universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Interaction partners of Brachyury and FoxA in *Nematostella vectensis* investigated by RIME"

verfasst von / submitted by

Laura Unger, B.Sc.

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna, 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 877
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Genetik und Entwicklungsbiologie
Betreut von / Supervisor:	Univ.-Prof. Dipl.-Biol. Dr. Ulrich Technau

Acknowledgements

Firstly, I would like to thank Uli Technau for giving me the opportunity to perform my master project in his lab and for his advice and feedback during my lab work and writing.
Also, I want to thank Wolfram Weckwerth for his advice within the cooperation and providing access to the mass spectrometer in his Molecular Systems Biology lab. Thanks to the MoSys lab members Stefanie Wienkoop for her advice, Sonja Tischler for measuring my samples and my special thanks to Martin Brenner for introducing me to the *Maxquant* software and giving me valuable insights for result analysis.
I am especially grateful towards Patricio Ferrer Murguia who mentored me during my lab work and supported me in any situation. Thanks to Grigory Genikhovich for sharing his expertise with various methods and matching of my results to the Nve database. Furthermore, I would like to recognize the invaluable assistance that all colleagues of the Technau Lab provided during my study and who made this thesis possible.
Also, I want to acknowledge the Core Facility for Cell Imaging and Ultrastructure Research for access to the confocal laser scanning microscope.
Last but not least thanks to my parents for their unconditional support during my studies and to my sisters and David for their encouragement.

Content

1. Introduction

1.1 Combinatorial control of gene regulation by transcription factors

Transcription factors (TFs) are gene regulatory proteins, that directly interact with specific regulatory elements on the DNA to regulate gene expression. Thereby they can act as activators or repressors and enable targeted control of single genes (Locker, 2001). They bind specific control sequences, also termed cis-regulatory elements, which are noncoding DNA sequences that navigate transcription of neighboring genes (Swarr et al., 2019). The TFs bind the DNA with their DNA-binding domain (e.g. helix-turn-helix or homeodomain) that categorize them into different families. Thereby the exact amino acid sequence determines which DNA motif is recognized. Furthermore, TFs possess an activation domain (AD) with which they interact with the transcription machinery. If the TF operates activating, it recruits the transcription machinery, including very abundant general transcription factors, to the promoter of the targeted gene. Thereby the cis-regulatory DNA sequence can lie very distant to the gene`s promoter upstream or downstream and a given target gene can have several of cis-regulatory elements that can be bound by several different gene regulatory proteins. In eukaryotes a single TF binding to a gene`s regulatory element is usually not sufficient, but it requires interaction of several TFs and other gene regulatory proteins in different extent, to activate or suppress transcription. TFs bind as homomers or homodimers to the DNA, but some can also bind as heterodimers, which is one way of combinatorial control of gene expression. Functioning heterodimerization increases the DNA binding possibilities but it can be used as inhibitory regulation too. Another method of combinatorial gene control is gene regulatory proteins binding in complexes to the DNA, even several different complexes can work together (Alberts et al., 2011). Thereby the TF`s activation domain or transactivation domain (TAD) is interacting with co-activators to consequently recruit them to the gene regulation site. Combinatorial gene control combines incoming information, ensuring high specificity and precision in gene regulation hence enabling higher complexity of an organism (Locker, 2001).

A gene regulatory network (GRN) collects the interplay between transcription factors and their binding sites on the DNA to regulate gene expression (Yachie-Kinoshita and Kaizu, 2019). These GRNs have a crucial role in setting up the body plan in a developing organism (Botman et al., 2014). Deciphering them, answers how any given developmental process is controlled by heritable DNA sequences. Ultimately differences in specific linkages can explain phenotypical divergence between organisms although possessing similar gene sets (Levine and Davidson, 2005). To solve a GRN, knowledge about the involved transcription factors and signal molecules needs to be acquainted, the timepoints and localization of their expression and their interactions with each other (Li and Davidson, 2009).

1.2 Differences in GRNs setting up the mesendoderm in Bilateria and Cnidaria

Cnidarians, comprising of sea anemones, corals, jellyfish and hydroids, are a basal metazoan phylum and the most closely related sister group to the Bilateria (e.g. insects, humans and worms) (Darling et al., 2005). Their evolutionary split dates back around 600–700 million years. Compared to bilaterians, cnidarians show a simple physiology and low biodiversity. A cnidarian model organism is the sea anemone *Nematostella vectensis* (*Nv*) which belongs to the supposedly most basally branching class of Anthozoa. Like the other cnidarian species *Nv* is a diploblastic animal, consisting of only two cell layers. The outer ectoderm provides a boundary to the environment while the pharyngeal ectoderm and the internal bifunctional (mes)endoderm are in charge of nutrient absorption and muscle movements. Cnidarian species were thought to have only the oral–aboral axis, giving them a radial symmetry but in many members, including *Nv*, a bilateral symmetry was discovered (Layden et al., 2016). On the contrary the bilaterians are triploblastic. They possess a third germ layer, the mesoderm and next to the oral–aboral, the dorsal–ventral axis, generating a bilateral symmetry (two symmetrical body halves) (Layden et al., 2016). The emergence of the third germ layer, the mesoderm, marks the evolutionary transition between diploblastic and triploblastic animals and it seems to be the crucial emergence that made bilaterian complexity and their broad biodiversity possible (Scholz and Technau, 2003). In support of this it has been argued that the mesoderm is important during embryogenesis to induce other tissues and lastly it promotes morphogenetic processes essential during development (Technau, 2001). The common germ-layer-homology-theory assumes that the cnidarian (mes)endoderm is homolog to the bilaterian endo- and mesoderm. During evolution a subpopulation of cells from the border between ecto- and endoderm conducting mesoderm-typical functions could have segregated into a separate cell layer – the mesoderm (Scholz and Technau, 2003; Technau, 2001). An opposing theory is that the mesoderm arose from the ectoderm (Servetnick et al., 2017). However, in 2017 Steinmetz et al. challenged these theories by proposing that the mesoderm is not a bilaterian-specific feature but that the cnidarian endoderm corresponds to the bilaterian mesoderm and the pharyngeal ectoderm to the bilaterian endoderm. Possibly those were already separate prior to the bilaterian-cnidarian ancestor (Steinmetz et al., 2017).

To find out, understanding the ancestral conditions of the ancestor of all bilaterians is of most interest. Therefore, an experimentally amenable, slowly evolving outgroup among the Cnidaria is needed (Layden et al., 2016). Comparisons between bilaterian and cnidarian development already revealed some traits that were thought to be novel bilaterian inventions but are older than the split between those phyla (Layden et al., 2016). For example it was thought that body plans and genome complexity go hand in hand which inferred that the evolutionary increase in complexity of gene regulatory networks is the reason why bilaterians became more complex than cnidarians (Layden et al., 2016). However, morphologically simple metazoans such as cnidarians have a genetically surprising complex genome. When the *Nematostella* draft genome was published (Putnam et al., 2007), it was found to contain a lot of gene families also present

in bilaterians (Genikhovich and Technau, 2009c). Furthermore, *Nematostella* possesses many genes necessary for mesoderm development in bilaterians including orthologs of *forkhead* and *brachyury* (Darling et al., 2005). This indicates that these groups share a common "genetic toolkit" but then evolved independently (Darling et al., 2005). Therefore, researchers investigate the GRN responsible for the specification of the cnidarian ecto- and endoderm and the bilaterian endo- and mesoderm. What are the differences and how did the bilaterian GRN evolve? It was found that the canonical Wnt pathway is a critical upstream regulator for the mesendodermal GRN in both (Rottinger et al., 2012). Also *Nematostella`s* pharyngeal ectoderm has genetic and functional similarities with the bilaterian endoderm and *Nv`*s endoderm has similarities to the bilaterian mesoderm (Steinmetz et al., 2017).
Still, further functional analysis of genes showing conserved mesodermal expression in cnidaria are needed (Technau, 2001). Finally understanding the genes and GRNs that led to the evolution of mesoderm will help to understand the evolution of bilaterian body plans (Scholz and Technau, 2003).

1.2.1 A key TF for bilaterian mesoderm development - Braychury- in the mesoderm-less-*Nematostella*

Nematostella possesses genes that are important players in mesoderm formation in bilateria. One key transcription factor for bilaterian mesoderm differentiation is *brachyury* (*bra*). It belongs to the T-box gene superfamily from which other members play important roles in mesendoderm differentiation too (Technau, 2001). They are transcriptional regulators involved in tissue specification, morphogenesis and organogenesis. T-box genes code for proteins with a 180 amino acids long DNA-binding domain, the T-domain. The Brachyury protein (Bra) consists of 393 amino acids with a molecular weight of 43,896 Da. In *Nematostella,* the Bra binding motif is a 20-base pair long palindromic sequence, to which the protein binds as a dimer interacting with major and minor grooves of a 24 nucleotide DNA duplex (Müller and Herrmann, 1997). As monomer it can also bind to motifs consisting only of half of the palindromic sequence to regulate gene expression (Casey et al., 1998). Furthermore, it can interact with other proteins or cofactors which consequently affects target gene recognition and regulation. For example several T-box proteins are known to bind homeobox factors and then synergistically activate target genes during development (Papaioannou, 2014). The amino acid sequence of the T-box domain is highly conserved and shows about ~80% identity to vertebrate Brachyury. This conservation of the DNA binding domain suggests that it binds to the same DNA motifs as in vertebrates and as a result regulates the same genes. The C-terminal activation domain is much less conserved (Scholz and Technau, 2003; Technau, 2001). However, a mutation in the transactivation domain leads to a strong phenotype that is very similar in zebrafish, *Xenopus* and mice. This supports the view that this region is indispensable for the function of the gene in mesoderm patterning and suggests that Bra interacts and regulates downstream mesoderm-specific genes by interaction with cofactors (Conlon et al., 1996).

Whole mount in situ hybridizations revealed the *bra* expression pattern during development (Scholz and Technau, 2003). The *Nematostella* embryo is divided into five expression domains along the oral–aboral axis (Rottinger et al., 2012). Most orally lies the central domain which is the pre-endodermal plate. Adjacent lies the central ring from which ectodermal tissue e.g. of the ingrowing pharynx, will form. Aborally to the central ring lies the external ring and most aborally located the apical domain. *bra* expression appears already in the blastula broadly in the oral hemisphere (E. Haillot, unpublished). At the transition to gastrula stage it becomes confined to the central ring around the blastopore. Cells from this region are precursors of the pharynx that will separate endo- from ectodermal cells during gastrulation (Servetnick et al., 2017). Expression remains in the ectoderm around the central domain in the planula larva but during metamorphosis into the primary polyp the expression is found in the endodermal part of the blastopore. In primary polyps *bra* expression persists around the mouth and presumably in the growing mesenteries (Scholz and Technau, 2003). It was also observed in the adult gonads (male and female) (P. Ferrer Murguia, unpublished).
CRISPR/Cas9 knockout of *bra* in early *Nv*-embryos does not affect the onset of gastrulation, but later gastrulation stages are severely disturbed (Servetnick et al., 2017). It leads to failure of pharynx formation; the embryos cannot elongate their oral–aboral axis and have disturbances in axis patterning. The endoderm is still specified but disorganized and cell polarity in ectodermal cells is disturbed. As a consequence of Bra knockout many genes are affected, in fact inside and outside of its expression domain, including some expressed at the aboral pole. It seems like its ancestral function is not the initiation of blastopore formation, but it is important in morphogenesis, cell polarity and patterning of ecto- and endodermal derivatives along the oral-aboral axis. Its presumably central role in the GRN in the early embryo makes it therefore a particularly interesting example to study the evolution of developmental mechanisms (Servetnick et al., 2017). A lot of information about *brachyury* in different species has been gathered already but even more information about its network, including cofactors and upstream/downstream components is needed (Technau, 2001).

1.2.2 Interaction of the transcription factor FoxA with Bra?

Another conserved role in mesendodermal patterning in vertebrates, has the transcription factor *forkhead*. It is co-expressed with *brachyury* and they act synergistically to define the dorsal mesoderm. The *Nematostella foxA* full-length clone is 1774 bp long and the resulting protein 286 amino acids with a molecular weight of 32,165 Da. The Fox family is defined by the Forkhead domain towards the N-terminal end, which is a 110 amino acids long winged helix. The Forkhead domain is highly conserved with more than 95% amino acid identity to vertebrate Forkhead. At the C-terminal end are two smaller transactivation domain motifs also conserved between *Nv* and vertebrates.
In *Nv*, *foxA* is co-expressed with *bra* in the central ring ectoderm around the blastopore marking the boundary between ecto- and endoderm and it is expressed in the developing pharynx.

The theory is that those two also interact in *Nv* to regulate gastrulation and the development of structures like the pharynx. The fact that they also work together in vertebrates, leads to the suggestion that their functional relationship is conserved throughout eumetazoan evolution (Fritzenwanker et al., 2004). However, their physical interaction in *Nematostella* has not been proven since.

1.2.3 Interaction of members of the Sox transcription factor family with Bra?

Members of the *sox* gene family are known to be involved in gastrulation, development of the nervous system and also in mesendodermal patterning. They are defined by their DNA binding domain, the high-mobility group (HMG), which is highly conserved across species. Although outside of this domain their sequences are very divergent, their activity is very specific. This is probably due to the temporal and spatial expression specificity and combinatorial interactions with cofactors and other transcription factors (Fritzenwanker et al., 2004).
Koch and colleagues performed ChIP-seq in mouse neuro-mesodermal progenitor (NMP) cells against Sox2 and Bra (Koch et al., 2017). They found that they co-occupy a large fraction of genes (NMP genes and lineage control genes) whereby they antagonize each other in cells undergoing the lineage choice towards neural or mesodermal fate.
Nematostella possesses at least 14 *sox* genes of which many are involved in neural development, e.g. *Nvsoxb(2)* was found to promote the development of neurons and nematocytes (the stinging cells of cnidaria) (Richards and Rentzsch, 2014). Hence it is reasonable to assume that Sox transcription factors interact with Bra in *Nv* too, maybe also antagonizing each other to direct the cells to a certain fate.

1.3 The sea anemone *Nematostella vectensis* as cnidarian model organism in EvoDevo biology

The sea anemone *Nematostella vectensis* (*Nv*) has become a popular cnidarian model organism. It naturally lives in brackish salt water e.g. in permanent pools or estuaries with varying salinity and temperature (Hand and Uhlinger, 1992). These variable environmental conditions make it robust and easy to maintain in a laboratory. Furthermore, it reproduces frequently sexually and asexually. Sexual reproduction takes place via external fertilization, generating a zygote which undergoes several rounds of cleavage until it becomes a blastula. The blastula gastrulates by invagination through the blastopore (Kraus and Technau, 2006; Magie et al., 2007). Future endodermal cells invaginate inwards as an epithelial layer forming the pre-endodermal plate. The pre-endodermal plate invaginates further until it meets the outer ectodermal epithelium. During late gastrulation the pharynx forms by ectodermal cells of the central ring invading the gastric cavity (Servetnick et al., 2017). Progressively the gastrula develops into a free-swimming planula larva whereby the oral–aboral axis elongates. 7–9 days after fertilization the planula metamorphoses into a four-tentacle primary polyp. It reaches the adult polyp stage after around 3–6 months depending on food resources (Genikhovich and Technau, 2009c).

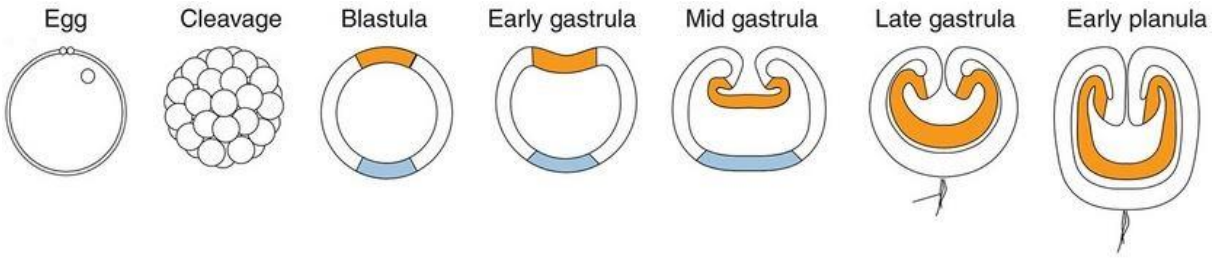

Fig. 1 Early development of *Nematostella vectensis*.
The orange domain resembles the (presumptive) endoderm. The blue domain marks the aboral ectoderm and white the remaining ectoderm. Oral is up in all images. Modified after (Layden et al., 2016).

All developmental stages are accessible to laboratory manipulation which makes it perfect for developmental studies (Darling et al., 2005). Asexual reproduction happens mainly by transverse fission and together with its regenerative capacity, underlying mechanisms of regeneration can be investigated (Hand and Uhlinger, 1992); (Layden et al., 2016).
As mentioned, *Nv* belongs to the Cnidaria and therefore, has an advantageous phylogenetic position.
These reasons make *Nematostella* a good model organism for EvoDevo research of early branching eumetazoans. It provides access to genomic, molecular, regenerative and developmental research. Combined with studies in other cnidarian organisms it can clarify cnidarian biology and gene regulatory networks and help to understand the formation of complex bilaterian body plans (Darling et al., 2005).

1.4 ChIP-seq against Brachyury in *Nematostella*

In an unpublished study by the Technau lab ChIP-seq against the transcription factor Brachyury in late gastrula *Nematostella* embryos (27 hours post fertilization (hpf)) was conducted. Chromatin-immunoprecipitation (ChIP) enables the identification of sites in the genome bound by a TF at a certain timepoint in vivo. The proteins become crosslinked to DNA in the living cells which are subsequently lysed and the DNA mechanically fragmented. Subsequently the TF of interest is immunoprecipitated by an antibody together with the bound DNA sequences. The precipitated DNA fragments are sequenced, and the underlying motifs occupied by the TF discovered by bioinformatic tools (Alberts et al., 2011). The binding sites identified for Bra overlapped with enhancers previously predicted by (Schwaiger et al., 2014) who had created a genome-wide map of gene regulatory elements in *Nv*.
The closest gene to the binding site was conventionally considered the target gene. Interestingly among the identified DNA motifs were not only Bra binding motifs (T-(half)-palindrome) but also DNA binding sites of other transcription factor families. 16.7% represented Sox motifs and 6% other motifs (R. Dnyansagar, B. Zimmermann, unpublished). This raised the idea that Brachyury does not always bind directly to the non-T-motifs but sometimes also indirectly via interaction with the corresponding (Sox) transcription factors.

1.5 RIME as proteomics tool to study gene regulatory networks

Since the turn of the millennium researchers of the "systems biology" discipline focus on exploring the dynamic interplay in a cell, investigating the dynamic networks between RNA, DNA and proteins. (Rehm and Letzel, 2010). The term "proteome" describes the whole set of proteins expressed by the genome of an organism, whereas "proteomics" covers also the structure, function and interactions of proteins in an organism. Proteomics revealed that not the number of genes but the interactions of proteins cause the large variety of observed life phenotypes (Mishra, 2010). An example is the human proteome which consists of about 25000 proteins but with 665000 estimated interactions between those (Stumpf et al., 2008).

To identify protein-protein interactions researchers used e.g. the Yeast two-hybrid system or more recently mass spectrometry (MS). MS revolutionized protein research, by substantially simplifying the identification of peptides and proteins. It facilitated an effective way to determine an amino acid sequence and with the help of genomics and bioinformatics match it to a certain protein and its underlying gene (Mishra, 2010). It is based on the principle that molecules can be ionized into charged ions which usually occurs by protonation or deprotonation. As the amino group of peptides/proteins tends to add an H^+, they become mainly positively ionized. These positively charged ions can be separated and analyzed based on their mass-to-charge (m/z) ratios. From that the ion masses can be concluded. Since each amino acid and as a result each peptide has a unique molecular weight (MW), the original polypeptide can be identified. The high precision allows it to detect even slight differences in MW and therefore changes in sequence or structure of a peptide(Mishra, 2010).

The mass spectrometer used in this project was the *Q Exactive™ Hybrid Quadrupole-Orbitrap™* Mass Spectrometer in the Molecular Systems Biology (mosys) Department of the University of Vienna. The following text gives a short technical overview of the MS procedure.

Before the samples are injected into the mass spectrometer itself, they are exposed to pre-mass spectrometry separation in this case by Reverse Phase Liquid Chromatography. The peptides bind reversibly to the hydrophobic matrix and dependent on their hydrophobicity are eluted sooner or later before being passed into the mass spectrometer (Mishra, 2010). The actual mass spectrometer consists of three main compounds: an ion source generating ionized molecules; one or more analyzers which separate the ionized molecules according to their m/z ratio and a detector that traces the separated ions (Rehm and Letzel, 2010).

The Q Exactive uses the Electrospray Ionization (ESI) as method of ionization. The protein/peptide solution becomes vaporized by high voltage and sprayed out of a capillary in tiny droplets which is generating ions. While being drawn to the opposite pole the droplets become smaller due to evaporation. Finally, they explode into tiniest droplets many containing just one ion. What stays is an ion gas (Rehm and Letzel, 2010). The ions come into the analyzer where they are separated by their mass to charge (m/z) ratio. In this case a Quadrupole mass filter. The ions pass the Quadrupole in an oscillating way as they are deflected and repelled by the changes in polarity. It is determined which m/z ratio enables an ion to pass the filter, all other

ions get lost. This way filtered they are eventually transferred into the C-Trap analyzer. There the ions lose their kinetic energy by collision with nitrogen gas. Afterwards the peptide ions are injected into the analyzer to receive the fragment mass spectra. However certain ions (20 most abundant ones) are exposed to Tandem-MS. In a HCD collision cell they are further fragmented into smaller ions. Those enter the analyzer a second time which provides higher sensitivity (creative-proteomics, 2019). The final Orbitrap mass analyzer is a spindle serving as a central electrode surrounded by outer electrodes. Depending on the ion`s mass (m) it goes narrower around the current or wider and depending on the charge (z) it goes quicker or slower. The outer electrodes detect the resulting image current of the oscillating ions. Via Fourier Transformation the frequencies of the oscillations and as a result the m/z ratios are obtained (Mishra, 2010). As mentioned from those the underlying molecular weight and consequently the amino acid sequence can be identified.

A method to identify (transient) protein complexes is the affinity purification of endogenous proteins coupled to mass spectrometry. Mohammed et al. developed a method called rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) with which protein complexes and in particular transcription factor complexes can be investigated. It promises to be a sensitive method, able to not only identify stable, highly affine interactions but also more transient and/or weak ones. It is divided into six main work steps. Firstly, the fixation of cells with formaldehyde leading to crosslinked protein complexes. This is followed by lysate preparation isolating the nuclear proteins, and immunoprecipitation against the protein of interest. The precipitate is trypsinized to produce peptides which are then measured by liquid chromatography-tandem mass spectrometry. Lastly the data is analysed computationally (Mohammed et al., 2016). With this "bottom-up" approach the sample protein(s) become proteolyzed and the molecular masses of the resulting peptide fragments determined. By alignment with peptide masses of known amino acid sequences from a protein databank the proteins can be identified (Mishra, 2010). With this tool protein interactions in an organism can be investigated. Together

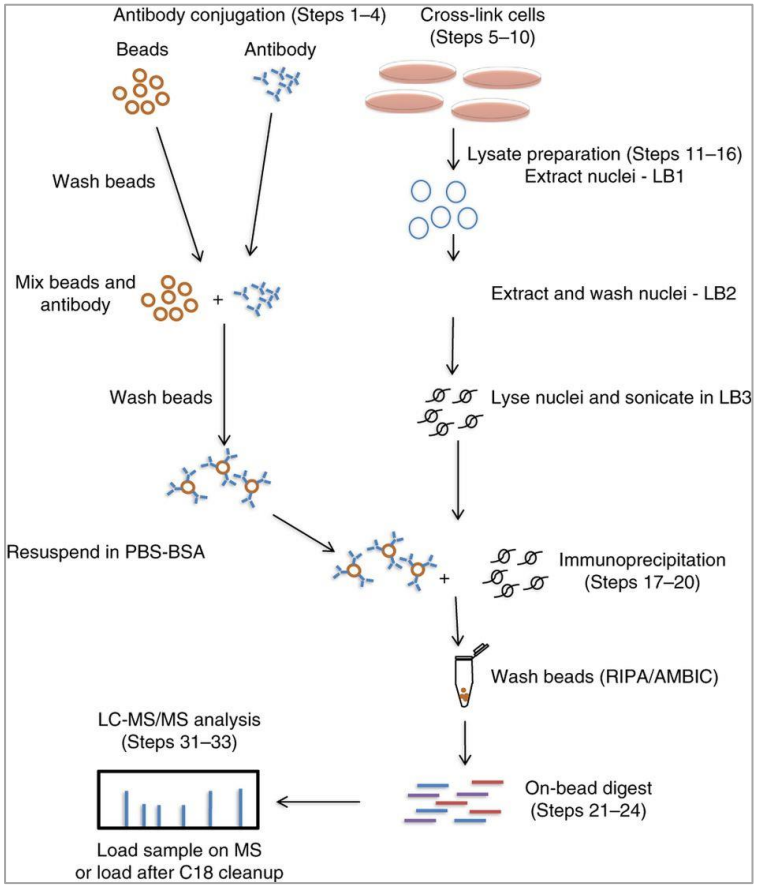

Fig. 2 Workflow of a RIME experiment.
Modified after (Mohammed et al., 2016).

with chromatin immunoprecipitation-sequencing (ChIP-seq) knowledge about interactions and cis regulation of a transcription factor can be combined, contributing to the exploration of gene regulatory networks (Mohammed et al., 2016).

1.6 Aim of my project

The aim was to establish a protocol for "rapid immunoprecipitation mass spectrometry of endogenous proteins" (RIME) in *Nematostella vectensis*. The approach was to combine the X-ChIP-seq protocol already established in the Technau laboratory with the RIME protocol of Mohammed et al., to develop a compatible protocol for *Nv*.
Eventually RIME against Bra and FoxA should be performed. Therefore, pre-experiments were necessary to determine a method that would reliably increase Brachyury expression in the embryos because it naturally occurs only in a fraction of cells, potentially not providing enough material for mass spectrometry. Hence the effects of two different inhibitors of Gsk3ß were tested. They mimic active Wnt-signalling which is ectopically activating the expression of the Wnt signaling target genes *bra* and *foxA*.
With that RIME was supposed to help identify interaction partners of the transcription factor Brachyury in *Nematostella* at 27 hours post fertilization, the same developmental stage at which the ChIP-seq against Bra had been performed. The results can be matched to potentially support the theory of Bra binding to cis-regulatory elements on the DNA in heterodimerization or complexes with certain transcription factors from other families. This would provide further information about Brachyury`s involvement in the gene regulatory network active during gastrulation and mesendodermal patterning. Ultimately comparisons with bilaterian GRNs during segregation of the germ layers can be refined. This finally leads to better understanding of the evolution of the mesoderm and the high bilaterian complexity.
As only little Bra antibody was left, the protocol should firstly be tested with the FoxA antibody to investigate the interactome of the transcription factor FoxA too. These results can be combined with the results of a ChIP-Seq against FoxA in *Nematostella*, which is planned for the near future in the Technau group.

2. Materials & Methods

2.1 Work with *Nematostella vectensis*

2.1.1 Animal culture, induction of spawning and in vitro fertilization

The animals are kept in boxes filled with artificial sea water with 16‰ salt (*Nematostella* medium/NM) at 18°C in the dark. The boxes are under regular low medium exchange made possible by a recirculation system. They are fed five times a week with *Artemia salina* larvae to allow sufficient gamete production. Males and females are kept in separate plastic boxes (Hand and Uhlinger, 1992; Genikhovich and Technau, 2009c).
For in vitro fertilization the protocol described in (Genikhovich and Technau, 2009b) is as followed. Spawning is induced by putting the animals for 10 hours under light and 25°C. After a span of 1.5-2 hours back at 18°C, the egg packages from the female boxes can be collected and then added to the sperm rich NM of the male boxes. With this in-vitro-fertilization it is made sure the embryos` development is synchronized. The zygotes are released from the surrounding jelly by treating them in 1% cysteine at pH 7.5 on a shaker for 20-40 minutes and washed afterwards to remove the cysteine (Fritzenwanker and Technau, 2002; Genikhovich and Technau, 2009b). Eventually they are put in a 21°C-incubator until they reach the desired developmental stage.

