTATGATATTTCTCCTCTCTAGGTCCTGGAATGAGGAGACATCTCTCATCTTTTACCCTG  
CTAAGACACAAGTCGTTACCTGTTTGTGCACACCTGAGATGTTAAATAGGGTAAAAATA  
TTTGCTTCTGGTTTACAAAAGGCTAAGTTTGTATGGGCAGCAGGAGGAGGAGCTAAAGGTC  
AGAAACATGTAGTAGGACTTGGACTCGGTGCATGCTGTCTGGAAACGTGCCGAAGTGAC  
TGGCAAAAAGCAAGTGCCTAAATGTTAGGTGCTGTGTTGCCGTGAAGGTCAGCCAGAGC  
ACACATAATGTGAGTGGATTAGCATGATCATTGACACTCAGGGCTATAGCAA  
TGTCCTACCCTCTCTCTGAGGATCTACAAGCCTGAACATCTATGGGAATCAAGAGACTT  
CAAAGGAAGCTGGTGGACAATGGTGCCCACTTTCTTGTCTACAGAGATCTGGGTTTTGTT  
GCTACATTGA

**>Gga-Eya2-ISH [802 bp; NM\_204915.1]**

ACATTTATGCCACGGCTCCAGATGACATAGAAGGCAACAGTAAAGCAGCACCACAGTGTG  
CTCTCCATCTTTACTCAACAAATGACAGTCCAGTCTTTCCCAACAATGGCAACGTATGGA  
CAGACTCAGTACAGTGCAGGAATCCAACAGGCTGCTGCATACACTGCCCTACCCTCCTCCA  
GCGCAGCCCTACGGCATACTTCCCTACAGCATCAAAACAGAGGACAGCTTGAGCCATTCC  
CCAGGACAGAGTGGGTTTCTTAGTTATGGATCCAGTTTTCAGTACCCCGACTGCTGGACAA  
GCACCGTATACCTACCAGATGCATGGCACAACAGGGATTTACCAGGGAGCCAATGGCCTG  
ACAAATCTGCTGGATTCAAGTGTGTGCATCAGGAATATTCATCATACCCAAGCTTTCCCT  
CAAAGCCAGTACTCACAGTATTACAGTTCCTCCTACAACCTCCTACATGTCCACAAC  
AGCATCAGCCCTTCAGCCATCCCAACCTCCACTTATTCTCTGCAGGAGTCTTCTCACAAC  
ATCACCAGTCAGAGCACAGAATCGCTGTCTGGAGAATATGGAACAACACCAGCAAAAGAT  
ATAGAAAACAGACAGACATCACAGAGGGTCCGGATGGCAAGGTACGAGCCCGATCAAAAAGA  
AGCAACGATCCTTCCCCACTGCTGACAGTGAGATTGAGCGTGTATTTGTGTGGGATTTG  
GATGAGACGATAATTAATTTTTCACTCCTTACTCACGGGAACCTTTGCATCCAGATATGG  
AAGGACACAACAACGCTCTGTGC

**>Gga-Eya3-ISH [476 bp; XM\_417715.3]**

CCCAGACCACCCAAACCTACGGACTACCTCCTTTTGGCATCAACTACCAATGCCAGTCCA  
GTCTCTACATCCTCAACTGTTGTCAATATTTCCACATCAGCAGTAGCCAGCATCTCACAG  
GAATATCCTACGTACACAATCCTTGGCCAGAGTCAAGTACCAGACGTGTACCCAAGTTCT  
GGCTTTGGAGTCATAACACCAGCAGACGCAACGCGGAGAGCACTGCATTAGCAACAGCT  
ACGTATCCATCTGAAAAACCAACGCCATGGTGCTACACGGACGGTGCAGAGACATTCC  
TCTGGAGATGCATCCCAAGTCCCTCATTTGTCAAGAGCAACAGCAAGTAAAGAGTCAGAT  
GAACAGGCAAGAAAAAATATCCCTGGGAAGAACAGAGGGAAAAGGAAAGCAGACACCTCT  
TCTTCAACAGACAGTGAACCTGGAGCGAGTGTCTCTGGGACCTGGATGAGACCAT

