UPPSALA
UNIVERSITET

UPTEC X 14 031
Examensarbete 30 hp
Augusti 2014

Formation of the musculoskeletal system during the craniofacial development of zebrafish

Henning Onsbring Gustafson

UPPSALA UNIVERSITET

Degree Project in Molecular Biotechnology

Masters Programme in Molecular Biotechnology Engineering,
Uppsala University School of Engineering

UPTEC X 14 031	Date of issue 2014-08

Author

Henning Onsbring Gustafson

Title (English)

Formation of the musculoskeletal system during the craniofacial development of zebrafish

Title (Swedish)

Abstract

The musculoskeletal system supports the internal structures of the body and consists of bones, ligaments, muscles and tendons. This system forms during early embryonic development, a process where many components today are unknown. In order to get a better understanding for those developmental steps, fluorescent *in situ* hybridisation has been performed on five genes. All five genes represent different transcription factors. These genes were selected based on the assumption that they could be important for the formation of the musculoskeletal system. After *in situ* hybridisation was performed, embryos were stained by immunohistochemistry to get a reference signal in the cartilage to enable easier interpretation of the expression pattern. In this study four of the selected transcription factors, *Scleraxis a*, *Scleraxis b*, *Mohawk a* and *Mohawk b* turned out to be expressed close to points where muscles are attached to the cartilage elements in the zebrafish head. Therefore, these genes are good candidates for future functional studies of muscle attachment development.

Keywords

Musculoskeletal system, Scleraxis, Mohawk, Early growth response 1, *in situ*, zebrafish

Supervisors

Dr. Tatjana Haitina
Department of Organismal Biology, Evolution and Development
Uppsala University

Scientific reviewer

Dr. Xesus Abalo
Department of Neuroscience, Pharmacology
Uppsala University

Project name	Sponsors
Language **English**	Security
ISSN 1401-2138	Classification
Supplementary bibliographical information	Pages **40**

Biology Education Centre Biomedical Center Husargatan 3, Uppsala
Box 592, S-751 24 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 471 4687

Formation of the musculoskeletal system during the craniofacial development of zebrafish

Henning Onsbring Gustafson

Populärvetenskaplig sammanfattning

Det muskuloskeletala systemet består av muskler, ben, senor och ligament. Senor är en del av den vävnad som binder muskler till skelett. Ligament binder samman ben med ben. Detta system är bland annat nödvändigt för att vi ska kunna röra på oss. Trotts detta systems viktiga betydelse är dess utveckling ur många aspekter av forskarvärlden fortfarande okänd. En bättre förståelse kan uppnås om gener som har betydelse för utvecklingen av denna vävnad kartläggs.

I detta arbete har genuttrycket för fem utvalda gener undersökts med zebrafisk som modellorganism. Dessa gener har valts ut under antagandet att de kan ha betydelse för bildandet av muskuloskeletala systemet. Via en teknik som kallas fluorescent *in situ* hybridisering (FISH) har delar av embryot där den undersökta genen är uttryckt identifierats. Via FISH går det även att avgöra vid vilken tidpunkt i utvecklingen som genen uttrycks.

När en gen uttrycks, bildas ett så kallat RNA-fragment som är unikt för just den genen. FISH bygger på att tillverka egna RNA-fragment som kan binda specifikt till RNA som uttrycks från den gen som är av intresse. Det tillverkade RNA-fragmentet är anpassat på ett sådant vis att när det senare binder till genprodukten så kan detta detekteras i ett mikroskop. På så sätt går det att avgöra vart i embryot som den analyserade genen är aktiv.

I den här studien var fyra av de fem olika generna som kartlades aktiva på de positioner i zebrafiskembryot där kända muskelfästen finns. Detta indikerar att dessa kan ha betydelse för utvecklingen av till exempel senor. Den typen av studie som den här rapporten beskriver är ett första steg inför senare studier som förhoppningsvis kan leda till nya behandlingsmetoder av skador på det muskuloskeletala systemet.

Examensarbete 30 hp

Civilingenjörsprogrammet i molekylär bioteknik

Uppsala universitet, augusti 2014

Table of contents

1. Abbreviations

Dpf	Days post fertilisation
Egr1	Early growth response 1
Hpf	Hours post fertilisation
IHC	Immunohistochemistry
ISH	*In situ* hybridisation
FISH	Fluorescent *in situ* hybridisation
GFP	Green fluorescent protein
HYB	Hybridisation solution
Mkxa	Mohawk A
Mkxb	Mohawk B
PBDTT	Phosphate buffered saline + DMSO + triton + tween
PBS	Phosphate buffered saline
PBSt	Phosphate buffered saline + tween
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
RT	Room temperature
Scxa	Scleraxis A
Scxb	Scleraxis B
TAE	Tris-acetate-EDTA
TWS	Tyramide working solution

2. Introduction

2.1. Craniofacial musculoskeletal system

The transition from jawless to jawed species represents a very important event in the vertebrate evolution. Some of the possibilities the establishment of jaws created were the ability to hunt and process motile prey that were earlier not accessible (1). In the craniofacial musculoskeletal system of the vertebrate head, jaws and the pharyngeal arches correspond to a segmented pattern. Those structures mainly develop during the embryonic development of vertebrates including zebrafish (*Danio rerio*) and are mostly derived from neural crest cells (2). The stability of this system depends on several different musculoskeletal tissues like tendons, ligaments and other connective tissues (3). To get insight in the early evolution of jawed vertebrate it is important to explore the function of these components. Better understanding of the musculoskeletal system can as well contribute to a knowledge needed for development of future medical treatments of damages related to this area.

2.2. Gene candidates

In this project zebrafish was used as a model to increase the understanding of the musculoskeletal system (Figure 1). The cells that represent the precursors of the muscle attachments were of specific interest. The five genes targeted in this procedure were selected based on assumptions that they were important during the development of the musculoskeletal system.

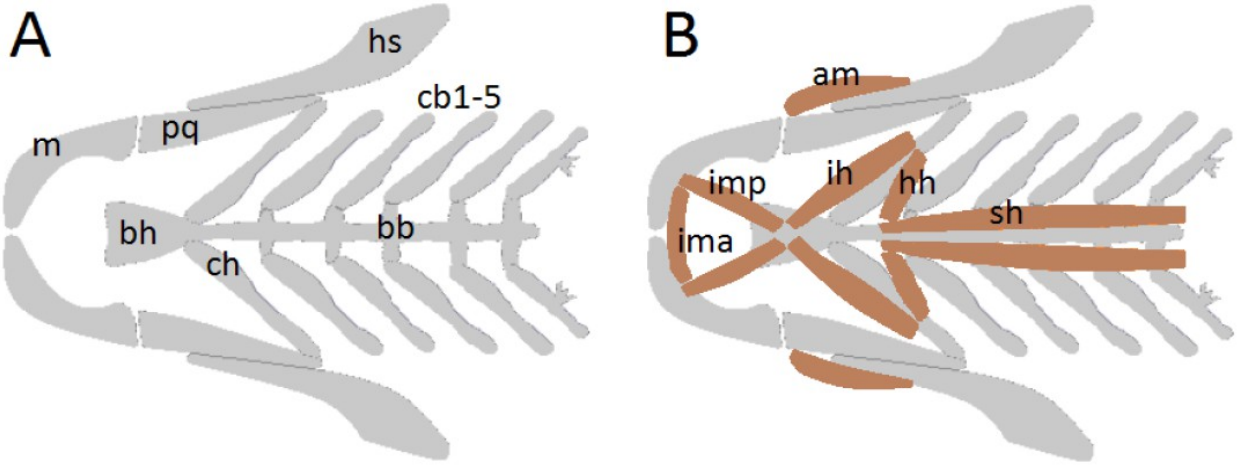

Figure 1. Anatomy of the cartilage and muscle system in zebrafish cranium at 6 dpf. a) Abbreviations: bb, basibranchial; bh, basihyal; cb, ceratobranchial; ch, ceratohyal; hs, hyosymplectic; m, Meckel's cartilage; pq, palatoquadrate. b) Abbreviations: am, adductor mandibulae; hh, hyohyoideus; ih, interhyodieus; ima, intermandibularis anterior; imp, intermandibularis posterior; sh, sternohyoideus. Sketch based on figures and data from (2,3).

The genes included in the study were both teleost specific duplicates of *mohawk* (*mkxa* and *mkxb*) (4) and *scleraxis* (*scxa* and *scxb*) (3). Also *early growth response 1* (*egr1*) (5) was included. All five represent transcription factors. Transcription factors are proteins that regulate gene expression by binding to a specific DNA sequence. Binding to this site will either promote or block the recruitment of RNA polymerase that is needed to start the transcription. Transcription factors were chosen as gene candidates to be characterized in this project because they will play an important role in early development.

2.2.1. Mohawk

In an article from 2010, Ito and colleagues suggested that *Mkx* had an important role in tendon differentiation based on knockout studies on mice. The mice included in the study had clearly affected tendon development in the limb, tail and platysma for the *Mkx* null mice (6). Also studies with zebrafish as model were done by Chuang and colleagues in 2010 where they showed that an morpholino knockdown of *mkxa* causes defects in the formation of craniofacial muscles (4). Those observations motivated the inclusion of both *mkx* duplicates among the gene candidates in this project.

2.2.2. Scleraxis

Several articles point out *Scx* as a marker of tendon differentiation (7,8). In the article published by Murchinson *et al.* in 2007 the researchers show how force-transmitting tendons are disrupted in $Scx^{-/-}$ mice (7). There were also no data based on *scx* expression in zebrafish available at the time when the gene candidates got selected, which made the gene highly interesting to include in the study. During the time of this project when the fluorescent *in situ* hybridisation (FISH) protocol was optimized, another study was published that confirmed the assumed importance for Scxa in the development of the musculoskeletal system in zebrafish (3).

2.2.3. Early growth response 1

Early growth response 1 (EGR1) was suggested in an article from 2011 by Lejard *et al.* to regulate collagen type I production and in this way affect tendon differentiation. Also, ISH results in the same publication, where the expression of *Egr1* and *Scx* in mouse limb tendons showed a similar pattern, made *Egr1* an interesting candidate for this project.

2.3. Fluorescent *in situ* hybridisation

To get an insight where and when the selected gene candidates are expressed, FISH has been used. FISH is a technique where a complementary RNA probe is binding to the transcript of the characterized gene. The probes will be made of digoxigenin labelled UTP that antibodies can bind to. Those antibodies can be detected upon addition of a fluorescent substrate that will enable the visualisation of regions in the embryo where the gene is transcribed.

Even though earlier *in situ* studies have been published on zebrafish for *mkxa* and *egr1* (4,5) before this project was initiated, and for *scxa* (3) during the work of this project, all gene candidates remain interesting to analyse with a new approach. In this project a reference stain by immunohistochemistry has been done. This is possible by using the transgenic zebrafish line

col2a1a:GFP and two sets of antibodies (called primary and secondary antibodies). First the primary antibodies bind to GFP protein in the cartilage. When the secondary antibodies are added they bind to the primary antibodies. The secondary antibodies will carry a fluorescent tag and can therefore be detected (Figure 2). This enables easier orientation of the spatial expression. This staining is detected by a fluorescent dye, therefore also fluorescent *in situ* was preferred before other *in situ* staining methods. Having both signals as a fluorescent signal makes it easier to analyse them simultaneously and to merge them both into one image.

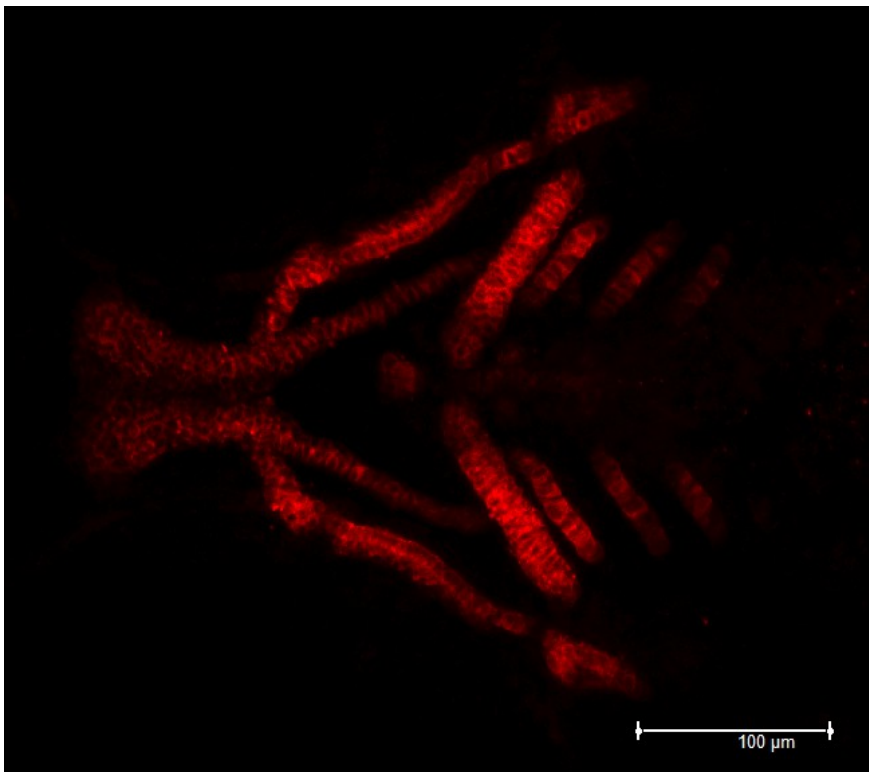

Figure 2. Maximum projection of confocal image stack of the cartilage system in zebrafish embryo at 3 dpf. The figure shows an embryo from the transgenic zebrafish line col2a1a:GFP. The cartilage system is stained by immunohistochemistry, were anti-GFP antibodies is used and the result is analysed by confocal imaging.