2.1.2 Treatment of embryos with inhibitors for ectopic activation of the canonical Wnt pathway

Because a relatively high amount of proteins is needed for mass spectrometry analysis, a lot of starting material must be available. Mohammed et al. advise 10^8 cells for a protein not previously optimized, corresponding to around 14000 *Nematostella* embryos to be harvested and preprocessed. In the template experiment conducted by Mohammed et al. they had used cell cultures with all the cells expressing the antigen. As *brachyury* and *foxA* are expressed only in a small subpopulation of cells around the blastopore, it would have required a higher number of embryos. Therefore, to achieve an ubiquitous expression of these transcription factors was the first experimental step of this project.
Wnt/ß-catenin signaling establishes the oral-aboral axis and defines the oral pole by activating oral marker genes in Cnidaria (Leclere et al., 2016). We took advantage of the fact that among these, *bra* and *foxA* are downstream of active Wnt/ß-catenin signaling in *Nematostella* (Rottinger et al., 2012). A known selective inhibitor of glycogen synthase kinase-3 beta (Gsk3ß) is 1-azakenpaullone (Azk). As Gsk3ß is part of the destruction complex which is degrading ß-catenin, its inhibition prevents the degradation of ß-catenin (see Fig. 3). Instead it accumulates and gets into the nucleus binding transcription factors (TCFs) to activate target genes. Hence Azk mimics active Wnt signaling (Kunick et al., 2004). Another Gsk3ß inhibitor is CHIR99021

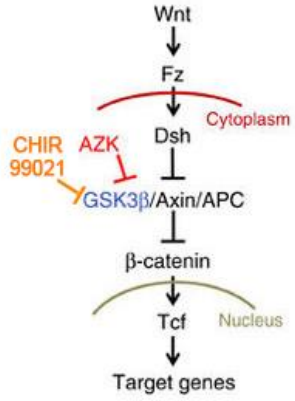

Fig. 3 The inhibitors AZK and CHIR99021 activate Wnt/β-catenin signalling by inhibiting GSK3β. Modified after (Kraus et al., 2016).

(CHIR) (Bennett et al., 2002), which has been reported to be the strongest and most specific Gsk3ß inhibitor (Bain et al., 2007). However, Azk has been used in *Nematostella* frequently, e.g. by (Marlow et al., 2013), (Kraus et al., 2016) but CHIR not. To decide which one to use for this experiment, their concentration-dependent effect was compared by applying three different ascending concentrations and analyzing the effects on *bra*, *foxA* and certain *sox* genes with in situ hybridization, immunohistochemistry and Western Blots.

After approximately 4 hpf at 21°C when most of the embryos had reached early cleavage, they were transferred in a petri dish with fresh NM. Different concentrations of the inhibitors 1-Azakenpaullone (9-Bromo-7,12-dihydro-pyrido[3′,2′:2,3]azepino [4,5-b]indol-6(5H)-one, Sigma-Aldrich, #A3734) or CHIR99021 (6-[[2-[[4-(2,4-Dichlorophenyl)-5-(5-methyl-1H-imidazol-2-yl)-2-pyrimidinyl]amino]ethyl]amino]-3-pyridinecarbonitrile, Sigma-Aldrich, #SML1046) were added to the medium. The concentrations tested were 5 µM, 10 µM and 15 µM. As control served the highest concentration of 53 µM of Dimethyl sulfoxide (DMSO) in which the inhibitors were solved in. The embryos were treated from early cleavage to late-gastrula (~27 hpf) at 21°C for approximately 23 hours, based on the setup used by (Kraus et al., 2016). Eventually they were washed four times with NM and ready for further processing. They were fixed for in situ hybridization (see 2.4) and immunohistochemistry (see 2.5) to visualize the inhibitors` effects on *brachyury* and *foxA*-expression. Furthermore, protein extracts from embryos undergone the different treatments were analyzed by Western Blots against Brachyury and FoxA (see 2.6.4). When the most efficient treatment had been determined, it was used to prepare embryos for RIME (see 2.7) and Magnetic Co-IP (see 2.8) for subsequent mass spectrometry (see 2.8).

2.2 Fixation of embryos

2.2.1 Fixation for Whole Mount in situ Hybridization

For whole mount in situ hybridization (ISH) embryos were fixed based on the "Fixations Nematostella Guide" from the Technau Lab in 10 ml falcon tubes or 2 ml Eppendorf tubes with all steps at 4°C. In the first rounds of ISHs they were fixed in fresh 3.7% formaldehyde/2.5% glutaraldehyde in 1x PBS with 0.1% Tween 20, Triton X-100 and DMSO. First, they were put into the glutaraldehyde solution until they sunk down and then immediately transferred to the formaldehyde solution. They were left on a rotator for 1 h at 4°C and then washed with PTw five times, followed by one wash step in 60% Methanol/PTw and two in 100% Methanol. After the second wash they were rotated another hour at 4°C and then stored in a fresh tube at -20°C ready for in situ hybridization (see 2.4). The second rounds of ISHs, embryos were fixed in 4%

PFA in 1x PBS for ~18 hours on a rotator at 4°C and then washed and stored the same way as described previously.

2.2.2 Fixation for Immunohistochemistry

For immunostaining embryos were fixed in fresh 3.7% formaldehyde dissolved in 1x PBS with 0.2% Triton X-100 and Tween 20. They were rotated for 1 h at 4°C and eventually washed ten times in 1x PBS with 0.2% Triton X-100 and 0.2% Tween 20. Subsequently the immunostaining protocol was followed (see 2.5).

2.2.3 Fixation for RIME

The RIME fixation procedure is based on the "Nematostella_embryo_X-ChIP_better_nuclei_110925withDTT-2" protocol from Schauer & Genikhovich adapted from (Schwaiger et al., 2014) where in brief, the embryos were fixed with 1.85% formaldehyde for 15 minutes. It was combined with the Nature protocol from Mohammed et al., 2016 where they suggest 1% for 8 minutes. Based on this, I tested different formaldehyde concentrations and different fixation times.

Paro Fixation Solution

HEPES pH 8.0 (NaOH)	50 mM	Concentration in Paro Fixation Solution but becomes diluted eventually in Crosslinking solution!
EDTA	1 mM	
EGTA	0.5 mM	
NaCl	100 mM	
Formaldehyde	1.85% or 1% or 0.5%	Must be final concentration in Crosslinking solution!
ELIX H2O	Fill up to total volume	

Crosslinking Solution (with final concentration of 1.85% or 1% or 0.5% formaldehyde)

Paro Fixation Solution	1/4	Becomes diluted 1:4 with PBS
1x PBS	3/4	

HEG Buffer

HEPES pH 7.5 (KOH)	50 mM
EDTA	1 mM
Glycerol	20 %

After the inhibitor treatment (see 2.1.2) the embryos were washed four times with *Nematostella* Medium. Then they were fixed in 15 ml falcon tubes with the Crosslinking solution, rotating slowly at room temperature. The different fixation conditions tested were: 1.85% for 15 minutes, 1.85%, 1%, 0.5% for 8 minutes and 1% for 5 minutes. In all conditions in the last 3 minutes of fixation, the embryos were given time to sink down. Afterwards they were immediately washed two times with ice cold 1x PBS on ice to remove the formaldehyde. Then they were washed once in PBS with 100 mM Glycine and 0.01% Triton X-100 for 10 minutes on a rotator at 4°C to stop the fixation, followed by two washes in pure PBS for 10 minutes also on

the rotator at 4°C. In case of not immediately proceeding with the protocol the embryos were aliquoted in 2 ml tubes and equilibrated in HEG Buffer. After taking off as much supernatant as possible they were snap-frozen in liquid nitrogen and stored at -80°C. If the embryos were sinking too slowly to the bottom, they were centrifuged at 400 g for 2 minutes at 4°C. The further steps of RIME are described further down (see 2.7).

2.3 Generation of anti-sense probes for ISH

The generation of anti-sense probes for the later whole mount in situ hybridizations followed the protocol from the practical course "Übungen in Zell- und Entwicklungsbiologie", 2017 with slight deviations as indicated below.

2.3.1 RNA extraction and reverse transcription into cDNA

First step was the RNA isolation from *Nematostella* in different developmental stages with Trizol, Chloroform and Isopropanol. Afterwards the RNA concentration and quality were measured with the *NanoDrop 2000 Spectrophotometer* (ThermoScientific) and controlled with an 1%-agarose gel using *SYBR Safe* (ThermoScientific) as nucleic acid stain. In 20 µl reverse transcription reactions the isolated RNA was turned into single-stranded cDNA by the *SuperScript™ III Reverse Transcriptase* (Invitrogen) with random hexamers serving as primers.

Random Hexamers (20 µM)	2.5 µl
dNTP Mix (10 mM each)	1 µl
RNA	5 µg
RNAse-free H2O	Add up to 13 ul

After denaturation at 65°C for 5 minutes in the PCR block the following components were added:

1st Strand Buffer (5x)	4 µl
Dithiothreitol (100 mM)	1 µl
RNAse OUT (40U/µl)	1 µl
SuperScript III RT (200U/µl)	1 ul

The reaction was incubated at 25°C for 5 minutes, then at 50°C for 90 minutes and finally inactivated at 70°C for 15 minutes in the PCR block. The finished cDNA was stored at -20°C.

2.3.2 PCR of gene of interest, ligation into cloning vector and *E. coli* transformation

Out of the cDNA, the genes of interest were amplified by Polymerase Chain Reaction (PCR). Therefore, the *Taq DNA Polymerase* (ThermoScientific) or the *Q5® High-Fidelity DNA Polymerase* (NEB) and gene-specific primers were used which were designed with help of the website *Primer3web* (http://primer3.ut.ee, version 4.1.0). The list of primers used for this work can be found in the supplement.

ddH2O	Add up to 25 µl
Gene-specific Forward Primer (10 µM)	2.5 µl
Gene-specific Reverse Primer (10 µM)	2.5 µl
dNTPs (each 10 mM)	0.625 ul
cDNA template	0.5 µl
Taq PCR Buffer (10x)/Q5 Reaction Buffer (5x)	2.5 µl/5 µl
Taq Polymerase (5U/µl)/Q5 Polymerase (2U/µl)	0.2 µl/0.5 µl

	Taq DNA Polymerase		**Q5 High Fidelity DNA**	
1 x	95°C	5 min	95°C	1 min
40 x	95°C	30 s	95°C	30 s
	Primer-dependent	30 s	Primer-dependent	30 s
	72°C	1 min per 1 kb	72°C	30 s per 1 kb
1 x	72°C	10 min	72°C	5 min
1 x	10°C	Pause	10°C	Pause

25 µl of the PCR product with 5 µl 6x loading dye were then loaded on a 1%-agarose gel and ran for 35 minutes at 120 V. The double stranded cDNA got purified and cleaned from the gel via the *peqGold Gel extraction Kit* (Peqlab).
The cDNA of the desired gene was cloned into cloning vectors with the help of *T4 DNA Ligase* (Promega). If Taq polymerase was used, into the *pGEM-T Easy* vector (Promega) with a single 3`-terminal thymidine at both ends because this polymerase synthesizes sticky ends in the PCR. If Q5 Polymerase was used, into the *pJET1.2/blunt* Vector (ThermoScientific) because Q5 synthesizes blunt ends.

Rapid Ligation Buffer (2x)	2.5 µl
pGEM-T Easy/pJET1.2 Vector	0.5 µl
PCR Product	1.5 µl
T4 DNA Ligase (3 U/µl)	0.5 ul

The reaction was left at room temperature for 30 minutes. With the ligated vectors, competent TOP10 *E. coli* got transformed. They were thawed on ice and 5 µl ligation reaction was added to 50 µl cell aliquot and left on ice for 10 minutes. This was followed by a heat-shock at 42°C for 45 seconds. Afterwards they were immediately put on ice again for 2 minutes. Then 250 µl of SOC-medium without antibiotics was added and the bacteria solutions were plated on petri dishes with ampicillin-containing medium and incubated over night at 37°C.

2.3.3 Insert orientation check and probe template linearization

The next day five colonies that had been able to grow were picked and transferred to new plates with a tip. To check the insert orientations from each picked clone`s insert, two colony PCRs per clone were done. Both with a gene-specific primer (either forward or reverse) plus the forward vector primer in one reaction and the reverse primer in the other one. If the insert

was in *pGEM-T Easy* the M13 Primers were used and if it was in *pJET1.2* the *pJet* forward and reverse primers were used.

ddH2O	Add up to 20 µl
Gene-specific forward or reverse Primer (10 µM)	1 µl
Vector-specific forward Primer (10 µM)	1 µl
Vector-specific reverse Primer (10 µM)	
dNTPs (each 10 mM)	0.3 ul
DNA template (transformed E. coli)	2 µl
Taq PCR Buffer (10x)	2 µl
Taq Polymerase (5U/µl)	0.08 µl

1 x	95°C	5 min
25 x	95°C	20 s
	50°C	30 s
	72°C	1 min per 1 kb
1 x	10°C	Pause

With gel electrophoresis the insert orientations were identified. The chosen and regrown clones were bred in 5 ml supermedium containing ampicillin overnight in a glass reaction tube shaking at 37°C. The next day the plasmid templates were purified from the bacteria with the *QIAprep Spin Miniprep Kit* (Qiagen). The plasmids were sent for Sanger-Sequencing to Microsynth Austria to make sure the desired sequence got inserted.

Before RNA-synthesis the plasmids had to be linearized by PCR with standard primers binding to the vector outside the insert. For *pGEM-T Easy* the M13F and R primers were used and for *pJET1.2* depending on the insert orientation, either pJetF plus sp6-outer primer or pJetR plus T7-outer primer because *pJET1.2* does not contain a sp6/T7 promoter site and it has to be added by these ”outer” primers.

ddH2O	Add up to 100 µl
M13F/pJetF/pJetR Primer (10 µM)	5 µl
M13R/sp6-outer/T7-outer Primer (10 µM)	5 µl
dNTPs (each 10 mM)	1.5 ul
Plasmid template	1 µl
Taq PCR Buffer (10x)	10 µl
Taq Polymerase (5U/µl)	0.4 µl

1 x	95°C	1 min
35 x	95°C	20 s
	50°C	30 s
1 x	10°C	Pause

In the follow the PCR products were run on a 1% agarose gel and the amplified sequences purified from the remaining vector by the *peqGold Gel Extraction Kit* (Peqlab).

2.3.4 RNA probe synthesis, precipitation and dilution

Finally, the Digoxigenin-labelled RNA probe got synthesized in an in-vitro transcription. Depending on the insert orientation the *MEGAscript SP6 Kit* (Ambion) or the *HiScribe T7 High Yield RNA Synthesis Kit* (NEB) were used to produce an antisense RNA probe. The reaction incubated over night at 37°C.

T7 or SP6 Transcription Buffer (10x)	1 µl
Purified PCR Product	5 µl
DIG-UTP Labeling Mix (10x)	2 µl
RNAse OUT (40U/µl)	1 ul
T7 or SP6 Enzyme Mix (10x)	1 µl

The next day 0.75 µl *DNase I* was added to stop the reaction and incubated for 20 minutes at 37°C. Eventually the synthesized RNA probe had to be precipitated with 5 µl RNase free H_2O and 5 µl Lithium chloride for at least 2 hours at -20°C. After a 10-minute centrifugation at maximum speed at 4°C, the supernatant was removed, and the pellet washed two times with ice-cold 70 % DEPEC Ethanol always followed by a centrifugation step. Eventually the supernatant was removed, the pellet air-dried in the PCR block and finally resuspended in 30 µl H_2O. The RNA concentration was measured with the *NanoDrop 2000 spectrophotometer* (Thermo-Scientifc) and a 1% gel was run to check the probe size. If alright, the RNA probe was diluted to a concentration of 50 ng/µl in 50% formamide and got stored at -20°C.

2.4 Whole mount in situ hybridization

Whole mount in situ hybridizations (ISHs) were performed against the mRNA of *foxA* and *brachyury* to reveal the effects of the different Azk- and CHIR99021-treatments on their expression and thus determine the most efficient treatment (see 2.1.2). Furthermore, ISH against 12 *sox* genes were performed. As mentioned in the introduction (see 1.4) Sox protein binding motifs have been identified in a Bra ChIP-seq, making them potential interaction partners of Brachyury ~27 hpf. The in situs can reveal potential co-expression which is the precondition for interaction. Also, in situs were conducted against the *sox* genes in treated embryos to check the effects of ectopic Wnt signaling on their expression too.
The procedure followed the "*In Situ Hybridization Protocol for Nematostella*", 2015 from Eduard Renfer based on (Genikhovich and Technau, 2009a), with some deviations. It was performed with the help of the *BioLane HTI 16Vx* (Intavis) in situ machine.

Hybridization-Buffer (Hybe)

Formamide (100%)	50%
SSC pH 4.5 (20x)	5x
SDS (10%)	1%
Heparin (20 mg/ml)	100 µg/ml
Tween 20	0.1%
Torula RNA	5 mg/ml
ELIX H2O	Fill up to total volume

All steps on the first day took place at room temperature unless stated otherwise. The embryos in methanol (see 2.2.1) had to be transferred to self-made sieves with cut tips or glass pipets. In a tray they came into the in situ machine. It started with rehydrating the fixed embryos from Methanol by one 5-minute wash in 50% Methanol/PTw and three PTw washes. Then they were digested by 80 µg/ml *Proteinase K* (20 mg/ml correspond to 1 U/µl, Ambion) for 3 minutes to facilitate the later penetration by the labeled RNA. The digestion was stopped by two glycine washes (4 mg/ml) in PTw. To block the charged residues, they were washed two times 5 minutes in 1% triethanolamine (TEA) in PTw, followed by one wash in 1% TEA plus 3 µl acetic anhydride/ml and one wash in 1% TEA plus 6 µl acetic anhydride/ml. After two 5-minute washes in PTw, the animals got refixed to stabilize them after the harsh treatment in 3.7% formaldehyde in PTw for 1 hour. They were washed five times 5 minutes in PTw, 10 minutes in 50% Hybridization Buffer (Hybe)/ PTw, 10 minutes in 100% Hybe and were finally pre-hybridized for 2 hours at 60–64°C in 100% Hybe to block all unspecific nucleic acid binding sites. The antisense RNA probes (see 2.3) were diluted to 1 ng/µl in 250 µl Hybe and pipetted in the wells of a flat rack. The sieves were added to the wells and the embryos finally hybridized with the Hybe-Mix containing the Digoxigenin-labeled probe between one to three nights at hybridization temperatures between 60–65°C in a water bath.

Maleic Acid Buffer

Maleic Acid	100 mM	
NaCl	150 mM	
ELIX H2O	Fill up to total volume	Adjust pH to 7.5 by NaOH!

Blocking Solution

Blocking reagent (Sigma Aldrich)	1%
Maleic Acid Buffer pH 7.5	1x

On day 2 of the protocol, followed the post-hybridization washes at hybridization temperature in which the reaction stringency got increased by increasing the pH-value and lowering the salt concentration stepwise. Once in 100% Hybe for 10 minutes, for 30 minutes in 60% Hybe/40% 2x SSCT pH 7.0, for 30 minutes in 30% Hybe/70% 2x SSCT and for 30 minutes in 2xSSCT. Then three times for 20 minutes in 0.075x SSCT. From here all further steps took place at room temperature. The embryos were washed for 10 minutes in PTw and then blocked for at least 1 hour in blocking solution. Just like 300 µl per sieve of the α-Digoxigenin antibody

coupled to alkaline phosphatase (AP) (Sigma-Aldrich) diluted to 1:2000. After blocking, the embryos were incubated over night at 4°C on a shaking table with the antibody.

Alkaline Phosphatase Buffer

NaCl (5 M)	100 mM
MgCl2 (1M)	50 mM
Tris pH 9.5 (1M)	100 mM
Tween20	0.1%
ELIX H2O	Fill up to total volume

On day 3 the embryos were washed ten times with PTw for 10 minutes, followed by two 5-minute washes in Alkaline-Phosphatase (AP) Buffer. Next the embryos were rinsed in well-plates and incubated in the staining solution, consisting of 4.5 µl nitro blue tetrazolium (NBT) and 3.5 µl 5-bromo-4-chloro-3-indolyl phosphate (BCIP) per 1 ml AP-Buffer in the dark. After staining the reactions were stopped with 1 ml 100% Ethanol for 1 hour at different time points between 1 hour to 24 hours. The Ethanol was replaced by 1 ml PTw and lastly the embryos were infiltrated in Glycerol over night at 4°C. The next day they could be mounted for microscopy.

2.5 Immunohistochemistry

Immunohistochemistry was also performed to visualize the expression patterns of *bra* and *foxA* and the effects of the different treatments (see 2.1.2) on their expression. However, immunohistochemistry gave a higher resolution and made the single nuclei plus the potential expression fluorescently visible.

Blocking Solution

Sheep Serum	20%
Bovine Serum Albumin	1%
PBS (1x)/0.2% Triton X-100/0.2% Tween 20	79%

The fixed embryos (see 2.2.2) and the primary antibodies were blocked separately for 2 hours in blocking solution at room temperature on a rocking table. For the FoxA and Bra antibody I used a dilution of 1:500. The embryos were incubated with the antibody over night at 4°C on a rocking table. The next day they were washed ten times for 10 minutes each with 1x PBS/0.2% Triton X-100/0.2% Tween 20. Eventually the animals and the secondary antibody together with the Phalloidin staining dye and 4′,6-Diamidin-2-phenylindol (DAPI) were blocked with blocking solution for 1 hour at room temperature in the dark on a rocking table. I used a 1:1000 dilution of anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 568, a 1:30 dilution of Alexa Fluor 488 Phalloidin Antibody and 1:1000 dilution of DAPI. The embryos were incubated over night at 4°C on a rocking table in the dark. On the last day the embryos were washed again ten times for 10 minutes with 1x PBS/0.2% Triton X-100/Tween 20. As much PBS as possible was removed and two to three droplets of Vectashield (Vector Laboratories) were added.

The samples infiltrated over night at 4°C in the dark before they could be mounted on the following day.

Pictures of the immuno-stained embryos (taken with confocal laser scanning microscopy) showed the nuclear localization of Brachyury and FoxA. With the *Cell Counter* plugin from Fiji (Schindelin et al., 2012) the occurrence of DAPI fluorescence (nuclei) colocalizing with fluorescent signal of the secondary antibody (Bra/FoxA transcription factors) was counted. The counts were compared between the untreated embryos and the embryos treated with the different treatments to identify the most efficient one.

2.6 Protein work

2.6.1 Protein extraction from *Nematostella*

To show the effect of the different treatments (see 2.1.2) at a protein level, Western Blots were performed with the protein extracts of untreated and differently treated gastrulae (27 hpf). The proteins were extracted from around 200 embryos per sample. This way it was made sure that the Blot was quantitatively evaluable. 250 embryos were counted directly after cysteine treatment (see 2.1.1) and then 200 that had developed properly at ~27 hpf directly before extraction. They were dissolved in 50 µl Cell Extraction Buffer + 1 mM phenylmethylsulphonyl fluoride (PMSF) by repeated pipetting on ice. The protein concentration was then measured by a Bradford protein assay (see 2.6.2) and adjusted to the same concentration. The extracts were frozen at -80°C and eventually used in Western Blots (see 2.6.4).

2.6.2 Bradford assay

To determine the relative amounts of proteins before Western Blots and after immunoprecipitations, Bradford protein assays were conducted. For each sample to be measured 400 µl ELIX H_2O were mixed with 100 µl 5x *Roti-Quant* (Roth) solution and 1 µl of sample. The mixture was incubated at least 5 minutes to maximum 45 minutes and then pipetted into a cuvette. The sample`s absorbance at 595 nm was measured in a *BioPhotometer* (Eppendorf). First a standard curve with known concentrations of BSA (0.1 mg/ml, 1 mg/ml and 10 mg/ml) was generated. With this the protein concentrations of the samples, based on their absorbance values, could be calculated.

2.6.3 Denaturing (SDS) discontinous PAGE: Laemmli method

In a sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) proteins become separated by their molecular size under denaturing conditions. By a following Coomassie Blue staining the protein composition in samples before and after immunoprecipitation could be compared (see 2.6.5). Furthermore, the protein separation is necessary for further investigation

in Western Blots. The protocol followed can be found in the book: *Short Protocols in Molecular Biology*, 5th edition, Wiley & Sons, Inc..

4x Laemmli Buffer

For 100 ml:

Tris-HCl pH 6.8	3 g	Store at room temperature
Glycerol (40%)	40 ml	
SDS (5%)	5 g	
Bromophenol Blue	5 mg	
ELIX H2O	Fill up to 100 ml	
ß-Mercaptoethanol	10%	Add fresh before use!

5x Protein Buffer pH 8.3

For 1 l:

Tris Base	15.1 g
Glycine	94 g
SDS (10%)	50 ml
ELIX H2O	Fill up to 1 l

The gels consist of a stacking gel on top with 3.9% acrylamide which stacks all proteins. As a result, they start migrating through the underlying separating gel at the same time. For the separating gel I chose 12% acrylamide. The recipes are enough for two SDS-Page gels:

	Stacking gel		**Separating Gel**	
30% Acrylamide/0.8% Bisacrylamide		0.65 ml		4 ml
Tris Cl/SDS (4x)	pH 6.8	1.25 ml	pH 8.8	2.5 ml
ELIX H_2O		3.05 ml		3.5 ml
Ammonium Persulfate (10%)		25 µl		33 µl
TEMED		5 µl		6.6 µl

Per gel two clean glass plates with a spacer in between were locked together with pressure clamps and placed onto a casting stand. First the separating gel solution was prepared and then carefully without bubbles, pipetted in the glass plate sandwich. The top of the gel was then covered with ELIX H_2O to generate a smooth surface. After 30 – 60 minutes of polymerization the water got removed and the interspace dried with *Whatman* Filter Paper. Now the stacking gel could be prepared and pipetted on top. Instantly the comb got placed in the stacking gel to form the wells for sample loading. After another 30 – 60 minutes of polymerization the comb could be removed, and the wells were rinsed with water and dried with *Whatman* paper again. 15 µl of protein extracts (with the approximate same amount of protein) were mixed with 5 µl 4x Laemmli Buffer and denatured at 99°C in the *Thermomixer comfort* (Eppendorf) for 5 minutes. The samples were given time to cool down before being pipetted into the wells. As a ladder served the *Color Protein Standard-Broad Range* (NEB). At least two sandwiches had to be placed upright in the electrophoresis chamber (Bio-Rad) which became filled with 1x Protein Buffer up to the marking. Also the space between the two sandwiches was filled

with 1x Protein Buffer so that the wells were covered. Finally, the samples could be carefully loaded with gel loading pipet tips and the chamber was connected to a *PowerPac Basic Power Supply* (Bio-Rad). The gel ran at 100 V and between 4 – 400 mA for 2 – 3.5 hours until the size marker of the smallest protein of interest reached the bottom of the gel. Subsequently the proteins in the gel were either blotted onto a membrane (see 2.6.4) or detected by a Coomassie Blue Staining (see 2.6.5).

2.6.4 Western Blot

Western Blots (WB) were performed to ascertain the differences in expression levels of Brachyury and FoxA between untreated embryos and embryos being subjected to different treatments (see 2.1.2). Also, they were used to survey the efficiency of the (co-)immunoprecipitations of Bra/FoxA during RIME or magnetic Co-IP (see 2.7.3.2, 2.8.2). The protocol followed can also be found in the book: *Short Protocols in Molecular Biology*, 5th edition, Wiley & Sons, Inc..

2.6.4.1 Protein blotting by tank (wet) electrotransfer

1x Transfer Buffer pH 8.3

For 2 l:

Tris Base	11.6 g
Glycine	5.8 g
SDS (10%)	5 ml
Methanol	400 ml
ELIX H2O	Fill up to 2 l

10x TBS pH 7.6

For 1 l:

Tris Base	24 g
NaCl	88 g
ELIX H_2O	Fill up to 1 l

1x TBST

TBS (10x)	1x
Tween 20	0.1%
ELIX H_2O	Fill up to total volume

With the help of a small applicator the sandwiches (see 2.6.3) were opened, the stacking gels cut off and the separating gels carefully transferred into 1x Transfer Buffer. The gels equilibrated for 30 minutes in the buffer. Per gel two Whatman Papers and one Nitrocellulose membrane were trimmed to the same size. The membrane also equilibrated in Transfer Buffer for 15 minutes. Eventually the transfer sandwich for blotting was assembled in a tray covered with Transfer Buffer to avoid trapping of air bubbles. On the black bottom of a plastic transfer cassette (will become cathode side) the components were placed in the following order: sponge, Whatman Paper, equilibrated separating gel, Nitrocellulose membrane, Whatman paper, second sponge. This sandwich was placed in a gel transfer cell with the black side adjacent to the black side of the cell. The cell was put in the electrophoresis chamber filled with 1x Transfer Buffer up to the marking. To cool the system during blotting, the chamber was surrounded by ice and ice-filled falcons were placed in the chamber itself. The chamber was connected to a

PowerPac™ HC High-Current Power Supply (Bio-Rad) and the proteins from the gel blotted to the membrane for 1 hour at 100 V.

2.6.4.2 Immunodetection with primary and secondary antibodies

Blocking Solution

TBST	1x
Bovine Serum Albumin	2.5%
Milk Powder	2.5%

Membranes with the blotted proteins on it got washed three times for 5 minutes in 1x TBST and then blocked in a tray with blocking solution for at least 1 hour at room temperature or at 4°C overnight on a rocking table. The antibodies used for immunodetection of the proteins of interest (α-FoxA and α-Bra) were diluted 1:5000 in 5 ml blocking solution in 15 ml falcon tubes. For finding the most efficient treatment I needed a loading control to see that the same amount of protein was loaded on the gel in the first place. As loading control for the FoxA WB served an antibody against a housekeeping protein, α-ß-Actin. However, α-ß-Actin and α-Bra could not be used on the same membrane as their molecular size is very similar and the bands would not have been distinguishable. Eventually the membranes were rolled and put in the falcon tubes with the blocked antibodies, in a way that the protein surface faced inside, and they were not overlapping themselves. This was left rotating over night at 4°C. Next day the membranes were washed three times for 10 minutes with 1x TBST and blocked again for 1 hour with blocking solution. Then they were incubated with the secondary antibody for 1 hour at room temperature on the rocking table. As secondary antibody served the α-rabbit coupled to horseradish peroxidase in a 1:100000 dilution. Finally, the membranes were washed again three times for 5 minutes in 1x TBST and were now ready for evaluation.