**>Gga-Eya4-ISH [873 bp; ENSGALT0000022662.4]**

CTCCG AACGCCAGGTCTATGGAAATGCAAGATCTAGCAAGTCTCATAATCTTGTGGAA  
GCAGCGATGCGCCGGTAGCTCTAAACTGGATAAGCCTAATCTCAGTAGTACATCAGTTA  
CAACAAATGGAACAGGAGTGTCTCTTCTTGCAGTCAAAACAGAGCCCATGAACAGCAATG  
AAACAACAACAACACTGGAGATGGATCGCTTGACACTTTTACTGGTAAATAACAAGTAGT  
GGCTATAGCCCAAGATCAGCGCATCAGTACTCTCCGCAGATATATCCCTCCAAGCCCTAT  
CCACACATTTCTTCTACACCAGCAGCTCAAACAATGTCTGCCTATGCCGGACAAAACCCAG  
TATTCAGGAATGCAGCAACCAGCAGTCTATACAGCCTACTCACAGACAGGACAGCCATAC  
AGCCTACTTACTTACGATTTGGGTGTAATGTTGCCAGGCATCAAGACGGAAGTGGGCTC  
TCGCAGACACAGTCAACACTGCAGAGTGGGTGCCTCAGTTACAGTCCAGGGTTTTCCACC  
CCACAGCCAGGCCAAACACCGTATTTCTATCAGATGCCAGGTTCTAGTTTTACACCATCA  
TCTACTATTTTATGCAAAACAATTTCTGTTTCAAATTTCTACGAACCTCAGTAGTTCACAACAG  
GATTATCCTTTCATACACAGCTTTTGGCCAAAACAGTATGCACAGTATTACTCGGCATCA  
ACATATGGTGCATATATGACCTCAAACAACACAGCCGATGGCACTTCATCATCATCA  
ACCTACCAGTTACAGGAATCTCTCCCTGGCCTGACTAGTCAACCAGGTACAGATCTACAT  
TCAGGGCAGTTTTGACACAGTACAGAGCCCTCC

**>Gga-Six1-ISH [597 bp; NM\_001044685.1] (from pGEMT(Six1-3'UTR) construct)**

CGGCTGTTTCGGGTACAATAGAAAAGAAACCTGAAAGCACACGCGAGGTCGGTAAACACAGA  
CTCGAGACAGACAGCTGCGGGCTCAC TGCGAGGCAGAGCTGGGGGTTGGGGGAGGCAGAC  
AGAGAGAGACGCTCCTCGGAAGGAAAGGGCCATCCGCATTTATTATCGCTGTTGTTACT  
GTCCTAAAGTGCAGCCGCATCGCTGCGTGTCCCGGCCGGGAGCAGCACCCGCCGCTTAA  
GACAGAGGCACCTCAGACTGCCGCTTCCATAATTATCTATTTTTTTTTCCCTCTTTTTTTT  
TTTTTTTTTTAATTTTAACTACATCGAAATCCTTCCGGTGAAGCTAAATATGAAAAAAA  
AGAAATGTAAACGAAAAAGTTCCCTCCTCCTCCCGCTGCCGCAAAGGGCCCCCGCTC  
CCTCTGCCTCTGGGCCCAACAGCCGGGATTTATGGGGCGCAATGGGATCCGGGGTTCCG  
GGTTGGGTTTTTTTTCCCTCACCTCTCCCTTCTATTTCTATTTCTGCTGCTGTAAT  
CCCCTCACTTGAACCTGAAAGGCCGAGGGCTTTCTGCCGCTCTCTGCATTTGGCTA

**>Gga-Six4-ISH [538 bp; XM\_003641442.2]**

TGGAGAGCCACAACCTTCGACTCGTCCAACCACCCGCTGCTGCAGGAGCTGTGGTACAAAG  
CTCGTACACCGAGGCGGAGCGAGCCCGGGCAAACCCCTGGGGGCGGTGGACAAGTACC  
GGCTGCGGAGGAAATACCCCTGCCCGCACCATCTGGGACGGCGAGGAGACGGTGTACT  
GCTTCAAGGAGAAGTCCCGCAACGCCCTCAAGGAGCTTACAAGCAGAACCCTACCCCT  
CGCCGGCCGAGAAGCGCAACCTGGCCAAGATCACCGGGCTGTCCCTCACGCAGGTCAGCA  
ACAGGTTCAAGAACCGCCCGCAGCGGGACCGCAACCCGTCGAGACGCAGTCCAAAAGCG  
AATCAGATGGCAACCCAGCACGGAGGATGAGTCCAGTAAAGGGCAGGAGGATTTATCCC  
CACATCCGCTCTCCAGCTCGTCCGACGGAGTCAACAGCCTCAGCCTTCCAGCCACATGG  
AGCCCGTCTACATGCAGCAGCTTGGAACAATAAAATAGCCTTGAGCTCATCTGGTGT