During this project, unlike earlier published data (3–5), both teleost specific duplicates of the *mkx* and *scx* genes have been studied. Some developmental stages have been included in the *in situ* runs that have not been investigated previously.

To control that the FISH protocol and all components used are working, *krox20* (4) was added as a positive control. This gene is very suited for this purpose because of the distinct expression pattern at 24 and 48 hpf where two stripes in the hindbrain are visible.

2.4. Aim

The aim of this project was to characterize the expression of *scleraxis a*, *scleraxis b*, *mohawk a*, *mohawk b* and *early growth response 1* by using fluorescent *in situ* hybridisation.

3. Methods

The following methods were used throughout the project. Additional information about reagents (9.1) and conditions for the zebrafish housing (9.2) can be found in the supplementary data.

3.1. Embryo fixation and dehydration

First, embryos were fixed in 4% PFA in PBS (pH 7.4) for 2 to 3 hours at room temperature (RT). Then the embryos were transferred into PBSt, i.e. PBS with a final concentration of 0.1% tween and pH 7.4. The embryos were then dechorinated. After dechorination the embryos were washed twice in PBSt for 5 min. The samples were then dehydrated by first a 30% methanol in PBSt wash, followed by a 60% methanol in PBSt wash and finally in 100% methanol. The three dehydration steps were done at RT for 5 min. Finally, the methanol in the test tubes with embryos were replaced with fresh methanol and stored at -20 ^{o}C.

3.2. RNA isolation

Zebrafish embryos were washed several times with sterile water. Then, 500 µl Trizol was added and the sample was homogenized by passing it through a 23 gauge needle 20 times and another 20 times through a 27 gauge needle.

In the next step, 500 µl was added again and mixed by pipetting up and down. Then the sample was incubated 5 min at RT. After the incubation, 200 µl chloroform was added and mixed by shaking for 15 sec. The test tube was incubated for 3 min at RT.

The sample was spun at 14 000 rpm for 15 min at 4 ^{o}C. The aqueous phase was after this step removed and mixed with an equal volume of 2-propanol by pipetting up and down 6 times. In the following steps the test tube was incubated for 10 min at RT, centrifuged for 30 min at 14 000 rpm at 4 ^{o}C and then the supernatant was discarded.

The pellet was washed with 1 ml 75% ethanol. This step was followed by a centrifugation for 5 min at 14 000 rpm at 4 ^{o}C. The supernatant was once again discarded. The remaining pellet was air dried at RT. Then the pellet was resuspended in 50 µl nuclease free H_2O followed by a 5 min incubation step at RT. After gentle vortexing of the test tube, the sample was stored at -80 ^{o}C. The RNA clean up was done using a Qiagen RNEasy Mini Kit according to the protocol published in Journal of visualized experiments from 2009 (9).

3.3. First strand cDNA synthesis reaction

To a test tube the following components were added: 1 µl of 50 µm oligo$(dT)_{20}$, 5 µg RNA, 0.5 µl dNTP mix and distilled water up to 13 µl. The mixture was incubated for 5 min at 65 ^{o}C and then put on ice for at least 1 min. After a short centrifugation of the test tube the following components were added: 4 µl First-strand buffer 5X, 1 µl of 1 M DTT, 1 µl RNaseOUT™ and 1 µl SuperScript™ III reverse transcriptase. The content of the tube was mixed by slowly pipetting up and down. The mixed sample was then incubated 30-60 min at 50 ^{o}C, followed by a 15 min incubation at 70 ^{o}C. Finally, the test tube was stored at -20 ^{o}C.

3.4. PCR

In order to amplify a single product, the annealing temperature and extension time were optimized for each gene candidate in corresponding PCR program (Table 1). The T7 polymerase sequence 5'CTGTAATACGACTCACTATAGGG3' was added to the 5' end of the reverse primer to enable direct transcription from the PCR product.

More details about the primer design and their sequences are listed in the supplementary data (9.3). *Mkxa*, *mkxb* and *egr1* had 60 ºC as annealing temperature. For *scxa* and *scxb*, 59 ºC and 50 ºC were used as annealing temperatures, respectively. The extension time for *mkxb* was 45 s while the rest of the gene candidates had an extension time of 60 s.

Table 1. A generalized PCR program used to amplify the different gene candidates.

Step	Temperature	Time	Number of cycles
First denature	94 ºC	5 min	1
Denature	94 ºC	30 s	
Annealing	50-60 ºC	40 s	30
Extension	72 ºC	45-60 s	
Last extension	72 ºC	5 min	1
Short term storage	4 ºC	∞	1

The PCR reaction mixture was prepared by adding 35 µl H_2O, 5 µl PCR buffer 10X, 2.5 µl DMSO, 0.5 µl dNTP (20 mM), 2 µl forward primer (10 µM), 2 µl reverse primer (10 µM), 1 µl Taq polymerase and 2 µl template cDNA to a test tube. The reaction mixture was pipetted up and down a couple of times before put in the thermal cycler. When the PCR program was done, a fraction of the amplified sample was verified by gel electrophoresis, and the rest was stored at -20 ºC.

3.5. Gel electrophoresis

Gel electrophoresis was performed on a gel made by 1.5% agarose in TAE buffer. 7 µl PCR product was mixed with 1.4 µl loading dye 6X. The gel was stained by incubation in a mixture of 200 ml TAE buffer and 30 µl GelRed 10 000X. This incubation step was performed on a rocker.

3.6. RNA probe preparation

In the positive control case for *krox20* a plasmid was ordered from Addgene (10). To a test tube, 4.4 µg plasmid DNA was added together with 1 µl Xba1, 2 µl Tango buffer 10X and water up to 20 µl. This mix was then incubated at 37 ºC for 2 h. After incubation, the volume was adjusted to 100 µl.

The linearised product or the PCR products that were amplified from the primers listed in 9.4 were cleaned by using a QIAquick PCR purification kit according to the manufacturer's instructions (11). The DNA concentration was measured by NanoDrop (NanoDrop 1000 Spectrophotometer manufactured by Thermo Scientific).

The transcription reaction contained 1 µg linearised DNA, 2 µl transcription buffer 10X, 2 µl of 0.1 M DTT, 2 µl dNTP labeling mix, 1 µl RNase inhibitor, 2 µl RNA polymerase (T3 for *krox20* and T7 in all other cases). Volume for the mix was adjusted to 20 µl by addition of water before the incubation step started. The reaction was incubated at 37 ºC for 2 hours.

The transcription reaction was purified by using a RNeasy mini kit according to the manufacturer's instructions (12) with the modification that the first three centrifugation steps were done at 10 000 rpm for 30 s. When the transcription product was purified the RNA concentration and quality was measured by NanoDrop and an equal amount of formamide was added. The RNA probes were stored at -20 ºC.

3.7. Whole mount fluorescent *in situ* hybridisation

All steps, if nothing else is mentioned, were performed at RT. FISH was performed on 15 somites stage and 24 hpf for the positive control *krox20*. For the other genes, 48 hpf, 3 dpf, 4 dpf and 5 dpf were analysed. The FISH was based on a protocol developed by Hauptman and Gerster (13), with modifications as follows.

3.7.1. Day 1, rehydration, Proteinase K digestion and hybridisation

The fixed and dehydrated embryos from 4.1 were rehydrated by the following steps:

- 2% H_2O_2 in metanol for 20 min on rocker
- 60% metanol in PBSt for 5 min
- 30% metanol in PBSt for 5 min
- 5x in PBSt for 5 min

The rehydrated embryos were then treated with Proteinase K (5 µg/ml) at RT. The incubation time was adapted according to the developmental stage of the embryos in the sample (Table 2).

Table 2. The different durations of the treatment with Proteinase K based on the time after fertilisation.

Developmental stage	Duration of Proteinase K treatment
Somite stages	4-5 min
24 hpf	8 min
48 hpf	12 min
3 dpf	15 min
4 dpf	22 min
5 dpf	30 min

After Proteinase K treatment, the embryos were washed three times in PBSt for 10 min. Then the embryos were post-fixed on a rocker in 4% PFA for 20 min. This step was followed by four washes in PBSt for 5 min. Then PBSt was removed and replaced with hybridisation solution (HYB). HYB was prepared according to Table 3. When 300 µl HYB had been added, the test tubes were incubated for 4 hours at 65 ºC with agitation.

Table 3. The components and their final concentrations in hybridisation solution.

Components	Volume	Final concentration
Formamide (deionized)	25 ml	50%
SSC 20X (pH 7)	12.5 ml	5X
Heparin sodium salt (50 mg/ml)	50 µl	50 µg/ml
0.1% Tween-20	250 µl	0.1%
tRNA (5 mg/ml)	2.5 ml	250 µg/ml
Denhardt's solution 50X	5 ml	5X
Salmon sperm DNA (10 mg/ml)	2.5 ml	500 µg/ml
Dextran sulfate	2.5 g	5%
Nuclease free water	up to 50 ml	

The probes prepared according to 4.6 were diluted to 300 ng/ml in HYB and then denatured at 80 ^{o}C for 5 min. This denaturing step was carried out when the four-hour incubation in HYB was done so that the old HYB could be replaced with HYB containing probes. The embryos were then incubated at 65 ^{o}C with agitation overnight in a HYB mixture containing probes.

3.7.2. Day 2, post-hybridisation washes and antibody incubation

First SSC 2X, SSC 0.2X and post-hybridisation solution (post-HYB) were prewarmed at 65 ^{o}C. Post-HYB was prepared according to Table 4.

Table 4. The components and their final concentrations in post-hybridisation solution.

Components	Volume	Final concentration
Formamide (deionized)	25 ml	50%
SSC 20X (pH 7)	12.5 ml	5X
Nuclease free water	up to 50 ml	

The HYB containing unbound probe was then removed before the following washes on rocker:

- 75% post-HYB and 25% SSC 2X for 20 min
- 50% post-HYB and 50% SSC 2X for 20 min
- 25% post-HYB and 75% SSC 2X for 20 min
- SSC 2X for 10 min
- SSC 0.2X, twice, for 30 min

After the last washing step the test tubes rested on the bench until they had cooled down to RT. Then new washes with increasing concentration of Tris-buffer (Tris-HCl-buffer 50 mM, NaCl 150 mM, Tween 0.1%) were performed at RT:

- 75% SSC 0.2X and 25% Tris-buffer for 10 min
- 50% SSC 0.2X and 50% Tris-buffer for 10 min
- 25% SSC 0.2X and 75% Tris-buffer for 10 min
- Tris-buffer, four times, for 5 min

In the next step the embryos were incubated in 8% sheep serum in Tris-buffer for 3 to 4 hours on rocker. Then the sheep serum in Tris-buffer was exchanged with anti-DIG-POD diluted 1:1000 in 8% sheep serum in Tris-buffer. Finally, the tubes were incubated on rocker overnight at 4 ºC.

3.7.3. Day 3, post antibody washes

The anti-DIG-POD antibody was removed followed by ten washes in Tris-buffer for 10 min on rocker. The next washing step was done with Tris-buffer at 4 ºC for 5 hours on rocker.

After 5 hours the embryos were incubated in room tempered amplification buffer for 15 min on rocker. Tyramide working solution (TWS) was prepared by diluting fluorescein amplification reagent 1:50 in amplification buffer.

The embyos were then incubated at RT in TWS protected from light. The signal was monitored after around 25 min of incubation in TWS. After the TWS incubation step, the embryos were washed in PBSt three times for 10 min on a rocker. Based on the background, an additional washing step overnight at 4 ºC in PBSt where done if necessary. Signal detection was done by fluorescent microscopy.

3.8. Fluorescent microscopy

The signal detection was done on a Leica M205 FA with a Plan APO 1.0x objective. Around 140x zoom were usually applied with a GFP filter.

3.9. Immunohistochemistry

All steps if nothing else is mentioned were performed at RT. All steps were protected from light during the whole protocol. A transgenic col2a1a:GFP zebrafish strain was used (14).

3.9.1. Day 1, primary antibodies

Following the *in situ* protocol the embryos were initially incubated in precooled acetone for 7 min at -20 ºC, followed by four washing steps in PBSt for 10 min. After that, the embryos were incubated in PBS containing 0.1% trypsin and 1 mM EDTA for 30 min at 37 ºC. Then the embryos were washed twice in PBSt for 2 min on rocker.