2.6.4.3 Chemiluminescent signal detection

As a last step the proteins of interest were made visible by the chemiluminescent reaction of the horseradish peroxidase coupled to the secondary antibody with the substrate provided. The first WB membranes were documented on X-Ray paper but later WBS digitally. As substrate for the horseradish peroxidase served the *Stable Peroxide Solution* and the *Luminol/Enhancer Solution* of the *SuperSignal™ West Femto Maximum Sensitivity Substrate* Kit (ThermoScientific). Those two were mixed in equal parts, whereby 1.5 ml was sufficient for one membrane. For X-Ray pictures I had to use a dark room. The membrane was covered with the substrate mix and wrapped in a transparent sheet. Then it was placed in a X-Ray film cassette and carefully covered with a piece of X-Ray film with forceps. The cassette was closed between 45 seconds to 2 minutes. Afterwards the X-Ray film was brought straight away into the Developer solution. When it had turned dark or respectively the staining did not strengthen any further, it was transferred to the Fixer solution where it was left for at least several minutes. Finally, the membrane was washed with water.

For later WBs, digital pictures were taken directly of the membrane. The membrane was also covered with the substrate mix and then wrapped in a transparent sheet. It was put in the *EpiChemi*^{3} *Darkroom* (UVP) and the door closed. With the camera on top pictures of the chemiluminescence coming from the reaction were taken. Additionally, to show the protein ladder a picture with the overhead-white-light was taken.

2.6.5 Coomassie Brilliant Blue staining

The Coomassie Brilliant Blue dye binds to the proteins separated in the SDS-PAGE. This way differences in protein content before and after IP could be visualized. As a result, the antibodies` efficiency in the immunoprecipitation could be evaluated. Furthermore, it could be compared if the different fixations influence the IP and which condition looks the most promising.

Coomassie Blue Staining Solution

For 1 l:

Tris Base	15.1 g
Glycine	94 g
SDS (10%)	50 ml
ELIX H2O	Fill up to 1 l

Destaining Solution

Methanol	40%
Acetic Acid	10%
ELIX H2O	50%

The Coomassie Blue Staining Solution was heated in the microwave on high power for 40 – 60 seconds. The heat would clearly shorten the staining time. Eventually the gel from the SDS-Page was placed in the solution and the box put on a rocking table for 30 – 60 minutes till the entire gel had turned dark purple. Afterwards the staining solution was filled back in the bottle for reuse. The gel was de-stained by washing it with ELIX water several times and finally with de-staining solution on the rocking table until the bands became clearly visible. Finally, the Destaining Solution was removed, and the gel was stored in ELIX water until it got photographed.

2.7 Rapid Immunoprecipitation Mass Spectrometry of Endogenous Proteins

A method to identify (transient) protein complexes is the affinity purification of endogenous proteins coupled to mass spectrometry (Mohammed et al., 2016). For that I combined two protocols: basically, I followed the protocol "nematostella_embryo_X-ChIP_better_nuclei _110925withDTT-2" (Alexandra Schauer) from point 1.1 to 1.3 including fixation of the embryos, preparation of the nuclei and shearing of the nuclei. However, I tried different deviations based on the *Nature* protocol from Mohammed et al., 2016 to make it compatible for *Nematostella*. I then followed the Mohammed et al. protocol, starting with immunoprecipitation over enzymatic digestion and peptide desalting to mass spectrometry but also tried different deviations

to find the optimal protocol.
To all buffers used in this protocol, phenylmethylsulphonyl fluoride (PMSF) (1 mM) was added freshly before use to avoid protein degradation. If crystals had formed in the 100 mM PMSF solution, it was put on the magnetic mixer and heated up to ~50°C until they had dissolved.

2.7.1 Preparation of embryos for RIME

For one RIME-run I used 10^8 cells according to Mohammed et al., 2016. Assuming that a single late gastrula embryo contains approximately 7000 cells (Kirillova et al., 2018): midgastrula median: 6934 cells) this corresponds to around 14000 embryos. As I analyzed (see 3.1.1.2), both proteins (Bra/FoxA) are present in 100% of the cells after inhibitor treatment. Therefore, calculations could be made based on these numbers. One female box produces on average 17000 eggs. Therefore, two boxes were used for one biological replicate, to have a buffer as not all eggs become fertilized, develop properly and some get lost during the procedure. For statistical power three biological samples are needed per run and condition. One male box plus one backup box was sufficient for six female boxes. The animals were induced and embryos generated (see 2.1.1) which were treated with the inhibitors as described before (see 2.1.2).

2.7.2 Preparation of chromatin and proteins

For preparing the chromatin and proteins I tried different versions of the two protocols mentioned above (see 2.7) and tested the different buffers.

2.7.2.1 Preparation of nuclei from fixed embryos

E1 Buffer (Lysis Buffer 1) (storable for 1 month at 4°C)

HEPES pH 7.5	50 mM	Add freshly: PMSF and Dithiothreitol with 1 mM concentration each!
NaCl	140 mM	
EDTA	1 mM	
Glycerol	10%	
NP-40-Igepal	0.5%	
Triton X-100	0.25%	
Milli-Q H2O	Fill up to total volume	

Lysis Buffer (storable for 1 month at 4°C)

HEPES pH 7.5	50 mM	Add freshly: PMSF with 1 mM!
NaCl	500 mM	
EDTA	1 mM	
Triton X-100	1%	
Sodium Deoxycholate	0.1%	
SDS	0.1%	
Milli-Q H2O	Fill up to total volume	

Lysis Buffer 2 (storable for 1 month at 4°C)

Tris-HCl pH 8.0	10 mM
NaCl	200 mM
EDTA	1 mM
EGTA	0.5 mM
Milli-Q H2O	Fill up to total volume

Add freshy: PMSF with 1 mM!

Lysis Buffer 3 (storable for 1 month at 4°C)

Tris-HCl pH 8.0	10 mM
NaCl	100 mM
EDTA	1 mM
EGTA	0.5 mM
Sodium Deoxycholate	0.1%
N-lauroylsarcosine	0.5%
Milli-Q H2O	Fill up to total volume

Add 1 mM of PMSF freshly!

The nuclei preparations steps were performed on ice. In all variants tested, the fixed embryos (see 2.2.3) were resuspended in a 15 ml falcon tube in 7 ml E1 Buffer which would lyse the cell membranes but not the nuclear membranes with its mild non-ionic detergents NP-40-Igepal and Triton X-100. The solution was poured in the grinding chamber of a *Glass 7 ml Dounce Tissue Grinder Set* (Wheaton) and homogenized with the tight pestle at least forty times to obtain the nuclei. The homogenized solution was poured back in the falcon tube, filled up to 10 ml with E1 Buffer and kept on ice for 5 min. Then it was spun in a *Centrifuge 5804R* (Eppendorf) in a swing out rotor at 1500 g for 5 minutes at 4°C. The supernatant was poured into a new falcon tube and spun again. Both pellets, containing the nuclear material, were resuspended together in 10 ml of either the E1 Buffer according to the X-ChIP protocol or the Lysis Buffer 2 according to the Nature protocol. The Lysis Buffer 2 would prepare the sample for subsequent transfer into Lysis Buffer 3 which is based on Tris-HCl instead of HEPES. To check the purity of the nuclei in the pellet, 20 µl of the sample were set aside and Hoechst was added in a 1:200 dilution and the sample checked under the fluorescence microscope. Blue nuclei without any cytoplasm attached to them should be visible. If it looked clean enough, the solution was again centrifuged at 1500 g for 5 minutes at 4°C and the resulting pellet resuspended in 2 ml of Lysis Buffer if it had been in E1 Buffer in the previous step or in 2 ml of Lysis Buffer 3 if it had been in Lysis Buffer 2. These buffers with the harsher denaturing ionic detergents sodium deoxycholate + SDS (Lysis Buffer) or N-lauroylsarcosine (Lysis Buffer 3) would solubilize the nuclear membranes and release chromatin and the nuclear proteins. The lysate was centrifuged at 2000 g for 2 minutes at 4°C to get rid of insoluble debris and then resuspended in 135 µl Lysis Buffer or Lysis Buffer 3 corresponding to what it had been solved in before.

2.7.2.2 Shearing of fixed nuclei

To shear the chromatin into workable fragments, the extracted genomic DNA was subjected to sonication. According to Mohammed et al., 2016 the fragments should be between 200 – 600 base pairs. I tested different sonication settings to to find those reliably producing this

fragment length.
First the *S2 Focused-ultrasonicator* (Covaris) was prepared as it needs 30–50 minutes to degas and cool down. Therefore, ELIX water was filled into the water bath up to the right *15* check mark, to make sure the water level was the same for every sonication. The sonicator and the cooling system were turned on and cooling set to 5°C. Now 130 µl of the resuspended pellet were pipetted into a *microTUBE AFA Fiber Pre-Slit Snap-Cap* (6x16mm, Covaris) without bubbles. The tube was placed in a metal holder in the water bath. The parameters that had to be set in the *SonoLab Software for Covaris Focused-ultrasonicators* were Duration, Cycles, Intensity, Duty Factor, Cycles/Burst and Mode. "Cycles" states how often the "Duration" of sonication was repeated, "Intensity" is a dimensionless proxy for the power emitted from the transducer during each burst (intensity 10 corresponding to approximately 350 Watts) and the "Duty Factor" gives the percentage of active burst time in the acoustic treatment. They are listed in Table 1 in the order in which they were varied to achieve changes in shearing size, starting with changes in time before changes in intensity. I started with the settings applied in the lab protocol and then tried the settings as suggested by the Covaris-guide for an outcome of 300 bp length.

Table 1 Different sonication settings
tested to receive optimal chromatin fragmentation.

Duration [s]	Cycles	Intensity	Duty Factor [%]	Cycles/Burst	Mode
60	16	5	20	200	Frequency Sweeping
50					
	10				
	5				
	2				
60	16	4	10		
40	10				
	5				
	2				

At the end of my project we received a new sonicator, the *S220 Focused-ultrasonicator* (Covaris) with a different setting arrangement. The final settings with the old sonicator correspond to the following new settings: Duration of 40s for 2 Cycles, Peak Incident Power of 175 W, Duty Factor of 20%, Average Power of 35 W, 200 Cycles/Burst and Frequency Sweeping.

After sonication the samples were transferred to a new 1.5 ml tube on ice. The metal holder and the water bath were dried off and the tubes washed with ethanol to be reused. 15 µl were taken to control the sonication efficiency (see 2.7.2.3). If it looked alright, the rest of the sample was centrifuged for 10 minutes at 4°C with 20000 g. The supernatant was taken into a new tube and filled with Lysis buffer or Lysis Buffer 3 to 330 µl (115 µl left, 215 µl added). If a SDS-

Page/Coomassie Staining/Western Blot was planned (to compare protein content before and after IP), the supernatant was filled up to 340 µl and 10 µl were set aside and frozen for later analysis. Regularly 1 µl of the sample was used to measure the protein concentration with a Bradford assay (see 2.6.2). The procedure had to be continued on the same day to avoid freezing the sample before the subsequent immunoprecipitation (see 2.7.3).

2.7.2.3 Checking sonication efficiency

To check the chromatin shearing efficiency namely the fragment sizes, 15 µl of the sonicated sample were set aside. They were centrifuged for 10 minutes at 4°C with maximum speed. The supernatant was taken to a new tube with 1% SDS, 100 mM NaCl and 50 µg/ml RNase A filled up with Milli-Q H_2O to a total volume of 50 µl. The sample was heated in a *Thermomixer comfort* (Eppendorf) at 55°C for at least 30 minutes to digest the RNA. Then 200 µg/ml Proteinase K (Ambion) were added, and the sample was incubated for at least 1.5 hours at 65°C to reverse the crosslinking and digest the proteins. Eventually the DNA was extracted and precipitated with the *QIAquick PCR Purification Kit* (Qiagen). The double stranded DNA concentration was measured with the *Qubit dsDNA High Sensitivity Assay Kit* (Invitrogen) on the *Qubit 2.0 Fluorometer* (Invitrogen). About 300 ng of ds DNA were run on a 2% agarose gel to check the fragment sizes. If the sonication was successful and protein concentration was sufficient, the IP was continued.

2.7.3 Immunoprecipitation

To immunoprecipitate the desired proteins out of the nuclear lysates, I basically followed the Nature protocol from Mohammed et al. 2016 with slight deviations. By immunoprecipitating from the enriched chromatin, it is made sure that transcription factors that might be identified by mass spectrometry are interacting with the target protein on the chromatin itself. This is a requirement to classify them as interaction partners in transcription.

2.7.3.1 Antibody conjugation to magnetic beads

To simplify the process of immunoprecipitation (IP) and to have less contamination with antibody in the mass spectrometry-analysis the magnetic *Dynabeads Protein A for Immunoprecipitation* were used. The bead stock solution was vortexed just before use and for one sample 100 µl of suspension were transferred to a 2 ml Eppendorf tube. The tube was placed in a magnetic stand which was kept on ice. When the beads had migrated to the side of the tube aligned against the magnet, the supernatant could be removed with a pipette and 1 ml of 1x PBS + 5 mg/ml Bovine Serum Albumin (BSA) added to wash and block the beads. To make sure the beads were evenly suspended, the tube was lifted from the magnetic stand and agitated by hand. Back on the stand the beads were given time to settle until the supernatant had cleared. This procedure was repeated three times for a total of four washes. Eventually the beads were suspended in 500 µl of 1x PBS + BSA and a sufficient amount of antibody for IP

was added. Mohammed et al. had optimized their antibody at 10 µg/100 µl beads for immunoprecipitating from lysates of ~6x 10^7 cells. They tested 2 to 20 µg and found that 2 µg was enough to identify the target protein but with reduced sensitivity, above 10 µg they could not detect an improvement in sensitivity but only an increase of background noise. Therefore, I used also 10 µg of the FoxA and Bra antibody (stock 1 mg/ml) on 100 µl beads as starting point. As an unspecific negative control, I used an antibody produced in the same host species (rabbit) but without any target in the embryos, the anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 568 and as further negative control only beads without adding any antibody. The beads with the antibody/control beads were gently mixed and the mixture panned at 4°C overnight on a rotating mixer. Afterwards the beads were placed on the magnetic rack and the supernatant removed. The antibody-bound beads were washed in 1 ml of PBS + BSA for a total of five washes to remove any unbound antibody. As last step they were resuspended in 100 µl of PBS + BSA.

2.7.3.2 Immunoprecipitation with magnetic beads

RIPA Buffer (storable for 1 month at 4°C)

HEPES pH 7.6	50 mM
EDTA	1 mM
Sodium Deoxycholate	0.7%
NP-40-Igepal	0.5 M
Lithium Chloride	0.25%
Milli-Q H2O	Fill up to total volume

Add 1 mM PMSF freshly!

The 100 µl of antibody-bound beads (from 2.7.3.1) were added to the 330 µl cell lysate sample (from 2.7.2.2) and rotated at 4°C overnight. The next day the sample was placed again on the magnetic stand, whereby in some cases the supernatant was kept for a Western Blot to check whether the targeted protein could still be found in the supernatant after IP. Then the beads were washed in 1 ml of RIPA buffer at 4°C for a total of ten washes. Eventually the beads were washed twice in 1 ml of cold 100 mM ammonium hydrogen carbonate (AMBIC) solution pH 7.5–8, which was freshly made and kept at 4°C, to remove detergents and salts and to have the beads in a buffer for subsequent tryptic digestion. For the second AMBIC wash, the beads were transferred to a new tube. If a SDS-Page/Western Blot was planned 10 µl of the 1 ml were set aside (or make it higher concentrated by using less AMBIC). Before continuing, as much supernatant as possible was removed because spare liquid would influence the digestion efficiency. Either the beads were now snap-frozen in liquid nitrogen and stored at -80°C or the peptide digestion followed directly. If they had been frozen and thawed also any residual buffer was digested to not lose any antibody-antigen complex that may have detached from the beads.

As co-elution of the antibody with the antigen may interfere with downstream analysis like Western Blots or mass spectrometry, I tried an alternative protocol in which the FoxA antibody was coupled to the magnetic beads before IP. For that I used the *Pierce BS^3 Crosslinker* ((Sulfo-

DSS) is bis(sulfosuccinimidyl)suberate, Thermo Scientific, #21580). For the preparation and crosslinking of the beads and the elution after IP, I followed the manual from Invitrogen but with slight deviations as indicated, that were derived from the RIME protocol.
According to the RIME protocol the beads were washed and incubated with the antibody in PBS + BSA (see 2.7.3.1), but I used only 5 µl of antibody. I also incubated the antibody with the beads for 1 hour instead of 10 minutes, to make it more comparable to RIME-IP. For crosslinking the antibodies to the beads I followed the *Immunoprecipitation Crosslinking Protocol using Dynabeads*. The beads were washed in Conjugation Buffer, incubated in BS^3 and the reaction quenched by Quenching Buffer. The difference was that I washed the beads with PBS/BSA again instead of only PBST. For IP I left all samples in 100 µl of PBS/BSA and added 330 µl of the cell lysates (see 2.7.2). After incubation over night at 4°C the supernatants were saved for analysis and the beads were now washed in RIPA Buffer (see 2.7.3.2). Now the samples and the supernatants were eluted and denatured in 2x Laemmli Buffer at 99°C for 5 minutes. Contrary to the manual I did not use elution buffer but just the Laemmli buffer. I let the samples cool down and evaluated them in a Western Blot (see 2.6.4).

2.7.4 Preparation for Mass Spectrometry

2.7.4.1 Enzymatic digestion with trypsin

In Bottom-up proteomics, proteins become digested into peptides which can be identified by mass spectrometry. The most common used protease to generate peptides is trypsin. Trypsin cleaves proteins at the carboxyterminal side of lysin and arginin-residues, except if they are followed by a proline. (Rehm and Letzel, 2010). As a result, it can be predicted into which peptides a protein will be cleaved.
To digest the bead-bound proteins, 10 µl (10 ng/µl) of trypsin in 100 mM AMBIC, were directly added to the washed beads from step 2.7.3.2, using an enzyme-to-protein ratio of 1:100 (wt/wt). I used frozen trypsin aliquots with 1000 ng/µl of trypsin in 1 M HCl, from which I mixed 0.1 µl with 9.9 µl AMBIC. The samples were vortexed for 15 seconds every 2–3 minutes for the first 15 minutes to ensure that the beads were evenly suspended and then they were left for digestion overnight at 37°C in an oven incubator without further agitation. After the overnight digest, 10 µl of trypsin at the same enzyme-to-protein ratio as in the last step were added additionally to the sample, and again left for digestion for 4 hours at 37°C. Finally, the tubes were put back on the magnetic rack and the removed supernatant containing the digested peptides was retained. The supernatant (~20 µl trypsin solution plus any residual buffer) was filled up directly to a volume of 100% (vol/vol) formic acid such that the final concentration of formic acid was 5% (vol/vol), so to 20 µl, 1 µl formic acid was added. If a precipitate is visible, it means that detergent is still present but can be removed by subsequent SPE (see 2.7.4.2).

2.7.4.2 Peptide desalting by solid-phase extraction

The resulting peptide samples had to be purified and concentrated for MS to prevent protein aggregation and the presence of high-molecular weight species. Therefore, I used spin columns that contain a porous C18 reversed-phase resin. The peptides bind via hydrophobic interactions, but salts and buffers are washed off. Thus, they remove interfering contaminants, resulting in increased sensitivity and a high-quality spectrum. Finally, peptides are released with organic solvents in MS-compatible solutions (Mohammed et al., 2016).
All desalting steps were performed using the *Pierce™ C18 Spin Columns* (ThermoScientific). The whole procedure was undertaken in a timely manner to avoid that the resin would dry. Firstly, the column was tapped to settle the resin, then the top and bottom cap were removed, and the column placed into a 2 ml Eppendorf receiver tube. To condition the cartridge, it was centrifuged two times with 100 µl of 50% (vol/vol) acetonitrile/Milli-Q H_2O at 1500 g for 1 minute. Then to equilibrate the cartridge, it was centrifuged two times with 100 µl of 0.1% (vol/vol) formic acid in Milli-Q water at 1500 g for 1 minute. Finally, the acidified peptides from step 2.7.4.1 were loaded onto the cartridge and centrifuged at 1500 g for 1 minute. To maximize the peptide binding the flow-through was reloaded three times. Then the peptide-loaded cartridge was washed four times with 0.1% (vol/vol) formic acid, each time centrifuging at 1500 g for 1 minute. The column was then placed in a new receiver 1.5 ml tube and the peptides eluted twice. Once with 50 µl of 60% (vol/vol) acetonitrile/0.1% (vol/vol) formic acid and the second time with 30 µl by centrifugation. The two eluates were combined and then dried by vacuum centrifugation with the *Concentrator Plus* (Eppendorf). For 80 µl it took around 80 minutes with the settings *V-AQ* at room temperature. The dried peptide samples could be stored at −20°C or −80°C for several weeks before MS.

2.7.5 Mass spectrometry

Before MS the dried sample needed to be reconstituted in 2% acetonitrile/0.1% formic acid in Milli-Q H_2O. Organic solvents help to spray the sample into droplets and weak acids help to produce positively charged molecules needed in MS (Rehm and Letzel, 2010). For one injection into the mass spectrometer, 20 µl were necessary as 15 µl are sucked up and 5 µl serve as a buffer. The facility team of the mosys-Department loaded the samples for me.
I conducted MS-runs on 4 different dates with in total 25 samples. They are listed in Table 2, together with their preparation procedure. As mentioned previously, I worked primarily with the FoxA antibody which is why the majority of the RIME experiments were conducted against FoxA and not enough replicated RIME experiments were performed against Brachyury. 13 RIME samples stemmed from immunoprecipitations with the FoxA antibody, whereby 8 were replicates prepared identically (but on different dates) and 5 were prepared differently as described in Table 2. With the Bra antibody 2 replicate RIME samples were prepared and 2 samples were prepared following the Co-IP protocol. Also, negative controls for the Co-IP were prepared, one using an α-mouse antibody and one incubated only with the magnetic beads. Also, 6

negative controls for the IPs were measured, 3 immunoprecipitated with the α-mouse antibody and 3 samples only incubated with the magnetic beads.
The third run was a targeted approach. Therefore, before starting the run I gave the mass spectrometer the molecular weights of peptides that are known to be produced by trypsinization of the proteins of interest. As a result, the machine detects all peptide ions with the respective mass spectra all the time while neglecting residual peptides in the samples. This increases sensitivity for the proteins of interest but simultaneously nothing "unexpected" can be recorded. For Bra RIME I targeted: Bra itself, FoxA, Lmx, FoxB as they are suspected interaction partners and the Sox proteins Nve3409, SoxF.1, SoxA, SoxE1, Sox9 and SoxB2 because of the previous Bra ChIP-seq. For FoxA RIME I targeted FoxA itself, Bra, Lmx and FoxB.

Table 2 List of mass spectrometry samples and their preparation procedure.

			Sample preparation		
Sample Name		**Mass spec run**	**CHIR99021 treatment**	**Formaldehyde fixation**	**(Co-)Immuno-precipitation**
UT_FoxA		1	None	1% for 8 minutes	FoxA IP
FoxA_1	FoxA replicates		10 µM		
FoxA_2		2			
FoxA_3					
FoxA_4					
FoxA_5		3 (Targeted approach)			
FoxA_6					
FoxA_7					
FoxA_8					
FoxA_nonfixed				no fixation	
FoxA_0.5%8m				0.5% for 8 minutes	
FoxA_1%5m				1% for 5 minutes	
FoxA_1,85%8m				1.85% for 8 minutes	
Bra_1	Bra replicates			1% for 8 minutes	Bra IP
Bra_2					
α-mouse_1					α-mouse IP
α-mouse_2					
α-mouse_3					
only beads_1					only beads IP
only beads_2					
only beads_3					
UT_CoIP_Bra		4	None		Bra Co-IP
CoIP_Bra			10 µM		
CoIP_α-mouse					α-mouse Co-IP
CoIP_only beads					Only beads Co-IP

2.7.6 Maxquant analysis

From the determined molecular weights the PC concluded the underlying amino acid sequences. The found peptide sequences were blasted against a protein database to identify the origin proteins. Therefore, the raw data generated by the mass spectrometer, were uploaded into the quantitative proteomics software *MaxQuant*. For protein-identification I grouped the samples and thus the identified proteins into seven different groups (Table 3 The groups "all Bra" and "all FoxA" include proteins from all samples that were prepared using the respective antibody. The "replicate Bra" group includes proteins only from the identically prepared RIME samples. The "replicate FoxA" group includes proteins from the identically prepared RIME samples but without proteins identified in "FoxA_2/3/4" which are also replicates but as described later, no FoxA was identified in them. The FoxA control group contains the proteins from the negative control samples α-mouse_1-3 and only beads_1-3. The Bra control group additionally includes proteins identified in CoIP_α-mouse and CoIP_only beads (Co-IP was only performed with the Bra antibody). The "all controls" group includes proteins from all controls combined.

Table 3). The groups "all Bra" and "all FoxA" include proteins from all samples that were prepared using the respective antibody. The "replicate Bra" group includes proteins only from the identically prepared RIME samples. The "replicate FoxA" group includes proteins from the identically prepared RIME samples but without proteins identified in "FoxA_2/3/4" which are also replicates but as described later, no FoxA was identified in them. The FoxA control group contains the proteins from the negative control samples α-mouse_1-3 and only beads_1-3. The Bra control group additionally includes proteins identified in CoIP_α-mouse and CoIP_only beads (Co-IP was only performed with the Bra antibody). The "all controls" group includes proteins from all controls combined.

Table 3 Grouping of samples respectively the identified proteins for analysis.

Group Name	Samples	Amount
all Bra	Bra_1, Bra_2, UT_CoIP_Bra, CoIP_Bra	4
replicate Bra	Bra_1, Bra_2	2
all FoxA	UT_FoxA, FoxA_nonfixed, FoxA_0.5%8m, FoxA_1%5m, FoxA_1,85%8m, FoxA_1, FoxA_2, FoxA_3, FoxA_4, FoxA_5, FoxA_6, FoxA_7, FoxA_8	13
replicate FoxA	FoxA_1, FoxA_5, FoxA_6, FoxA_7, FoxA_8	5
Bra control	α-mouse_1, α-mouse_2, α-mouse_3, CoIP_α-mouse, only beads_1, only beads_2, only beads_3, CoIP_only beads	8
FoxA control	α-mouse_1, α-mouse_2, α-mouse_3, only beads_1, only beads_2, only beads_3	6
all controls	α-mouse_1, α-mouse_2, α-mouse_3, CoIP_α-mouse, only beads_1, only beads_2, only beads_3, CoIP_only beads	8

I ran the program for each sample group once with the UniProt database FASTA for *Nematostella* which contains 25162 proteins including 126 reviewed (Swiss-Prot) and 25036 unreviewed (TrEMBL). And once with the FASTA "nveGenes.good.130208.longCDS. protein.corrected

_inframe_stops" from the Technau lab generated amongst others by B. Zimmerman which contains 26235 protein sequences.
Here I state the settings that I altered from the default in *Maxquant*:
- Digestion: Trypsin, Maximum missed cleavages: 2; - Modifications: Fixed modifications: no alkylation & reduction of Cysteine, no Carbamidomethyl, Variable modifications: oxidation and acetylation; - Global parameter: Advanced: Decoy mode: Revert, Protein quantification: Label min. ratio count: 1 (1 peptide of 7 amino acids length), Discard unmodified counterpart peptide: no.
The minimum peptide count was set to 1 which means the list of identified proteins includes proteins identified by only 1 peptide. We decided to not apply the usual minimum 2 peptide count as this project`s aim was identification of proteins and not quantification. 1 peptide is more sensitive but also increases the risk of identifying false positives.
As a result, I received lists of identified proteins once specified by UniProt identifiers and once by Nve numbers. For completeness and sensitivity, I included both lists into the evaluation and counterchecked identified proteins between those. For analysis, I focused on certain values given for the identified proteins. On the "intensity" of a protein which is the sum of all individual peptide intensities belonging to a protein. It is specified in arbitrary units since it is a relative intensity (for calculation of the exact concentration you would have to create a standard curve). On the "LFQ intensity" which is obtained by a label-free quantification based on the normalizing LFQ algorithm (Cox et al., 2014). LFQ determines the relative amount of protein in a sample and is normally used for protein quantification to compare protein abundance between different sample types (e.g. before and after drug treatment), and not for sheer identification. However, as I did not prepare all replicates in parallel and measured the samples in different MS runs it is reasonable to use LFQ. It does not assume that samples are measured under uniform conditions and within a narrow time frame. Furthermore, different peptides are ionizable to different degrees. As a result, rare proteins with easily ionizable peptides can have high intensities and by contrast abundant proteins with badly ionizable peptides low intensities. LFQ considers these differences.