**>Gga-Dach1-ISH [1084 bp; ENSGALT0000027373.3] (from pGEMT(Dach1.2-3'UTR) construct)**

TGCCTCATTTGTTGGCTTGGTACATAAATGGAAATGTTAAGGTTTAAAGGGGAACCAATT  
TATAAGCTGGATGTTTAGAAAGTATCTTGCTAAAACCAAGTGAATATTTACAGACCATGA  
GATGTTAACGTAAGTTGAATTTTTTGCCCTCTTTAGTTATACAGGTTTTGGGTTGGTAT  
TTTGTTTTTTATTTCAAATTTTGTTTTGTTTTGTTTTGTTTTTCATAGATGATGAAGAAA  
GTTGTGCTCATGTTATTTGTTTATATGCTTTTGTAAATCTTAAAGATATTAATGTCTAGTTG  
TTCTATATTATAACCATATTTGCGCTCTATGCAAGCCCTTGAACAGAACATACTCATCT  
TCATGTAGGACCTATGAAAATTTGCTATTTTTATCTATATATTTAAAGTTTTCTAAAAAT  
GAAAAAAGGTTATTACGAATTTTGTGTACAAAATCTGTACAAAATCTGTTTTTACAT  
CATAATGCAAGAATTTGAAATTTTTCTATGGTAGCCTAGTTATTTGAGCCTGGTTTCAAT  
GTGAGAACCACGTTTACTGTTATTGTATTTAATTTCTTTTCTTTTCAACAATCTGCTAA  
TAAACTGTCTGAAATCTCCCTGTGACTTCATTTACAGTTTCTTTTAAATTTTCTG  
AAATGTGTAACATCAGAGGAATATTTACTTTCTAATGGGAGGCATCTGAAAACAACACAA  
GTCAGCTCTTTGTAATGTGAGGAGAACAATGCTGAATGATTTTATTTAACATGCAACTGC  
TTCTATCCCTAATATGAATTAAGTGTGGTGA AAAACATCATAAAAGCACACTTTGTGGTTA  
TTGTTTAAACAAGCAGATTTCCATATTTTTTTTCTTCCAGCTAAGCAAACCTGCCCCAA  
TCTACATGATTTTATTTATGTACATTTTCTCAGTTTATGAGCTGTTATTTGTACCTTTTT  
TATATATTGTGATATTTCCATGATTTCTATTTTCAGAAAGCTTTGTGCTGAATAATGTAAA  
GTGGACACATTTGATGGAACAAAACATATAATTTCCCTAGCTACAGAGAGGGTAATCCAGT  
TGTT

**List 2: Sequences of 3'UTR fragments cloned in order to perform luciferase reporter assays.** Predicted miRNA sites are indicated in bold, and the primers, containing BglIII (grey background) and NheI (dark background) sites, have been underlined in the sequences.

**>Gga-Eya1.1-UTR [1291 bp; ENSGALT00000025181.4]**

AGATCTCAGCTCAGCAGCACTTTGAAACCCAGAGCCTCCTTCTGCCCATGGACGGTACG  
 CCTGTGCTTGTGTGTCAGCATTGGACTACAGAACTTTGTGATTTCAACATGTTGACGTACA  
 GCTGCAATTGGTCTTAACCCTTGCCCTTTCAGTAAACGGAGGAGCATGTCTTTTTCTTCA  
 GAACAGCTGTTGACTCTGGT**ACTGCGAG**TCCAACAAAAATAAGCCATGCGAATGTTTTAA **miR-27/128 site**  
 (1)  
 CAGTCATCTTTACTATATTTGCTACCAAAGAAATGGAGAAGGAAGAAAAGGAAATAAA  
 GGTTCGTAC**CTGTGA**AAAAAGGACCTGCAAGTGCTTTGTAGTTTTTAACTCTTCAATGT **miR-27/128 site**  
 (2)  
 GACACACGCCGTTTCTTCAACACAGCAAACCTGATTGCACAATGAAGACTGAGATTTTTTC  
 AAAATACCAGTGGAGTAATTTCTTCTAAAGAAGGTTTACTTTTGGTTTCTCATACCCAG  
 GGTACTCTGTACATCTTACTTATTTATGAACAGACTGTATTTAACATCATATAACTGA  
 GGATATGTGTAATAGGAATAAAGGCTATTATAAGCCTTTGCCTTACGATACAGCAACTAC  
 TTTTGTATTTAGCACATTACAGAGTAGTTTAAATATGTCTAATTTAACTAATAGGTAC  
 ATCACTGAGACAATCATGTACAGGAAGAAATTTTGTGTAAATTTGTAATAATGAATGATT  
 CTTTTACATATCGTTAAGGTAAATGCTATTGAAAGATAGTAATGCCTTGTGGTGAAGAA  
 TGAGGCTACGTGTGCACAAGATGTGCAGTGCCTTGTCAACACATTGGATATAAATATGTA  
 GATAATGGATTTTTTTAGATAAATTTGTCA**AGACC**AAAGCATGGATGTCAAGTGTCAAT **miR-133a site**  
 ATGAATGGGTTTTGTCTTTTTCAGCTATTTCTCTGCCTTTTTCTCTCATCTGTTCT  
 GATTATGAAAAGATTTCTTTCCCCCATCAAGGAAATACAGATGAAACACAACCGAAGA  
 GGAGTACTTTGCTGCTCTCTGTTGCTCTTAAACACTTTTTTAGAGTATTGACAAATGA  
 ATTAGCAGATTCCATAAAGAAATAGAGAAAAACAATATATTTTAAAGACATGACTTAAACT  
 GGAATCTAGGTACCGACTAGTTTATGTCTCATCCAAAAGTAAAGGAAATTTATTGTGATCG  
 TTATTTTAGAAATCAGAAATGGTAATATTTGGAGAAATTTAGCAAGATACTCGTATAC  
 ATTTAAGTTTTTTTAAATCCTCTTCCACCAGGAGTCAG**GCTAGC**