In the next step, the embryos were incubated in 0.5% hyaluronidase in PBSt for 30 min at 37 ºC. After hyaluronidase treatment, two washes for 2 min in PBSt were done on rocker. The washed samples were then incubated 75 min in 1% Blocking reagent on rocker.

The blocking step was then followed by an incubation step overnight with anti-GFP rabbit diluted 1:200 in 1% Blocking reagent on rocker at 4 ºC.

3.9.2. Day 2, secondary antibodies

First PBDTT was prepared by diluting Triton 100X to 0.5%, Tween-20 to 0.1% and DMSO to 1% in PBS. Then, six washes were performed, each for 25 min on rocker. After the washes, the embryos were incubated for 75 min in 1% Blocking reagent on rocker. In the next step, the embryos were incubated with 1:500 dilution of Alexa Fluor 594 F(ab')2 Fragment of Goat Anti-Rabbit IgG (H+L) in 1% Blocking reagent and incubated overnight at 4 ºC.

3.9.3. Day 3, detection

The embryos were washed four times for 30 min in PBDTT on rocker and then mounted in low melting agar for fluorescent and confocal microscopy. At this point the stored embryos were kept in PBSt at 4 ºC, still protected from light.

3.10. Confocal microscopy

Confocal microscopy was done on a Leica TCS SP5 II system with a 20X PL APO N.A. 0.7 objective. 488 nm and 615 nm laser were used. The confocal microscope was mainly used to merge the two fluorescent signals from both FISH and immunohistochemistry.

4. Results

4.1. RNA isolation

RNA was successfully isolated from the 2 cell stage up to 4 dpf (Table 5).

Table 5. The concentration and quality of the isolated RNA. The "260/280"-value represents the absorption at 260 nm divided by the absorption at 280 nm. For a pure RNA sample the 260/280-value should be close to 2 and the 260/230-value should be in the 1.8-2.2 range (15).

Developmental stage	Concentration (ng/µl)	260/280	260/230
2 cell + 24 hpf	621.5	2.07	2.05
2 cell	59.4	1.83	0.73
15 somites	221.3	2.07	1.84
24 hpf	421.4	2.11	2.16
48 hpf	581.9	2.09	1.93
3 dpf	733.7	2.14	2.13
4 dpf	1007.4	2.11	2.24

4.2 Gel analysis of PCR products

For all targets, single PCR products of correct length were obtained (Figure S1). The PCR product length is displayed in Table 6.

Table 6. PCR product length calculated for the different gene candidates by an *in silico* PCR software available online (16).

Target amplified from cDNA	PCR product length
mkxa	844 bp
mkxb	400 bp
scxa	503 bp
scxb	532 bp
egr1	820 bp

4.3. RNA probe preparation

After the transcription reaction, purified RNA concentration and quality were measured immediately before addition of formamide and storage in -20 ^{o}C (Table 7).

Table 7. The concentration and quality of RNA probes.

Gene	Concentration after addition of formamide (ng/µl)	260/280	260/230
krox20	203.4	2.07	1.89
mkxa	167.6	2.03	1.65
mkxb	41.0	1.92	1.70
scxa	148.4	1.99	1.86
scxb	192.9	1.97	1.92
egr1	270.7	2.01	1.83

4.4. Whole mount fluorescent *in situ* hybridisation

4.4.1. *Krox20*

When observing the positive control with the fluorescent microscope, the expected two bands in the hind brain were visible (Figure 3). Compared to the reference (4), it seemed to be higher levels of unspecific signal which indicated that the protocol needed further optimization. Despite that, the result could be considered as successful since no major troubleshooting had to be done for the protocol or used components.

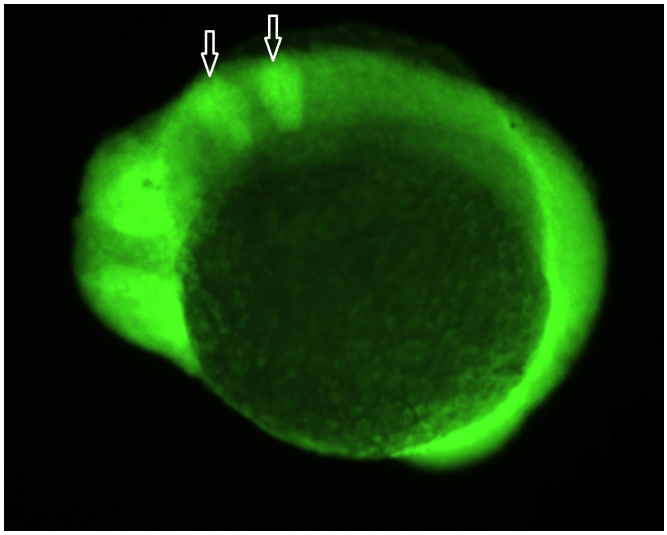

Figure 3. Expression pattern of *krox20* at 24 hpf. Expression can be observed as two stripes in the hindbrain (arrows).

4.4.2. *Mohawk A* and *Scleraxis A* before optimization of FISH protocol

During this project the optimization of the FISH protocol was a very time consuming part. After the analysis of *krox20*, two rounds of FISH were done for *mkxa* (Figure 4a) and an additional round for *scxa* (Figure 4b) before the result for *egr1* finally was satisfying (Figure 4d). Still the probe was changed for *egr1* which could explain that details could be seen at a higher resolution and with less background. But since there were so much signal, for example in the neuromasts and in the epithelium (Figure 4c) which could not be found in earlier published data (4), it seemed like the FISH-protocol caused too high background.

The change made in the protocol to achieve a lower background (Figure 4d), was to replace Western Blocking reagent with 8% sheep serum in Tris. The antibody concentration was also changed from 1:500 dilution to 1:1000 dilution. Both of those changes were mentioned in an article about ISH trouble shooting from 2006 (17). Earlier attempts to reduce the background staining by using lower concentrations of TWS or extending the last PBSt wash over the weekend were not successful.

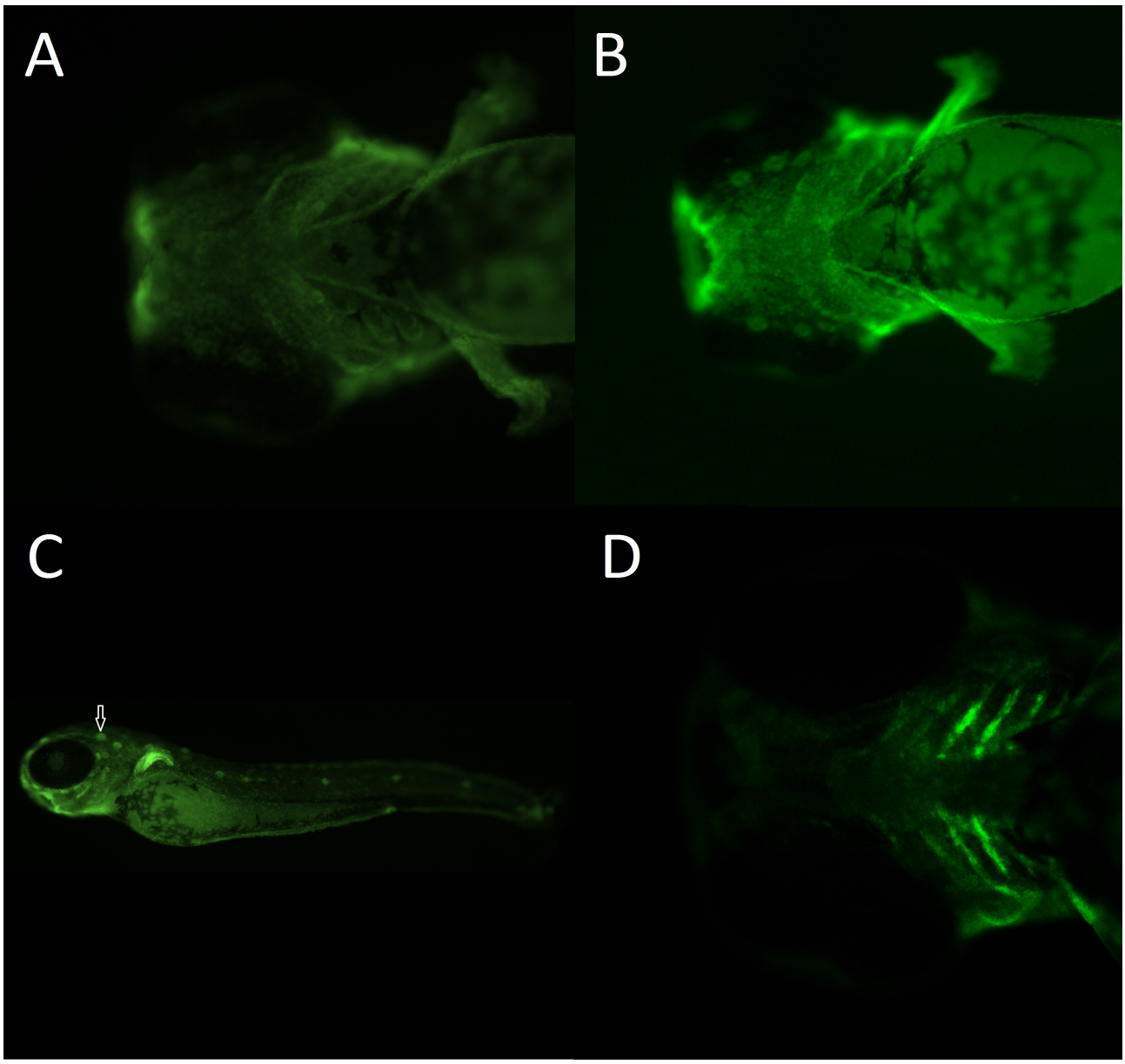

Figure 4. Optimization progress of FISH protocol. a) The expression pattern of *mkxa* 4 dpf ventrally, details hard to distinguish. A lot of signal in the epithelium. b) The expression pattern of *scxa* 4 dpf ventrally, details hard to distinguish as well. Not clear what signal that is specific. c) The expression pattern of *mkxa* 3 dpf laterally, strong signal in the neuromast (arrow) which contradicts earlier published results. d) The expression pattern of *egr1* 4 dpf ventrally, details more clear. No neuromasts or strong signal in the skin.

4.4.3. *Mohawk A*

Of the analysed developmental stages, 3 and 4 dpf where the ones that showed clear specific signal in distinct areas of the embryo. The expression pattern for 3 dpf (Figure 5) was very similar to the 4 dpf embryo.

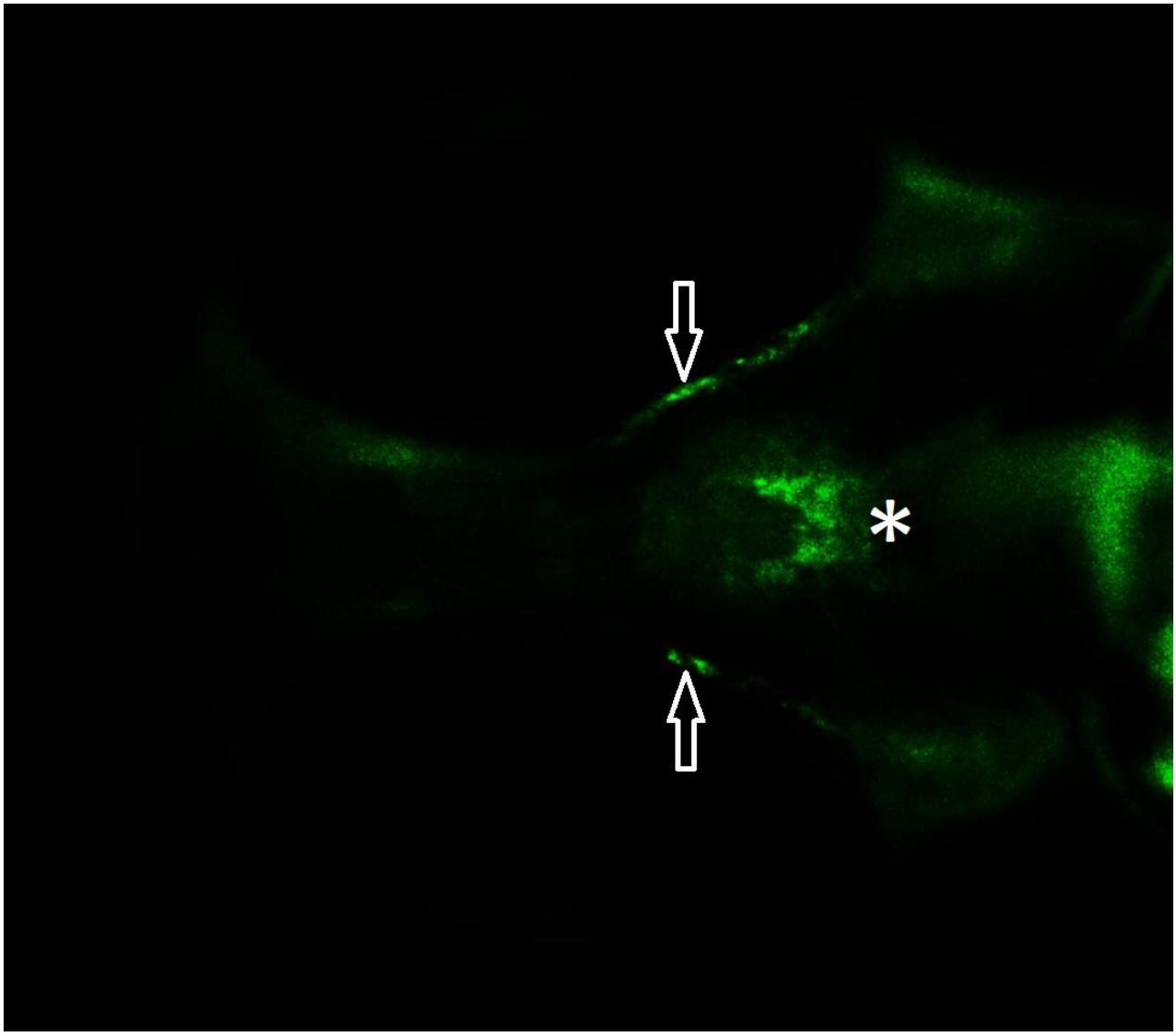

Figure 5. Ventral view of *mkxa* expression in 3 dpf embryo. Signal indicate expression ventromedial to the palatoquadrate (arrows). There is also signal visible on the anterior part where the sternohyoideus muscle attach to the ceratohyal cartilage (asterisk).