2.8 Magnetic co-immunoprecipitation

Another attempt to identify interaction partners of Bra/FoxA was to use the *Universal Magnetic Co-IP Kit* (Active Motif). With this (Benitez et al., 2017) had identified the interaction of PTEN with DAXX in human glioma cells. The important difference to RIME is that Co-IP does not need fixation of protein complexes with formaldehyde. Instead the identification of protein interactions is based on IP with an antibody against the target antigen which thereby co-immunoprecipitates any proteins bound to it. The protein complexes in this method can disintegrate easier during the nuclear extraction procedure, however the Active Motif kit promises to facilitate a more specific but less stringent procedure than conventional Co-IP methods, to preserve the protein complexes. For RIME I headed to work with 10^8 cells. For the magnetic Co-IP a

maximum of 2 x 10^7 cells is supposedly required, however this protocol is not geared for MS. Therefore, I first examined with how many cells I would obtain ~1000 µg of protein after nuclear fraction digestion which is an approximate value for successful MS. I tested the protocol with one sample each of~2.8 x 10^7, ~7 x 10^7 and ~14 x 10^7 untreated cells.

2.8.1 Nuclear extraction from cells

The manual`s section "Nuclear Extraction from Cells" was followed. When working with 2.8 x 10^7 cells, the volumes of reagents as proposed by the kit manual for 8.8 x 10^6 cells were used and when working with 7 x 10^7 and 14 x 10^7 cells, the amount of reagents proposed for 2 x 10^7 cells was used. I only mention deviations from the manual.
For cell collection the embryos were brought into a 15 ml falcon tube with Ca^{2+}- and Mg^{2+}-free *Nematostella* medium and pipetted up and down until they had dissolved in single cells. Eventually, they were centrifuged for 5 minutes at 400 g to exchange the medium with the ice-cold PBS/inhibitor solution and again centrifuged before the second PBS wash. They were transferred to a new falcon and in the further course the manufacturer manual was followed. The nuclei were isolated, and a nuclear fraction digestion and collection performed. Finally, Bradford assays were done for protein quantification (see 2.6.2).

2.8.2 Co-immunoprecipitation

The co-immunoprecipitations (Co-IP) were finally performed with ~14 x 10^7 cells following the kit`s manual. The following antibody/extract mixtures were prepared: untreated cell extract + α-Bra-antibody (UT_CoIP_Bra), CHIR-treated cell extract + α-Bra-antibody (CoIP_Bra) and the negative controls CHIR-treated cell extract + α-Mouse (rabbit) secondary antibody coupled to Alexa Fluor 568 and CHIR-treated cell extract + only beads.
To the (frozen) 200 µl sample in Complete Digestion Buffer, 300 µl of Complete Co-IP/Wash Buffer and 2 µl of antibody were added. I left it incubating for 18.5 hours at 4°C on a rotator. Then the magnetic Protein G beads were incubated with the samples and afterwards washed as described in the manual. Contrary to the manual, the bead pellets were not resuspended in Reducing Loading Buffer but washed twice in AMBIC and 10 µl set aside for Western Blot. From then on it was proceeded as written above (see 2.7.4) until mass spectrometry. Unfortunately, only one version of each of the four different sample types made it to MS, because after that the Q Exactive mass spectrometer had been shut down for an unknown time period.

2.9 Documentation systems and programs

For analysis and documentation of results different microscopes, cameras and programs were used to obtain the best possible results.

Table 4 Microscopes, cameras and software programs

	Microscope	Camera	Software
Animal Observation	Nikon SMZ800 and Nikon SMZ745	-	-
Whole Mount in Situ Hybridization	Nikon Eclipse 80i and Nikon (?)	Nikon DS-Fi1 and	NIS-Elements BR 4.20.01 (64-bit) and
Immunohistochemistry	Leica TCS SP5 Laser Scanning Confocal Microscope	Color CCD camera Retiga-2000R	Qcapture Pro 6.0 acquisition
Nuclei-Check for RIME	Nikon Eclipse 80i	Nikon DS-Fi1	NIS-Elements BR 4.20.01 (64-bit)
Western Blot membrane	-	UVP EC3 Chemi HR 410	UVP VisionWorks LS Analysis

Fluorescent images taken by the confocal microscope, were edited with *Fiji* (Schindelin et al., 2012). For editing and designing of photograph-collages *Adobe Photoshop CC* 2015 was used. Primers were designed with help of the web tool *Primer3web*. For working with DNA-/RNA- and amino acid sequences the programs *ApE - A plasmid Editor* and *CLC Sequence Vie*wer were used.

3. Results

3.1 Effects of enhanced Wnt-signaling on expression patterns of *bra*, *foxA* and 12 *sox* genes in late gastrulae

3.1.1 *bra/foxA*-expression increases depending on inhibitor concentration

To observe which treatment conditions (see 2.1.2) would achieve an ubiquitous *brachyury* and *foxA* expression in late gastrulae 27 hpf, histological methods and Western Blots were applied.

3.1.1.1 ISH reveals most efficient treatment

The expression patterns of *bra* and *foxA* in control, DMSO treated control and treated embryos at ~27 hpf were compared by whole mount in situ hybridizations (see 2.4).

In late-gastrulae, *bra* is expressed orally around the blastopore in ectodermal cells while the pre-endodermal plate has finished invagination (Fig. 4a, b). *foxA* is also expressed orally around the blastopore in ectodermal cells but reaches further inside and is also expressed in the pharynx anlagen (Fig. 4c, d). The DMSO control embryos show that DMSO in which the inhibitors are solved in, has no visible effects on the expression of *bra* or *foxA* and neither on the embryos morphology (Fig. 4e-i).

Firstly, I analyzed the effects of three different Azakenpaullone (Azk) and CHIR99021 (CHIR) concentrations in treatments from early cleavage to 27 hpf (for ~23 hours) on *bra* expression. Treatment with 5 µM of Azk lead to an aboral expansion of the oral expression ring around the blastopore. The embryos start gastrulation but fail to finish (Fig. 4j). 10 µM lead to an ubiquitous expression in most of the embryos but with many of them still forming a blastopore and at least starting the process of gastrulation (Fig. 4k). With 15 µl the expression appears ubiquitous in all embryos with an intense staining and they are not able to start invagination (Fig. 4l). 5 µM of CHIR results in many cases in a faint staining of the whole embryo with most of the embryos struggling to gastrulate (Fig. 4o). 10 and 15 µM lead to a strong ubiquitous expression of *bra* and prevented the onset of gastrulation (Fig. 4p, q). However, differences in intensity can also be observed, showing a stronger staining with 10 µM CHIR (Fig. 4p) than with 15 µM (Fig. 4q).

Next, I investigated the effects on *foxA*, whereby I left out 5 µM of both inhibitors as it had not shown sufficient effects on *bra* expression. Basically, the same could be observed as for *bra* (Fig. 4m, n, r, s).

Treatment with Azk as well as with CHIR leads to non-gastrulating enlarged and round shaped embryos (Fig. 4j-s) compared to their untreated counterparts (Fig. 4a-i).

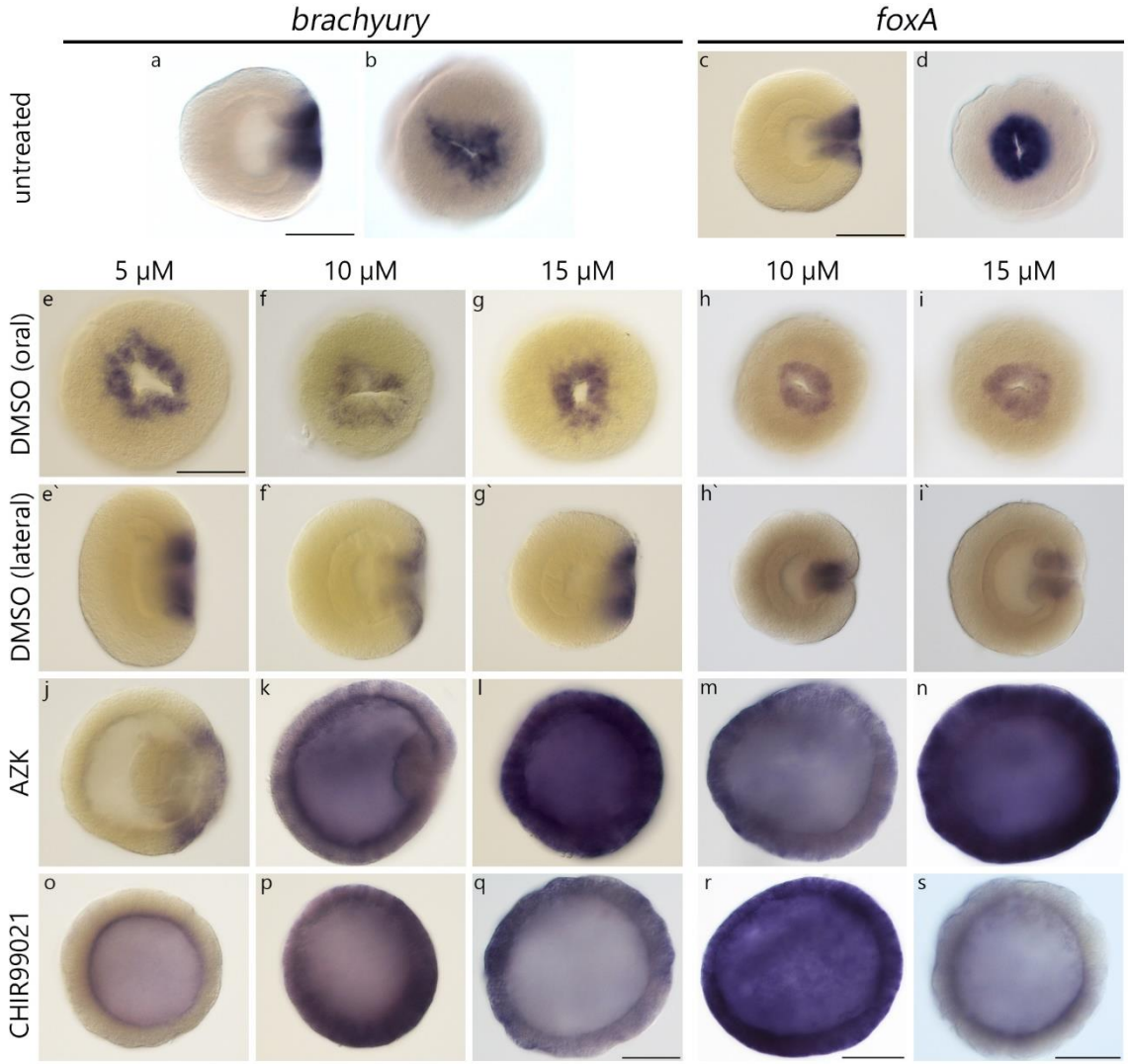

Fig. 4 Effect of Gsk3ß inhibitors on *bra/foxA*-expression patterns
in late-gastrulae (27 hpf) visualized by whole-mount-in-situ-hybridizations. (a, c & e-i) oral views. (b, d, e`-k) lateral views with oral pole to the right. (l-s) view not determinable. First row shows the untreated controls. For the others the chemicals used are written on the left, their concentration applied in treatment from early cleavage to late gastrula on the top, except for DMSO: DMSO controls correspond in all cases to 53 µM DMSO in *Nematostella* medium. The animals are 27 hpf. DMSO controls show no DMSO impact on *bra* & *foxA*-expression or on morphology. The effect of Azk on both genes becomes stronger with increasing concentration, peaking at 15 µM with an ubiquitous strong expression pattern of both genes and no success in gastrulation. CHIR99021`s effect peaks at 10 µM, decreasing with 15 µM. Azk = 1-Azakenpaullone. Scale bar: 100 µm.

3.1.1.2 Immunohistochemistry confirms most efficient treatment

Furthermore, by verifying the presence of the transcription factors in the nuclei it was made sure that the genes were not only expressed but also translated into proteins. Therefore, immunohistochemistry was performed against Bra/FoxA in green, DAPI in blue and filamentous Actin in red (see 2.5).

The Brachyury protein is normally found orally around the blastopore as can be seen in the DMSO controls (Fig. 5a-f). The merge of Bra fluorescence (green) with DAPI (blue) proves its

location in the nucleus. Treatment with 5 µM Azk expands *brachyury* expression aborally (Fig. 5g). With 5 µM CHIR Bra can be detected at the aboral side of the embryos, but it is not detected in every nucleus (Fig. 5h). 5 µM of both inhibitors affect gastrulation, in the case of Azk the pre-endodermal plate still forms but fails to invaginate fully and seems arrested in early gastrula stage (Fig. 5g`), whereas with CHIR the embryo struggles to form the pre-endodermal plate (Fig. 5h`). 10 µM Azk lead to presence of Brachyury in nuclei all over the embryo (Fig. 5i) just like 10 µM of CHIR (Fig. 5j). 10 µM of both inhibitors completely prevent gastrulation and the embryos remain round shaped (Fig. 5i`, j`). 15 µM of Azk seems to enhance the expression further with Brachyury being present ubiquitously (Fig. 5k). 15 µM of CHIR do not lead to an increase of the effect but it remains the same as with 10 µM (Fig. 5l). Like with 10 µM, the embryos fail to gastrulate and their morphology reminds of a ball (Fig. 5k`, l´).

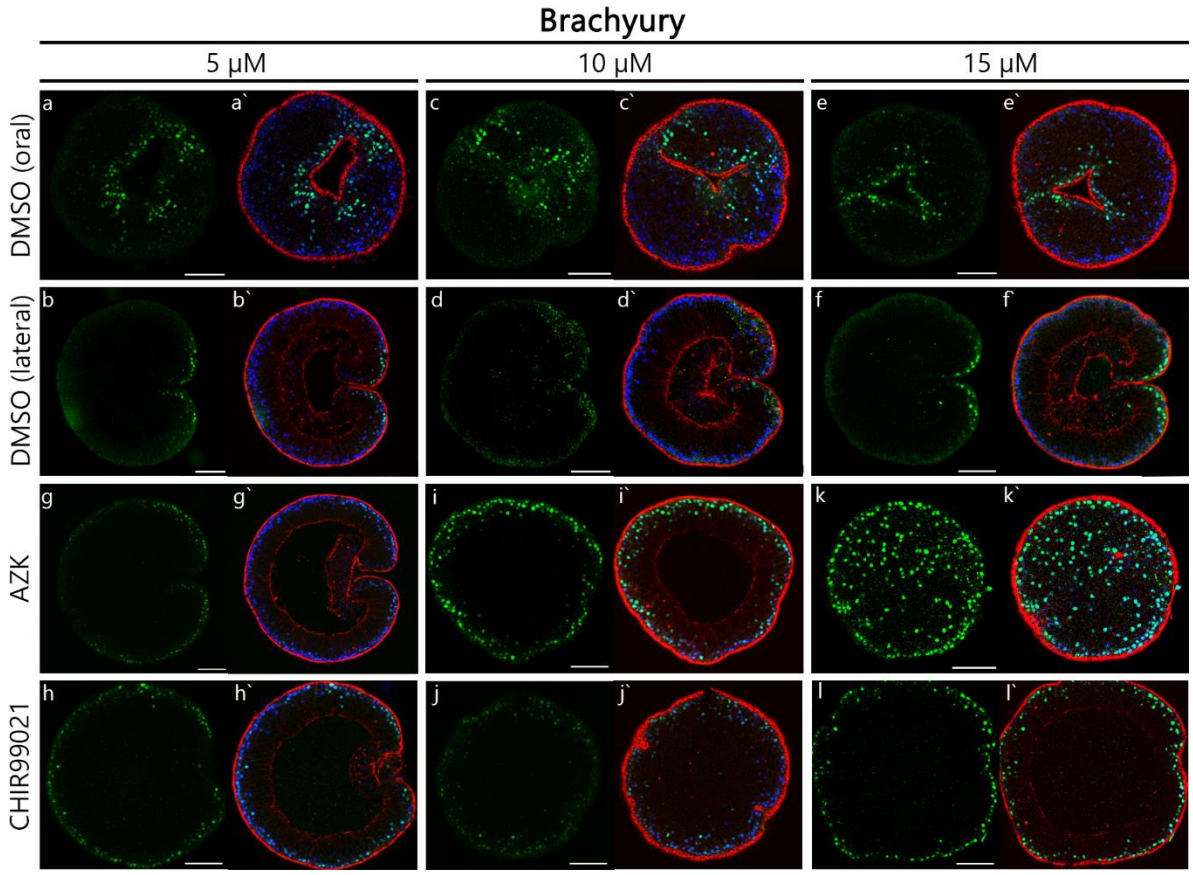

Fig. 5 Effect of Gsk3ß inhibitors on Bra protein expression

in late-gastrulae (27 hpf) visualized by immunostainings. Two images of the same embryo are shown, left hand Bra in green and right hand the merge with nuclei (DAPI, blue) and f-actin (phalloidin, red). (a, c & e) oral views. (b, d, f, g & h) lateral views with oral pole to the right. (i-l) view not determinable. Chemicals used are written on the left, their concentration applied on the top, except for DMSO: DMSO controls correspond in all cases to 53 µM DMSO in *Nematostella* medium. Azk effect becomes stronger with increasing concentration, peaking at 15 µM. CHIR99021 effect peaks at 10 µM, unchanged with 15 µM. Azk = 1-Azakenpaullone. Scale bar: 100 µm.

Basically, the same effects could be observed in the immunostainings against FoxA. At late gastrula stage FoxA wild typically is present in nuclei of cells around the blastopore and theectoderm of the developing pharynx (Fig. 6a, a`, b, b`). After treatment with 10 µM Azk it is found in nuclei all over the embryo (Fig. 6c). With 10 µM of CHIR FoxA is present in nuclei ubiquitously (Fig. 6d).

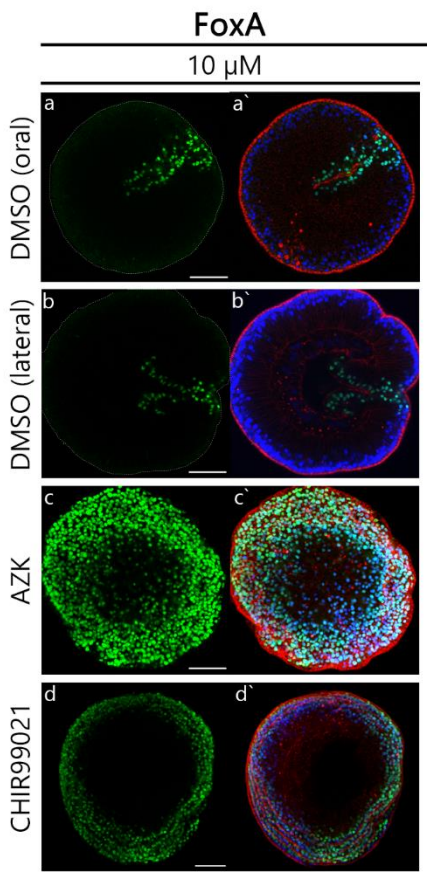

Fig. 6 Effect of Gsk3ß inhibitors on FoxA protein expression
in late-gastrulae (27 hpf) visualized by immunostainings. Two images of the same embryo are shown, left hand FoxA in green and right hand the merge with nuclei (DAPI, blue) and f-actin (phalloidin, red). (a & b) lateral views with oral pole to the right. In the left pictures the outlines of the embryos are traced with a white dashed line. (c & d) view not determinable, images consist of several pictures arranged as a z-stack with *Fiji* whereby (c) provides also a surface view and (d) consists only of cross sections. Chemicals used are written on the left, their concentration applied on the top, except for DMSO: DMSO control for 15 µM corresponds to 53 µM DMSO in *Nematostella* medium. With 10 µM of Azk and CHIR99021 FoxA is expressed ubiquitously in the nuclei. Azk = 1- Azakenpaullone. Scale bar: 100 µm.

With the higher resolution of the confocal laser scanning microscopy images, I could count the fluorescing nuclei, thus cells producing the transcription factors Bra/FoxA. I counted Brachyury in untreated embryos on average in 400 cells (n = 8). With a total cell number on average of 7000, this corresponds to ~6% of the cells containing Brachyury. I counted FoxA on average in 622 cells (n = 7) corresponding to ~9%. As I need 10^8 cells containing my protein of interest (according to Mohammed et al., 2016), this would require ~250000 embryos for investigating Brachyury and ~161000 for FoxA. Instead, after inhibitor treatment with 10 µM of CHIR99021 both proteins are present in 100% of the cells and only ~14000 embryos are required for one RIME experiment.

3.1.1.3 Western Blots confirm treatment effects on FoxA but WBs against Bra struggle

To further verify the effects of the different treatments, I conducted Western Blots, as they provide a quantitative analysis to help assess whether a saturation-effect is reached.
The Western Blots were each conducted with protein-extracts from: untreated, DMSO-treated, 10 and 15 µM Azk-treated and 10 and 15 µM CHIR-treated embryos 27 hpf. The extracts protein concentrations were assimilated.
The first Western Blot against Brachyury (44 kDa) has a weak band at the height of ~45 kDa in all samples (not shown) which would meet the expectations regarding size. However, no differences in intensity between the samples are observable. The second Western Blot against Brachyury shows only one band at the height of ~58 kDa for all samples (Fig. 7a, yellow

arrowhead). Apparently, Bra could not be successfully detected.
The Western Blot against FoxA (32 kDa) plus ß-Actin (45 kDa) as loading control, shows several bands in all samples (Fig. 7b). The band at ~32 kDa in all samples presumably represents FoxA (pink arrowhead). By comparison of the untreated and DMSO-only treated samples with the treated samples it becomes obvious that the FoxA concentration is highest in the 10 and 15 µM Azk and 10 µM CHIR samples. Unexpectedly, the 15 µM CHIR sample contains less FoxA but still more than in the untreated and DMSO extracts. Furthermore, all samples show a small band at the height of ~40 kDa which could represent ß-Actin, even so it is not exactly on the expected height (green arrowhead). If so, it supports that the treatments increase the FoxA concentration as ß-Actin shows no difference in intensity between the samples. Except for 15 µM CHIR, where the band is considerably weaker and hence the difference in FoxA concentration would be due to loading a less concentrated protein extract. Also, all samples show a band at ~57 kDa which could be of same origin as the band seen in the second Bra Western Blot (Fig. 7a and b, yellow arrowhead). It might represent unspecific binding of the primary or secondary antibody on the western-blotted-membranes. Lastly, they all show a band between a 135 and 100 kDa with strongest intensity in the DMSO sample and same intensity in the other samples. It probably also represents unspecific binding of the primary or secondary antibody.

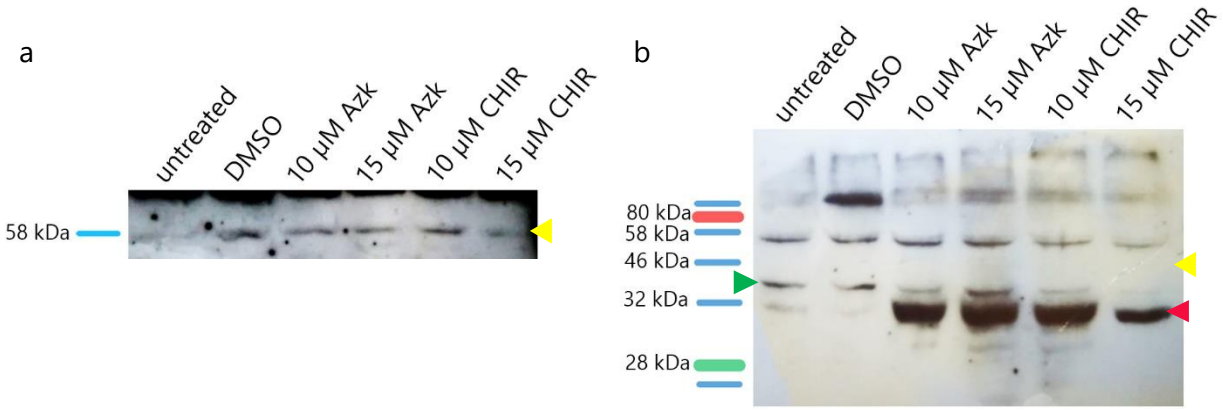

Fig. 7 Effect of Gsk3ß inhibitors on Bra and FoxA protein levels
Western Blots with cell extracts from late gastrulas (27 hpf) treated differently as specified on top. (a) Detection with Bra (44 kDa) antibody only revealed a band at 58 kDa (yellow arrowhead). (b) Detection of FoxA (32 kDa, pink arrowhead) and ß-Actin (45 kDa, green arrowhead). Please note the dramatic increase of FoxA protein in response to ectopic Wnt signaling. The first lane each represents the *Color Protein Standard-Broad Range* (NEB).

3.1.2 Effect of CHIR99021-treatment on expression patterns of 12 *sox* genes 27 hpf

I performed further in situ hybridizations against 12 of the *Nematostella sox* genes (see Table 5). As Sox proteins are suspected to physically interact with Bra, I checked for co-expression with *bra* (and *foxA*). Furthermore, I validated the effect of the Gsk3ß inhibitor treatment with 10 µM CHIR99021 on their expression. I matched the expression patterns with single cell RNA-seq data from 24 hpf embryos, provided by J. Steger (see supplement 6.4).

Table 5 *sox* Nve numbers and corresponding gene names used in the Technau lab.

Nve number	Gene name	Nve number	Gene name
None (v1g235335)	soxE1	Nve23841	sox3
Nve20190	soxE2	Nve15777	sox5
Nve16845	soxF.1	Nve23709	soxB1
Nve17897	soxF.2	Nve13400	sox1 after Magie et al. (soxB2 after (Royo et al., 2011))
Nve2102	sox2	Nve24655	soxB2 after Magie et al. (soxBa after (Royo et al., 2011))
Nve12762	soxC	Nve426	soxA

The *soxE1* probe leads to an overall slight staining of the untreated embryos with a slightly stronger staining in the oral ectoderm (Fig. 8c). It has been reported to be expressed in the oral domain (Magie et al., 2005) and was detected by the single cell RNA-seq in ecto- and endodermal cells and cells of the blastopore lip which supports the weak signal to be true. After treatment the embryo shows the same slight staining everywhere (Fig. 8c`). *sox9* also results in an overall slight staining but with some accumulation at the blastopore (Fig. 8d). Single cell data showed it in few ectodermal cells. The treated ones are also slightly stained everywhere (Fig. 8d`). *soxF1* and *soxF2* both show no staining neither in the untreated nor in the treated embryos (Fig. 8e, e`, f, f`). This has already been observed by (Magie et al., 2005) and appears also in the single cell data. *sox2* exhibits a salt and pepper pattern in the whole ectoderm (Fig. 8g). After treatment the pattern remains the same just appearing in less cells (Fig. 8g`). *soxC* shows also a salt and pepper pattern in the ectoderm plus expression at the blastopore reaching in the pharyngeal ectoderm (Fig. 8h). The salt and pepper pattern remains in the treated embryos but with disappearance of the developing pharynx the pharyngeal expression also disappears (Fig. 8h`). *sox3* gives no staining at all (Fig. 8i, i`). However, it has been shown to be expressed broadly in the oral ectoderm (Magie et al., 2005) and single cell data states that at least 24 hpf, it is expressed in the whole ectoderm including the blastopore lip. Hence, my probe must be non-functional. *sox5* shows few single cells in the ectoderm in a belt-like way around the middle of the embryos (Fig. 8j) supported by single cell data to be ectodermal and neuronal. After treatment no staining occurs anymore (Fig. 8j`). *soxB1* appears in a domain around the blastopore, the pharyngeal ectoderm and broadly in the aboral half (Fig. 8k). With 10 µM CHIR it stains the whole embryo quite intense (Fig. 8k`). The probe *sox1* stains the oral ectoderm, a bit further than halve of the embryo (Fig. 8l). The treated ones also are completely stained (Fig. 8l`). *soxB2* gives a salt and pepper pattern in a few ectodermal cells (Fig. 8m), which remains in the treated embryos (Fig. 8m`). *soxA* stains the outer lining of the blastopore but only with 60°C of hybridization temperature (Fig. 8n), with 64°C it showed no staining at all, just like in the treated embryos (Fig. 8n`) which is supported by single cell data.

