universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

Lineage tracing of *SoxC* expression in *Nematostella vectensis* using the KAEDE photoconvertible fluorophore

verfasst von / submitted by

Robert Reischl, BSc

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 066 834
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Molekulare Biologie
Betreut von / Supervisor:	Univ.-Prof. Dipl.-Biol. Dr. Ulrich Technau
Mitbetreut von / Co-Supervisor:	

Acknowledgements

I would like to express my cosmic gratitude to Uli Technau for giving me the opportunity to conduct my Master's thesis research project in his lab, as well as to Andreas Denner for the supervision and guidance during my practical work and Alison Cole for valuable insight and support especially during the writing process. Further credits go to Julia Steger, Tatiana Bagaeva and Patricio Ferrer and many other lab members for professional support on topics in their respective fields. I would also like to acknowledge the Core Facility for Cell Imaging and Ultrastructure Research (CIUS).

Thanks to my whole family for their unconditional support. Honourable mentions go to the friends I had and made along the way, Christoph Scharf, Ronja Reinhardt, Clara Conrad-Billroth, Fernanda Martinez, Julia Hagauer and last but not least my love Marion Maier.

Abstract

Cnidarians have gained popularity as a model phylum to investigate the evolution of stem cells and developmental biology. In the medusozoan *Hydra,* i-cells are dedicated stems cells giving rise to neuronal cell-types, gland cells as well as gametes. In the basally-branching anthozoan *Nematostella* a stem cell population remains yet to be identified. Cell lineage tracing of conserved stem-cell and pluripotency factors found in the *Nematostella* using fluorescent reporter lines is an established method.

Common fluorophores like GFP and mCherry being co-expressed with a putative stem cell transcription factor can be used to deduce the origin of a differentiated cells due to the fluorophore outliving the activation and suppression of the promotor. While the stability of the fluorophore might be an advantage for certain questions in lineage tracing might by a disadvantage for others due to the lack of temporal resolution. The photoconvertible KAEDE fluorophore states a novel approach for lineage tracing by enabling the non-reversible conversion of the fluorophore from green to red emission. Therefore, transcription factors can be investigated regarding their current activity in cells adding temporal information to lineage tracing.

To investigate the usability of the KAEDE fluorophore for lineage tracing in *Nematostella* we chose the promoter of the *NvSoxC* gene to drive KAEDE expression. Latest single-cell data have shown that the *Hydra* homolog *HvSoxC* is expressed during the differentiation of cells which originate from i-cells. In *Nematostella NvSoxC* is active in neuronal and glandular cells which belong to the same cell types as those generated by i-cells in *Hydra*.

Our results show that the KAEDE system is a feasible approach for lineage tracing in *Nematostella*. The brightness and stability of the native as well as of the photoconverted fluorophore offer a variety of *in vivo* experiments in the rather translucent animals. The additional temporal resolution gained by the photo-convertibility of the fluorophore allowed us to hypothesise a developmental time course for neuronal cell types in *Nematostella*. These data augmented with single cell data allowed a refinement of the lineage tracing. We conclude that *NvSoxC* is likely to have a role as differentiation transcription factor in neuronal precursor cells in *Nematostella*. These data differ from the *HvSoxC* as *NvSoxC* seems to be expressed also in mature neurons instead of differentiating neurons only.

Zusammenfassung

Cnidaria werden zunehmend beliebter als Modellorganismen zur Erforschung der Evolution von Stammzellen und in der Entwicklungsbiologie eingesetzt. In der Medusozoan *Hydra* gibt es i-cells, dedizierte Stammzellen, aus denen neuronale Zelltypen, Drüsenzellen sowie Gameten hervorgehen. In der **basalverzweigten** Anthozoa *Nematostella* wurde bisher noch keine Stammzellpopulation identifiziert. Zum cell lineage tracing von konservierten Stammzell- und Pluripotenzfaktoren, die in *Nematostella* identifiziert wurden, werden häufig fluoreszierende reporter lines eingesetzt.

Um den Ursprung von differenzierten Zelltypen herzuleiten werden transgene Linien erzeugt welche fluorophore wie GFP und mCherry zusammen mit einem putativen Stammzelltranskriptionsfaktor exprimieren. Auf Grund der Stabilität dieser Fluorophore tragen die differenzierten Zellen auch nach der Repression des Transkriptionsfaktors noch ein Fluoreszenzsignal mit sich. Die Stabilität der Fluorophore erlaubt so ein sichtbar machen der Zelltypen die aus einer bestimmten Vorläufer-/Stammzellpopulation hervorgehen.

Diese Stabilität vermindert die zeitliche Auflösung weshalb in dieser Studie der KAEDE-Fluorophor getestet und eingesetzt wurde. Dieser Fluorophor ist mittels UV-Licht irreversible von grüner zu roter Emission konvertierbar. Daher können Transkriptionsfaktoren hinsichtlich ihrer aktuellen Expression/Repression in den Zellen untersucht werden.