4.4.4. *Mohawk B*

For *mkxb*, clear specific signal was only observed in the 3 dpf embryo. The expression pattern consisted of two small areas at a similar position where signal was observed in *mkxa* at 3 dpf. Those two small areas of expression were located on the anterior part of the sternohyoideus muscle where it attaches to the ceratohyal and basihyal cartilage (Figure 6).

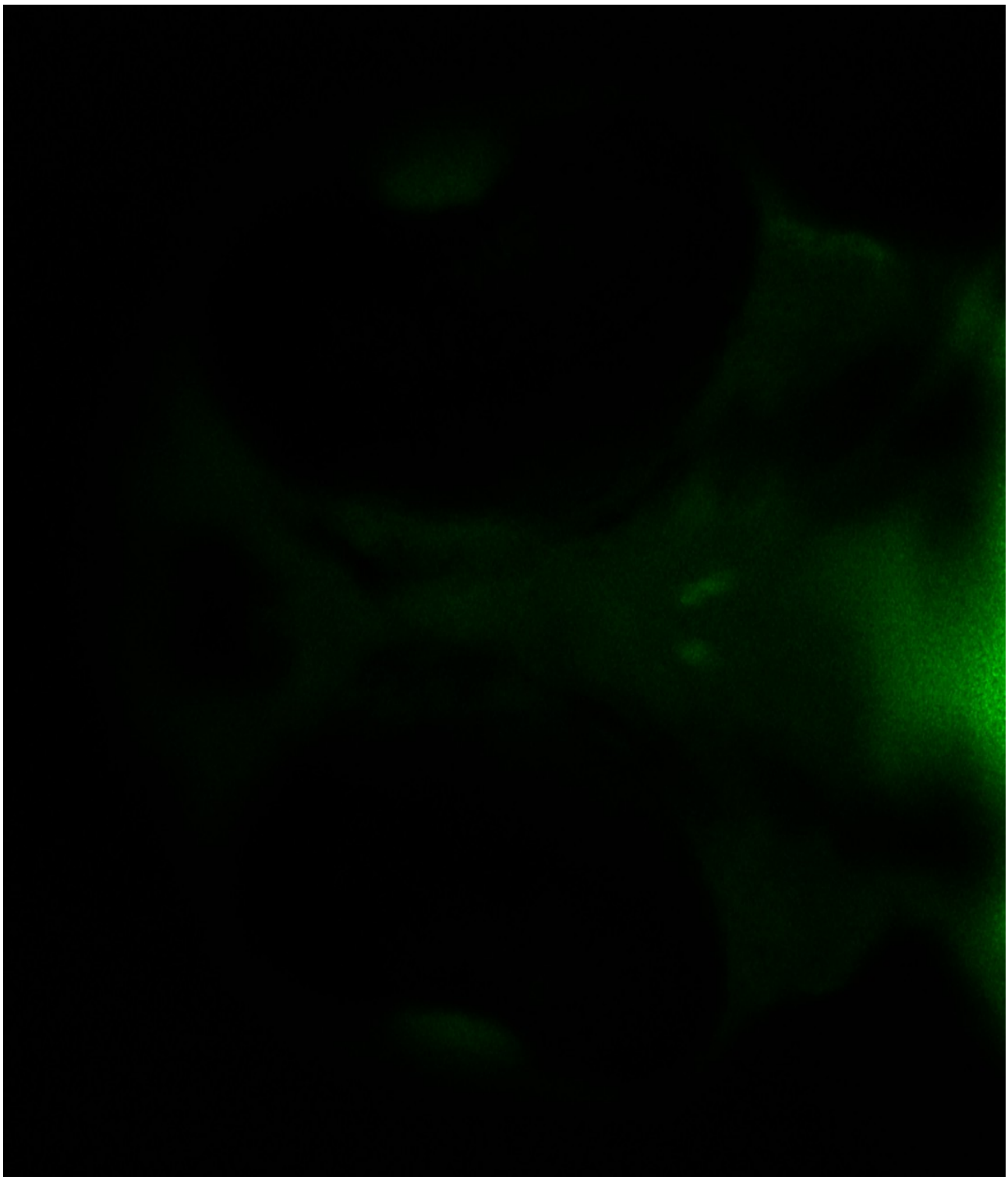

Figure 6. Ventral view of *mkxb* expression in 3 dpf embryo. Two areas are visible at the anterior part of the sternohyoideus muscle. These expression signals could also be confirmed in Figure 12.

4.4.5. *Scleraxis A*

At 48 hpf, a lot of signal was visible in the whole head and elevated levels were indicated along the midline (Figure 7a). The 3 dpf expression pattern is similar to *mkxa* (Figure 7b). A similar elongated area of expression could be observed ventromedial to the palatoquadrate. Expression was also detected in the anterior part of the sternohyoideus. Two small areas of expression were also visible on the side of the Meckel's cartilage close to Meckel's-palatoquadrate cartilage joint. Embryos at 4 dpf showed a very similar expression pattern to the embryos at 3 dpf. However, no specific expression signal could be detected at 5 dpf.

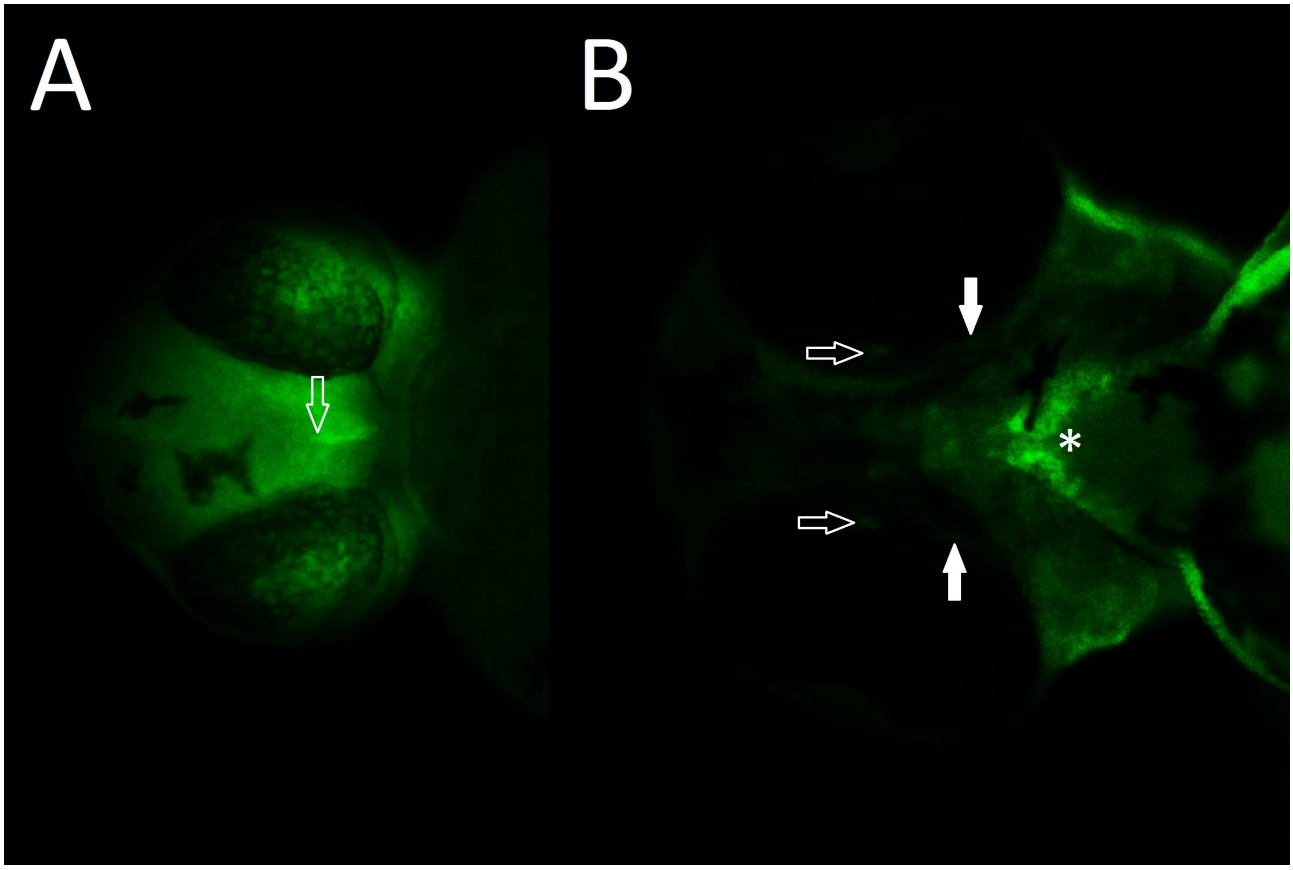

Figure 7. Ventral view of *scxa* expression in 48 hpf and 3 dpf embryos. a) Expression at 48 hpf was detected in the whole head with a strongly labeled midline (open arrow). b) Expression at 3 dpf was detected as a stripe ventromedial to the palatoquadrate (filled arrows). Anterior to the end of those two stripes, two small areas of expression could be observed (open arrows). Expression signal was also observed at the anterior part, where the sternohyoideus muscle is attached (asterisk). These expression signals could also be confirmed in Figure 13.

4.4.6. *Scerlaxis B*

The expression pattern of *scxb* was very similar to the pattern of *scxa*. Expression at 48 hpf was almost identical with high levels in the head and labeled midline across anterior-posterior axis. At 5 dpf, the signal was too diffuse to distinguish the specific signal from the non-specific like in the other cases for *mkx* and *scx*. The pattern of *scxb* was very similar to *scxa* both at 3 dpf and 4 dpf, whereas the expression signal close to the sternohyoideus muscle attachment position was broader in scxb (Figure 8).