**>Gga-Eya1.2-UTR [970 bp; ENSGALT00000025181.4]**

AGATCTCAACTACCTGCAAAGCTGCGAAGCACCAATGTAGCTGCGAAGCAACACCATTCT  
 TTATTGGAGGAATCATGTTGCTGCTAGAATTAATGTTTGTGGATTGCTTTATTACCGAC  
 AAAGAGGTTCTTATCCTGCCAAGACTTTCACATGTGCTTTACTTTACCCAGTATGAGTAC  
 TGGAATTCACACAGTACTCACATACATACACGAAGCACATTGTGCATATCTTTTAAACC  
 TAGACTCTAGCCCTTCTTTCTTCAAGCAAAGTAGGCAGAAATGGAGGTTGGTTGTCTTT  
 TTTTTTTCCCATCTCTTTCTATCTTTAGTCACTTGACAGCTGAAAAAACAGTTTAT  
 CTGGGCTTATTGTACAAAAAGCGTGTGTGTCCACAATTGTGTACAGAATTTTTCTTCA  
 TTAATTTGTTTTAAATTAATAAAATGATTTGTGAACATATTAAGCAACTCTCTTGAT  
 GCAAGTCTATGATATTTCTCTCTCTAGGTCTGGAATGAGGAGACATCTCTCATTTTTC  
 ACCGTCCTAAGACACAAGTCGTTCCACCTGTTTGTGCACACCTGAGATGTTAAATAGGGTA  
 AAAATATTTGCTTCTGGTTTACAAAAGGCTAAGTTTGTATGGGCAGCAGGAGGAGGCTA  
 AAGGTCAGAAACATGTAGTAGGACTTGACTCGGTGCATGCTGTCTGGAACGTCGCCGA  
 AGTGACTGGCAAAAAGCAAGTGCCATAATGTTAGGTGCTGTGTTGCCGTGAAGGTCAGC  
 CAGAGCACACATAATGTGAGTGGATTAGCATGATCATTACAGACATTGCACTCAGGGCTAT  
 AGCAATGTCTACCCTCTCTCTGAGGATCTACAAGCCTGAACATCTATGGGAATTCAG  
 AGACTTCAAAGGAAGCTGGTGGACAATGGTGCCACTTTCTTGTCTACAGAGATCTGGGT  
 TTTGTTGCTACATTGA**GCTAGC**

**>Gga-Eya3-3UTR [487 bp; ENSGALT0000001127.4]**

AGATCTGTAGTCTCCAGAGGGAGGGGTAACACAGCTGAGAAGGCTCTTACAGATACTTC  
 TGCTTTTCTATTAGTTTATAGAACCAAGTAAACAGAAAACCTTATTTTTATAGA  
 AAAATACTGATGGCAGAGCTGAACCTCCCTTGTGTTTGCAGCAAAAAGAGCTATTGTT  
 TGGTTGGTTTTTTTTTCCATGGGAAATATTAATGAAAATATCAAAAATACCTCTACTGCT  
 GTGAAAATGTGTCTCTCTCTCTGGGTGTTCAAAAAGCAGTTAATTTATTATGATAT  
 CCTTACATTTCTTCAACGTGGGATTTATCCATTCTGGGATAAGTGGTTCCCTGTGATG  
 AGGGAGGTGTTGCTGTTGGTTGTTCCCTATTGGCTTCTAGGCAGTGCCTGTGCAATGCATG  
 TGACTGAGCCCTGTTAATGAGC**AGGGGATCA**TTAGATTCCAGCTGTCCATGACTCAGACC **miR-133 site**  
 ACAGATGCCACAG**GCTAGC**