Mit dem Promotor des *NvSoxC*-Gens als Treiber, soll der KAEDE-fluorophor hinsichtlich seiner Eignung für cell lineage tracing untersucht werden. Neueste single-cell Daten haben gezeigt, dass das *Hydra*-Homolog *HvSoxC* während der Differenzierung i-cells aktiv ist. In *Nematostella* ist NvSoxC in neuronalen und Drüsenzellen aktiv, die zu denselben Zelltypen gehören wie diejenigen, die von i-cell in *Hydra* stammen.

Unsere Ergebnisse zeigen, dass das KAEDE-System ein praktikabler Ansatz für lineage-tracing in *Nematostella* ist. Die Helligkeit und Stabilität des nativen sowie des photokonvertierten Fluorophors ermöglicht eine Vielzahl von *In-vivo* Experimenten. Die zusätzliche zeitliche Auflösung ermöglichte es, einen Entwicklungszeitverlauf für neuronale Zelltypen in Nematostella zu modellieren. Single-Cell Datenermöglichten eine Verfeinerung dieses Modells. Wir schließen daraus, dass *NvSoxC* eine Rolle als Differenzierungstranskriptionsfaktor in neuronalen Vorläuferzellen in *Nematostella* hat. Diese Daten unterscheiden sich von *HvSoxC*, da *NvSoxC* auch in reifen Neuronen exprimiert wird, anstatt ausschließlich in differenzierenden Neuronen.

Content

1. Introduction

1.1. Biological significance of stem cells

Sexually reproducing organisms arise from a totipotent zygote and develop into a multicellular organism mainly consisting of differentiated cells. During this development from a cell that is a totipotent cell, the potency of its offspring declines until the differentiated cells do not divide or divide only symmetrically giving offspring to the same cell type. Since a cells' environment poses a multitude of stressors cell death is inevitable. This makes it crucial for organisms to be able to regenerate tissue. Regeneration and growth of organisms rely on stem cells which are self-renewable and can differentiate into certain cell types. These stem cells are often multipotent lineage-specific progenitors which reside in stem cell niches. The stemness of these cells is maintained by a dynamic microenvironment of signal cues. These cues called stem cell factors influence translational and transcriptional processes of these cells. Under the influence of these stem cell factors differentiated mouse fibroblasts can be reprogramed into so called induced pluripotent stem cells (iPS cells) (Takahashi and Yamanaka, 2006; Liu et al., 2007; Ferraro et al., 2010; Voog and Jones, 2010; Shen et al., 2012).

One of the most basally branching metazoan phyla, the ctenophores, have dedicated, non-motile stem cell populations expressing pan-metazoan stem cell markers like Piwi and Vasa. Therefore, early branching metazoans could help to further elucidate basic stem cell concepts, their regulation and evolutionary origin (Dunn et al., 2008; Alié et al., 2011).

1.2. The three types of stem cells in *Hydra*

Many cnidarians are known to possess immense regenerative potential allowing them to reorganise the original body plan from a suspension of single, dissociated cells. Some species even lack signs of senescence, making them potentially immortal (Bosch, 2007; Denner, 2017 unpublished; Kirillova et al., 2018). Additionally, the interesting position of cnidarians as an eumetazoan sister group to bilaterians (**Figure 1**) puts them in a prime spot for investigating metazoan stem cells by comparing results of these two sister groups and deducing similarities and evolutionary mechanisms.