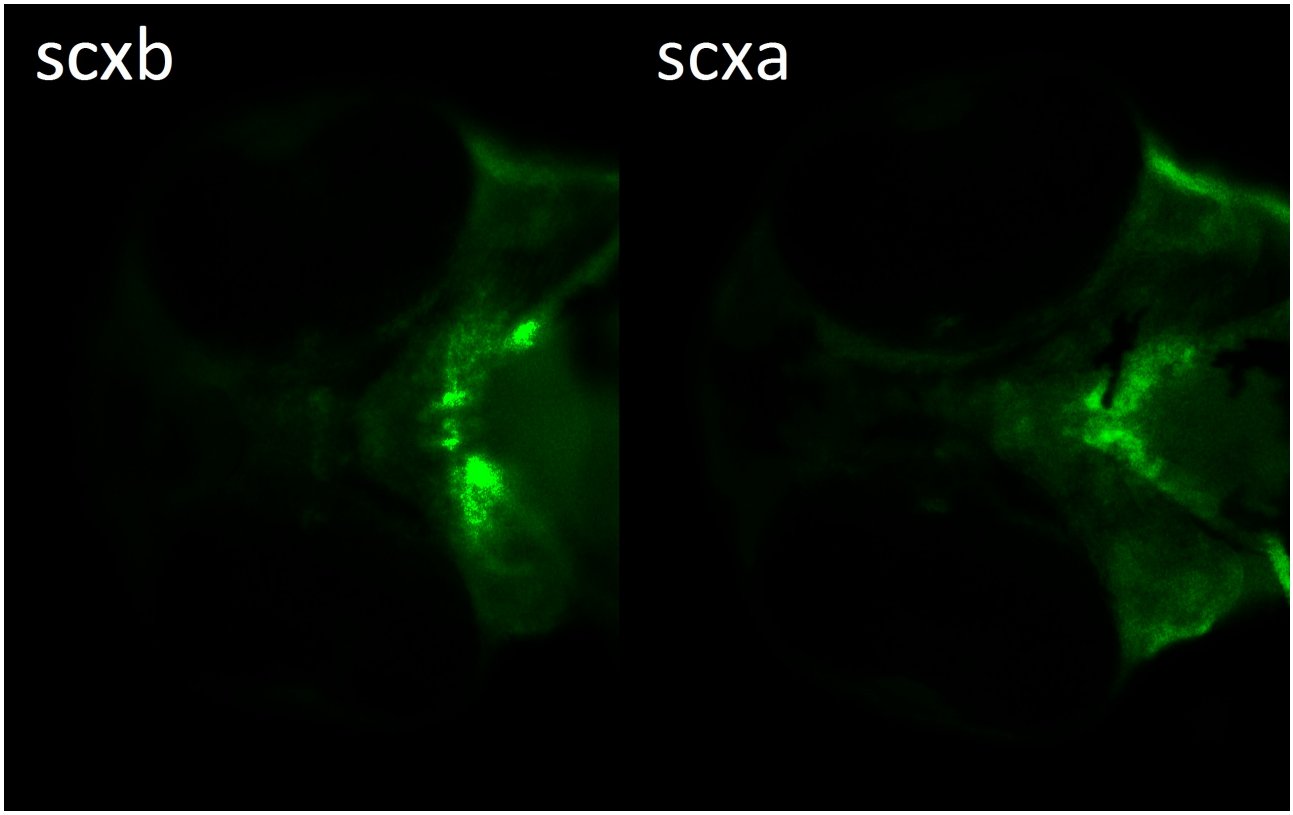

Figure 8. Ventral view of *scxa* and *scxb* expression at 3dpf. A notable difference is the broader expression pattern for the anterior part of the sternohyoideus muscle attachment position in scxb.

4.4.7. *Early growth response 1*

At 48 hpf, signal is visible in the area where the jaws will later appear. Also, two small areas of expression in the anterior part of the head are visible (Figure 9a). 3 and 4 dpf is once again two very similar stages, and at 5 dpf no clear expression pattern can be seen. At 3 and 4 dpf an expression pattern can be seen around the pharyngeal arches 3 to 5 (Figure 9b).

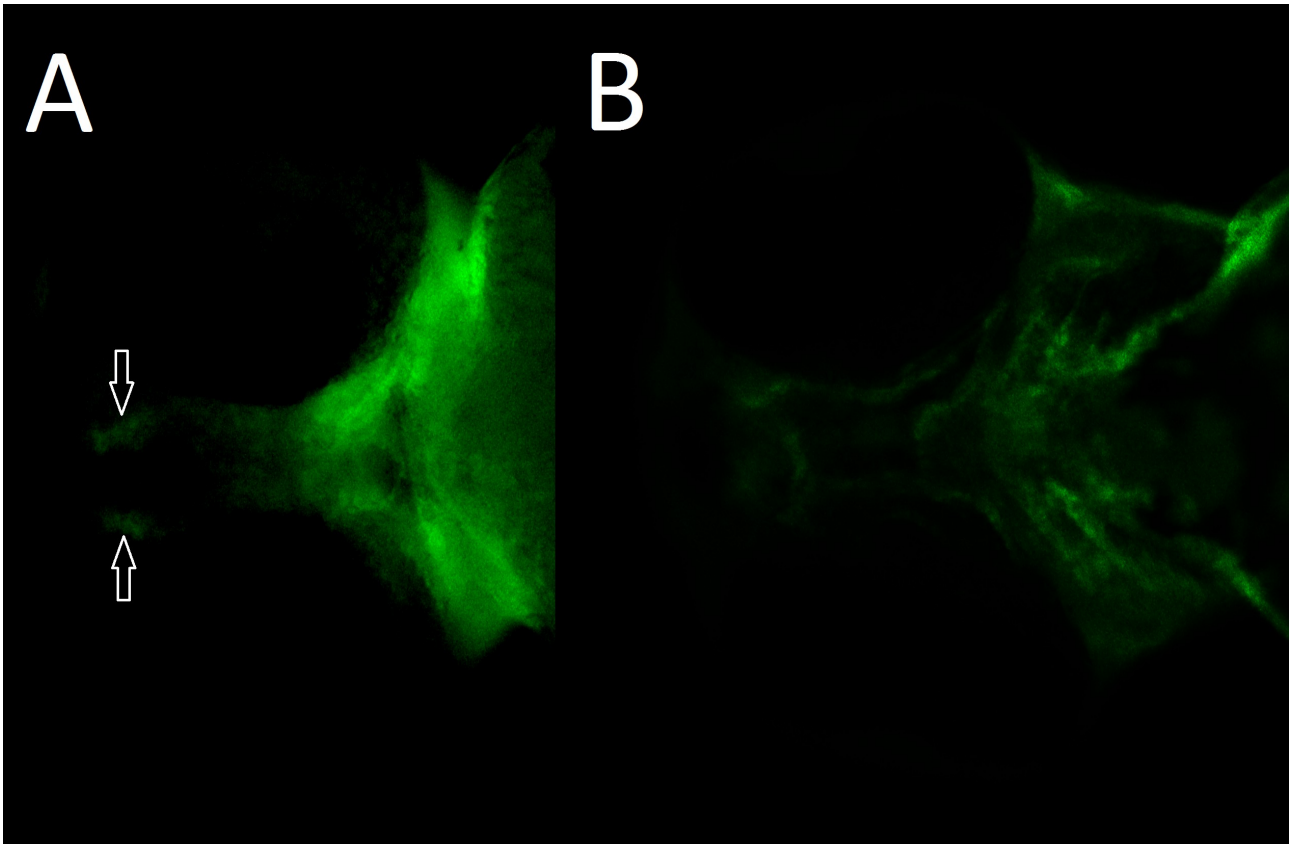

Figure 9. Ventral view of *egr1* expression in 48 hpf and 3 dpf embryos. a) In the 48 hpf embryo, expression could be detected where the jaws will later appear. In the anterior part of the head, two short stripes of expression could also be detected (arrows). b) A distinct expression pattern around the pharyngeal arches could be seen at 3 dpf.

4.4.8. Negative control

All FISH steps were repeated on a negative control, with the only difference that no probes were added. Even though the iris and exposure settings in the detection step were at high levels, no specific signal could be observed (Figure 10).

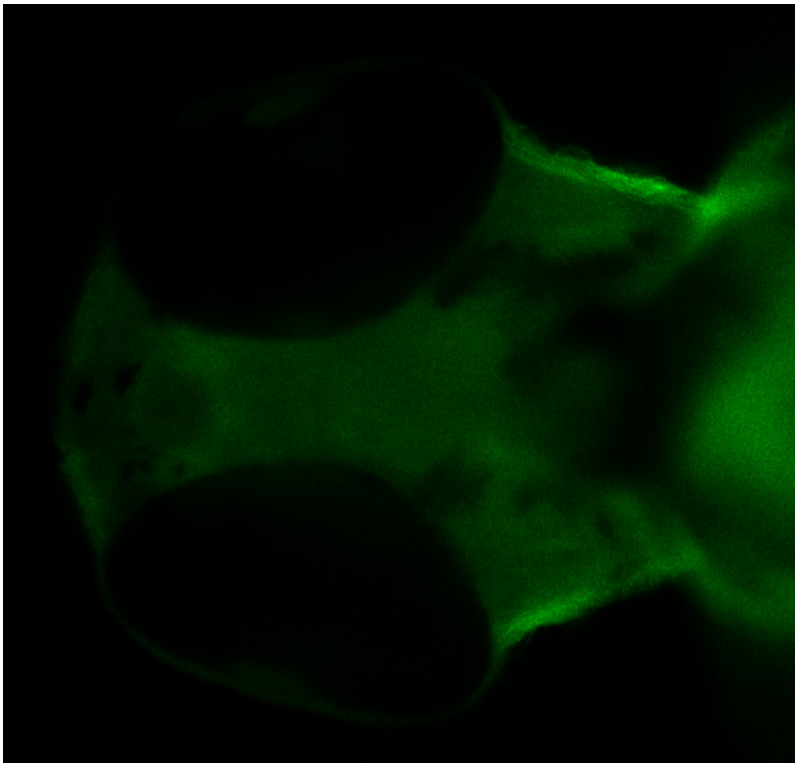

Figure 10. Ventral view of 3 dpf embryo where no probes have been added during the ISH process. No expression pattern was visible when probes were excluded from the FISH procedure.

4.5. Immunohistochemistry

The developmental stage that seemed most relevant to this project based on the result from the fluorescent microscopy were embryos at 3 dpf. On that stage of development, the most distinct expression patterns were observed in the areas of interest, i.e. close to muscle attachment sites. The embryos at 4 dpf often had a similar expression pattern, but with weaker signal. The embryos at 48 hpf often showed a lot of expression in the whole head, or no expression at all. While the embryos at 5 dpf did not show any clear signal. Therefore, the immunohistochemistry was performed on embryos at about 3 days post fertilisation.

4.5.1. *Mohawk A*

The reference signal from the immunohistochemistry localised the elongated area of expression to be ventromedial to the palatoquadrate (Figure 11). The merged image could also show that the expression observed under the jaw was slightly posterior to the sternohyoideus muscle attachment site.

Similar areas of expression compared to observations in *scxa* and *scxb* where also detected close to the Meckel's-palatoquadrate cartilage joint (Figure 11). To be able to observe that pattern more clearly with the fluorescent microscope, the focus had to be shifted from the settings used in Figure 5. The more detailed analysis by confocal microscopy revealed expression in smaller populations of cells as well. Signal could be detected where the intermandibularis anterior muscle attaches to Meckel's cartilage.

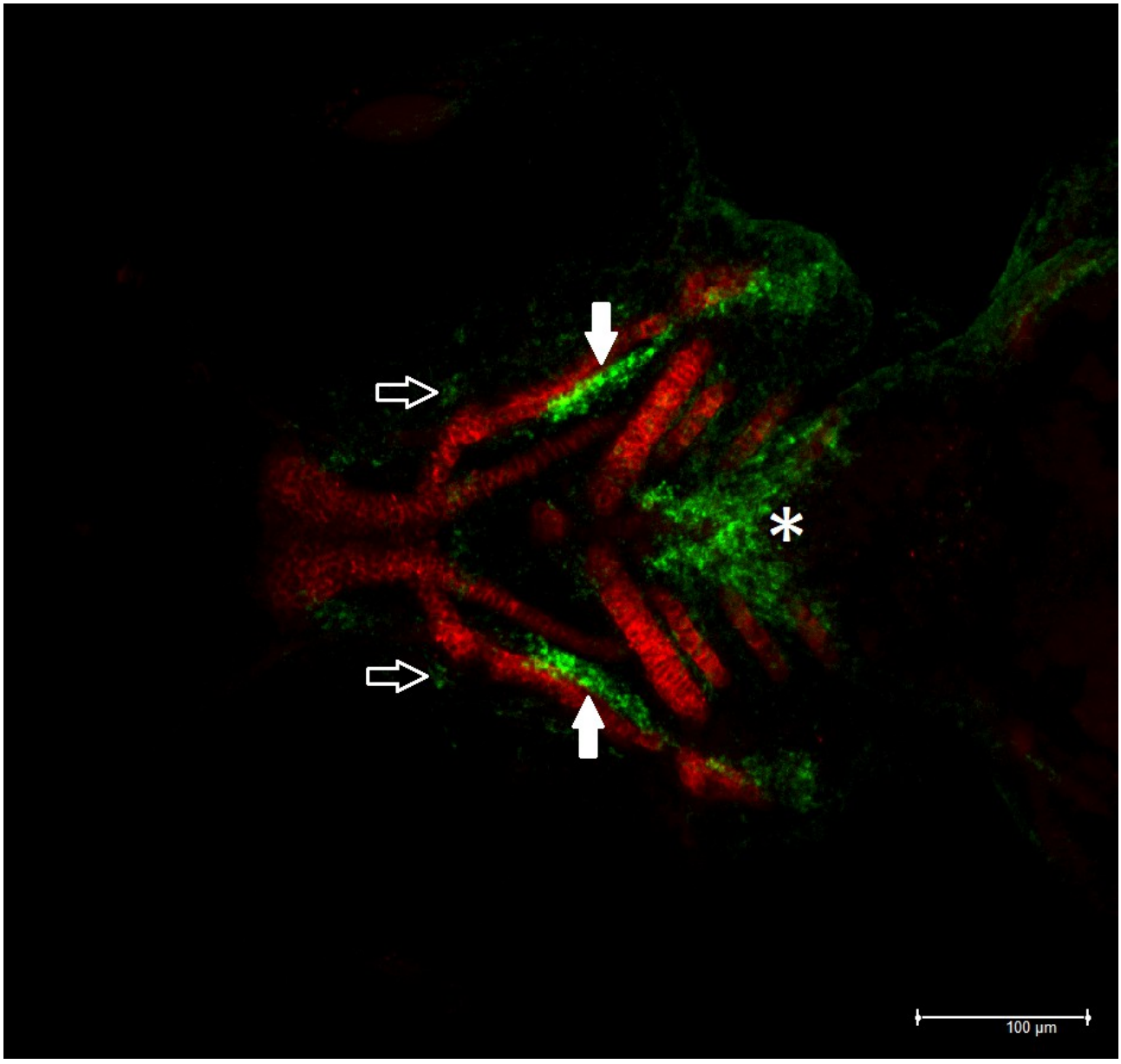

Figure 11. Maximum projection of confocal image stack showing expression of *mkxa* at 3 dpf. The expression pattern of *mkxa* (green) relative to the cartilage stain in a transgenic col2a1a zebrafish embryo (red). Now the two small areas of expression close to the Meckel's-palatoquadrate cartilage joint, which were not detected in Figure 5, can be observed (open arrows). Signal could also be detected ventromedial to the palatoquadrate (filled arrows) and in the anterior part of the sternohyoideus (asterisk).