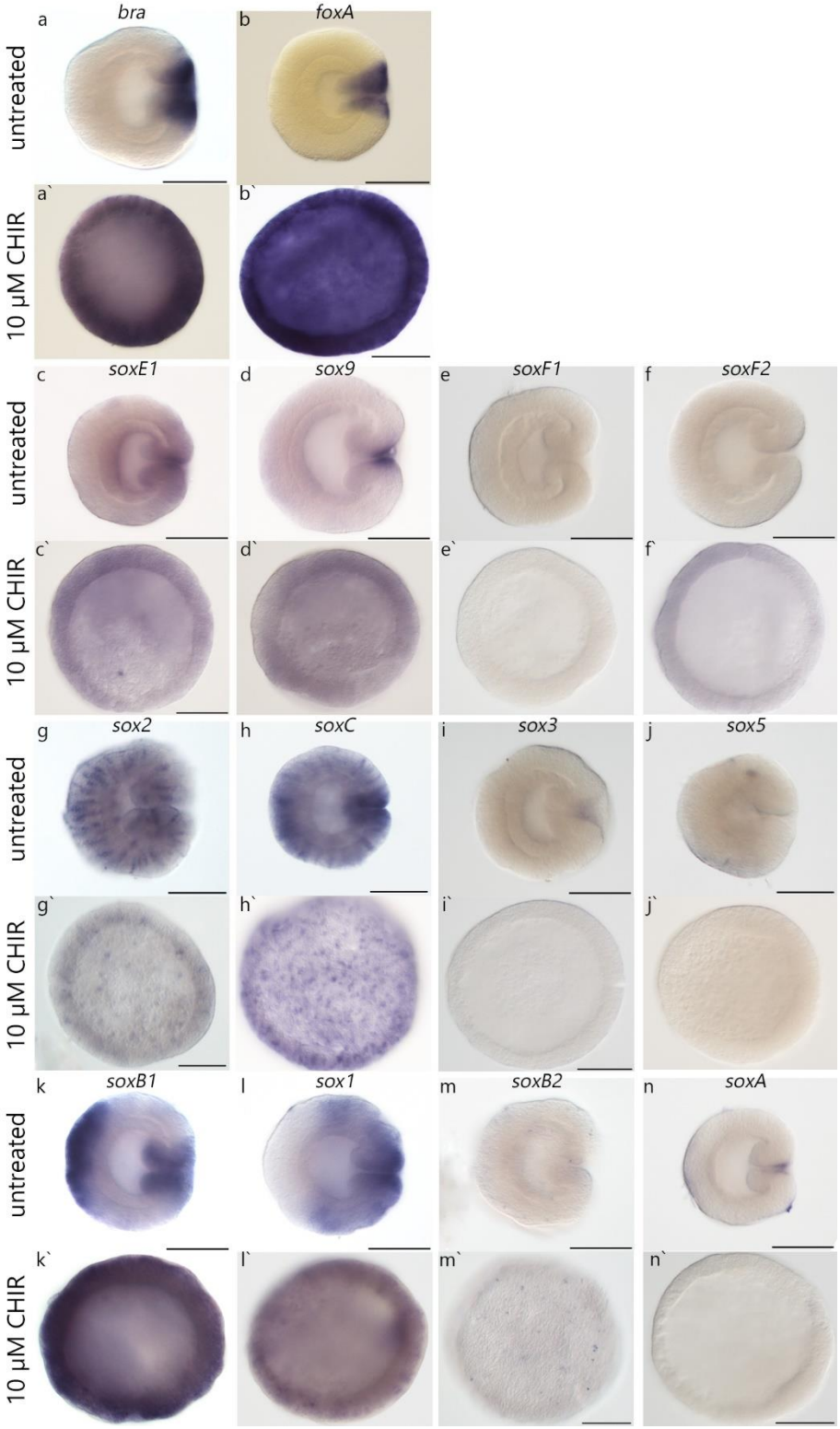

Fig. 8 Gene expression patterns of 12 *sox* genes in untreated and CHIR-treated embryos (in comparison to *bra/foxA*)

whole mount in situ hybridizations in late-gastrulae (27 hpf). The upper of each picture pair shows the control embryo with the untreated gene pattern: lateral views with oral pole to the right. The bottom picture shows the gene pattern after treatment with 10 µM CHIR: view not determinable. The gene name is written on top of the embryo pair. (a) *brachyury* expression. (b) *foxA*. (c-n) *sox* genes 10 µM CHIR = 10 µM CHIR99021 in *Nematostella* medium. Scale bar: 100 µm.

3.2 Parameter variations in the RIME protocol

I tested different parameters changes to obtain a protocol for RIME in *Nematostella*. To test this for me and the Technau laboratory new method(s), I started to work with the FoxA antibody, as it was available in large amounts whereas the Brachyury antibody was running out.

3.2.1 Formaldehyde fixation conditions

To stabilize potential interactions between Brachyury/Foxa with other cofactors (e.g. transcription factors) and to precipitate those during immunoprecipitation, the embryos were fixed with formaldehyde. To optimize the fixation conditions, I tested different parameters of time and concentrations: 1.85% for 15 minutes, 1.85%, 1%, 0.5% for 8 minutes and 1% for 5 minutes (see 2.2.3).

To decide which are the best conditions, I compared the workability of the fixed embryos during the post-fixation washes, i.e. if they would sink down after washing steps, their optical appearance after the post-fixation washes and the nuclei purity after their isolation out of the cells (Table 6). Also, I compared the ds DNA concentration after sonication as well as the total protein content by Bradford assay (Table 6). As I applied the 1.85% for 15 minutes fixation during my first RIME approaches, I have no data on protein content as at this time point, I did not perform Bradford essays yet. I had also not done sonication and therefore the nuclei purity evaluation is missing as well. Since stronger fixations might impair mass spectrometry, I excluded this fixation condition from further testing anyway.

Table 6 Effects of different formaldehyde fixation conditions on animal material.
Changes in formaldehyde concentration [%] and fixation time [minutes]. n = number of samples

Fixation condition	**1.85% 15 min**	**1.85% 8 min**	**1% 8 min**	**1% 5 min**	**0.5% 8 min**
Workability during post-fixation washes	Very convenient; sinking quickly after washing	Very convenient; sinking quickly after washing	Convenient; sinking autonomously after washing	Inconvenient, sinking only after slight centrifugation	Inconvenient, not sinking at all
Optical evaluation of embryos after post-fixation washes	Clear washing solution; white embryos	Clear washing solution; whitish embryos	Slightly more turbid washing solution; yellowish embryos	Turbid washing solution; yellow embryos, partially clumped	Turbid & slimy washing solution, very yellow and dissolving embryos
Optical evaluation of nuclei purity after cell lysis (Hoechst staining)	NA	High nuclei density; clean but many agglomerated	High nuclei density; clean and very few agglomerated	High nuclei density; clean and very few agglomerated	Low nuclei density; clean but many agglomerated
ds DNA concentration [ng/µl]	5 (n=2)	15 (n=1)	25 (n=6)	12 (n=1)	5 (n=1)
Total protein content [µg]	NM	373 (n=1)	611 (n=2)	921 (n=1)	2 (n=1)

Furthermore, I compared the total nuclear protein extracts from the differently fixed samples and a non-fixed sample on SDS PAGES stained with Coomassie Blue. Once before immunoprecipitation (IP) and once after IP with the FoxA antibody coupled to magnetic beads. The samples from before IP show a smear representing all proteins contained in the nuclear protein extracts (Fig. 9a). The smear is strongest in the 1% 8 min and 1% 5 min sample. In 1.85% 8 min it is weaker, in 0.5% 8 min even less and without fixation nearly no proteins are present. According to this, fixation with 1% formaldehyde for 8 or 5 minutes, results in the highest protein yield. However, the quantitative evaluation of stained gels is not completely reliable since the stainability is dependent on the particle density of the protein (Neuhoff et al., 1990). After IP with the FoxA antibody all samples, including the non-fixed sample, show only one band between 46 and 32 kDa at approximately 40 kDa (Fig. 9b). This could represent FoxA even though it is not on the expected height of 32 kDa. It is also noticeable that there is no difference in protein concentration. However, as this band was not detected by an antibody (like in a Western Blot) there is no clear evidence that the band does represent FoxA. Furthermore, this band is at the same height as the one observed in all Western Blots against Bra and FoxA as well when I only loaded the antibodies themselves. Hence it is more likely that it represents the FoxA antibody itself.

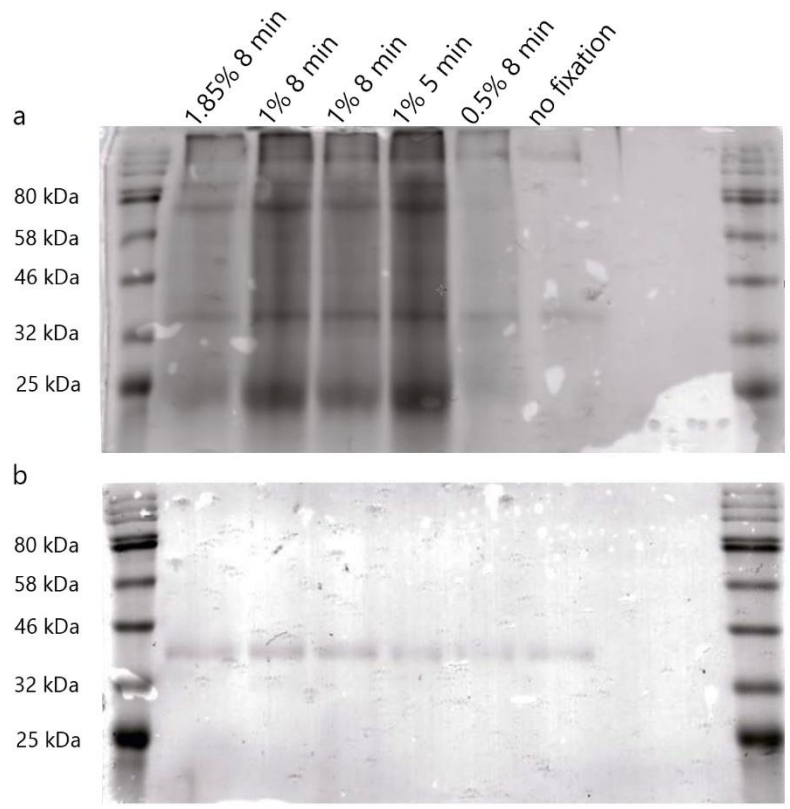

Fig. 9 Protein content in extracts from embryos fixed under different conditions before and after immunoprecipitation against FoxA on Coomassie-Blue-stained-SDS Page (12% bis-acrylamide). The first and the last lane each represent the *Color* Protein *Standard-Broad Range* (NEB). The fixation conditions are written on top: formaldehyde concentration in % and fixation duration in minutes. (a) samples from before IP. Smear represents all proteins contained in the nuclear extract. (b) same samples after IP against FoxA. Only a single band remains in all samples, probably representing the FoxA antibody itself.

As negative controls, I performed IPs without adding any antibody to the magnetic beads but following the exact same protocol ("only beads IP") and IPs with an α-mouse antibody ("α-mouse IP") (Fig. 10). Samples from before the IP show as well the smear representing all extracted proteins. However as expected, after the only beads and α-mouse IPs there is nothing visible on the SDS-PAGE. This suggests that no proteins are bound by the magnetic beads or an unspecific antibody.

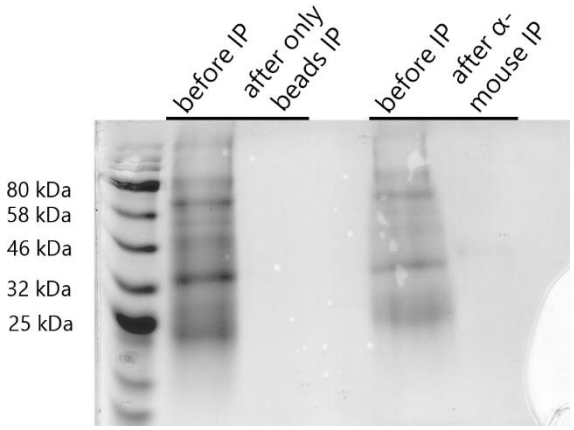

Fig. 10 Protein content in negative controls before and after immunoprecipitation
on Coomassie-Blue-stained-SDS Pages (12% bis-acrylamide). "only beads" = cell extract incubated only with the magnetic beads; α-mouse = cell extract immunoprecipitated with antibody against mouse antigen. The first lane represents the *Color Protein Standard-Broad Range* (NEB). The bars each mark the same sample, just once loaded before IP and once after IP.

Lastly, I also compared the results from the different fixation conditions with mass spectrometry which is presented in chapter 3.4.2.

3.2.2 Buffers for extraction of nuclear material

To optimize the method to isolate nuclei and eventually the chromatin from *Nematostella* embryonic cells, I compared the two buffer variants as described under 2.7.2.1. I fixed the embryos for 8 minutes with 1% formaldehyde, lysed the cells and isolated the nuclei and checked their purity under the fluorescent microscope by Hoechst-staining.
In the Technau lab protocol only one buffer called E1 buffer is applied for nuclei isolation. Following this protocol, the nuclei in the sample appear mostly clean and scattered without cytoplasm (blue scraps) attached to them (Fig. 11a, b, c). Whereas the Nature protocol from Mohammed et al. uses two different buffers for isolation, firstly the E1 and then a second LB2 buffer which is based on Tris-HCl instead of HEPES. In my hands this leads to less pure nuclei where you can often see remains of the cytoplasm. Also, nuclei are less scattered but often in small clumps (see Fig. 11d, e, f).

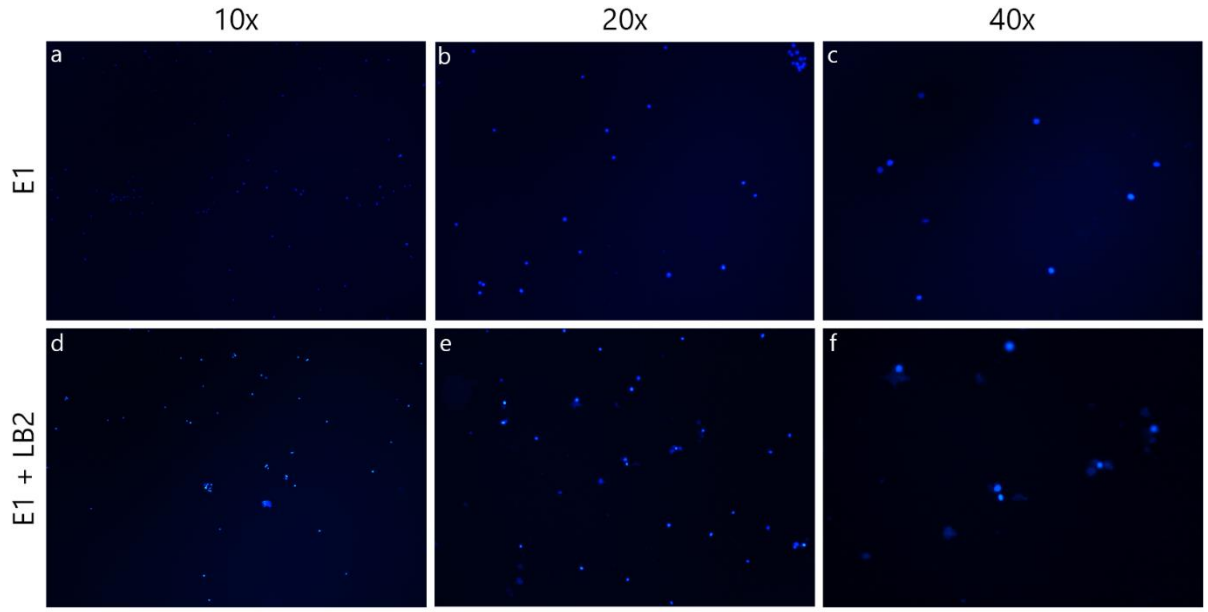

Fig. 11 Nuclei appearance after isolation with two different buffer combinations
during RIME procedure under the fluorescence microscope. Blue fluorescence = Hoechst stained nuclei. Written on top: magnification of pictures. The pictures are independent from each other. (a, b & c) isolated nuclei from cells lysed with E1 Buffer according to Technau Lab protocol. (d, e & f) with E1 and subsequently E2 Buffer after Mohammed et al., 2016.

3.2.3 Sonication settings for chromatin fragmentation into 200 – 600 bps

To shear the crosslinked chromatin into fragments between 200 – 600 bps, I tested different sonication settings (see 2.7.2.2). I then loaded around 300 ng of de-crosslinked precipitated DNA on a 2% agarose gel to check the fragment size.
In my first trials (Fig. 12a) the long sonications (16 cycles with 60 – 40 seconds) with 20% Duty Factor and an Intensity (I) = 5 lead to too short fragments between 50 – 200 bp. Similarly, using 10% and I = 4 lead to fragments around 200 bp. With severely reduced sonication time of 2 cycles for 40 s (10%/4 I), the chromatin is not completely sonicated, visible as a smear on the gel, whereas 2 cycles for 40s using 20%/5 I looked most promising with fragments around 300 – 400 bps. Therefore, I conducted more trials with short sonication durations, always with 40s but different amount of cycles (Fig. 12b). 10 cycles (both 20%/5 I and 10%/4 I) lead to around 100 bp long fragments just like 5 cycles (20%/5 I). However, 5 cycles with 10%/4 I lead to fragments around 300 bps, just like 2 cycles with 20%/5 I. 2 cycles with 10%/4 I result in a peak of fragments at 400 bps. Ultimately shearing of the crosslinked DNA into fragments between 200 – 600 bps was achieved with the following settings: Duration of 40 s for 2 cycles, Intensity 5, 20% Duty Cycle, and 200 Cycles/Burst. They reliably produce fragments mainly between 200 – 300 bps.

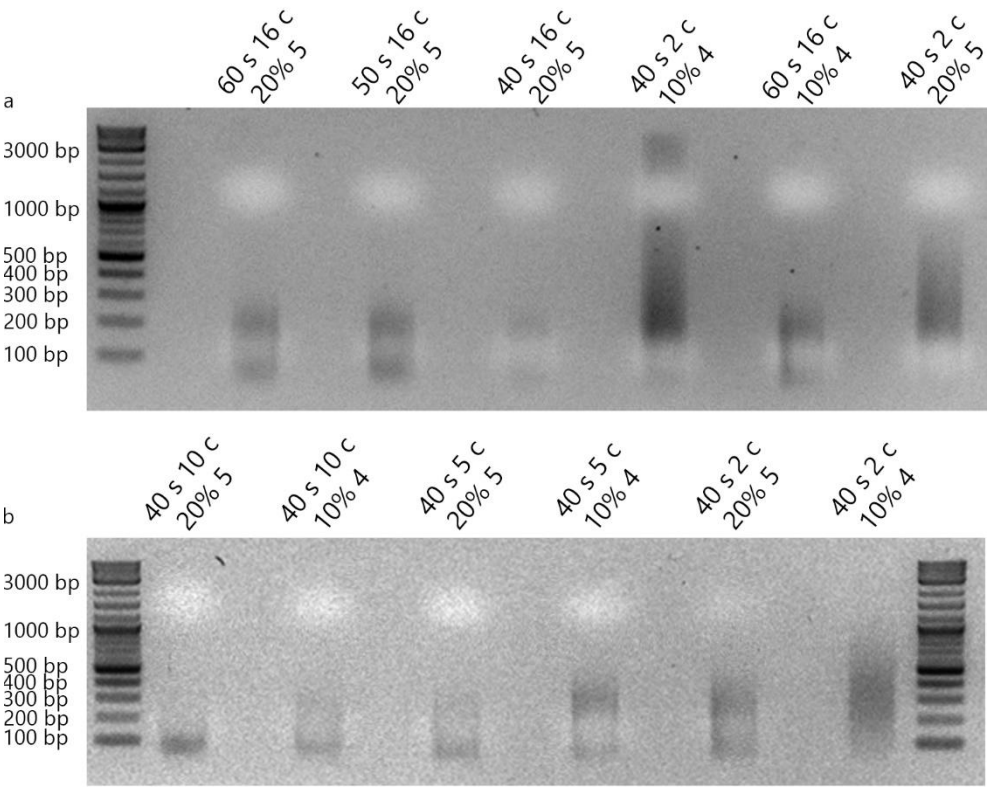

Fig. 12 Chromatin fragment size after different sonication treatments on 2% agarose gels. First (and last) lane: 100 bp DNA ladder. (a) first trials. (b) later trials. Written on top: applied sonication settings: duration [s]; number of cycles [c]; duty factor [%] and intensity *x*. 40 s 2 c 20% 5 led consistently to fragments between 200 – 400 bps.

3.2.4 Validation of immunoprecipitation efficiency by Western Blots

The right amount of antibody in an immunoprecipitation (IP) is crucial to identify the target protein plus any bound proteins and simultaneously keep the background, in form of unspecific binding, low. As mentioned (see 2.7.3.2), according to Mohammed et al., 2016, RIME needs between 2 – 20 µg of antibody per 100 µl magnetic bead suspension which needs to be optimized for every antibody depending on its specificity and affinity. Mohammed et al. optimized their antibody at 10 µg/100 µl bead suspension when immunoprecipitating from lysates of~6x 10^7 cells. I started working with the FoxA antibody in the RIME experiments to become acquainted with the protocol. I used 10 µg of the FoxA antibody (stock 1 mg/ml) on 100 µl beads. To check the immunoprecipitation efficiency, I conducted Western Blots afterwards.

Samples from after the FoxA-IP show a broad band between 46 and 32 kDa, in fact after incubation of the blotted membrane with the FoxA antibody (Fig. 13, lane 1, yellow arrowhead) as well as with the Bra antibody (Fig. 13, lane 5, yellow arrowhead) which leads to the suggestion that not the protein itself is detected but the antibodies that were used in the IP for RIME. The same observation can be made for the Bra-IP samples, which also have a band between 46 and 32 kDa after detection with the FoxA (Fig. 13, lane 3, yellow arrowhead) as well as with the Bra antibody (Fig. 13, lane 6, yellow arrowhead). This is one more hint that the FoxA/Bra antibodies themselves were detected on the western-blotted-membranes.

To check whether this omnipresent band between 46 and 32 kDa actually arises from the antibodies used in the IPs, I conducted WBs, loading only the FoxA and Bra antibody without

protein extract on the SDS-PAGE. Indeed, the same band can be observed for both. It did not matter whether I incubated the blot with the pure antibodies with the primary and secondary or only the secondary antibody. Furthermore, I incubated the blots with the sample from the FoxA- and Bra-IP just with the secondary (α-rabbit) antibody and it also resulted in a band between 46 and 32 kDa. This confirms that the FoxA and Brachyury antibody used for IP are detected in the Western Blots and not the precipitated proteins.

The supernatant (SN) taken after the FoxA-IP incubated with the FoxA antibody shows a band at ~33 kDa which might actually represent the protein itself, as it would be on the expected height (Fig. 13, lane 2, pink arrowhead). However, that means that the IP did not work or at least was not very efficient. The SN after Bra-IP shows nothing after incubation of the blot with the Bra antibody which is expected (Fig. 13, lane 7). But, the SN from the Bra-IP sample shows a band after incubation with the FoxA antibody at ~33 kDa (Fig. 13, lane 4, pink arrowhead), like the SN from the FoxA IP. This suggests that it could be FoxA itself as it should not be precipitated in the Bra-IP and therefore must still be in the supernatant.

As an unspecific negative control, I used the α-mouse (rabbit) antibody. This also leads to the band between 46 and 32 kDa (Fig. 13, lane 8, yellow arrowhead) underlining that this band arises from the detection of the antibodies used in the IPs.

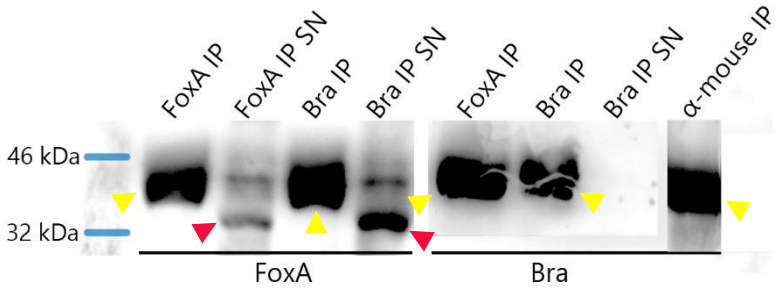

Fig. 13 Western blots after immunoprecipitations against FoxA, Bra & α-mouse with cell extracts from late gastrulas (27 hpf) prepared after RIME including CHIR99021 treatment. IP = immunoprecipitation, SN = supernatant after IP. Detection on the left membrane with FoxA and on the right with Bra antibody. The yellow arrowhead points at the omnipresent band probably representing the antibodies used in IP and the pink arrowhead points at FoxA.

To prevent the antibodies from the IP being detected in the Western Blots, I tried to prevent their appearance in the samples after IP, in the first place. Therefore, I worked with cross-linking the beads to the FoxA antibody with the *BS*^{3} *Crosslinker* before conducting IP in the RIME procedure (see 2.7.3.2).

However, those samples and their supernatants show the same band between 46 and 32 kDa at ~40 kDa, just much weaker (WB not shown)). Additionally, they show a faint band at ~55 kDa. Hence the cross-linking was not successful or the precipitation itself is not working.

3.3 Co-immunoprecipitation was not successful

I also conducted co-immunoprecipitations (Co-IPs) that do not need formaldehyde fixation to obtain the protein complexes for mass spectrometry (see 2.8). Testing different amounts of starting material to receive approximately 1000 µg of protein showed that the highest number of cells is needed, namely ~14 x 10^7 cells.

The samples from Bra-Co-IP and UT-Bra-Co-IP (untreated embryos) show a broad but weak band between ~46–36 kDa when assayed with the FoxA-antibody (Fig. 14, lane 1 and 2, yellow arrowhead). It appears on the same height as the band from the FoxA- and Bra-IP, so apparently the antibodies themselves are detected here as well. The two Co-IP samples being incubated with the Bra antibody, show blurry bands between ~48 to 40 kDa (Fig. 14, lane 6 and 8, orange arrowhead). As size determination is not necessarily distinct, this could also be on the same height as in the other samples. The supernatant (SN) taken after the Bra-Co-IP has a band at ~44 to 36 kDa, on the same height as the FoxA- and Bra-IP samples, possibly antibodies that dissolved from the beads and remained in the SN (Fig. 14, lane 7, yellow arrowhead). As negative control served also the α-mouse (rabbit) antibody in a Co-IP. This gives in both cases (incubated with FoxA and Bra) the same bands as the FoxA- and Bra-Co-IPs (Fig. 14, lane 3 and 9, yellow arrowhead), supporting the interpretation that the antibodies themselves are detected. As further negative control served only the magnetic beads without adding any antibody. It shows no signal as expected (Fig. 14, lane 4 and 10). The input control (6 µg protein measured by Bradford) from the blot incubated with the FoxA antibody, gives a slight undefined band at ~32 kDa which could correspond to FoxA (Fig. 14, lane 5, pink arrowhead). Whereas the blot incubated with the Bra antibody shows nothing Fig. 14, lane 10). This indicates like in chapter 3.1.1.3, Fig. 7, that the Bra antibody was not working in Western Blots under these circumstances.

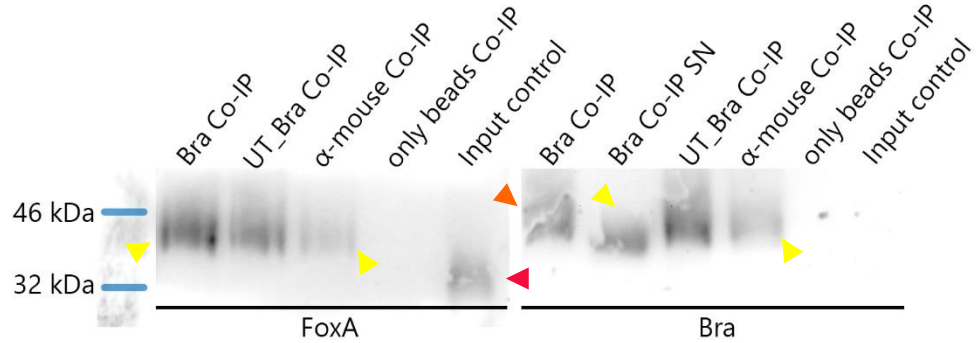

Fig. 14 Western blots after co-immunoprecipitations against FoxA, Bra & α-mouse with cell extracts from late gastrulas (27 hpf) prepared after the magnetic Co-IP protocol including CHIR99021 treatment except for UT = untreated embryos, Co-IP = Co-immunoprecipitation, SN = supernatant after Co-IP. Detection on the left membrane with FoxA and on the right with Bra antibody. The yellow arrowhead points at the omnipresent band probably representing the antibodies used in IP, the orange arrowhead points at a higher band possibly also representing the antibodies from IP itself and the pink arrowhead points at potential FoxA.

3.4 Mass spectrometry

The immunoprecipitated samples (see Table 2) were analyzed in the mass spectrometer to detect the m/z ratios of the contained peptides. The raw data coming out of the Q Exactive are displayed as chromatograms in combination with spectrograms. In the chromatogram (Fig. 15a) each peak represents peptides with the same retention time, meaning they needed the same time in the organic solvent and must therefore have the same features. The spectrogram (Fig. 15b) shows a segment out of the chromatogram at a certain timepoint whereby each peak represents ions with a specific m/z ratio.
Eventually the underlying proteins were determined and analyzed computationally.