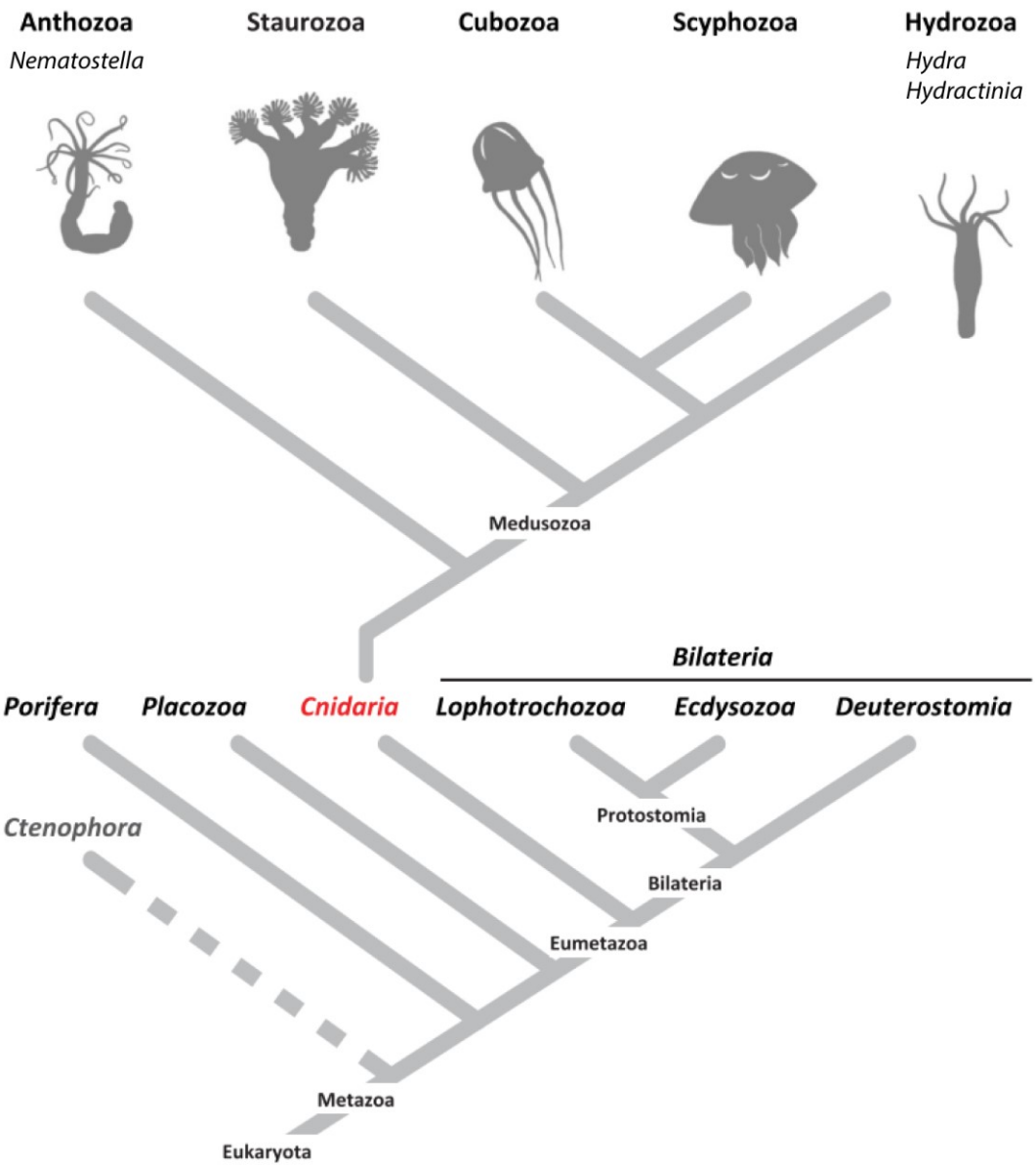

Figure 1: Schematic phylogenetic tree of Cnidaria and other Metazoa. *The relationship of the Anthozoan Nematostella to the four Medusozoan classes within the Cnidaria phylum can be seen on top. The lower half is a depiction indicating the Cnidarian position as a sister group to Bilateria. Modified from (Bosch et al., 2017)*

The first commonly used cnidarian model system was *Hydra* (**Figure 1**). The body plan of the polyp can be described as a two-layered tube with four to seven tentacles arranged around the oral opening (**Figure 2A**). Hydra can reproduce sexually and asexually through budding which takes place in a specific aboral area of the animal (Budding region)). This fresh water polyp also possesses an immense regenerative capacity and is suspected to undergo no senescence suggesting immortality (Martínez, 1998).

Three independent stem cell lines can be distinguished in the *Hydra* polyp (**Figure 2B**). The ectodermal and endodermal epithelial lineages are located in the gastric region of the polyp. The continuously reproducing epithelial cells are displaced towards the tentacles or the aboral

end where excess cells are sloughed off at the tentacles and the foot. The majority of new cells bud off to form new polyps in the aboral budding region. Depending on positional cues the ectodermal epithelial cells, which are displaced towards the extremities can differentiate into tentacle battery cells or basal disk cells (**Figure 2B**). The third lineage is located in the interstitial space between the endodermal and mainly ectodermal epithelial cells. These multipotent interstitial-cells (i-cells) can give rise to gland cells, gametes, neurons and stinging cells (nematocytes or cnidocytes). The neuronal cells make up about 70% of all differentiation products of this stem cell lineage. (Watanabe et al., 2009; Technau et al., 2015). I-cells express marker genes typical for germ cells and stem cells, such as *eed, nanos, piwi* and *vasa*, but no homologs of the pluripotency marker genes *sox2, klf4, nanog* and *oct4* have been identified in the *Hydra* genome. (Genikhovich et al., 2006; Takahashi and Yamanaka, 2006; Nishimiya-Fujisawa and Kobayashi, 2012; Juliano et al., 2013).

The two major derivatives of i-cells, neurons and nematocytes, are terminally differentiated and arrested in G0/1. The two major types of neurons in *Hydra* are the ganglion neurons defined by possessing multiple dendrites and sensory neurons, forming together a diffuse nervous system. (Rentzsch et al., 2017).