4.5.2. *Mohawk B*

The cells that express *mkxb* show the same morphology as many of the cells *mkxa* were expressed in. The two areas of expression from *mkxb* are at similar position as the anterior parts of the sternohyoideus muscle attachment site (Figure 12).

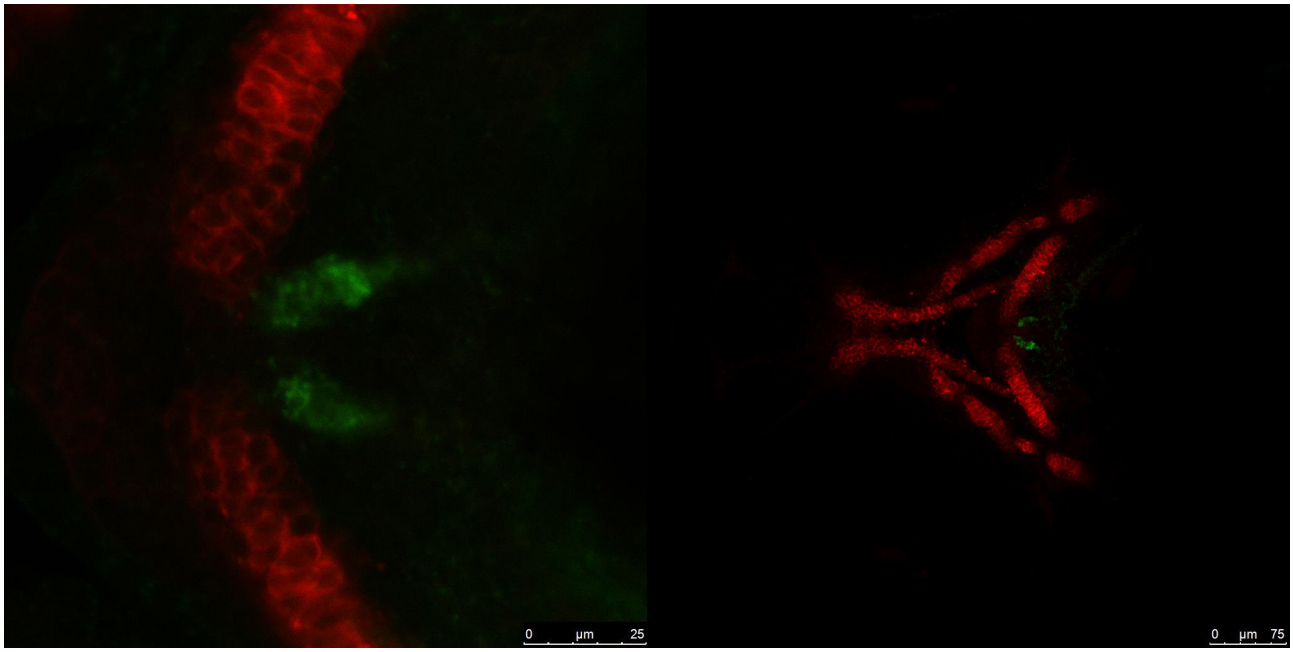

Figure 12. Maximum projection of confocal image stack showing expression of *mkxb* at 3 dpf. Expression pattern can be observed posterior to the ceratohyal, close to where the sternohyoideus muscle attach.

4.5.3. *Scleraxis A*

With data from the confocal image, the two small areas of expression close to the Meckel's-palatoquadrate cartilage joint, both from *mkxa* and *scxa*, can be confirmed to be at similar positions. But the two areas of expression ventromedial to the palatoquadrate are not as extended in *scxa* as in the *mkxa* case. There is less signal in the area of the sternohyoideus muscle attachment as well (Figure 13).

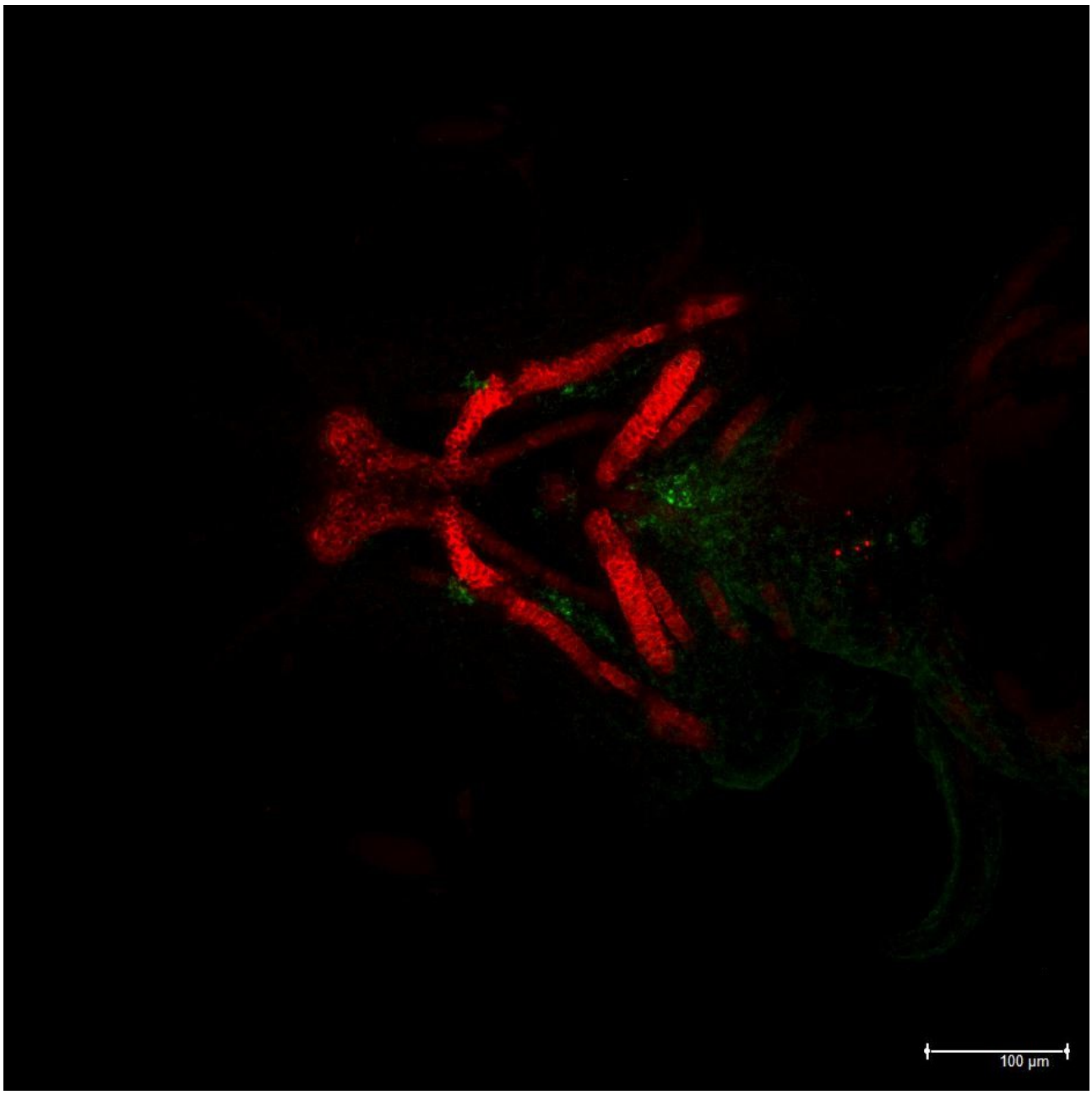

Figure 13. Maximum projection of confocal image stack showing expression of *scxa* at 3 dpf. The expression pattern for *scxa* is similar to *mkxa* (Figure 11).

4.5.4. *Scerlaxis B*

The same observations done with the fluorescent microscope were also seen in the confocal microscope for *scxb* (Figure 14). Expression patterns of the *scxa* and *scxb* were very similar. The two areas of expression close to the Meckel's-palatoquadrate cartilage joint and ventromedial to the palatoquadrate are at similar positions. But in the *scxb* case the expression pattern in the area of the sternohyoideus muscle attachment position is broader. Analysis of more specific areas by confocal microscopy detected expression in small populations of cells at several muscle attachments sites. The sites observed were the attachment of intermandibularis anterior to Meckel's cartilage, intermandibularis posterior to Meckel's cartilage and intermandibularis posterior to basihyal cartilage.

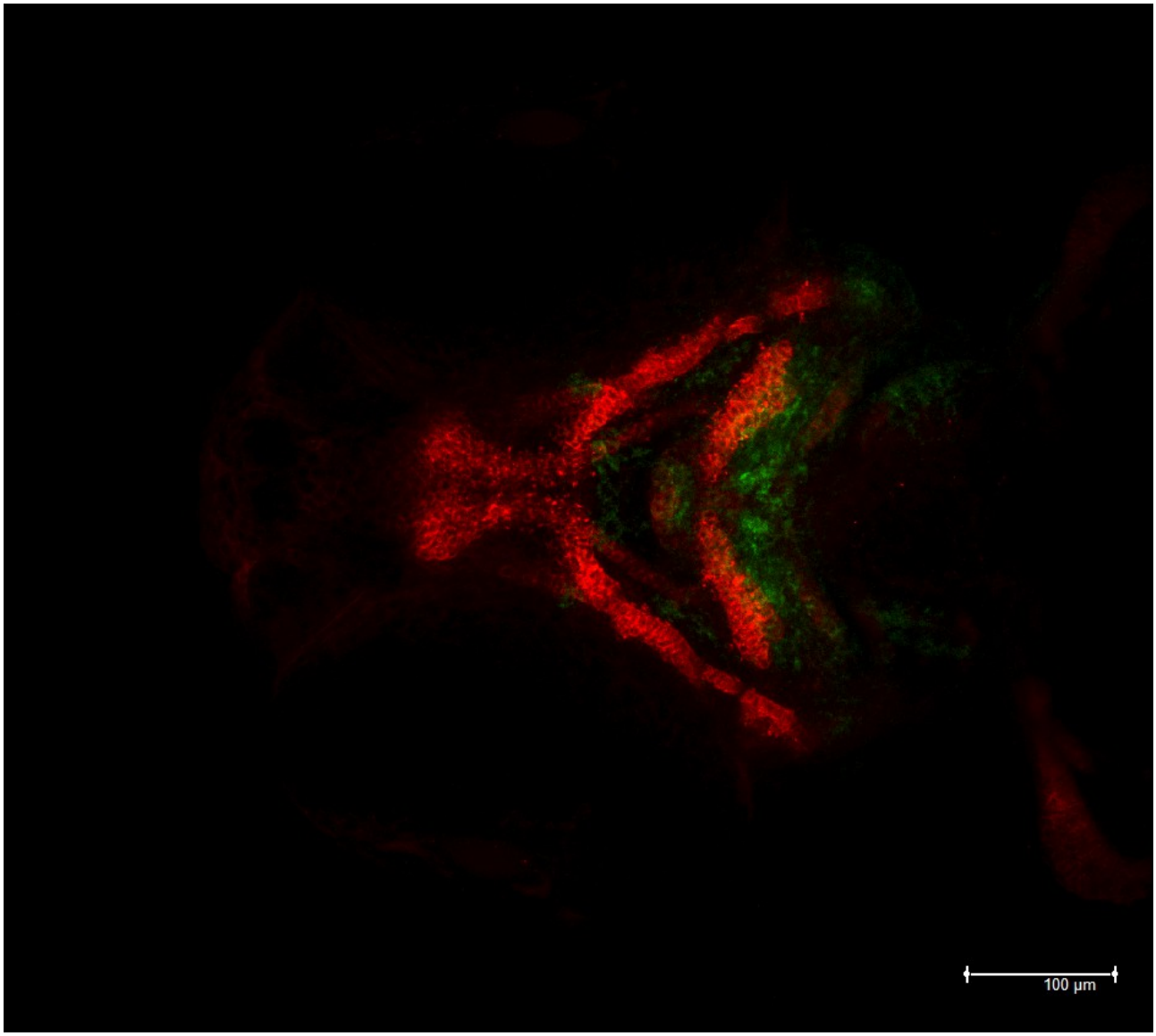

Figure 14. Maximum projection of confocal image stack showing expression of *scxb* at 3 dpf. The repeated signal detection done with other embryos that went through another round of FISH is consistent with the fluorescent microscope data.