>Gga-Eya4-UTR [1043 bp; ENSGALT00000022662.4]

AGATCTTGTCTAAAGTTGGCGATCCTTTTTTTTTATATATATATATTTCAAGTACACTGAA  
TTTTTATGTGTGATTCAATGCCTCTGGCTTTACACATATGAATTTGCTTTAAGAAGGGAAG  
AAATATTTGGAATTAATAAATCCAAACTGAAGAATTCAGATTGCTGAATGGAGTTAAAAAC  
ATTAGTGCTACATAAGGAAGCTCTATGGTCTTATATATATGCAACGTTTTTAAATGGATTAA  
AACTGTGGAGGTTGCTGGTACACACCCGAATGAGCCCTGACAGGAGTGAACAAAGACTCG  
AACTGGCAAAGCACCAACACGCGTTTTTTAACCACAAAGTGGTGTCAACAACATTCCCT  
CAAAATGGGATATATTTCTCAGCACTGAGGTTTGAACCAGACTTTAGCCTACCTAACCCAG  
AAAATCTGAATTGGAATGCACTCAGACTGTATAATGACAATCCTGTCTAGACCTGTAATT  
TGTGTAATTTATGATGAAAATAACTACTGTGACTTTATTAGCAGCTGATTTTGAAGT  
GGATGCAATTTTTCTTTCTTTTTGGGGGGTGGGGGAGGGGAGAGGGTTATATAATATTA  
TCTCTTTTATAAGTTTGGCAAACAGAAATGTCATAATGATGTGTGCTTAAGGAGAA  
GACTGTGTTTGTGTGTTATAATGTAACCTTTGGTTAAAAACTATGTAGATAAACAAACAAA  
AAAAGCCTTTGTGATAATTTTTGACATGACCAAATTTGAAATTCAGAGAAATCAAAGAGA  
AGGGCTGCACCAAAGCATTTAAGTTTTTGTGTCAGTAAAAAAAAAAAAATAATAATAAG  
GAAAGTTGTGTTTTTATTTGGATTCTCAATAATCCACTGACTGAGGAAAGTTGAGAGT  
ATGTTGGGAGGTTGATAGTGGGGACCGTACCTTGAAAGCTAAAAAAGAGTGTTTAACA  
CCTTCAACATCGTTTTCTGATTCTCTCATGAGAGAACTAGGAGCCATTTTCATGAATTAC  
TTAGTCCACTGAAAGCCATGCAGTGAGTGCTAGC

miR-499 site

miR-1a/206 site

miR-27/128 site

miR-133 site

>Gga-Six1-UTR [594 bp; NM\_001044685.1]

AGATCTTAGCCAAATGCAGAGAGCGGCAGAAAGCCCTCGGCCTTTTCGAGTTCAAGTGAGC  
GGGAATTACAGCAGCAGAAATAGAAATAGAAAGGGGAGAGGGTGAGGGGAAAAAAAAACCCA  
ACCCGAACCCCGGATCCCATTGCGCCCATAAATCCCGCTGTTCGCGCCAGAGGCCAG  
AGGGAGCGCGGGCCCTTTGCGGCAGCGGGAGGGAGGGGAACCTTTTCGTTTACATT  
TCTTTTTTTTTCATATTTAGCTTCACCGGAAGGATTTTCGATGTAGTTTTAAAATTA  
AAAAAAAAAGAGGGGAAAAAAAAATAGATAATTAGGAAGCGGCAGTCTGAGTGCCTCTGTC  
TTAACCGCGCGGTGCTGCTCCCGCGCCGGGACACGCAGCGATGCGGCTGCACTTTAGGAC  
AGTAACAACAGCGATAATAAATGCGGATGGCCCTTTCTTCCGAGGAGCGTCTCTCTCT  
GCTGCTCCCCCAACCCAGCTCTGCCTCGCAGTGAGCCCGCAGCTGTCTGTCTCGAG  
TCTGTGTTACCGACCTCGCGTGTGCTTTCAGGTTTTCTTTCTATTGTACCCGAACAGCCG  
GCTAGC

miR-1a/206 site

>Gga-Six4.1-UTR [1566 bp; XM\_003641442.2]