Although i-cells can be found in other hydrozoans they have not yet been found outside of this class. Hydrozoans are a rather derived class compared to other cnidarians, hence, i-cells might be a derived feature only occurring in Hydrozoa. The early branching anthozoans diverged from medusozoans more than 550 million years ago and thus roughly 50 million years after the cnidarian lineage arose. The comparison between anthozoans with medusozoans and further with bilaterians might help to elucidate ancient features of the eumetazoan ancestor (Rentzsch et al., 2017).

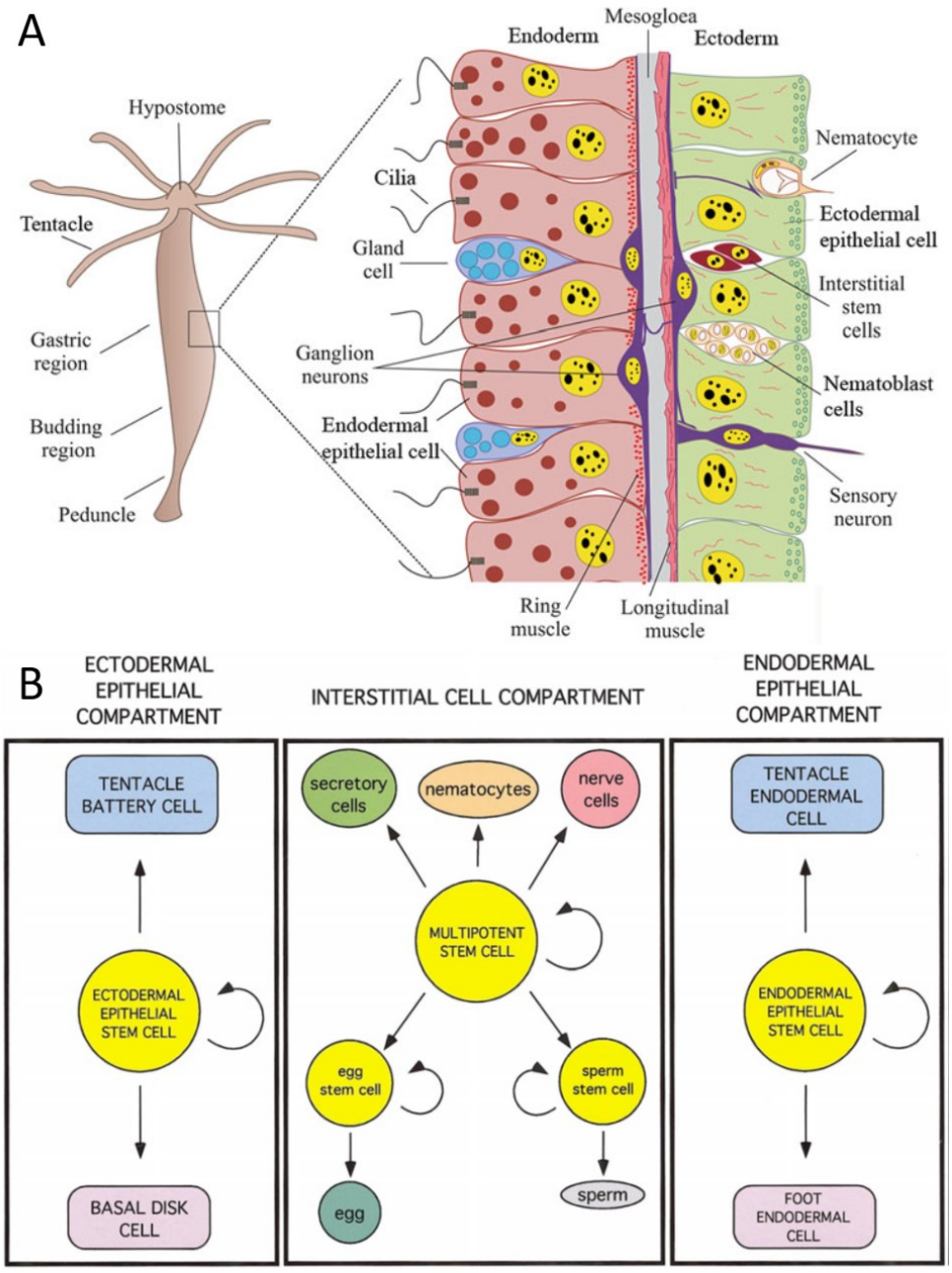

Figure 2: Morphology of Hydra and its Endo- and Ectoderm and the cell lineage compartments in Hydra A (Technau et al., 2015) B (Steele, 2002). *(A) Body plan and Epithelial organisation of an adult Hydra polyp. (B) The three cell lineage compartments of an adult Hydra polyp.*

1.3. The anthozoan *Nematostella vectensis* as model organism

Over the last two decades the anthozoan *Nematostella vectensis* (**Figure 1**) gained popularity as an alternative cnidarian model organism. A number of molecular tools have been established in *Nematostella,* such as *in situ hybridization* (ISH), morpholino- or RNAi-mediated gene knock-down and knock out using CRISPR/Cas9, as well as transgenesis (Genikhovich and Technau, 2009a; Renfer et al., 2010; Ikmi et al., 2014; Layden et al., 2016).