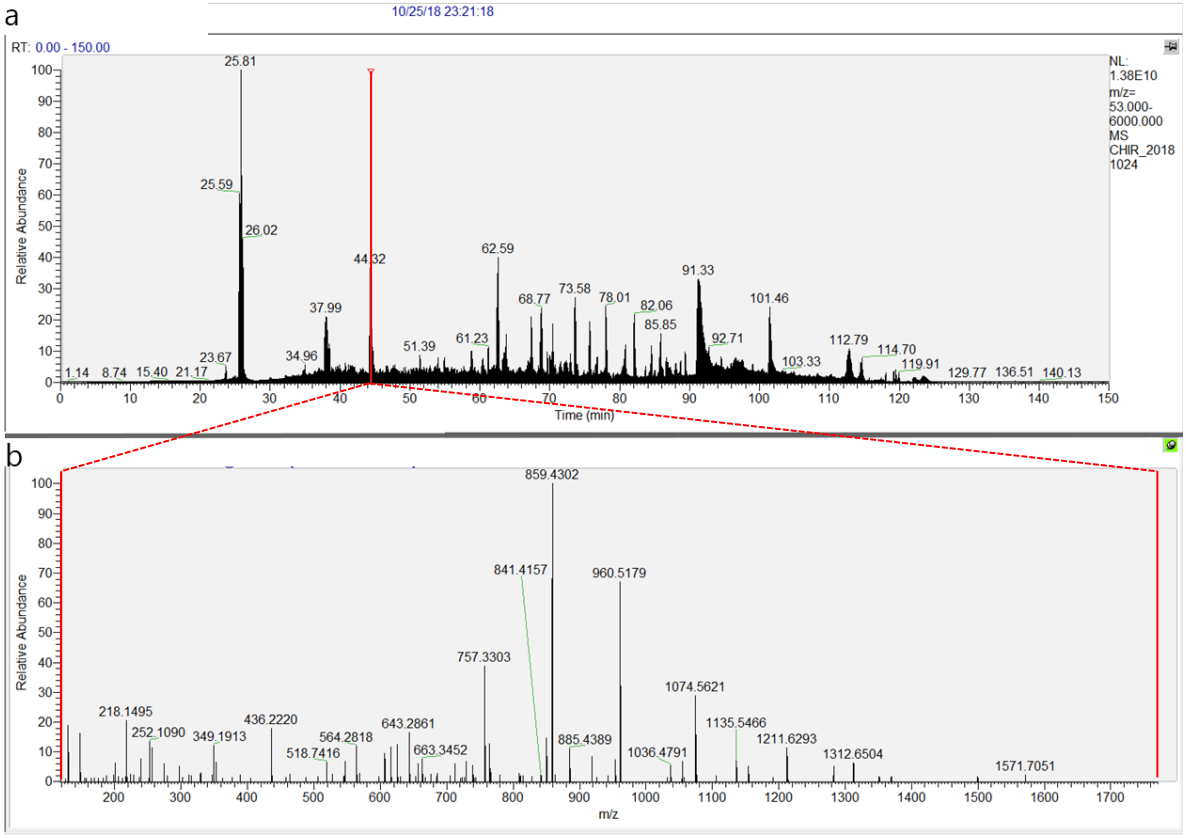

Fig. 15 Total ion current, full chromatogram of FoxA_1 RIME sample and an adherent spectrogram segment showing a FoxA peptide.
(a) chromatogram, the red line in the chromatogram specifies which segment is shown in b. (b) zoom into spectrogram showing a doubly charged ion from the mother sequence "VSSMTHNPVTVQSMAK" (peptide from FoxA protein) with a molecular weight of 859.4228 Da. "Relative abundance" = the amount of an ion produced in relation to the amount of the most abundant ion.

3.4.1 Half of the identified proteins in Bra- and FoxA-RIME are nuclear

To find out about the subcellular location of the identified proteins, I checked the Gene Ontology information on UniProt from the proteins identified with the UniProt-FASTA. Based on this I sorted them into proteins from the cytosol and proteins mainly present in the nucleus. The

distribution gives information on how pure the samples were and hence how well the protocol had worked isolating proteins out of the nucleus.
In the samples from all Bra-RIME experiments (Bra_1, Bra_2, UT_CoIP_Bra, CoIP_Bra; see Table 2) in total 463 proteins were identified. Half of them are cytosolic and the other half nuclear (Fig. 16a). In all FoxA samples (UT_FoxA, FoxA_1-8, FoxA_nonfixed, FoxA_0.5%8m, FoxA_1%5m, FoxA_1,85%8m; see Table 2) in total 2323 proteins were identified of which almost half are cytosolic and a bit more than half nuclear (Fig. 16b).

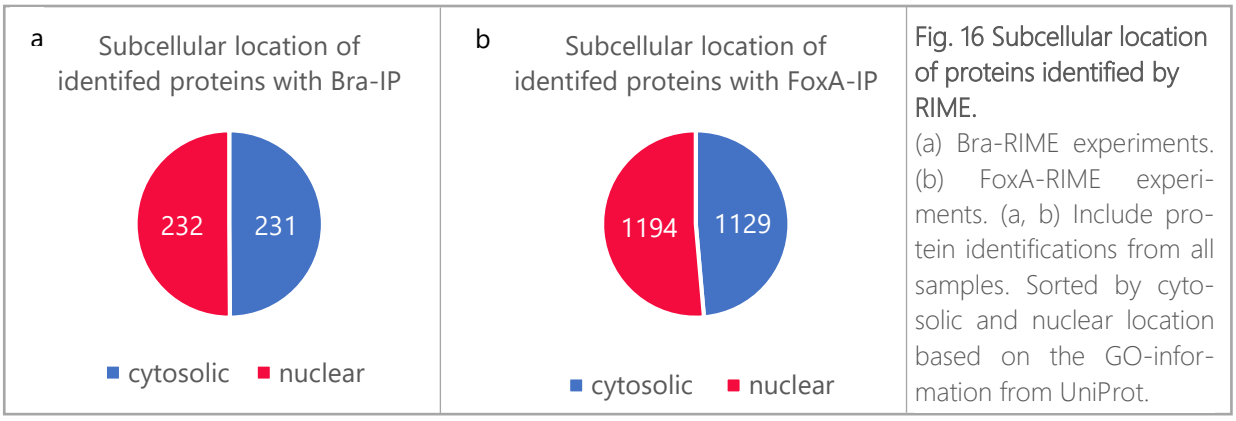

Fig. 16 Subcellular location of proteins identified by RIME.
(a) Bra-RIME experiments. (b) FoxA-RIME experiments. (a, b) Include protein identifications from all samples. Sorted by cytosolic and nuclear location based on the GO-information from UniProt.

3.4.2 Target protein (LFQ) intensity in differently prepared samples

As written in chapter 3.2.1, I additionally analyzed by mass spectrometry which way of sample preparation is the most compatible for RIME in *Nematostella*. Simultaneously I validated whether the single samples were successful and are of use for the identification of potential interaction partners. The protein that was targeted in the preliminary IP (in this case Bra or FoxA) must be identifiable by MS to validate that the antibody is compatible with RIME and to prove whether RIME worked with the specific sample. Hence, I ascertained for every sample whether the target protein (Bra or FoxA) could be detected. To compare the efficiency of the different sample preparations, I compared the LFQ intensity of FoxA between the differently prepared FoxA RIME samples. For Bra RIME I compared the unnormalized intensities instead of LFQ because according to LFQ intensity Bra would have been identified only in two samples versus according to intensity in three samples. This discrepancy could be explained by higher protein identifications and quantity by the LFQ determination and consequently an underestimation of protein abundance in the concerned sample.
As mentioned, looking at the unnormalized intensity Bra was detected in three out of four Bra RIME samples (Fig. 17a). It was also found in three out of four negative controls with an average intensity of 15 million and in one of the four beads-only controls with an average intensity of 1 million. However, at least in one of the two replicates, namely in the Bra_1 sample, Bra`s intensity was seven times more abundant than in the negative controls. In the other replicate and in the untreated Co-IP sample it was detected at similar intensities as in the anti-mouse antibody negative control. In the Co-IP sample intensity was zero. However, the fact that it was

found in the UT_CoIP_Bra sample suggests that IP worked at least in one of the two Co-IPs, interestingly with the non-treated embryos. Among the identified proteins in the Co-IP samples we found mainly histones and some cytoskeletal proteins but no transcription factors (apart from Bra in the UT sample) as potential cofactors which suggests that co-immunoprecipitation of proteins did not work.
Looking at the LFQ intensities of the FoxA RIME (Fig. 17b), FoxA has been detected in the negative control (α-mouse antibody) with an average of 30 million but in the beads-only control only with 1 million. In five out of the eight replicates it was detected with an average LFQ intensity of 103 million. In the samples FoxA_2/3/4 FoxA could not be detected at all (neither with LFQ nor with unnormalized intensity). Hence, these samples did not work. From the samples where I tried different fixation conditions, only 1% for 5 minutes detected FoxA but with lower LFQ intensity than in the replicates. This suggests apparently 1% formaldehyde for 8 minutes works best. No formaldehyde fixation leads to no FoxA LFQ intensity. Also, embryos without CHIRR99021 treatment do not contain enough target protein thus, treatment with CHIR is necessary.

3.4.3 Overlap of identified proteins between Bra- & FoxA-RIME experiments and negative controls includes Bra and FoxA

Specifically interacting proteins identified in a RIME experiment should not be found in the negative controls. Therefore, I evaluated how many of the proteins identified in the Bra- and FoxA-RIME samples were also found in the controls. Furthermore, I evaluated the differences in proteins found in all samples and in the replicates. For this analysis I used the protein list generated with the UniProt-FASTA.
150 from the total 356 proteins identified in all Bra samples were also identified in the controls (Fig. 18a). Of the 356 proteins, 309 were identified in the replicates of which 113 were identified also in the controls. So, 196 proteins are found exclusively in the Bra replicates and are supposedly specifically interacting.
Of the 2323 proteins found in all FoxA samples, 95 proteins were also found in the controls (Fig. 18b). Of the 2323 proteins, 2050 were identified in the replicates, of which 94 were also found in the controls. Here 1956 proteins are supposedly interacting specifically.
Overlap between all FoxA and all Bra sample includes 282 proteins which are either nonspecific proteins or could be cofactors of both transcription factors (Fig. 18c). The overlap between all Bra and all controls remains 150 proteins and the overlap between all FoxA and all controls is 217. 120 proteins were found in all three groups. Those are supposed to be unspecific proteins, but as they include amongst others, Bra and FoxA, they are not compulsory unspecific. Thus, I did not exclude these proteins from further analysis.

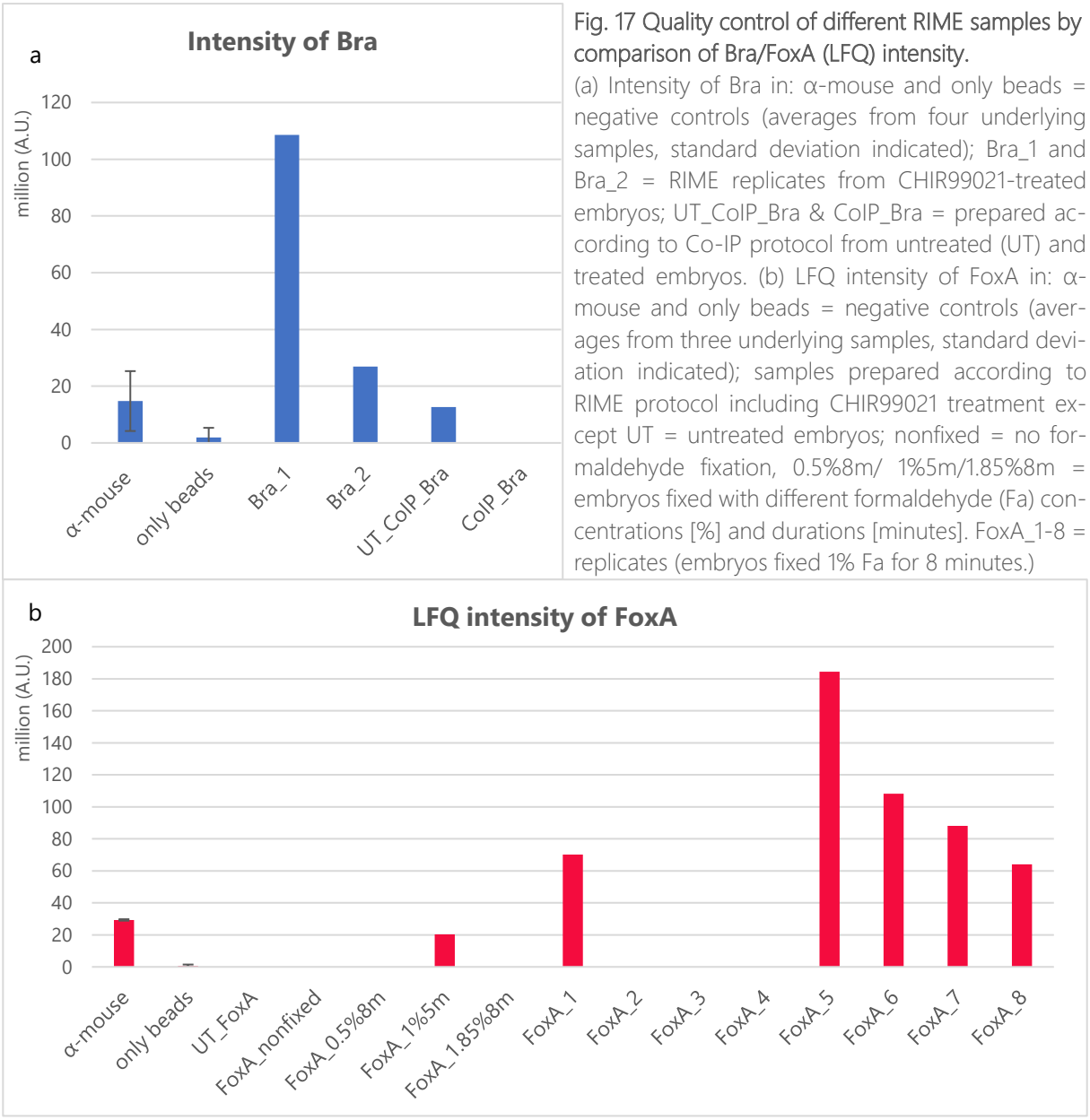

Fig. 17 Quality control of different RIME samples by comparison of Bra/FoxA (LFQ) intensity.
(a) Intensity of Bra in: α-mouse and only beads = negative controls (averages from four underlying samples, standard deviation indicated); Bra_1 and Bra_2 = RIME replicates from CHIR99021-treated embryos; UT_CoIP_Bra & CoIP_Bra = prepared according to Co-IP protocol from untreated (UT) and treated embryos. (b) LFQ intensity of FoxA in: α-mouse and only beads = negative controls (averages from three underlying samples, standard deviation indicated); samples prepared according to RIME protocol including CHIR99021 treatment except UT = untreated embryos; nonfixed = no formaldehyde fixation, 0.5%8m/ 1%5m/1.85%8m = embryos fixed with different formaldehyde (Fa) concentrations [%] and durations [minutes]. FoxA_1-8 = replicates (embryos fixed 1% Fa for 8 minutes.)

Fig. 18 Venn diagrams representing the overlap of identified proteins between different sample groups.
In brackets: the total number of proteins identified in the group. (a, b) Overlap of proteins between all Bra/all FoxA samples with the identically prepared Bra/FoxA replicates and with the unspecific controls α-mouse and only beads. 196 proteins identified exclusively in the Bra replicates and 1956 in the FoxA replicates. (c) Overlap of proteins between RIME experiments with all FoxA and all Bra and all unspecific controls. 120 proteins were identified in all three experiments including Bra and FoxA.

3.4.4 Several transcription factors under identified proteins

Next, I identified all transcription factors (TFs) out of the proteins identified in the replicate RIME samples. From the list generated with the Nve-FASTA, the proteins Nve numbers were matched to our annotated reference transcriptome. Like that I acquired a list of all contained DNA-binding-proteins for Bra and FoxA which is depicted in the supplements (see 6.4, Table 11 and Table 12). Within the framework of the master thesis I present only TFs of interest as potential interaction partners. Of those either knowledge about (potential) interaction exists already or they belong to a transcription factor family from which members are known to interact with members of the T-box/Fox family. Others were excluded either because they are general transcription factors or because information about potential interaction is not known (to me).

For analysis I firstly checked in how many samples they were found and by how many unique peptides they were identified in total, meaning non-identical peptides that only match to this protein in the proteome of *Nv*.

In Bra-RIME in total 396 proteins were identified of which 16 DNA are binding proteins (Table 12) of which five are transcription factors. This is a fraction of ~1.2%. The average fraction of TFs in the *Nematostella* genome is around 2%. I present three identified transcription factors in more detail (Table 7). Brachyury itself was found in both replicates with two different unique peptides. FoxA was also found in those samples with three unique peptides in total. One unique peptide identified LMX in one sample. Unfortunately, the *Nv* proteome does not contain the correct Lmx peptide sequence and could not be used for mapping of the peptides. The Lmx identification was taken from the UniProt database as it contains at least a 57 amino acids long sequence aligning to a part of the 239 amino acids long Nv Lmx sequence. Thus, possibly more peptides in more samples could have been identified, if it had contained the whole sequence.

Table 7 Partial list of transcription factors identified in Bra-RIME experiments.

Nve number	Protein name	TF family	Samples	Unique peptides (in total)
Nve20630	**FoxA**	Forkhead-box	2	3
Nve3568	**Brachyury**	T-box	2	2
Nve16579	**Lmx**	Homeobox	1	1

In FoxA-RIME in total 2591 proteins were identified of which 135 are DNA binding proteins (Table 11) and 63 are transcription factors. This makes a fraction of ~2.4% of the identified proteins being TFs and as mentioned the average fraction of TFs in the *Nematostella* genome is around 2%. I present 22 identified TFs in more detail (see Table 8). FoxA itself was found in five out of the eight replicates and was identified in total by three unique peptides. Bra was also found in the same five samples as FoxA and was identified by four unique peptides in total. We also detected FoxB in four samples with two unique peptides. The remaining proteins

were either found in two or only one sample but identified by different amounts of unique peptides. In FoxA_1 all proteins were detected. Some TFs were identified only in one sample by one peptide. Like in Bra-RIME, we detected also Lmx despite the poor mapping of Lmx to the database.

Table 8 Partial list of transcription factors identified in FoxA-RIME experiments.

Nve number	Protein name	TF family	Samples	Unique peptides (in total)
Nve20630	FoxA	Forkhead-box	5	3
Nve3568	Brachyury	T-box	5	4
Nve26195	FoxB	Forkhead-box	4	2
Nve15807	Grainyhead-like 2_mouse	GRH/CP2	2	13
Nve13458	Tcf/Lef	Sox-like HMG domain	2	4
Nve16579	Lmx	Homeobox	2	2
Nve16657	Smad4-like	TGFβ signalling	2	2
Nve15986	Noc	Zinc finger	2	2
Nve20136	COE	Zinc finger	1	4
Nve9807	Lhx6/8	Homeobox	1	4
Nve21766	TBX2	T-box	1	3
Nve5430	Pax-B_*Anthopleura japonica*	Homeobox	1	3
Nve4829	Forkhead box K1	Forkhead-box	1	3
Nve24711	Dachshund	DACH	1	3
Nve5479	Mothers against dpp 1_*Hydra vulgaris* (Smad1)	TGFβ signalling	1	2
Nve5105	Gli-Kruppel type zinc finger_*Daphnia pulex*	Zinc finger	1	2
Nve6473	Mesoderm induction early response 1_ *Zootermopsis nevadensis*	Myb/SANT	1	2
Nve26088	Mothers against dpp 2_human (Smad2)	TGFβ signalling	1	1
Nve613	REPO/REVPOL	Homeobox	1	1
Nve23709	SoxB1	HMG	1	1
Nve2485	HD032/NK-like 7	Homeobox	1	1
Nve19355	NFX1-type	Zinc finger	1	1

3.4.5 LFQ intensities of target proteins and potential cofactors

LFQ intensity can be used to compare the abundance of a certain protein between different samples as the peptides have the same features and hence comparison is not biased. Therefore, I compared the average LFQ intensities of each of the previously chosen transcription factors, between the replicates and the negative controls to check the differences in abundance (see Fig. 19a and Fig. 20a). Also, I compared the LFQ intensities of the different TFs between each other. This approach is only semiquantitative as the proteins consist of different peptides and

hence measurement is biased by their distinct features. However, it can at least provide hints about differences in abundance. I also present the intensity (without normalization) for comparison (see Fig. 19b and Fig. 20b).

For the Bra-RIME Bra itself, FoxA and Lmx were analysed. The average LFQ intensity of Bra shows its presence in the negative controls (Fig. 19a, blue bars). However, LFQ intensity in the replicates is around 20 times higher. For FoxA it is similar (Fig. 19a, pink bars). The average intensities show similar relations (Fig. 19b). According to the LFQ intensity, Lmx was not found in the controls but according to the intensity analysis it was detected with equal levels in the replicates and controls. Like the discrepancy between LFQ intensity and intensity for Bra in section 3.4.2, it could be explained by the fact that the LFQ determination leads to higher protein identifications and consequently an underestimation of protein abundance in the concerned sample.

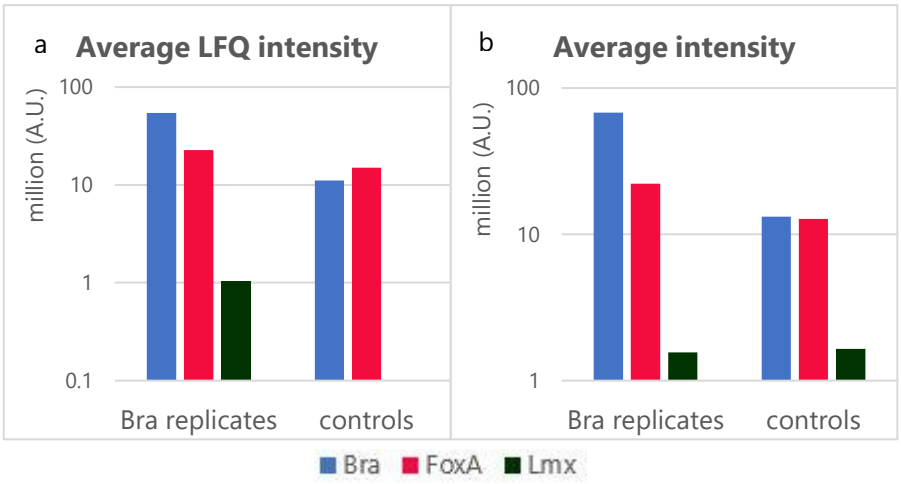

Fig. 19 Average (LFQ) intensities of the identified transcription factors in Bra-RIME replicates and negative controls.

(a) Average LFQ intensities. (b) Average intensities. Each bar represents the average value from all samples of the corresponding group: Bra replicates include only the samples that were identically prepared (Bra_1 and Bra_2), α-mouse and only beads samples are the unspecific negative controls. Y-axis has logarithmic scale.

In the FoxA-RIME FoxA and Bra were detected with 190-5 times higher LFQ intensities than all other identified transcription factors (except for Grainyhead-like 2 whose LFQ intensity is only around two times lower) (Fig. 20a). The same is reflected looking at the intensities (Fig. 20b). FoxA and Bra were also detected in the negative controls when analysed by their average LFQ intensity and intensity, yet considerably lower (around 5 times) (Fig. 20a and b). By contrast, FoxB and FoxK1 were measured with similar LFQ intensities and intensities in the negative controls as in the replicates. Lmx was detected in the controls with similar values as in the replicates when analysed by intensity but was not identified when analysed by LFQ intensity (same potential explanation as above). For the remaining TFs neither LFQ intensity nor intensity was detected in the negative controls.

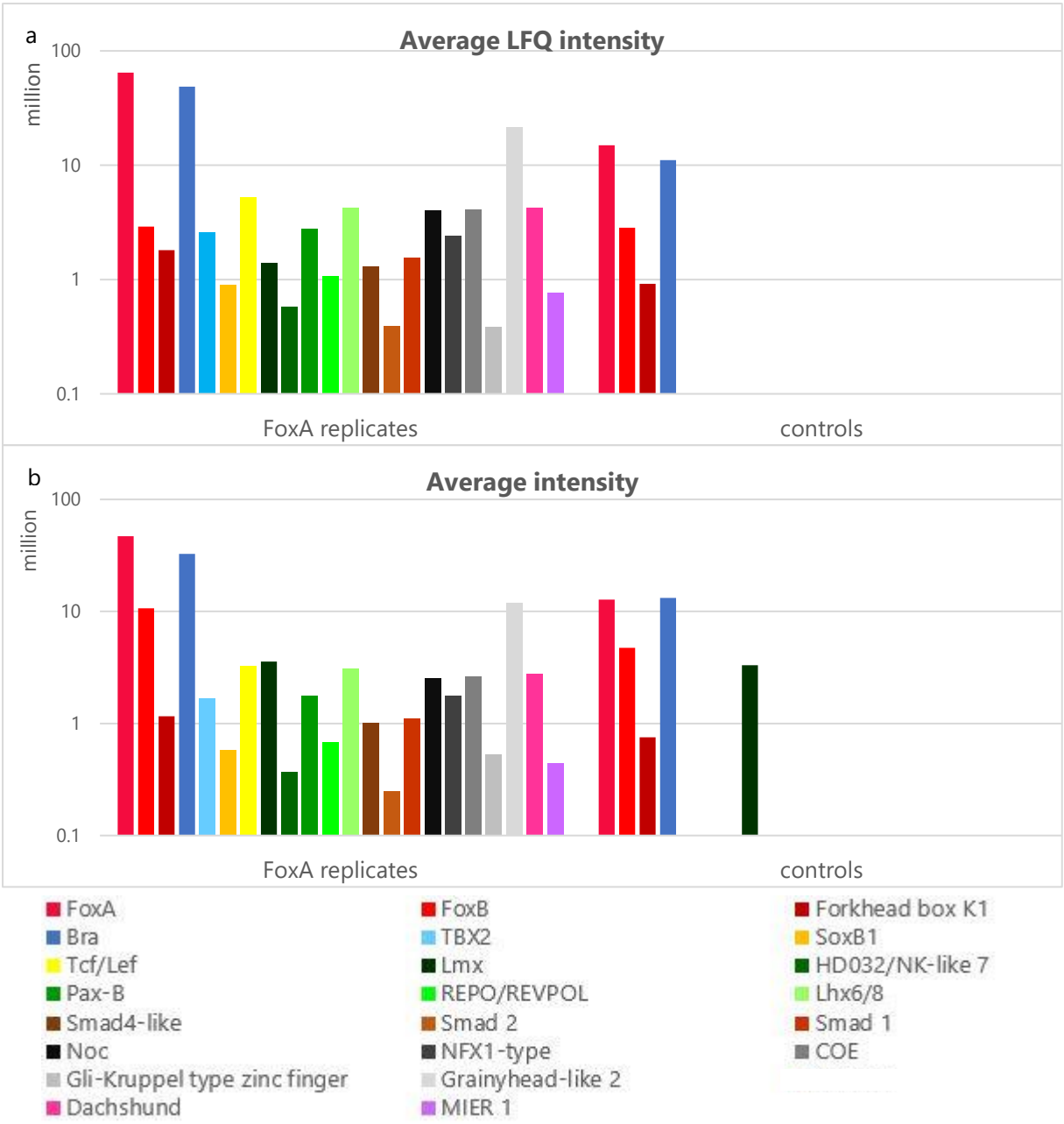

Fig. 20 Average (LFQ) intensities of the identified transcription factors in FoxA-RIME replicates and negative controls.

(a) Average LFQ intensities. (b) Average intensities.

The TFs are sorted after their family membership: red: Forkhead-box; blue: T-box; yellow: HMG; green: Homeobox; brown: Smad; black: Zinc finger; purple/pink: others. Each bar represents the average value from all samples of the corresponding group. FoxA replicates includes only the samples that were identically prepared but excluding samples FoxA_2/3/4. The unspecific negative controls include α-mouse and only beads samples. Y-axis has logarithmic scale, starting at 100000.

4. Discussion

4.1 RIME protocol for *Nematostella*

The first aim of this master project was to elaborate a RIME protocol for *Nematostella vectensis*. Rapid immunoprecipitation against endogenous proteins coupled to mass spectrometry provides a great tool to explore gene regulatory networks in an organism by identifying protein-protein interactions. As mass spectrometry is an unambiguous method to determine peptides, however, the reliability of the protein identification depends on sample preparation and on the quality of the protein database.

4.1.1 Early treatment with 10 µM CHIR99021 to obtain ubiquitous *bra*- and *foxA*-expression for subsequent RIME

To be able to perform RIME against Bra/FoxA in *Nematostella,* it was necessary to have enough protein material for mass spectrometry. To keep the amount of starting material as low as possible, we wished to increase the number of cells expressing Bra and FoxA in the embryo and all its potential interaction partners. One way is to ectopically activate the canonical Wnt signaling pathway, of which *bra* and *foxA* are targets (Röttinger et al., 2012). To this end, we tested Azakenpaullone and CHIR99021, two inhibitors of Gsk3ß, which are known to activate Wnt-signaling. In order to find the most efficient treatment leading to an ubiquitous *brachyury* and *foxA* expression in late gastrulae 27 hpf, I compared the effects of different concentrations of the two inhibitors.

Experiments revealed that 15 µM of Azk and 10 µM of CHIR as early treatment (from early cleavage to late-gastrula (~23 hours from 4 to 27 hpf at 21°C)) lead to a ubiquitous *bra*- as well as *foxA*- expression. Although in situ hybridisation is not fully quantitative, the differences in staining intensity were a robust observation. Immunohistochemistry confirmed that 15 µM of Azk and 10 µM of CHIR leads to the presence of both transcription factors in all nuclei. The ball-like morphology of the treated embryos is due to the fact that stabilization of ß-catenin in the whole embryo leads to a severe oralization and gastrulation cannot proceed anymore (Leclere et al., 2016).

Western Blots could confirm the effect of treatment on FoxA in general but failed to prove differences between the individual treatments and therefore remained inconclusive. Thus, an option to improve the WBs would be perhaps to shorten the incubation time with the primary and secondary antibody (Rehm and Letzel, 2010). Furthermore, the WB samples should include knockdown embryos as well. The Brachyury antibody seems to not work in WBs in my hands. However, as it has been successfully used in ChIP seq and immunohistochemistry, it should work in immunoprecipitation.

Although WBs could not support it, the other experiments indicate that CHIR is more efficient in lower concentrations than Azk, hence I suggest 10 µM of CHIR99021 as standard treatment for early embryos used in Bra- or FoxA-RIME experiments.