AGATCTGGAGAAAGGAAACGCCAGGGGAAGTACCAAGCAAATGGGTACAGGTTTGGCTAC  
TGCTACTGCGCTGATAATGTAACATAGATCTCTGTAGTGCAACGTACTATCAACAGTAAG  
TATACACTCCTGACAGTAGATAGAAAGCAAAAATCTCGCTCTGGAAAGGTAGAAACCATG  
CAGTTGTTTTCGCTGTTACTTCTAGTCCGTGGCATTGAAAAGTTTTTTAAGAACTCTAT  
AGACGATACCTCACAGGCTACTTTTTGTGTTGTTGGAGGTTAAGAAACAAAACGAAACCT  
AATGTTGTTACTGGTGTGTTTGTGTTTTTTTTTTCCAGCAGACTTGCCCTTCATACC  
AAGAATCTCGCATCAACATTGACAGGTTTTCAAAGTGTGCAGATAGCATTTTGTGAGTG  
CACCAGAAAAGCATCATCTGAACTCTAACTATGCAACTTGAAGCTTCTCTCCAGTA  
TGAATGTAATAATTTGTTTCAGTTCTTATAAAGGAGACTTTTTTTTTTTTTTAATTAGTA  
CGATGGATTTTATTTATGGTTTATGTTCCCTTTCACTTTAGACACAATGCATTGGAAAGG  
AAAACGTTTTCTGGCTGTTAATTTATTTTTATTGTTTGGAGCAGGAAGTAAACGTAGTTA  
TTTTCTGACAAGGTTTTACTTTTTTTGTAGGCTTTCAAGCAATAGACTGTAAAAGTAAATG  
TTAATCACACTGAAACTAGAAAATGAGCTCCAGTCTCACAAACCCAGGAGGTCCTGTGTC  
TCTACCCCTCTGTCTTTGCGTTTCAGGGTTTTCTGTTGTTTTGTTAGTAATTTGCAAA  
TTAGTTGAACCTTTTTCTTGGTGTATTTTTACGGAATCCATGTGCCTTTCTTCTCTAG  
TGGNCTTATCTTTGTTACTAAAGCACAGTGGCAGGAGGAATAAAAACACTGACTTGCAAA  
TCTCCGTCCTCAGGGCTGAATCCTGATGCCTTACTTTTGAAGTAGCCCTGTGACCCCG  
GAGTCCCTACTTGCACGAGTAAGAAGAGTTGAGTCTTTTGAAGAGCTGAGTCTTTTGGT  
CACGCTGGGTGACACTGAGTGGTGGTGCAGGGACCTGGAGGCAGAAGGATTTGTGCCGAGG  
GACCACTGACAGAGCAGTGCAGAGCAGTGCAGGAGGGAAGATCTTGGCTCTGGCCTCGGT  
CTGATGAGTGTGGAAGGTTGGTGGGCAGGCCAGGGAAGGAGTCCCATGTTTCCCCACAG  
CCGTTGGCCAGCCCTTGTGGGAGGCTTCAGCTGCTGTGAGCAAGCAGGCCCCCA  
TCCAGGATGCACCTAAGCATGGATTTATAAATCCTTCCAGAATAAGCAGCGCGGGGCC  
AGTGGTGGAACTGCAGCTGCCAGAAGATCCATGGAGCAGCCAGCTCCCTGCCTGCCT  
GGGGATGGGTGTCTCTCCCTTACCCTCTGCTCTCTCTCAGGGACAGCATAGGTAGGC  
AGGCTGTGGGCAGGTGAAAAGACCTTACCTTTCTCCAGAGAACTTTGTAGCGTGGAA  
AAAGCACATCCTGCTAGC

miR-133a site

>Gga-Six4.2-UTR [1594 bp; XM\_003641442.2]