4.5.5. *Early growth response 1*

In the tissue around newly formed pharyngeal arches, *egr1* expression can be detected (Figure 15). But no expression was detected close to any muscle attachment sites.

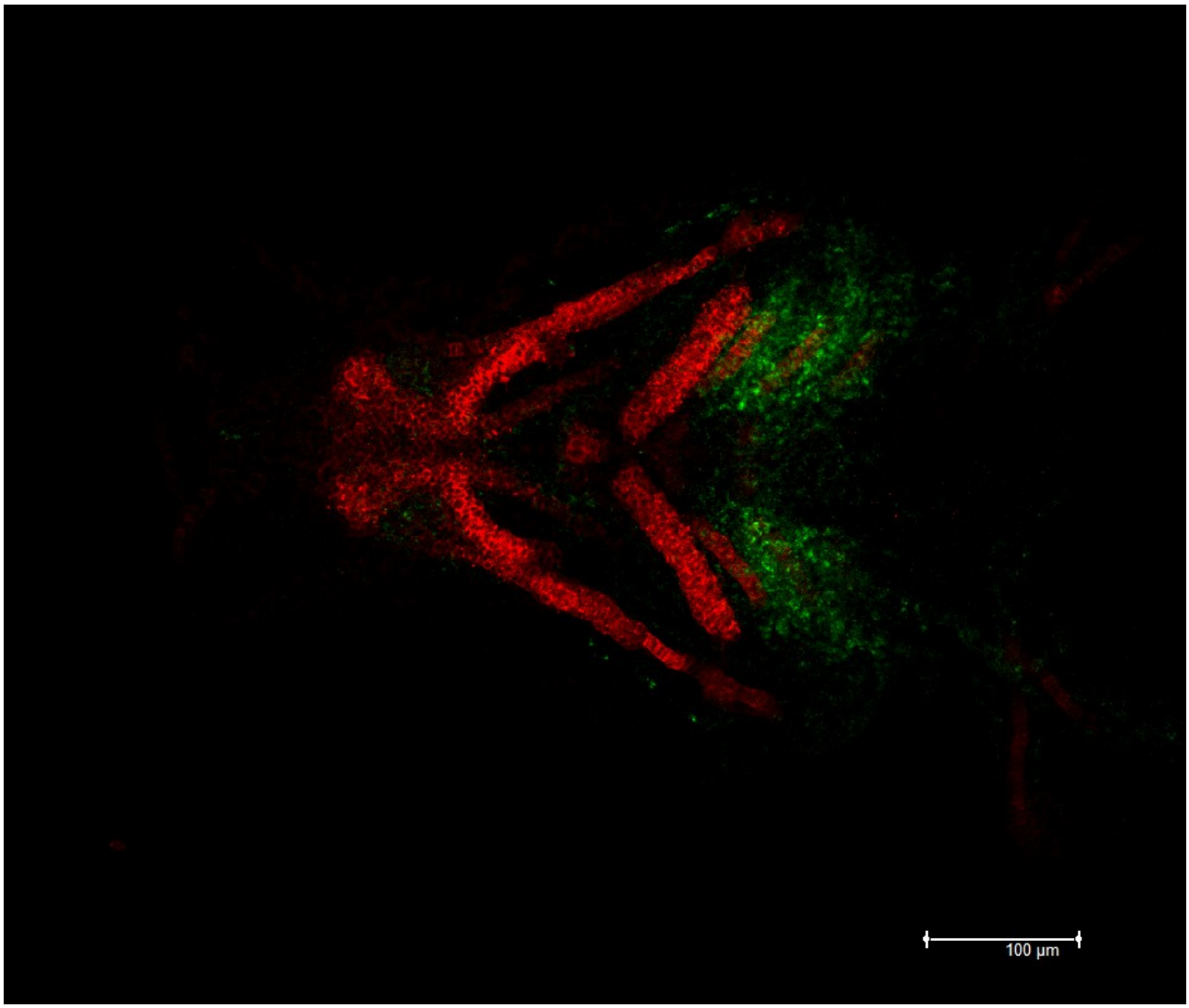

Figure 15. Maximum projection of confocal image stack showing expression of *egr1* at 3 dpf. The expression pattern can be observed around the pharyngeal arches 3 to 5.

5. Discussion

In this project, the expression of five transcription factors was investigated. The transcription factors *mkxa*, *mkxb*, *scxa* and *scxb* had expression detected close to muscle attachment sites. Therefore, they will be possible candidates in future functional studies.

No specific signal was observed in the negative control, and the other results were consistent with earlier studies (3,4). This indicates that the elevated signal detected from the FISH procedure corresponds to expression of the studied gene. Even though a lot of effort was put into eliminating background and optimizing the FISH protocol there was still a lot of autoflorescence in the skin. Since autoflourescense in the skin of embryos is inevitable (18), and the time was limited, the procedure first applied to *egr1* was followed throughout the project.

The developmental stage with the most interesting signal was 3 dpf. For this developmental stage, signal was detected in the vicinity of multiple muscle attachment sites. Expression signals at 4 dpf were also in most of the cases very similar to 3 dpf, but weaker. In the case for *mkxb*, no signal was visible at 4 dpf. It is unclear whether the weaker signal at 4 dpf depended on less expression or on the method procedure. At 4 dpf it could be the case that less probe was able to penetrate the embryo. All embryos at 5 dpf showed similar diffuse expression pattern from the FISH (Figure S2). Because of very different expression pattern compared to earlier developmental stages, this could most likely be explained by the fact that the protocol used for FISH was not optimized for 5 dpf. This might depend on insufficient Proteinase K digestion.

The only gene analysed that had no expression close to a muscle attachment site was *egr1*. Therefore, *egr1* is probably the least interesting gene candidate to investigate further for information about which cells differentiate into muscle attachments.

Interestingly, *mkxb* was detected during a shorter period of time compared to *mkxa* and in a much smaller area. The limited time *mkxb* expression could be detected indicate that timing of this gene expression might be important. This also indicates a very specific function for Mkxb in a small subset of cells.

The muscle attachment site from the anterior part of the sternohyoideus muscle is very long and extends all the way to the basihyal (3). The expression of *mkxb* seem to correspond to the anterior part of this attachment (Figure 12). The expression pattern of *scxa*, *scxb* and *mkxa* seem to have a more diverse positioning. Expression signal could both be detected in the vicinity of muscle attachment sites, but also as a stripe of expression ventromedial to the palatoquadrate. This stripe is not close to where any muscle attach. The area rather corresponds to where the posterior end of Meckel's cartilage is connected by ligament to the hyoid arch. It seems more likely that *mkxa* and both the *scx* duplicates have a more diverse function.

Similar expression pattern could be found in two small areas in *mkxa*, *scxa* and *scxb*. Those two areas of expression are located where adductor mandibulae attaches to Meckel's cartilage. All three genes also share expression at the point where the sternohyoideus attach to the basihyal cartilage. Therefore, they all seem to be important for muscle attachment differentiation. Expression ventromedial to the palatoquadrate is shared by *mkxa*, *scxa* and *scxb* as well.

The importance of *mkxa* for muscle attachment development is demonstrated by Chuang and colleagues (4). In this study, morpholinos is injected and the expression of *myoD*, a muscle specific marker (19), is analysed. At 48 and 58 hpf the development of muscles are disrupted, probably because of their inability to attach properly.

Even though *mkxa*, *scxa* and *scxb* shared a lot of similarities, differences could be observed. The shape of the expression pattern observed in the area of the sternohyoideus muscle attachment to the ceratohyal is shared between *mkxa* and *scxa*, but not *scxb*. The area of expression is broader in *scxb* and extends towards a more distal part of the ceratohyal cartilage. Another difference is the stripe ventromedial to the palatoquadrate. The stripe is more extended towards hyosymplectic in *mkxa* compared to observations in *scxa* and *scxb*. A more detailed analysis by confocal microscopy revealed expression of *scxb* that was not detected for *mkxa* or *scxa*. The expression of *scxb* could be observed where the intermandibularis posterior muscle attach both to Meckel's and basihyal cartilage. In a smaller population of cells, expression where the intermandibularis anterior muscle attach to Meckel's cartilage could be detected in both *mkxa* and *scxb* but not *scxa*.

Although *scxa, scxb* and *mkxa* had very similar expression patterns at 3 dpf, the genes did not share the similarity at 48 hpf. *Scxa* and *scxb* had similar expression detected in the whole head with a labelled midline, while *mkxa* only had some expression detected around the structures where the jaws will appear (data not shown). Those observations could indicate that Mkxa, Scxa and Scxb have different functions at 48 hpf and act from a different level in their pathways. But still, since the transcription factors share a similar expression pattern at 3 dpf, they probably act in the same type of cells. This is actually the case according to earlier published data. According to earlier studies, Mohawk controls tendon differentiation (6), and Scleraxis is suggested to be needed for early tendon differentiation (7).

To reveal more about the relationship between *mkxa*, *scxa* and *scxb*, additional developmental stages should be included in the ISH between 48 hpf and 3 dpf. That experiment would show which one of the transcription factors that is expressed first according to the pattern observed at 3 dpf.

Based on the similarity observed at 48 hpf, 3 dpf and 4 dpf between the *scx* duplicates, they could share a similar function or even act redundantly. In mice that only have a single *scx* gene, several defects could be observed in the *Scx*$^{-/-}$ mice (7). An attempt to knockdown *scxa* and *scxb* expression in zebrafish by morpholino injection was made by Chen and Galloway (3). No change in jaw morphology could be observed in this experiment. This could probably be explained by the possibility that *scx* is expressed too late for the injected morpholino to be effective. In future functional analysis, a double knockout of *scxa* and *scxb* should be made.

6. Acknowledgements

I would like to sincerely thank my supervisor Tatjana Haitina for incredible supervision throughout this project. All the effort she has put into teaching me new techniques have been invaluable and all the time she has spent on answering my questions have been very appreciated. I am happy to have worked with her during my undergraduate years and she has inspired me to continue with research at academic level.

I would like to thank Xesus Abalo as well, for critically reviewing this report and for all the rewarding discussions along the time of the project. I would also like to express my thanks to Cécile Jolly who has been available for questions regarding the *in situ* protocol.

7. References

1. Cerny R, Cattell M, Sauka-Spengler T, Bronner-Fraser M, Yu F, Medeiros DM. Evidence for the prepattern/cooption model of vertebrate jaw evolution. Proc Natl Acad Sci U S A. 2010 Oct 5;107(40):17262–7.

2. Schilling TF, Piotrowski T, Grandel H, Brand M, Heisenberg CP, Jiang YJ, et al. Jaw and branchial arch mutants in zebrafish I: branchial arches. Dev Camb Engl. 1996 Dec;123:329–44.

3. Chen JW, Galloway JL. The development of zebrafish tendon and ligament progenitors. Dev Camb Engl. 2014 May;141(10):2035–45.

4. Chuang H-N, Cheng H-Y, Hsiao K-M, Lin C-W, Lin M-L, Pan H. The zebrafish homeobox gene irxl1 is required for brain and pharyngeal arch morphogenesis. Dev Dyn Off Publ Am Assoc Anat. 2010 Feb;239(2):639–50.

5. Close R, Toro S, Martial JA, Muller M. Expression of the zinc finger Egr1 gene during zebrafish embryonic development. Mech Dev. 2002 Oct;118(1-2):269–72.

6. Ito Y, Toriuchi N, Yoshitaka T, Ueno-Kudoh H, Sato T, Yokoyama S, et al. The Mohawk homeobox gene is a critical regulator of tendon differentiation. Proc Natl Acad Sci U S A. 2010 Jun 8;107(23):10538–42.

7. Murchison ND, Price BA, Conner DA, Keene DR, Olson EN, Tabin CJ, et al. Regulation of tendon differentiation by scleraxis distinguishes force-transmitting tendons from muscle-anchoring tendons. Dev Camb Engl. 2007 Jul;134(14):2697–708.