In non-treated embryos, no FoxA could be detected by MS, suggesting that untreated embryos do not contain enough FoxA and overexpression is necessary. However, as I measured only one untreated sample this should be repeated with more replicates to be confident. In the Bra-RIME the untreated sample did contain Bra but with much lower intensity. Here as well more replicates should be examined. It is important to validate the necessity of the treatment. Although CHIR99021 does activate an endogenous pathway and is therefore more natural than e.g. an overexpression of Bra or FoxA it increases gene expression and hence side effects cannot be excluded. For example, proteins might interact with proteins that they normally would not and consequently RIME would identify false positive interaction partners. Or it could miss true cofactors that are not expressed anymore because they are suppressed by enhanced Wnt signaling.

4.1.2 1% formaldehyde for 8 minutes fixation of embryos for RIME

Since formaldehyde fixation leads to intramolecular crosslinks and masking of epitopes, mass spectrometry-sensitivity is impaired by formaldehyde used in RIME (Mohammed et al., 2016). This can severely reduce sequence coverage of proteins (Sutherland et al., 2008). Also, formaldehyde treatment can lead to loss of nuclear proteins during sample preparation. The nuclear proteins crosslinked to the DNA can form insoluble complexes, which become precipitated during lysis and consequently removed (Klockenbusch and Kast, 2010). Hence the risk of impairing sensitivity should be minimized by keeping the formaldehyde concentration low and fixation time short.
While stronger fixations with 1.8% formaldehyde for 15 or 8 minutes as conducted for ChIP-seq are more convenient, the LFQ intensity for the target protein FoxA was zero. Thus, for RIME this fixation protocol should be excluded. By contrast, samples fixed with 1% formaldehyde for 8 minutes were optimal, indicated by the detection of FoxA in five out of eight samples. In conclusion, these fixation parameters were used as the standard for RIME experiments.

4.1.3 Isolation of nuclear material with E1 Buffer followed by a fractionation efficiency control

The protocol was also optimized for the isolation of nuclei. The Hoechst staining showed that the use of E1 Buffer alone leads to cleaner and more scattered nuclei in *Nematostella* than the combined use of E1 and LB2 buffer as done by (Mohammed et al., 2016).
Still, half of the proteins identified in the Bra as well as the FoxA-RIME are supposedly from the cytosol. How meaningful this is can only be said if the ratio of cytosolic to nuclear proteins is significantly lower in the whole proteome which I could not compare. But, as mentioned, around 2% of the *Nematostella* genome are transcription factors. Of the proteins identified in Bra RIME only around 1.2% were TFs and in the FoxA RIME 2.4%. On the basis that nuclear material should be enriched and hence especially transcription factors this is low. Thus, a further step that should be included in future RIME experiments is control of the fractionation

efficiency before the actual IP against the target protein. Therefore, the nuclear fraction and the remaining cytoplasmic fraction are incubated with an antibody against an exclusively nuclear protein (e.g. histone H3) and separately against an exclusively cytosolic protein (e.g. GAPDH). Eventually these IPs are checked on a Coomassie-Blue-stained SDS PAGE or a Western Blot to detect any contamination of cytoplasmic protein in the nuclear fraction or vice versa (Lo Piccolo et al., 2015). This way uncleanly fractioned nuclear material is identified or if a lot of nuclear material remained in the disposable cytoplasmic fraction. It reduces the risk of measuring poor quality samples by MS, possibly saving money and time. Also, it assures that identified potential interaction partners are interacting with the target protein in the nuclear fraction and therefore on the chromatin. This is the requirement for them to interact in transcriptional regulation.

4.1.4 Western Blots could not confirm successful IP in RIME experiments but suggest lower antibody concentrations

Immunoprecipitation and consequently the used antibody are crucial for the success of RIME experiments. Therefore, I aimed to validate the Bra- and FoxA-IPs with Western Blots against the corresponding protein.
Unfortunately, the WBs could neither detect FoxA nor Bra in the samples after IP. A dominant band detected represents most likely the Bra and FoxA antibodies. Denatured antibodies leftover from the IP, bind the secondary antibody nearly as good as the assaying antibodies sitting on the antigen (Rehm and Letzel, 2010). Crosslinking the antibodies to the magnetic beads with BS^3 for IP reduced the dominant band, yet did not remove it completely, indicating that crosslinking worked but not sufficiently.
However, in the supernatants both after the FoxA and Bra IP and in the input control, potentially FoxA could be detected. As FoxA could also be detected in the protein extract samples, the FoxA antibody seems to generally work in WBs and it is suspicious that FoxA could not be detected in the IP samples. Bra by contrast, could not be detected in any WB. However, it is possible that antibodies work well in IP but not in WBs. As written previously, the Bra antibody worked in ChIP-seq which supports the assumption that the WBs did not work rather than the IP itself. For the future to reveal if a band corresponds to Bra/FoxA, samples with knockdown of Bra/FoxA should be run in comparison on the Western Blot.
In general, the WBs with BS^3 crosslinking could be repeated and further optimized. Alternative to BS^3 would be biotinylation of the Bra/FoxA antibody for the immunodetection on the WB, and detection of biotin with avidin coupled to horseradish peroxidase. Also, I washed the samples after IP with RIPA buffer before bringing into Laemmli buffer. Instead it may be better to wash only with PBS and elute directly in Laemmli buffer. This less stringent elution could prevent large amounts of antibodies being eluted from the beads and mask real signal. Anyway, the very intense band representing the antibodies used in IP, indicates that the amount of antibody is excessive. Lower amounts than 10 µl/100 µl beads could be tested in WBs and

optimized for subsequent RIME. This would probably also lower unspecific binding and reduce the amount of identified false positive proteins. Lastly, an input sample (whole cell lysate from before IP, prepared equally) as positive control should be included in every WB (BIO-RAD, 2016).

4.1.5 Magnetic Co-IP not successful

The co-immunoprecipitation detected Bra at least in one of the two samples. But as only Bra was detected and besides histones and some cytoskeletal proteins, no transcription factors that could represent potential interaction partners, it seems like Co-IP did not work and fixation with formaldehyde is indispensable. However, to really assess whether this method works or not, more replicates need to be measured. Because in the end, Co-IP would provide a more convenient, quicker and less impairing method.

4.1.6 (Potential) RIME improvements

Several points should be considered and improved in future RIME experiments.
For mass spectrometry there should always be a minimum of three biological replicates per sample type prepared in parallel and measured in the same MS run. The higher the number of replicates the better, because it increases the number of peptide spectrum matches (total number of identified peptide spectra matched for the protein) and the confidence in identified proteins (Mohammed et al., 2016). Thus, RIME against Bra should be repeated with at least three biological replicates in parallel. Although I prepared negative controls, I did not prepare them from the same lysate preparation which should be considered in future experiments. Furthermore, the controls should be measured in the same MS run with the replicates. Given that in the negative controls, in particular α-mouse controls, Bra and FoxA were identified, it can be concluded that unspecific binding to the magnetic beads but especially to the antibodies themselves can occur. Consequently, protein identifications appearing in replicates and controls without substantial differences in (LFQ) intensity, cannot be trusted. A way to reduce unspecifically precipitated proteins, would be to preclear the cell lysates with magnetic beads for ~2 hours before adding the antibody coupled beads.
Another important point is the possibility that the sheared fragment length of 200 – 300 bps is too long. Identified proteins could just be neighbours and not physical interaction partners of the target protein. Hence the sonication settings could be adjusted to try working with shorter chromatin fragments
Ultimately, I would recommend for further RIME experiments to prepare the samples in a mass spectrometry-oriented laboratory with their equipment and devices being qualified for MS. My samples contained some contaminants which do overlie true signals and, in the end, lower sensitivity. If salts and other additives do not have mass spectrometry purity, disturbed evaporation and formation of complexes might hinder proteins to access the machine and signals become suppressed (Rehm and Letzel, 2010). Also, storage devices and working materials

should be MS-compatible to prevent contaminations with polymers e.g. polyethylene glycol or detergents e.g. Triton X-100.

Other approaches could be tested for optimizing the investigation of interaction partners. The alternative to in-solution-digestion of samples with trypsin is in-gel-digestion. After separation of the sample on a SDS Page, the protein band is cut out from the gel and digested with trypsin. In parallel a WB should be carried out to confirm that the band does represent your target protein. This way less unspecific proteins get into MS and contaminants remain in the gel but on the other hand it decreases throughput (Mohammed et al., 2016). A method to confirm identified proteins is to use other peptide cutting enzymes than trypsin e.g. asparagine peptide lyase. As they produce different peptides, identification of the same proteins is a valuable countercheck and the protein sequence coverage can be increased (Mohammed et al., 2016).

4.2 Transcription factor interactions in the mesendodermal GRN

The second aim was to further investigate the GRN in *Nematostella* with RIME. Bra as well as FoxA are essential for mesendoderm development in Bilateria but are involved in endodermal patterning in the mesoderm-less cnidaria. Thus, they may interact with each other and coregulate target genes in a conserved manner. My results are consistent with a physical interaction between Brachyury and FoxA. Furthermore, the data support physical interactions between a group of transcription factors supposedly setting up the oral domain in *Nematostella*, namely Bra, FoxA, Lmx and FoxB. Lastly, the results point to interactions of FoxA with several other transcription factors that should be further examined.

4.2.1 Co-expression of *bra*/foxA with *sox* genes

The ChIP-Seq against Bra determined binding sites with and without T-box binding motifs. When peaks without a Bra motif were analysed for other TF binding motifs, Sox transcription factor binding motifs were the most abundant ones. This suggested that in these cases, Bra does not bind directly to the DNA but may interact with other TFs, in particular with TFs of the Sox family. Hence, I investigated the gene expression patterns of 12 *Nv sox* genes to see whether any is co-expressed with *brachyury*. As I performed RIME against FoxA as well, I also checked co-expression with *foxA*. I have not conducted in situ hybridizations against the *Nv sox* genes *soxJ* (v1g120772), *sox4* (v1g1122786) and *soxH* (Nve3409) because their ORF sequence was not clear at that time point and probe generation was not possible.
The ISHs showed that *soxE1*, *sox9*, *soxC*, *soxB1*, *sox1* and *soxA* (names after (Magie et al., 2005)) have overlapping expression domains with *bra* as well as with *foxA* and hence are candidates for an interaction with those. As mentioned, *sox3* was shown previously to be expressed in the oral ectoderm of the gastrula too and is therefore also co-expressed (Magie et al., 2005). After treatment with the Gsk3ß inhibitor CHIR99021, the expression of *soxE1* and *sox9* seems to be upregulated like *bra* and *foxA*. However, staining of the two *soxE* genes was in general very

weak and as I conducted only two successful ISHs, it must be repeated to confirm this effect. *soxB1* and *sox1* were clearly upregulated as a result of CHIR treatment and showed ubiquitous expression like *bra* and *foxA*. In contrast the salt and pepper pattern of *soxC* remained but expression at the blastopore and in the pharyngeal ectoderm where it is partly co-expressed with *bra* and *foxA*, disappeared.

4.2.2 RIME could not identify any (Sox) proteins as Bra interaction partners

To our surprise, Bra-RIME identified no Sox protein at all. Therefore, I cannot verify physical interactions between Bra and Sox TFs as suggested by the Bra ChIP-seq. Therefore, there is also no evidence that Bra and Sox2 might act together in *Nematostella* as they do in mice in the neuromesodermal progenitor cells (Koch et al., 2017). However, the results do not reject the hypothesis that Bra interacts with Sox protein(s). Firstly, I only prepared four Bra samples, among which only two were prepared in accord with the optimal sample preparation. Another issue to be considered is the previous treatment with CHIR99021. For Azakenpaullone it has been shown that genes react differently to overactivation of Wnt/ß-catenin signaling (Kraus et al., 2016). Some genes show a "saturating" expression behavior; originally oral expression expands aborally with increasing Azk concentration until expression is saturated in the entire embryo. But some genes are "window" genes; expression also expands aborally but from a certain concentration on, their expression domain is shifted to the aboral pole and eventually their expression vanishes. Indeed, 10 µM of Azk makes expression of the investigated genes disappear, as CHIR seems to be even more efficient, it most likely has the same effect. It was shown in this thesis, that *sox5* and *soxA* expression (and *soxC* expression partly), are abolished after 10 µM CHIR treatment. Hence those could not have been identified by RIME anyway. The same could be true for other candidate Sox genes against which I did no in situ hybridizations in CHIR-treated embryos. However, this can be easily examined by treating embryos with different ascending concentrations of CHIR and subsequent control in situ hybridizations.
Also, no other transcription factors were identified except for Lmx with one peptide in one sample. However it is for example known from several T-box proteins that they bind homeobox factors to then synergistically activate target genes during metazoan development (Papaioannou, 2014). Hence, as mentioned, Bra RIME should be repeated to consolidate a potential interaction with Lmx and other homeobox proteins and members from other transcription factor families too.

4.2.3 Network between Bra, FoxA, Lmx and FoxB setting up the oral domain is supported

The ISHs and immunostaining show that Bra and FoxA are co-expressed around the blastopore in the ectoderm of the developing embryo. Bra ChIP-seq showed that *foxA* is a transcriptional target of Bra in *Nematostella* (Technau lab, unpublished). It has been suggested by many researchers and studies that there is a conserved interaction between them, working together in

transcriptional regulation. My results using RIME strongly support a direct physical interaction between FoxA & Bra. RIME against Bra identified FoxA and the other way around Bra was identified in the FoxA-RIME. Thereby, in both cases they were identified in the same samples as the target protein. Also, in both cases they were the second most frequently detected protein after the target protein and measured within the same (LFQ) intensity range.
As the sample size for FoxA-RIMEs was 13, the result is trustworthy. However, I did not prepare all the samples on the same day and did not measure them in the same MS run. This should be done for validating the results and for Bra-RIME more replicates should be measured.
Double whole mount in situ hybridizations have already been performed and confirm overlapping expression of *bra* and *foxA* (T. Bagaeva and G. Genikhovich, unpublished).
Recent findings established a network between the transcription factors FoxA, Bra, Lmx and FoxB as crucial for patterning the oral domain in *Nv* (G. Genikhovich + T. Bagaeva, unpublished). They are all co-expressed in the central domain orally around the blastopore. As written, my results support an interaction between Bra and FoxA. Also, I identified FoxB in the FoxA-RIME, which either states it as direct interaction partner of FoxA or it could mean that Bra interacting with FoxA is also binding FoxB. Then they bind the DNA as a complex of Bra, FoxB and FoxA. FoxB was not identified in the Bra-RIME at all. But as mentioned, the Bra-RIME must be repeated to be reliable. The genetic interactions of Bra and FoxA with Lmx have been demonstrated but physical interactions have not yet been proven. My results can support a potential physical interaction of FoxA with Lmx. The physical interaction between Bra and Lmx cannot be confirmed as Lmx was only detected in one replicate with low intensity and it was also found in four of the eight controls.

4.2.4 Interaction of FoxA with several transcription factors needs validation

In the FoxA-RIME, SoxB1 was identified in one sample with 1 unique peptide only. Still, this result would match up with the fact that *soxB1* is expressed in the same domains as *foxA* orally around the blastopore and in the developing pharynx. *soxB1* is also upregulated upon enhanced Wnt signaling like *foxA*. In the sea urchin, SpSoxB1 is needed in ectodermal differentiation and has to be downregulated to allow nuclear ß-catenin localization for endodermal and mesenchymal specification (Kenny et al., 2003). Possibly in *Nematostella*, FoxA and SoxB1 exhibit combinatorial transcription control to promote ectodermal development.
Apart from Bra, SoxB1, FoxB and Lmx one further Forkhead-box TF and T-box TF have been identified in the FoxA RIME. Also, several zinc-finger containing TFs, homeobox TFs, two Smads, and members from other families were identified. For all transcription factors to be validated, further RIME experiments against FoxA must be performed. Especially the ones only identified with 1 peptide in 1 sample need further validation. As mentioned above, possibly the chromatin fragment length is too long meaning that identified proteins could just be neighbours and not physical interaction partners. For the candidates where it has not been done yet, whole mount in situ hybridizations at 27 hpf should be performed to check for overlapping expression

domains with FoxA. Another important next experiment is ChIP-seq against FoxA to detect FoxA binding sites in the genome. Maybe this will show whether FoxA binds to the same sites as some of the identified transcription factors, e.g. Brachyury. Combinatorial analysis of the RIME and ChIP-seq results will bring more clarity. Next, antibodies could be generated against new candidate interaction partners to perform RIME and to obtain independent evidence for their interaction (ten Have et al., 2011).
Furthermore, alternative methods that can identify physical interaction partners should be applied for validation. One possibility is the yeast-2-hybrid-screen (Fields and Song, 1989), however, this is a fairly artificial approach as interaction is not determined in the original cells but in yeast cells.

An alternative method to ChIP would be a method called Cut & Run (Skene and Henikoff, 2017). It works by antibody-targeted controlled cleavage by a micrococcal nuclease (MNase) that releases specific protein-DNA complexes into the supernatant. It can be performed in vivo and without the disadvantages of fixation. It could then be followed by the MS procedure to identify interaction partners of the targeted protein. The question is whether the protein-protein interactions are strong enough to be conserved throughout the procedure.
Another method to investigate spatial localization and interactions of proteins in cells is the so called TurboID (Branon et al., 2018). The protein of interest becomes labelled with an enzyme (the *E. coli* biotin ligase BirA) which tags all proteins in close proximity. Subsequently all tagged proteins are purified and analysed by mass spectrometry. The big advantage is the insurance that identified proteins must have been near neighbours of your protein of interest.

5. References

ALBERTS, B., JOHNSON, A., LEWIS, J., RAFF, M., ROBERTS, K. & WALTER, P. 2011. Molekularbiologie der Zelle, 5., *Wiley-VCH*.

BAIN, J., PLATER, L., ELLIOTT, M., SHPIRO, N., HASTIE, C. J., MCLAUCHLAN, H., KLEVERNIC, I., ARTHUR, J. S., ALESSI, D. R. & COHEN, P. 2007. The selectivity of protein kinase inhibitors: a further update. *Biochem J,* 408**,** 297-315.

BENITEZ, J. A., MA, J., D'ANTONIO, M., BOYER, A., CAMARGO, M. F., ZANCA, C., KELLY, S., KHODADADI-JAMAYRAN, A., JAMESON, N. M., ANDERSEN, M., MILETIC, H., SABERI, S., FRAZER, K. A., CAVENEE, W. K. & FURNARI, F. B. 2017. PTEN regulates glioblastoma oncogenesis through chromatin-associated complexes of DAXX and histone H3.3. *Nat Commun,* 8**,** 15223.

BENNETT, C. N., ROSS, S. E., LONGO, K. A., BAJNOK, L., HEMATI, N., JOHNSON, K. W., HARRISON, S. D. & MACDOUGALD, O. A. 2002. Regulation of Wnt signaling during adipogenesis. *J Biol Chem,* 277**,** 30998-1004.

BIO-RAD. 2016. *Western Blot: 10 Tips for the western blot detection of IP samples* [Online]. Available: https://www.bio-rad-antibodies.com/10-tips-for-the-western-blot-detection-of-ip-samples.html [Accessed 03.01.2020].

BOTMAN, D., RÖTTINGER, E., MARTINDALE, M. Q., DE JONG, J. & KAANDORP, J. A. 2014. A computational approach towards a gene regulatory network for the developing Nematostella vectensis gut. *PLoS One,* 9**,** e103341.

BRANON, T. C., BOSCH, J. A., SANCHEZ, A. D., UDESHI, N. D., SVINKINA, T., CARR, S. A., FELDMAN, J. L., PERRIMON, N. & TING, A. Y. 2018. Efficient proximity labeling in living cells and organisms with TurboID. *Nature Biotechnology,* 36**,** 880-887.

CASEY, E. S., REILLY, M. A., CONLON, F. L. & SMITH, J. C. 1998. The T-box transcription factor Brachyury regulates expression of eFGF through binding to a non-palindromic response element. *Development,* 125**,** 3887.

CONLON, F. L., SEDGWICK, S. G., WESTON, K. M. & SMITH, J. C. 1996. Inhibition of Xbra transcription activation causes defects in mesodermal patterning and reveals autoregulation of Xbra in dorsal mesoderm. *Development,* 122**,** 2427.

COX, J., HEIN, M. Y., LUBER, C. A., PARON, I., NAGARAJ, N. & MANN, M. 2014. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics,* 13**,** 2513-26.

CREATIVE-PROTEOMICS. 2019. *Q Exactive Hybrid Quadrupole-Orbitrap Mass Spectrometer* [Online]. Available: https://www.creative-proteomics.com/support/q-exactive-hybrid-quadrupole-orbitrap-mass-spectrometer.htm [Accessed 06.12.2019].

DARLING, J. A., REITZEL, A. R., BURTON, P. M., MAZZA, M. E., RYAN, J. F., SULLIVAN, J. C. & FINNERTY, J. R. 2005. Rising starlet: the starlet sea anemone, Nematostella vectensis. *Bioessays,* 27**,** 211-21.

FIELDS, S. & SONG, O. 1989. A novel genetic system to detect protein-protein interactions. *Nature,* 340**,** 245-6.

FRITZENWANKER, J. H., SAINA, M. & TECHNAU, U. 2004. Analysis of forkhead and snail expression reveals epithelial-mesenchymal transitions during embryonic and larval development of Nematostella vectensis. *Dev Biol,* 275**,** 389-402.

FRITZENWANKER, J. H. & TECHNAU, U. 2002. Induction of gametogenesis in the basal cnidarian Nematostella vectensis(Anthozoa). *Dev Genes Evol,* 212**,** 99-103.

GENIKHOVICH, G. & TECHNAU, U. 2009a. In situ hybridization of starlet sea anemone (Nematostella vectensis) embryos, larvae, and polyps. *Cold Spring Harb Protoc,* 2009**,** pdb.prot5282.

GENIKHOVICH, G. & TECHNAU, U. 2009b. Induction of spawning in the starlet sea anemone Nematostella vectensis, in vitro fertilization of gametes, and dejellying of zygotes. *Cold Spring Harb Protoc,* 2009**,** pdb.prot5281.

GENIKHOVICH, G. & TECHNAU, U. 2009c. The starlet sea anemone Nematostella vectensis: an anthozoan model organism for studies in comparative genomics and functional evolutionary developmental biology. *Cold Spring Harb Protoc,* 2009**,** pdb.emo129.

HAND, C. & UHLINGER, K. R. 1992. The Culture, Sexual and Asexual Reproduction, and Growth of the Sea Anemone Nematostella vectensis. *Biol Bull,* 182**,** 169-176.

KENNY, A. P., OLEKSYN, D. W., NEWMAN, L. A., ANGERER, R. C. & ANGERER, L. M. 2003. Tight regulation of SpSoxB factors is required for patterning and morphogenesis in sea urchin embryos. *Dev Biol,* 261**,** 412-25.

KIRILLOVA, A., GENIKHOVICH, G., PUKHLYAKOVA, E., DEMILLY, A., KRAUS, Y. & TECHNAU, U. 2018. Germ-layer commitment and axis formation in sea anemone embryonic cell aggregates. *Proc Natl Acad Sci U S A,* 115**,** 1813-1818.

KLOCKENBUSCH, C. & KAST, J. 2010. Optimization of Formaldehyde Cross-Linking for Protein Interaction Analysis of Non-Tagged Integrin β1. *Journal of biomedicine & biotechnology,* 2010**,** 927585.

KOCH, F., SCHOLZE, M., WITTLER, L., SCHIFFERL, D., SUDHEER, S., GROTE, P., TIMMERMANN, B., MACURA, K. & HERRMANN, B. G. 2017. Antagonistic Activities of Sox2 and Brachyury Control the Fate Choice of Neuro-Mesodermal Progenitors. *Dev Cell,* 42**,** 514-526.e7.

KRAUS, Y., AMAN, A., TECHNAU, U. & GENIKHOVICH, G. 2016. Pre-bilaterian origin of the blastoporal axial organizer. *Nat Commun,* 7**,** 11694.

KRAUS, Y. & TECHNAU, U. 2006. Gastrulation in the sea anemone Nematostella vectensis occurs by invagination and immigration: an ultrastructural study. *Dev Genes Evol,* 216**,** 119-32.

KUNICK, C., LAUENROTH, K., LEOST, M., MEIJER, L. & LEMCKE, T. 2004. 1-Azakenpaullone is a selective inhibitor of glycogen synthase kinase-3 beta. *Bioorg Med Chem Lett,* 14**,** 413-6.

LAYDEN, M. J., RENTZSCH, F. & RÖTTINGER, E. 2016. The rise of the starlet sea anemone Nematostella vectensis as a model system to investigate development and regeneration. *Wiley Interdiscip Rev Dev Biol,* 5**,** 408-28.

LECLERE, L., BAUSE, M., SINIGAGLIA, C., STEGER, J. & RENTZSCH, F. 2016. Development of the aboral domain in Nematostella requires beta-catenin and the opposing activities of Six3/6 and Frizzled5/8. *Development,* 143**,** 1766-77.

LEVINE, M. & DAVIDSON, E. H. 2005. Gene regulatory networks for development. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 4936.

LI, E. & DAVIDSON, E. H. 2009. Building developmental gene regulatory networks. *Birth defects research. Part C, Embryo today : reviews,* 87**,** 123-130.

LO PICCOLO, L., BONACCORSO, R. & ONORATI, M. C. 2015. Nuclear and Cytoplasmic Soluble Proteins Extraction from a Small Quantity of Drosophila's Whole Larvae and Tissues. *International journal of molecular sciences,* 16**,** 12360-12367.

LOCKER, J. 2001. Transcription Factors, BIOS Scientific Publishers Ltd, 1., *Academic Press*.

MAGIE, C. R., DALY, M. & MARTINDALE, M. Q. 2007. Gastrulation in the cnidarian Nematostella vectensis occurs via invagination not ingression. *Dev Biol,* 305**,** 483-97.

MAGIE, C. R., PANG, K. & MARTINDALE, M. Q. 2005. Genomic inventory and expression of Sox and Fox genes in the cnidarian Nematostella vectensis. *Dev Genes Evol,* 215**,** 618-30.

MARLOW, H., MATUS, D. Q. & MARTINDALE, M. Q. 2013. Ectopic activation of the canonical wnt signaling pathway affects ectodermal patterning along the primary axis during larval development in the anthozoan Nematostella vectensis. *Developmental Biology,* 380**,** 324-334.

MISHRA, N. C. 2010. Introduction to Proteomics - Principles and Applications, *John Wiley & Sons, Inc.*

MOHAMMED, H., TAYLOR, C., BROWN, G. D., PAPACHRISTOU, E. K., CARROLL, J. S. & D'SANTOS, C. S. 2016. Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nat Protoc,* 11**,** 316-26.

MÜLLER, C. W. & HERRMANN, B. G. 1997. Crystallographic structure of the T domain-DNA complex of the Brachyury transcription factor. *Nature,* 389**,** 884-8.

NEUHOFF, V., STAMM, R., PARDOWITZ, I., AROLD, N., EHRHARDT, W. & TAUBE, D. 1990. Essential problems in quantification of proteins following colloidal staining with Coomassie Brilliant Blue dyes in polyacrylamide gels, and their solution. *ELECTROPHORESIS,* 11**,** 101-117.

PAPAIOANNOU, V. E. 2014. The T-box gene family: emerging roles in development, stem cells and cancer. *Development,* 141**,** 3819.

PUTNAM, N. H., SRIVASTAVA, M., HELLSTEN, U., DIRKS, B., CHAPMAN, J., SALAMOV, A., TERRY, A., SHAPIRO, H., LINDQUIST, E., KAPITONOV, V. V., JURKA, J., GENIKHOVICH, G., GRIGORIEV, I. V., LUCAS, S. M., STEELE, R. E., FINNERTY, J. R., TECHNAU, U., MARTINDALE, M. Q. & ROKHSAR, D. S. 2007. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science,* 317**,** 86-94.

REHM, H. & LETZEL, T. 2010. Der Experimentator: Proteinbiochemie/Proteomics, 6., *Spektrum Akademischer Verlag*.

RICHARDS, G. S. & RENTZSCH, F. 2014. Transgenic analysis of a SoxB gene reveals neural progenitor cells in the cnidarian Nematostella vectensis. *Development,* 141**,** 4681-4689.

ROTTINGER, E., DAHLIN, P. & MARTINDALE, M. Q. 2012. A framework for the establishment of a cnidarian gene regulatory network for "endomesoderm" specification: the inputs of ss-catenin/TCF signaling. *PLoS Genet,* 8**,** e1003164.

ROYO, J. L., MAESO, I., IRIMIA, M., GAO, F., PETER, I. S., LOPES, C. S., D'ANIELLO, S., CASARES, F., DAVIDSON, E. H., GARCIA-FERNÁNDEZ, J. & GÓMEZ-SKARMETA, J. L. 2011. Transphyletic conservation of developmental regulatory state in animal evolution. *Proceedings of the National Academy of Sciences,* 108**,** 14186-14191.

SCHINDELIN, J., ARGANDA-CARRERAS, I., FRISE, E., KAYNIG, V., LONGAIR, M., PIETZSCH, T., PREIBISCH, S., RUEDEN, C., SAALFELD, S., SCHMID, B., TINEVEZ, J.-Y., WHITE, D. J., HARTENSTEIN, V., ELICEIRI, K., TOMANCAK, P. & CARDONA, A. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Meth,* 9**,** 676-682.