AGATCTATGGGTGTCCTCTCCCTTCACCCCTCTGCCCTCCTTCAGGGACAGCATAGGTAG
GCAGGCTGTGGGCAGGTGGAAGACCTTACCTTTCCCTCCAGAGAACTTTTGTAGCGTG
GAAAAGCACATCCTTAAGCTTTAAAATAATAATAATGCAGATGTATAAAAATGACAT
GCTGTTTATCTTTGTGTTTCTGTTTTGTTTTATTCTTATTCTAATGTGAATATATACA
GTAGCTAAGGCATGTATTCATCCTCGTCAGGGGCATGGGCTGAAGCCTTTGGAAGACAAT
GGGAAATTTTTCCATAACTTCTGTGGGCTCCACACCAGGCCAACCCAAAATCACTTGG
GCAAAAAGAGTGGACAGGAAAGTGAATAGAGAGAGCACATCCCAGTCTGTTTCTGTCTTG
GTTATTACCTGGTTCGGTGGAAAGCCTTAAGGAATGAAGGTTTCTCCTGTATGGCACAG
ACATCCTTGCTGTTGTTTTGCACTATATTTGAATGCAGCATGTTTTGCAAGGCAGTGTGTG
AAACCTCAGAGCTGAGCTCTGCTGGTAAATGCTGTTTACACTGCAAGCTGGAGTTCTGAA
TGTTGACGGGAAAAGTGCAGGGGAAGGGAAGAGCAGGCGTGCAGGCTTGCACCTTGGCAG
CTGCAAGGAGTCCCTGGTGCTTTGCAACAGGAGCAGCCCTGAGCAGTTGGCTGTCCCC
ATGTGCACCCGAACAAAAGGGTGTTTAAGGCAAGTGCCTTAATGTGGAGATTATCGCAA
GCAGAGCATGGAGCTATATGTTCCCTGGCACAACCTTCTTGGCATTCGGGGGATCTCCA
TGTGAGCAGGGAATCGGATGGTGCAGAACTCTGCTCCAGTGGTGTATCTGAGTGCCTTGC
TGGGCTCCAGCATGAACAGTTTGTGTTTTCCGCTCTCCTTCACCCCTCAAAAAGAGATGAT
GATTATGATAATAGTGTCTGTATCTGATCTGCATGAATTGCTCAACTTCTTGGAAAAGA
AGGATCCTGTATTGTTGATTAATGGAGAACGTATGCACAGAGCAGGAGCTCTGTAGCTC
TGGGGGAGTTGTTCCGGGCTTAGTGTGTTGGTTGTGAAGCTCAGACTGTTTCTAGAGGCA
AGCTAGAGAATATTTTCCAGCCCTCCGTCATTTGTTGAACCTCATCAGGTTCCATAACTGA
AAAGACTGTGGTTTTTATGCCAGTACATGTTACTTTTTTCTCTCCTCCCTTGCATTTGTG
GTGTTACTGAGGCTCTGAATGCCATGGCTTCGCAGCACAGTTGATGTGAAAACCTGTGTAT
GCTGTTACAGTGGTGCATCTCGGGGTGCCTGGTACTTCTAATAGGAGACCTCCCGTTC
TTCTGTGGGACGGAGGATGTATATAAAAGCAGTGGTTGGGACAAAAGTAAGATTAATCT
ATTCATCTTCCAGGTATCAAGTAACCTATTTTCGTTATTATGTTTTACTCTTTCTTCTCC
CACTGAAAATAAGTGTCAATCCCTCCAGGTGATGTAGTTAGAACCTCATATGTTTTTGC
AGGTACATATGCACAATATAAAGACTCGTTGCACGGTACAGCTAGC

>Gga-Six4.3-UTR [1448 bp; XM\_003641442.2]

AGATCTACCTCCCGTTCTTTTCGTGGGAGCGGAGGATGTATATAAAAGCAGTGGTTGGGA
CAAAGTAAGATTAATCTATTTCATCTTCCAGGTATCAAGTAACCTATTTTCGTTATTATGTT
TTACTCTTTCTTTCTCCCACTGAAAATAAGTGTCAATCCCTCCAGATGATGTAGTTAGAA
CCTCATATGTTTTTTGCAGGTACATATGCACAATATAAAGACTCGTTGCACGGTACAGTG
TTTTCTAAAATAGATATTTAAATAGTACTTTTACTCCTACTGTAAATTAAGGGTTGGTTTT
TATTATGTGGCCCTCAAGCGAGAAGGCTGACTGCTCTACGTAGACAGGACACAAAGCCCC
ATCGAGCATCTGGGGCCAGATCTTCAGCTGCTGTAATAAAGGCATTGCTCCACTGTTTCC
CAACTGAAGATTTACATGGTGTTCAGGGGTACGGATGCTCATCTCCACAGATACCGA
TGGCAGTGAGCATCTGAACCTCACGGACGGTCTTACTCTCACGTTCTGTGTTTCAATG
GAAGATAAGTCTATCCAGCACAACTACTACTTCTCCCTCCAGCTGTGGATGAGAATAAA
GGATCATTTCAACTTCACCTATGTTTGAACCTATAAGTCCACTGGAAATCTTACAGGAA
ACGTTGCCAAATTTACTGCTTTTTCTGGTCCCAGACTCAGCTTTGAAAAATACGCTTAAA
GTTTTAAGCAAATTCCTATAAAGCATGAGTGCTAAAAAACAACAGAAAAAATTTATTC
TTCACTACCTTTAATTAAGCTGCCTCTATTTACCAGACTTTTCGCAGAGAACCAGAAAT
ATACTGAATCAAATGCTCTGTCTGGACTAATGTGCAGCATTATTCCTGTGGTCTGACTTA
ATTGGACATTTACGCCACATCAGTCTGGAGCGCTATCGGCCAGGTCCTTGGCTGGGATA
AACGAGCACAGTTCACTGACTTTTTCGTGTCAATTTACACCAGCTGAGAAATGTGGTCCATT
CATTTCACTGGAGCTGGATCGGGGATGGATTTGGCCAGTTTGCAGTGACACTACGAACC
GGTTCACAAGGCCGATGCTCAGACAGTTGCATTGAACGGCACCAATCTATCCATCCAAGC
AAAATACCTCTTTCCGTTTATCTCTTACCCTGCCTGTGGATTTCAGGAAGATGAGTTTAG
AAAGCTGTTTTCCCTAGAAGTAAGGCTGGAACCGTCTTTATTGTTGTTATCCTTATGAGA
TCAATGACAACAATGGATGGATTACGGAAACGATTTGCAGTTAAATGAAGCAAGGCAGA
ATTTGGCCCCGCTGCACAGGAACCTTTCTGTATGTAGATTTCCAGCACCTGTGCACTT
TTCTTAACGAGGACAGCAAAACAGCCAACAGGGCACATCCCTCCAGCTCTGCATGAAACA
GACATGCAGGGTCCAGCTAGC