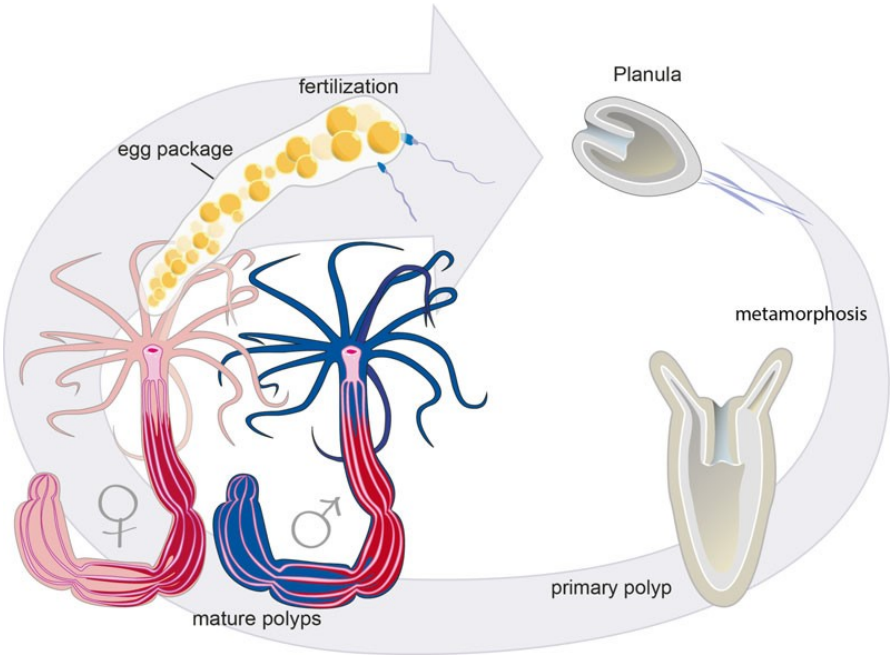

Figure 3: Schematic life cycle of Nematostella (Technau et al., 2015)

Nematostella vectensis is a sea anemone which can be found in the brackish waters of the American Atlantic and -Pacific coast in estuaries. The animals can reproduce sexually and asexually. Spawning can be induced conveniently under lab conditions by light and temperature cues (Fritzenwanker and Technau, 2002). The externally fertilized zygote starts to divide after 2h, although this can be delayed by cooling the medium to extend the time-window for the application of molecular methods. At the standard temperature of 21°C gastrulation starts 20 hours post fertilization (hpf) and is completed 30hpf. By 48hpf the embryo has reached the stage of a motile planula larva.

During this stage 8 endodermal folds, called mesenteries, are formed as well as four knob-like structures around the pharynx anlage at the oral pole. Additionally, a clear oral-aboral axis can be distinguished. The later structure will further develop into tentacles when metamorphosing into a sessile primary polyp after 5-10 days post fertilization. Under laboratory conditions

sexual maturity is reached after 6 months. In the wild, *Nematostella* lives in colonies buried in sediment reaching a length of up to 15mm although under lab conditions they can grow up to 60mm in length. (**Figure 3**, **Figure 4**) (Technau et al., 2015; Layden et al., 2016).