SCHOLZ, C. B. & TECHNAU, U. 2003. The ancestral role of Brachyury: expression of NemBra1 in the basal cnidarian Nematostella vectensis (Anthozoa). *Dev Genes Evol,* 212**,** 563-70.

SCHWAIGER, M., SCHONAUER, A., RENDEIRO, A. F., PRIBITZER, C., SCHAUER, A., GILLES, A. F., SCHINKO, J. B., RENFER, E., FREDMAN, D. & TECHNAU, U. 2014. Evolutionary conservation of the eumetazoan gene regulatory landscape. *Genome Res,* 24**,** 639-50.

SERVETNICK, M. D., STEINWORTH, B., BABONIS, L. S., SIMMONS, D., SALINAS-SAAVEDRA, M. & MARTINDALE, M. Q. 2017. Cas9-mediated excision of Nematostella brachyury disrupts endoderm development, pharynx formation and oral-aboral patterning. *Development,* 144**,** 2951-2960.

SKENE, P. J. & HENIKOFF, S. 2017. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife,* 6.

STEINMETZ, P. R. H., AMAN, A., KRAUS, J. E. M. & TECHNAU, U. 2017. Gut-like ectodermal tissue in a sea anemone challenges germ layer homology. *Nat Ecol Evol,* 1**,** 1535-1542.

STUMPF, M. P., THORNE, T., DE SILVA, E., STEWART, R., AN, H. J., LAPPE, M. & WIUF, C. 2008. Estimating the size of the human interactome. *Proc Natl Acad Sci U S A,* 105**,** 6959-64.

SUTHERLAND, B. W., TOEWS, J. & KAST, J. 2008. Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions. *J Mass Spectrom,* 43**,** 699-715.

SWARR, D. T., E.WERT, S. & A.WHITSETT, J. 2019. Molecular Determinants of Lung Morphogenesis Kendig's Disorders of the Respiratory Tract in Children, 9., *Elsevier Inc.* .

TECHNAU, U. 2001. Brachyury, the blastopore and the evolution of the mesoderm. *BioEssays,* 23**,** 788-794.

TEN HAVE, S., BOULON, S., AHMAD, Y. & LAMOND, A. I. 2011. Mass spectrometry-based immuno-precipitation proteomics - the user's guide. *Proteomics,* 11**,** 1153-9.

YACHIE-KINOSHITA, A. & KAIZU, K. 2019. Cell Modeling and Simulation, Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics, *Elsevier*.

6. Supplement

6.1 Buffers, Solutions and Media

10x PBS

NaCl	1.75 M
Na2HPO4	84.1 mM
NaH2PO4	18.6 mM
ELIX H2O	Fill up to total volume

Adjust pH to 7.4!

PTw

PBS	
Tween 20	0,1%

Agarose Gels

Agarose	1%/1.5%/2%
1x TAE-Buffer	50 ml/100 ml

50x TAE-Puffer

For 2l:

Tris Base	484 g
Acetic Acid	114.2 ml
EDTA (0.5 M)	200 ml
ELIX H2O	Fill up to 2 l

100 mM PMSF

Phenylmethylsulfonyl Fluoride	0.871 g
Isopropanol	50 ml

6.2 Antibodies

Table 9 Primary & secondary antibodies and fluorescent dyes.

Antibody/Fluorescent dye	Species	Stock	Source
A-Brachyury	Rabbit	1 mg/ml	PRIMM (custom made)
A-FoxA	Rabbit	1 mg/ml	BioGenes (custom made)
α-F-Actin, Phalloidin Alexa Fluor 488	-	300 U in 1.5 ml	Invitrogen #A12379
DAPI	-	5 mg/ml	Sigma Aldrich
α-Mouse, Alexa Fluor 568	Rabbit	2 mg/ml	Invitrogen #A11061
α-Rabbit, Alexa Fluor 568	Goat	2 mg/ml	Invitrogen #A11011

6.3 Oligonucleotides

Table 10 Primers for in situ probes.

Name	5` Nucleotide-Sequence 3`
Nve23841_Fw	AAA GTT AAA AGG CCA ATG AAC GC
Nve23841_Rv	GTC TAC AAA TAC AAC CGC CCG

Nve15777_Fw	TCC TGG GAG ATG AGT GGA GA
Nve15777_Rv	AAT TCG CTG TTC TTT GCC GT
Nve426_Fw	TGA CTT CGG ACA GAG TTC ACA
Nve426_Rv	TGA ATT GAG TTA CCT ACA CAG CT

6.4 Additional figures and tables

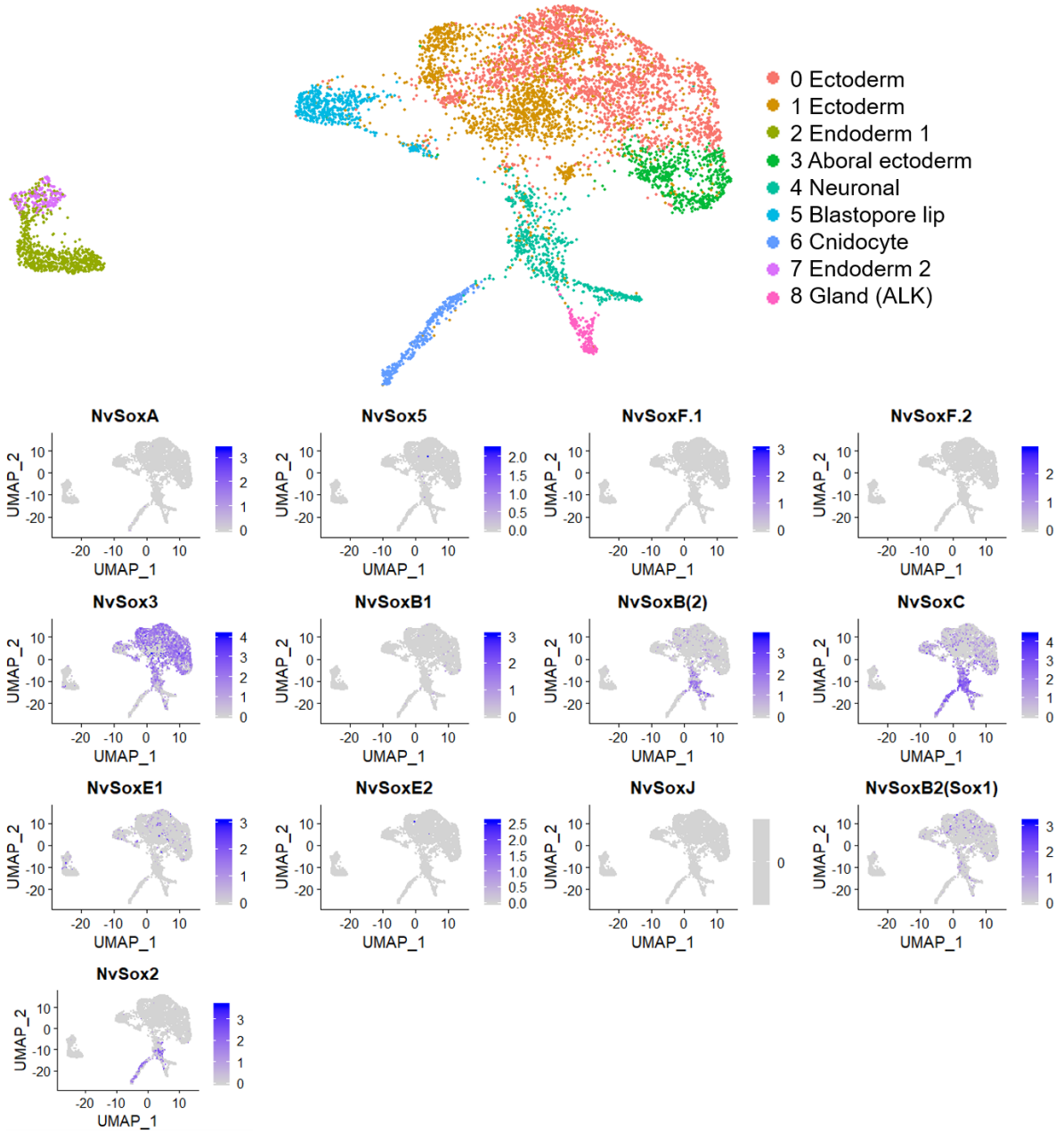

Fig. 21 Single cell RNA-seq data of *sox* genes from 24 hpf gastrula library.

Table 11 List of all DNA binding proteins identified in the FoxA RIME experiments (samples FoxA_1-8).

Gene ID	Curated ID	BLAST Hit	Trinotate Name
NVE10101		eukaryotic translation initiation factor 3 subunit M [Ictalurus punctatus]	EIF3M_NEMVE
NVE10331		maelstrom-like protein [Crassostrea gigas]	MAEL_XENTR
NVE10530		Protein polybromo-1 [Acromyrmex echinatior]	PB1_HUMAN
NVE10570		26S proteasome non-ATPase regulatory subunit 3-like protein [Callorhinchus milii]	PSMD3_MOUSE
NVE11233		TPA: zinc finger protein [Ciona intestinalis]	

NVE11306		Sjogren syndrome antigen B (autoantigen La) [Macaca fascicularis]	LA_HUMAN
NVE11734			ZN608_MOUSE
NVE11783			
NVE11839		REST corepressor 2, isoform CRA_a [Homo sapiens]	RCOR2_MOUSE
NVE12602		Origin recognition complex subunit 1 [Balearica pavonina gibbericeps]	ORC1_MOUSE
NVE17017		NFkB protein [Ciona intestinalis]	NFKB1_MOUSE
NVE12856		SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily E member 1 [Clonorchis sinensis]	SMCE1_RAT
NVE12910		Zinc finger CCCH domain-containing protein 14 [Cricetulus griseus]	ZC3HE_CHICK
NVE13032		hypothetical protein AaeL_AAEL004358 [Aedes aegypti]	SMRC2_HUMAN
NVE13092		Methionine aminopeptidase 2 [Crassostrea gigas]	MAP2_HUMAN
NVE1317			ZN358_HUMAN
NVE13317		GK16809 [Drosophila willistoni]	CHSP1_HUMAN
NVE13458	NvTcf/Lef	HMG box transcription factor Tcf [Hydra vulgaris]	PANG1_DROME
NVE13718		RING finger protein 113A [Harpegnathos saltator]	R113A_HUMAN
NVE13902		DNA methyltransferase 1 associated protein 1 [Xenopus (Silurana) tropicalis]	DMAP1_MOUSE
NVE14009		eukaryotic translation initiation factor 3 subunit K [Xenopus (Silurana) tropicalis]	EIF3K_NEMVE
NVE14231	rere		RERE_RAT
NVE14454		deformed epidermal autoregulatory factor 1 variant 1 [Mus musculus]	DEAF1_DROME
NVE14464		COP9 signalosome complex subunit 4 [Pongo abelii]	CSN4_PONAB
NVE14485		dynamin 1, isoform CRA_a [Homo sapiens]	DYN1_RAT
NVE22326			H10_PONAB
NVE14907			BAZ2B_HUMAN
NVE14970		Pre-mRNA-splicing factor RBM22 [Cuculus canorus]	RBM22_CHICK
NVE15521		tRNA (uracil-O(2)-)-methyltransferase-like [Danio rerio]	TRM44_MOUSE
NVE15807			GRHL2_MOUSE
NVE15870			TE2IP_MOUSE
NVE15913		PREDICTED: 26S proteasome non-ATPase regulatory subunit 11 [Xenopus (Silurana) tropicalis]	PSD11_XENTR
NVE15986	Noc	NocA-like transcription factor [Saccoglossus kowalevskii]	NOC_DROME
NVE16575			PP1RA_XENLA
NVE16657	NvSmad4-like	MAD homolog 4 (Drosophila), isoform CRA_c [Mus musculus]	SMAD4_MOUSE
NVE1669			CUL5_PONAB
NVE1670		Cullin-5 [Zootermopsis nevadensis]	CUL5_HUMAN
NVE16808			ZNF98_HUMAN
NVE16882		eukaryotic translation initiation factor 3, subunit E [Callorhinchus milii]	EIF3E_NEMVE
NVE17345		zinc finger protein [Ciona intestinalis]	KAT6A_MOUSE
NVE17368		PREDICTED: transcription factor RFX3 isoform X1 [Merops nubicus]	RFX3_XENTR
NVE18355		26S proteasome non-ATPase regulatory subunit 6 [Danio rerio]	PSMD6_MOUSE
NVE18452		general transcription factor IIE, polypeptide 2, beta 34kDa [Xenopus (Silurana) tropicalis]	T2EB_HUMAN
NVE19092		Histone-lysine N-methyltransferase PRDM9 [Crassostrea gigas]	ZN665_PONAB
NVE19122		cullin-2 [Bos taurus]	CUL2_PONAB
NVE19355	NFX1-type	NFX1-type zinc finger-containing protein 1 [Heterocephalus glaber]	ZNFX1_MOUSE
NVE1940		26S proteasome non-ATPase regulatory subunit 12 [Crassostrea gigas]	PSD12_MOUSE
NVE19823			ZC3HF_DROPS
NVE19829		signal transducer and activator of transcription [Fenneropenaeus chinensis]	STA5A_HUMAN
NVE20136	COE	PREDICTED: transcription factor COE3 isoform 1 [Pan troglodytes]	COE3_HUMAN
NVE20185			
NVE20261		BolA-related protein [Ornithodoros coriaceus]	BOLA2_MOUSE
NVE20630	NvFoxA	budhead [Hydra vulgaris]	FOXA2_XENTR
NVE20662		putative zinc finger protein [Daphnia pulex]	ZN622_HUMAN
NVE20721		doublesex- and mab-3-related transcription factor A2 [Xenopus (Silurana) tropicalis]	DMTA2_XENTR
NVE21002		PREDICTED: uncharacterized protein LOC100194365 isoform X2 [Zea mays]	NFYA_HUMAN
NVE21093		general transcription factor IIF subunit 1 [Xenopus laevis]	T2FA_XENLA
NVE21181		YY1 transcription factor, isoform CRA_c [Mus musculus]	TYY1_MOUSE
NVE21220		tRNA-dihydrouridine(47) synthase [NAD(P)(+)]-like [Xenopus (Silurana) tropicalis]	DUS3L_XENTR
NVE21410		replication protein A 32 kDa subunit [Bos taurus]	RFA2_MOUSE
NVE21411			ZN431_MOUSE
NVE21450			ZFR_PONAB
NVE21766	TBX2		TBX3_CHICK
NVE21786		Eukaryotic translation initiation factor 3 subunit C [Crassostrea gigas]	EIF3C_DANRE
NVE21973			
NVE22032		RNA-binding protein 27 [Chelonia mydas]	RBM27_HUMAN
NVE2230			CSN2_XENLA
NVE22444			PYC_HUMAN
NVE22445			PYC_HUMAN
NVE22695			YBOX1_XENLA
NVE22862		cAMP responsive element modulator, isoform CRA_a [Mus musculus]	CREB1_BOVIN
NVE2314		PREDICTED: splicing factor U2AF 35 kDa subunit isoform 2 [Orcinus orca]	U2AF1_BOVIN
NVE23207		Doublesex- and mab-3-related transcription factor A2 [Crassostrea gigas]	DMRT1_DANRE
NVE23340	CCCH10	Zinc finger CCCH domain-containing protein 10 [Crassostrea gigas]	ZC3HA_HUMAN
NVE23413		lysine-specific histone demethylase [Aedes aegypti]	KDM1A_HUMAN
NVE23539		Cold shock domain-containing protein E1 [Crassostrea gigas]	CSDE1_RAT
NVE23581		putative high mobility group 20A isoform 1 [Danaus plexippus]	HM20A_HUMAN
NVE23709	SoxB1	SoxB1 [Acropora millepora]	SOX2_CHICK
NVE2385		PREDICTED: cell division cycle 5-like protein [Pteropus alecto]	CDC5L_NEMVE
NVE24006			H10_PONAB
NVE24023			H2B3_DANRE

NVE24092		Limkain-b1 [Pteropus alecto]	MARF1_XENTR
NVE24168		SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5 [Crassostrea gigas]	SMCA5_HUMAN
NVE24711	Dachshund		DACH1_HUMAN
NVE24817		Vacuolar protein sorting-associated protein 72-like protein [Zootermopsis nevadensis]	VPS72_XENLA
NVE2485	NvHD032/NK-like 7		TLX2_HUMAN
NVE24950		Metastasis-associated protein MTA1 [Crassostrea gigas]	MTA3_HUMAN
NVE25080		COP9 signalosome complex subunit 3 [Danio rerio]	CSN3_DANRE
NVE25411			
NVE25474			ZKSC8_HUMAN
NVE26088		mothers against decapentaplegic homolog 3 isoform 2 [Homo sapiens]	SMAD2_DANRE
NVE26195	NvFoxB	Forkhead box protein B1 [Tupaia chinensis]	FOXB2_XENLA
NVE2817			
NVE3107		transcription factor Dp-1 [Homo sapiens]	TFDP1_HUMAN
NVE3110		zinc finger protein 782 [Homo sapiens]	ZN782_HUMAN
NVE3272		Phenylalanyl-tRNA synthetase beta chain [Salmo salar]	SYFB_MOUSE
NVE3360			NOL4_HUMAN
NVE3568	NvBrachyury	brachyury protein [Acropora millepora]	BRAC1_BRAFL
NVE3616			YLPM1_RAT
NVE3632		cullin-1 [Pan troglodytes]	CUL1_MOUSE
NVE3653		Upstream-binding protein 1 [Picoides pubescens]	TFCP2_XENLA
NVE3695		Myc-induced nuclear antigen with a molecular mass of 53 kDa-like 1 [Ciona intestinalis]	PRDM5_MOUSE
NVE3976		aryl hydrocarbon receptor nuclear translocator [Bos taurus]	ARNT_DROME
NVE4158		transcription factor CBF/NF-Y/archaeal histone -2 [Ciona intestinalis]	DPOE3_MOUSE
NVE4301			H1D_STRPU
NVE4781			ZFR2_HUMAN
NVE4829			FOXK1_MOUSE
NVE5018		COP9 signalosome complex subunit 1 [Crassostrea gigas]	CSN1_XENLA
NVE5105		Gli-Kruppel type zinc finger protein [Daphnia pulex]	TYY1_HUMAN
NVE5220		pre-mRNA-splicing factor cwc2, putative [Phytophthora infestans T30-4]	CWC2_YARLI
NVE5430		Pax-B [Anthopleura japonica]	PAX6_XENLA
NVE5479		Smad1 [Hydra vulgaris]	SMAD1_MOUSE
NVE5517		vacuolar protein-sorting-associated protein 36 [Danio rerio]	VPS36_DANRE
NVE5765		26S proteasome non-ATPase regulatory subunit 8 [Crassostrea gigas]	PSMD8_MOUSE
NVE613	NvREPO/NvREVPOL	Aristaless [Euperipatoides kanangrensis]	ARX_DANRE
NVE6378		DNA repair protein RAD51 homolog 3 [Danio rerio]	RA51C_CRIGR
NVE6473		Mesoderm induction early response protein 1 [Zootermopsis nevadensis]	MIER1_MOUSE
NVE6843	TRERF1-like		TREF1_MOUSE
NVE708		FACT complex subunit SSRP1 [Cuculus canorus]	SSRP1_CHICK
NVE7114		proliferation-associated 2G4 [Xenopus (Silurana) tropicalis]	PA2G4_RAT
NVE7115		orthodenticle A [Acropora millepora]	OTX1_HUMAN
NVE7278		cullin-4B [Rattus norvegicus]	CUL4A_HUMAN
NVE7308		PREDICTED: transcription factor E2F5 isoform X1 [Bos taurus]	E2F4_XENLA
NVE7747		Ab1-108 [Rattus norvegicus]	RPAB5_MOUSE
NVE7875		C. briggsae CBR-MET-2 protein [Caenorhabditis briggsae]	STB1B_DANRE
NVE7969		PREDICTED: cullin-3 isoform X1 [Microtus ochrogaster]	CUL3_RAT
NVE8532		acetylserotonin O-methyltransferase [Danio rerio]	ASMT_MACMU
NVE862			ARID2_HUMAN
NVE8905		Hnf4a protein [Danio rerio]	HNF4A_XENLA
NVE9145			ZN207_XENLA
NVE9280		endothelial differentiation-related factor 1 isoform alpha [Homo sapiens]	EDF1_HUMAN
NVE9319		methyl-CpG-binding domain protein 2 [Gallus gallus]	MBD2_MOUSE
NVE9411		chromatin accessibility complex protein 1 [Mus musculus]	CHRC1_MOUSE
NVE9807	Lhx6/8	LIM class homeodomain transcription factor, Lhx6/8 subclass [Branchiostoma floridae]	AWH_DROME
NVE9907		RecName: Full=DnaJ homolog subfamily C member 2; AltName: Full=Gliosarcoma-related antigen MIDA1; AltName: Full=Zuotin-related factor 1; Contains: RecName: Full=DnaJ homolog subfamily C member 2, N-terminally processed	DNJC2_RAT

Table 12 List of all DNA binding proteins identified in the Bra RIME experiments (samples Bra_1 & 2).

Gene ID	Curated ID	BLAST Hit	Trinotate Name
NVE5765		26S proteasome non-ATPase regulatory subunit 8 [Crassostrea gigas]	PSMD8_MOUSE
NVE613	NvREPO/NvREVPOL	Aristaless [Euperipatoides kanangrensis]	ARX_DANRE
NVE6378		DNA repair protein RAD51 homolog 3 [Danio rerio]	RA51C_CRIGR
NVE6473		Mesoderm induction early response protein 1 [Zootermopsis nevadensis]	MIER1_MOUSE
NVE6843	TREF1-like		TREF1_MOUSE
NVE708		FACT complex subunit SSRP1 [Cuculus canorus]	SSRP1_CHICK
NVE7114		proliferation-associated 2G4 [Xenopus (Silurana) tropicalis]	PA2G4_RAT
NVE7115		orthodenticle A [Acropora millepora]	OTX1_HUMAN
NVE7278		cullin-4B [Rattus norvegicus]	CUL4A_HUMAN
NVE7308		PREDICTED: transcription factor E2F5 isoform X1 [Bos taurus]	E2F4_XENLA
NVE7747		Ab1-108 [Rattus norvegicus]	RPAB5_MOUSE
NVE7875		C. briggsae CBR-MET-2 protein [Caenorhabditis briggsae]	STB1B_DANRE
NVE7969		PREDICTED: cullin-3 isoform X1 [Microtus ochrogaster]	CUL3_RAT
NVE8532		acetylserotonin O-methyltransferase [Danio rerio]	ASMT_MACMU
NVE862			ARID2_HUMAN
NVE8905		Hnf4a protein [Danio rerio]	HNF4A_XENLA
NVE9145			ZN207_XENLA
NVE9280		endothelial differentiation-related factor 1 isoform alpha [Homo sapiens]	EDF1_HUMAN
NVE9319		methyl-CpG-binding domain protein 2 [Gallus gallus]	MBD2_MOUSE
NVE9411		chromatin accessibility complex protein 1 [Mus musculus]	CHRC1_MOUSE
NVE9807	Lhx6/8	LIM class homeodomain transcription factor, Lhx6/8 subclass [Branchiostoma floridae]	AWH_DROME
NVE9907		RecName: Full=DnaJ homolog subfamily C member 2; AltName: Full=Gliosarcoma-related antigen MIDA1; AltName: Full=Zuotin-related factor 1; Contains: RecName: Full=DnaJ homolog subfamily C member 2, N-terminally processed	DNJC2_RAT

6.5 Abstract

Gene transcription is regulated by the combined actions of transcription factors. The interplay between those and cis regulatory elements on the DNA is summed up in a gene regulatory network (GRN). Those GRNs are instructing the development of an organism.
Brachyury (Bra) is an important transcription factor involved in the development of the third germ layer, the mesoderm, in the triploblastic Bilateria. It also plays a role in the patterning of the bifunctional mesendoderm in the supposedly mesoderm-less diploblastic Cnidaria. Another transcription factor important in bilaterian mesoderm development as well as in cnidarian mesendoderm development is FoxA. Despite genetical commonalities of these eumetazoan groups they differ drastically in their phenotypes and biodiversity. Therefore, it is of interest to dissolve the GRNs that set up the preliminary mesendoderm in bilaterians and the bifunctional mesendoderm in cnidarians. Determination of differences and similarities helps understand the evolution of the complex bilaterians. An advantageous cnidarian model organism is the sea anemone *Nematostella vectensis* (*Nv*).
In *Nematostella* a ChIP-seq against Bra was performed at an important developmental time point, the late gastrula stage, to reveal transcriptional targets of Bra. Besides, it revealed DNA binding sites that do not represent the Brachyury T-box motif, but motifs of other transcription factor families, in particular of the Sox family. The theory originated that Bra binds to the DNA together with interaction partners like Sox TFs, with whom it is coregulating transcription of target genes, potentially involved in mesendodermal patterning.
A way to investigate gene regulatory networks or more precisely transcription factor interactions is Rapid Immunoprecipitation Mass Spectrometry of Endogenous Proteins (RIME). I tested different protocol variations to establish a RIME protocol suitable for *Nematostella*. Eventually I used it to identify transcriptional interaction partners of Bra and FoxA. With my results a direct physical interaction between Bra and FoxA themselves could be supported which is known to be a conserved interaction in eumetazoans. Furthermore, a supposed network setting up the oral domain in *Nv* consisting of Bra, FoxA, FoxB and Lmx is supported by the results. Also, further potential interaction partners of FoxA could be identified but will need further verification. RIME against Bra did not lead to conclusive results hence it should be repeated to identify interaction partners like Sox proteins.
Ultimately, the proteomics tool -RIME- could be introduced for usage in *Nematostella* and was already successfully applied to further explore the GRN of FoxA.

6.6 Zusammenfassung

Gentranskription wird reguliert durch das Zusammenspiel verschiedener Transkriptionsfaktoren. Ihr Zusammenspiel untereinander und mit cis-regulatorischen Elementen auf der DNS wird zusammengefasst in einem Genregulationsnetzwerk (GRN). Diese GRNs stellen die Anleitung für die Embryonalentwicklung eines Organismus.
Ein wichtiger Transkriptionsfaktor für die Entwicklung des dritten Keimblatts, des Mesoderms, in den triploblastischen Bilateria ist Brachyury (Bra). Es spielt aber auch eine wichtige Rolle in der Musterbildung des bifunktionalen Mesendoderms der angeblich Mesoderm-losen Cnidaria. Ein weiterer Transkriptionsfaktor der in die Mesoderm-Entwicklung der Bilateria sowie der Mesendoderm-Entwicklung in Cnidaria involviert ist, ist FoxA.
Trotz dieser und weiterer genetischer Gemeinsamkeiten sind diese beiden Gruppen innerhalb der Eumetazoa, grundverschieden in ihren Phänotypen sowie ihrer Biodiversität. Daher ist es von großem Interesse die GRNs zu entschlüsseln, die das vorläufige Mesendoderm in Bilateria und das bifunktionelle Mesendoderm in Cnidaria etablieren. Die Erforschung von Unterschieden, aber auch Gemeinsamkeiten wird helfen zu verstehen wie es zur Evolution der komplexen Bilateria kam. Ein geeigneter und vorteilhafter Model Organismus innerhalb der Cnidaria ist *Nematostella vectensis* (*Nv*)
In einer ChIP-seq Studie gegen Bra in *Nv* wurden die Transkriptions Ziele von Bra zum Zeitpunkt der späten Gastrulation, bestimmt. Dabei wurden DNA-Bindestellen identifiziert, die nicht dem Bra-üblichen T-Box Motiv entsprechen, sondern DNA-Bindemotiven anderer Transkriptionsfaktor Familien, insbesondere solche der Sox Familie. Daraus entstand die Theorie, dass an diesen Stellen Bra nicht selbst an die DNA bindet, sondern durch Interaktion mit anderen Transkriptionsfaktoren. Mit diesen reguliert es schließlich die Transkription von Zielgenen, die möglicherweise in die Musterbildung des Mesendoderms involviert sind.
Eine Methode, um Genregulationsnetzwerke bzw. Interaktionen von Transkriptionsfaktoren zu untersuchen ist das sogenannte RIME (*Rapid Immunoprecipitation Mass Spectrometry of Endogenous Proteins*). Um diese Methode in *Nematostella* zu etablieren, wurden verschiedene Versionen und Paramater-Änderungen getestet. Schließlich konnte das Protokoll genutzt werden, um Transkriptions-Interaktionspartner von Bra und FoxA in *Nv* zu identifizieren. Meine Ergebnisse deuten auf eine Interaktion zwischen Bra und FoxA hin, welche als konservierte Interaktion in Eumetazoa gilt. Sie können außerdem ein vermutetes Netzwerk bestehend aus Bra, FoxA, FoxB und Lmx unterstützen, welches die orale Domäne in *Nv* organisieren soll. Zusätzlich konnten weitere potenzielle Interaktionspartner von FoxA identifiziert werden, welche allerdings weiterer Überprüfung bedürfen. Bra RIME führte zu keinen schlüssigen Ergebnissen, weshalb dies wiederholt werden sollte.
Schlussendlich konnte das *Proteomics*-Werkzeug -RIME- für den Gebrauch in *Nematostella* angepasst werden und bereits eingesetzt werden, um das GRN von FoxA weiter zu erforschen.