miR-499 site

miR-1a/206 site

>Gga-Dach1.2-UTR [1073 bp; ENSGALT00000027373.3]

AGATCTTGCCCTCATTGTTTGGCTTGGTACATAAAATGGAAATGTTAAGGTTTAAGGGGAA  
 CCAATTATAAGCTGGATGTTTAGAAAGTATCTTGCTAAAACAGTGTAATATTACAGA  
 CCATGAGATGTTAACGTAAGTTGAATTTTTGCCCCCTCTTAGTTATACAGGTTTTGGGT  
 TGGTATTTGTTTTTATTTCAAATTTGTTTTGTTTTGTTTTGTTTTGTTTTCATAGATGATGA  
 AGAAAAGTTGTGCTCATGTTATTGTTTATATGCTTTTGAATCTTAAAGATATTAATGTC  
 TAGTTGTTCTATATTATAACCATATTTGCGCTCTATGCAAGCCCTTGGAACAGAACATAC  
 TCATCTTCATGTAGGACCTATGAAAATTGTCTATTTTTATCTATATATTTAAAGTTTTCT  
 AAAAAAGAAAAAAGGTTATTACGAATTTGTTGTACAAAATCTGTACAAAATCTGTTT  
 TTACATCATAATGCAAGAATTGGAAATTTTCTATGGTAGCCTAGTTATTTGAGCCTGGT  
 TTCAATGTGAGAACCACGTTTACTGTTATTGTATTTAATTTCTTTTCTTTTCAACAATC  
 TGCTAATAAAACTGTCTGAAATCTCCCTGTGACTTCATTTACAGTTCATCTTTATTAAT  
 TTTCTGAAATGTGTAACATCAGAGGAATATTTACTTTCTAATGGGAGGCATCTGAAAACA  
 ACACAAGTCAGCTCTTTGTAATGTGAGGAGAACAATGCTGAATGATTTTATTTAACATGC  
 AACTGCTTCTATCCCTAATATGAATTACTGTGGTAAAAACATCATAAAAGCACACTTTG  
 TGGTTATGTTTTAACAGCAGATTTTCCATATTTTTTTTTTCTTGCCAGCTAAGCAAAC  
 GCCCAATCTACATGATTTATTTATGTACATTTCTCAGTTTATGAAGCTGTTATTTGTAC  
 CTTTTTTATATATTGTGATATTTCCCATGATTCTTATTTTTCAGAAAGCTTTGTGCTGAATA  
 ATGTAAGTGGACACATTTGATGGAACAAAACATATAAATCCCTAGCTACAGAGAGGGTA  
 ATCCAGTTGTTGCTAGC

miR-1a/206 site

**Table 1: *In vitro* transcription conditions for the synthesis of the RNA oligonucleotides used to perform RNA *in situ* hybridisation.**

Genes	RNA polymerase	Plasmid details
Eya1	SP6 antisense T7 sense	pGEMT(Gga-Eya1.2-3UTR) (3'UTR; 970 bp)
Eya2	SP6 antisense T7 sense	pGEMT(Gga-Eya2-ISH) (coding region; 802 bp)
Eya3	SP6 antisense T7 sense	pGEMT(Gga-Eya3-ISH) (coding region; 476 bp)
Eya4	T7 antisense SP6 sense	pGEMT(Gga-Eya4-ISH) (coding region; 873 bp)
Six1	T7 antisense SP6 sense	pGEMT(Gga-Six1-3UTR) (3'UTR; 597 bp)
Six4	SP6 antisense T7 sense	pGEMT(Gga-Six4-ISH) (coding region; 538 bp)
Dach1	T7 antisense SP6 sense	pGEMT(Gga-Dach1-3UTR) (3'UTR; 1084 bp)