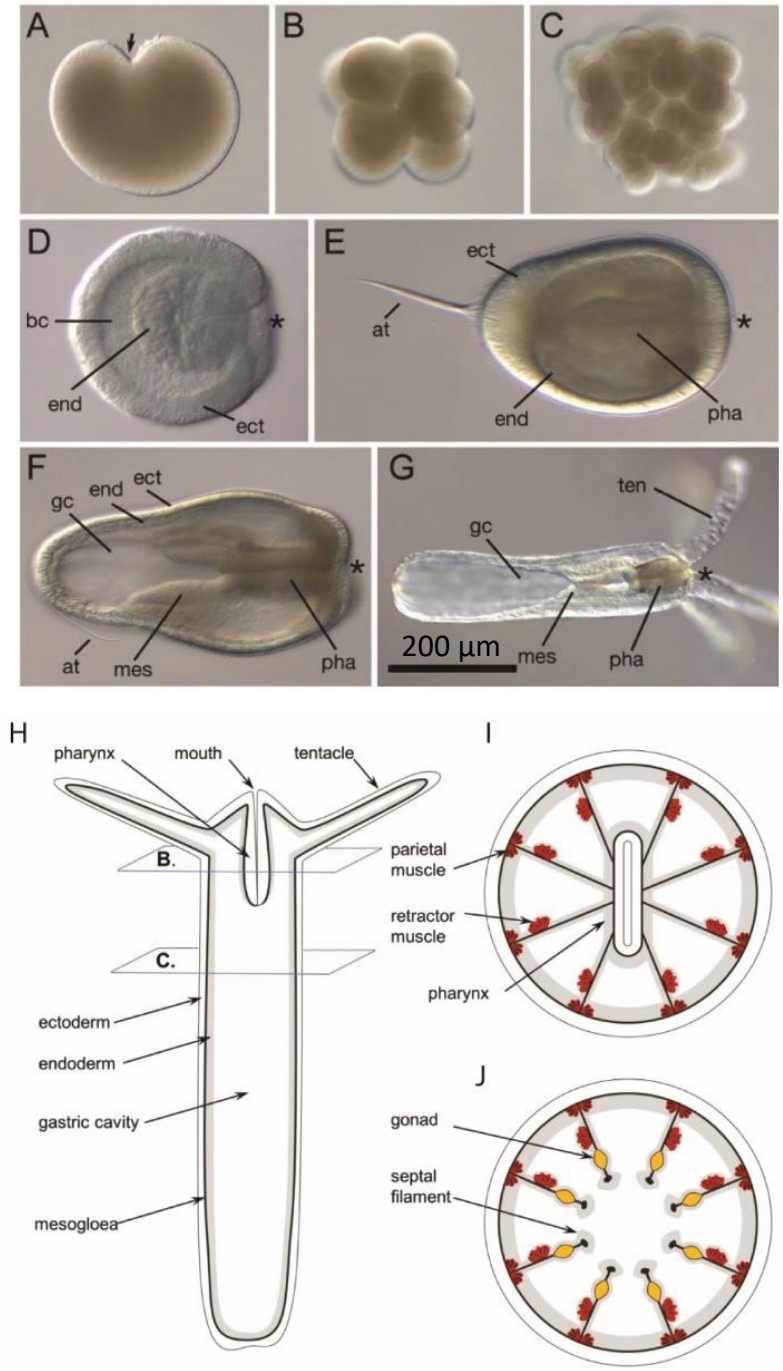

Figure 4: Developmental stages and body plan of Nematostella. ***A****: Zygote during first cleavage.* ***B****: 8-cell stage.* ***C****: 64-cell stage.* ***D****: Gastrula 24h post fertilization (pf).* ***E****: Planula 3dpf.* ***F****: Metamorphosing planula 6dpf.* ***G****: Primary polyp 8dpf. H: Schematic body plan of Nematostella.* ***I****: Section through the pharynx.* ***J****: Section through the gastric region.* ***Indications****: Arrow: first cleavage furrow. bc: Blastocoel, end: Endoderm, ect: Ectoderm. at: apical tuft, pha: Pharynx, gc: Gastrocoel, mes: Mesentery, ten: Tenctacle. *: Oral pole.* ***A-G*** *(Genikhovich and Technau, 2009a),* ***H-J****: Image courtesy of Grigory Genikhovich*

Consisting of two germ layers, commonly called endoderm and ectoderm, the morphology of *Nematostella* is superficially simple. The single oral opening which corresponds to the site of gastrulation leads into a sac-like gut. This mouth is surrounded by 16 tentacles which are used to catch prey. From the pharynx eight radially repeated mesenteries run along the oral-aboral axis. The mesenteries contain the gonads, cnidocytes, gland cells and myoepithelial cells, which allow a quick contraction of the animal along the oral-aboral axis. The arrangement of these retractor muscles, as well as the siphonoglyph, a ciliated groove on one side of the pharynx, establish a secondary body axis called directive axis. This subtle bilateral symmetry deviates from the radial symmetry around the oral-aboral axis often described in cnidarians (**Figure 4I**) (Layden et al., 2016).

In contrast to most bilaterians, in *Nematostella* specialized cells are not organized into organs but intermixed with differentiated epithelial cells in the ectoderm and endoderm. However, particular cell types are enriched in certain regions of the animal. To perform their assigned function of catching and transporting prey to the mouth, tentacles show a high number of nematocytes. These specialized sensory neurons contain a unique organelle called nematocyst. This organelle contains a harpoon like structure which unloads with extreme high speed upon mechano-sensory or chemical input, to envenomate the prey (Layden et al., 2016). Since the cell gets disrupted and destroyed by the discharge, this physiological function demands a high turnover rate of nematocytes. The cnidarian-specific stinging cell called nematocytes show neurophysiological properties, express neurogenic genes in their precursors and possess neurites capable of establishing synaptic contacts with other cells (Rentzsch et al., 2017).

Nematocytes make up one of the three general types of neurons in the cnidarian nervous system together with sensory cells and ganglion cells. These main classes are comprised of several subpopulations as molecular analyses have revealed. *NvElav1*-positive neurons e.g. form a network in the ectoderm and the endoderm. In the endoderm prominent longitudinal tracks are formed by these neurons following the parietal muscles on either side of the mesenteries (**Figure 5**) (Kelava et al., 2015). Besides this structure and a ring-like accumulation of a certain subpopulation of neurons in the aboral and oral body regions, no indication of centralization of the nervous system can be found. These observations can be made throughout the cnidarian phylum (Technau et al., 2015).