8. Brent AE, Schweitzer R, Tabin CJ. A somitic compartment of tendon progenitors. Cell. 2003 Apr 18;113(2):235–48.

9. Peterson SM, Freeman JL. RNA Isolation from Embryonic Zebrafish and cDNA Synthesis for Gene Expression Analysis. J Vis Exp JoVE [Internet]. 2009 Aug 7 [cited 2014 Jul 31];(30). Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3152201/

10. Addgene - ZP17 Krox20 Plasmid Data [Internet]. [cited 2014 Aug 3]. Available from: https://www.addgene.org/16924/

11. QIAGEN - Sample & Assay Technologies [Internet]. [cited 2014 Aug 3]. Available from: http://www.qiagen.com/resources/resourcedetail?id=05356d52-ff29-4ac2-be49-48cba2381cfa&lang=en

12. QIAGEN - Sample & Assay Technologies [Internet]. [cited 2014 Aug 3]. Available from: http://www.qiagen.com/resources/resourcedetail?id=f9b2e5ef-9456-431a-85ed-2a2b9fbd503d&lang=en

13. Hauptmann G, Gerster T. Multicolor whole-mount in situ hybridization. Methods Mol Biol Clifton NJ. 2000;137:139–48.

14. Dale RM, Topczewski J. Identification of an evolutionarily conserved regulatory element of the zebrafish col2a1a gene. Dev Biol. 2011 Sep 15;357(2):518–31.

15. Support for NanoDrop Products - www.nanodrop.com [Internet]. [cited 2014 Aug 4]. Available from: http://www.nanodrop.com/Support.aspx?Type=User%20Guides&Cat=NanoDrop%202000c

16. UCSC In-Silico PCR [Internet]. [cited 2014 Aug 3]. Available from: http://genome.ucsc.edu/cgi-bin/hgPcr?command=start

17. Welten MCM, de Haan SB, van den Boogert N, Noordermeer JN, Lamers GEM, Spaink HP, et al. ZebraFISH: fluorescent in situ hybridization protocol and three-dimensional imaging of gene expression patterns. Zebrafish. 2006;3(4):465–76.

18. Shi X, Shin Teo L, Pan X, Chong S-W, Kraut R, Korzh V, et al. Probing events with single molecule sensitivity in zebrafish and Drosophila embryos by fluorescence correlation spectroscopy. Dev Dyn. 2009 Dec 1;238(12):3156–67.

19. Bentzinger CF, Wang YX, Rudnicki MA. Building Muscle: Molecular Regulation of Myogenesis. Cold Spring Harb Perspect Biol. 2012 Feb 1;4(2):a008342.

20. Ensembl Genome Browser [Internet]. [cited 2014 Aug 13]. Available from: http://www.ensembl.org/index.html

21. Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. Bioinforma Oxf Engl. 2007 May 15;23(10):1289–91.

22. Zebrafish BLAT Search [Internet]. [cited 2014 Aug 14]. Available from: https://genome.ucsc.edu/cgi-bin/hgBlat?command=start

8. Supplementary data

8.1. Reagents

The reagents and kits used in alphabetic order together with their reference number and manufacturer:

Agarose	A9539-250G	Sigma-Aldrich
Alexa Flour 594 goat anti-rabbit fab2 (H+L)	A11072	Life technologies
Anti-DIG-POD	11 207 733 910	Roche
Anti-GFP rabbit	A11122	Life technologies
Blocking reagent	11 096 176 001	Roche
Chloroform	32211	Sigma-Aldrich
Denhardt's solution 50X	D2532	Sigma-Aldrich
Dextran sulfate	D8906-106	Sigma-Aldrich
DIG labeling mix	11 277 073 910	Roche
DMSO	D8418	Sigma-Aldrich
dNTP mix:		
Adenosin	D4788-25UMO	Sigma-Aldrich
Cytosine	D4913-25UMO	Sigma-Aldrich
Guanine	D5038-25UMO	Sigma-Aldrich
Thymidine	T9656-25UMO	Sigma-Aldrich
EDTA	E9884-500G	Sigma-Aldrich
Ethanol	200-578-6	Solveco
Formamide	47670-1L-F	Sigma-Aldrich
GelRed 10 000X	41003	Biotium
Heparin sodium salt	H-3393	Sigma-Aldrich
Hyaluronidase	H3506-100MG	Sigma-Aldrich
Hydrogen peroxide 30%	23615.248	VWR
Loading dye 6X	#R0611	Fermentas
Methanol	20847.295	VWR
NaCl	71376-1KG	Sigma-Aldrich
PFA	P6148-500G	Sigma-Aldrich
QIAquick PCR purification kit	28104	Qiagen

RNaseOUT™	10777-019	Invitrogen
RNeasy mini kit	74104	Qiagen
Salomon sperm DNA	15632011	Life technologies
Sheep serum	S-2263	Sigma-Aldrich
SuperScript™ III reverse transcriptase	18080-044	Invitrogen
T3 polymerase	11 031 163 001	Roche
T7 polymerase	10 881 775 001	Roche
TAE buffer 40X	V4281	Promega
Taq polymerase	D6677-1.5KV	Sigma-Aldrich
Transcription buffer 10X	11 464 384 001	Roche
Triton 100X	T-9284	Sigma-Aldrich
Trizma® base	T1503	Sigma-Aldrich
Trizma® hydrochloride	T5941	Sigma-Aldrich
Trizol	15596-026	Invitrogen
tRNA	R6625-256	Sigma-Aldrich
Trypsin	T-7409	Sigma-Aldrich
TSA™ Plus Fluorescein System	NEL741	PerkinElmer
Tween-20	822184	Merck
PBS	P4417-100TAB	Sigma-Aldrich
Proteinase K	03 115 887 001	Roche
RNEasy mini kit	74104	Qiagen
SSC 20X	161-0775	Bio-Rad
XbaI	#ER0681	Fermentas

8.2. Zebrafish housing

Zebrafish (*Danio rerio*) AB and a Tg(col2a1a:membr EGFP) strain embryos were obtained by natural spawning and maintained at 28.5 °C in system water with methylene blue. Experiments performed were approved by ethical permission C 262/11.

8.3. Primers

The first step in the primer design process was to identify the exons that were expressed in the most common splice variants from data in the ensemble database (20). The identified exon sequences were then used in the online software primer3 (21) that can design primers based on desired product length and the template sequence used. In the next step, the primer sequences were entered into a BLAT search (22). This was done to make sure that the paralogue would not be amplified by the primer pair.

A T7 5' overhang (5'CTGTAATACGACTCACTATAGGG3') was added to the reverse primer to enable direct transcription from the PCR product. The primers were ordered from Sigma-Aldrich. The following primers were used to amplify the parts of the gene candidates that would be used as probes:

mkxa-forward: 5'AGGCGAATGAGGTGGAAAGA3'

mkxa-reverse: 5'CTGTAATACGACTCACTATAGGGTCTCCCTCCAGTACGTCTCA3'

mkxb-forward: 5'AGGAAAGCGGACAAAATGTGG3'

mkxb-reverse: 5'CTGTAATACGACTCACTATAGGGAGTCTCTCTGCGTTGCCTT3'

scxa-forward: 5'CGCTCATTTCACCTGGACAC3'

scxa-reverse: 5'CTGTAATACGACTCACTATAGGGTAACTCCTCAGGGCGGATTT3'

scxb-forward: 5'CATGTCTTTTGCGATGGTGC3'

scxb-reverse: 5'CTGTAATACGACTCACTATAGGGGATCTGTCTGGGCTGTGAGT3'

egr1-forward: 5'CACGTCTTCCATCCCCTCTT3'

egr1-reverse: 5'CTGTAATACGACTCACTATAGGGCACTGGTGAGGAAGCTGAGA3'

8.4. Gel analysis

To be able to run the transcription reaction directly from the PCR product it need to be free from off target effects. In general this means that the annealing temperature should be put as high as possible but still at a level where the primer can hybridised to the template. This were successfully done for all gene candidates and the predicted fragment length where attained (Figure S1).

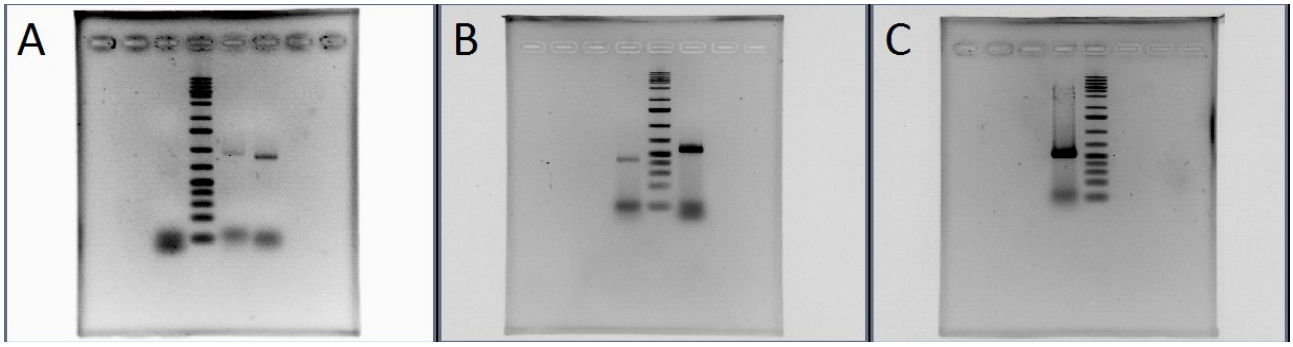

Figure S1. Gel analysis of PCR product. a) Well 1 – *scxb* (532 bp), well 2 – ladder, well 3 – *mkxa* (844 bp), well 4 – *egr1* (820 bp). b) Well 1 – *mkxb* (400 bp), well 2 – ladder, well 3 *scxb* (532 bp). c) Well 1 – *scxa* (503 bp), well 2 – ladder. The length of the first seven bands for the ladder used, from the shortest to the longest: 75 bp, 200 bp, 300 bp, 400 bp, 500 bp, 700 bp, 1000 bp.

8.5. FISH 5 dpf results

The protocol seemed to not be optimized for embryos at 5 dpf. Similar diffuse signals were obtained in all embryos (Figure S2), even for *mkxb* that showed no expression pattern at 4 dpf from the FISH.

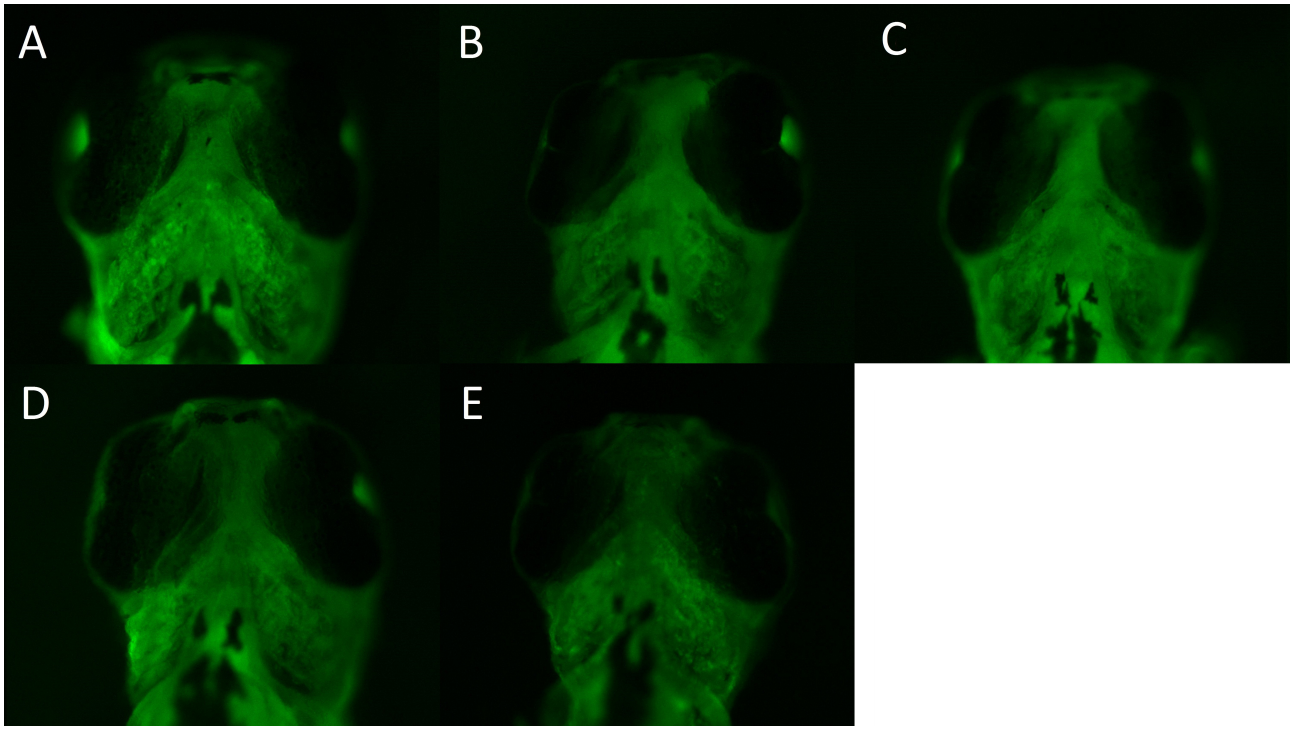

Figure S2. Fluorescent microscopy of *mkxa*, *mkxb*, *scxa*, *scxb* and *egr1* at 5 dpf ventrally. a) *mkxa* b) *mkxb* c) *scxa* d) *scxb* e) *egr1*