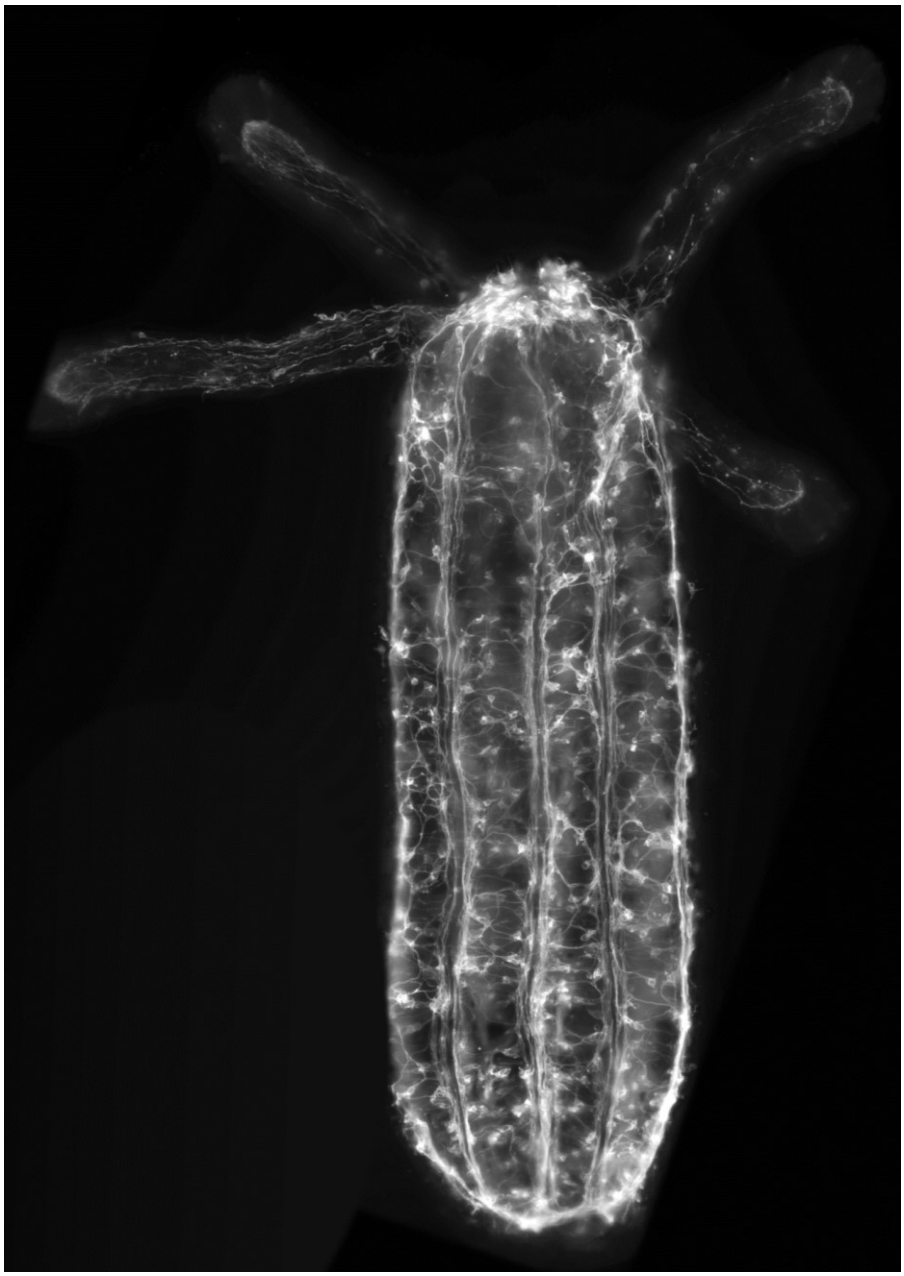

Figure 5: NvELAV1 positive cells in a Nematostella vectensis polyp.

The neurons in *Nematostella* generated throughout the ectoderm and the endoderm subject to a remarkably similar control of neurogenesis as it is present in vertebrates. The number of neural progenitor cells is controlled by *Notch* signalling whereas *SoxB* and *bHLH* transcription factors, amongst others, regulate their further development. *Wnt* signalling gradients are involved in the nervous systems patterning (Richards and Rentzsch, 2015; Bosch et al., 2017).

1.4. Single cell data reveal expression of conserved transcription factors

Nematostella shares the widely spread regeneration capabilities of other cnidarians such as *Hydra* although to this date no i-cell equivalent or stem cell niche was found in *Nematostella*. The genome of *Nematostella* reveals a genomic organization and regulatory landscape of comparable complexity as those of vertebrates. This enabled the search for homologues of known stem cell markers and the screening of these putative stem cell markers in *Nematostella* via *in situ hybridization* (ISH) and transgenesis (Schwaiger et al., 2014; Denner, 2017 unpublished; Renfer and Technau, 2017). A single cell (sc) approach by Alison Cole and Julia Steger was used to investigate germline-, i-cell-, and pluripotency-factors in 11 cell populations of *Nematostella* with collected data from 10 sequencing libraries *(***Figure 6***)*. These data were then analysed for germline-, i-cell- and pluripotency-factors found in the *Nematostella* genome.

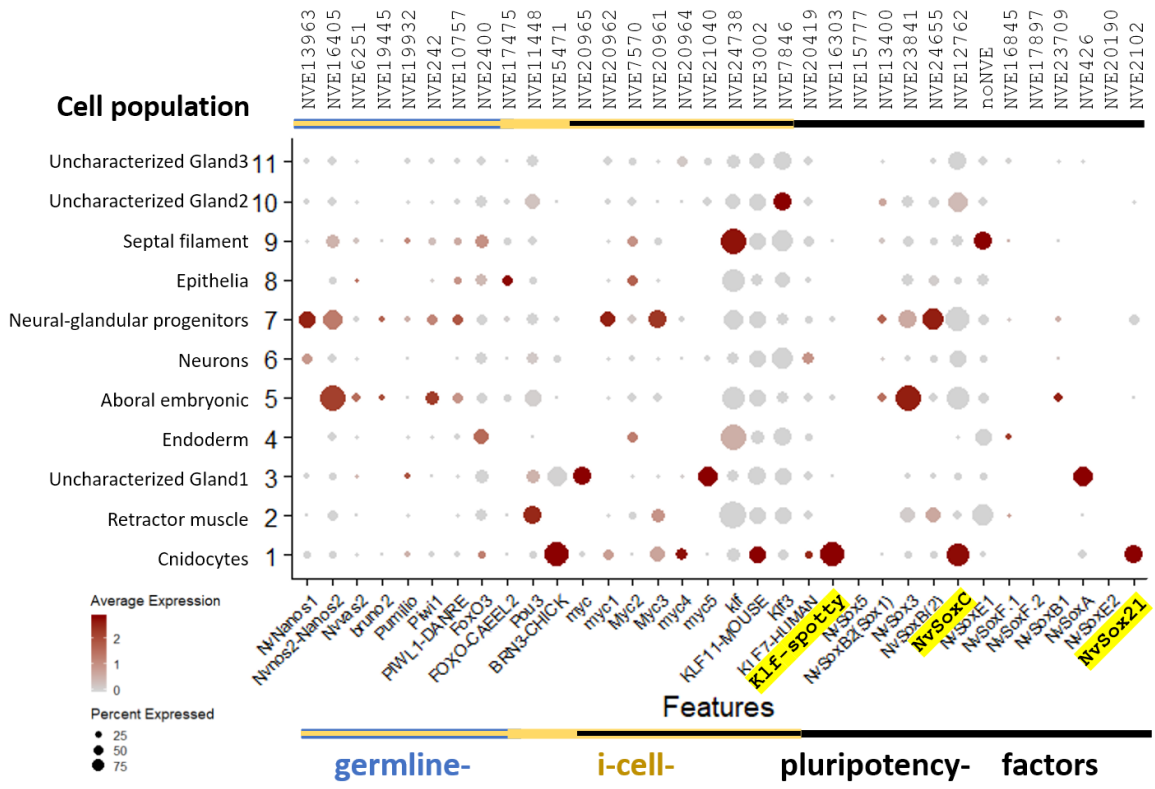

Figure 6: Expression of different transcription factors in cell populations of Nematostella. *The relative expression as well as the percentile expression of selected transcription factors associated with stemness. Genes that were subject to this study are highlighted.* ***Image courtesy of Alison Cole***

These preliminary data turn the spotlight on one gene, which is the main subject of this study, *NvSoxC (*Nve12762*)*. According to the single-cell data this gene is expressed mainly in

cnidocytes, neurons, and glandular cell types and to a small extent in endoderm and retractor muscles. These cell types overlap with those deriving from the i-cells in Hydra (**Figure 2B**) (Steele, 2002).

Two other genes that were selected due to their restricted expression pattern are *NvSox21* (Nve2102) and *Klf-spotty* (Nve16303). Transgenic reporter lines of these transcription factors were generated in a previous study (Ries, 2019 unpublished). Therefore, their known expression pattern can be compared with the new KAEDE lines for verification.

1.5. The Klf-Family

As part of the zinc-finger transcription factors called Krüppel-like factors (Klf) these transcription factors have 3 highly conserved Cys2/His2-domains on the C-terminal, which facilitates DNA binding. The N-terminal has a highly variable sequence with various functions such as protein-protein interactions, gene activation and gene repression. The suspicion that this gene family plays a key role in the emergence and diversification of cell types originates from phylogenetic analysis which revealed an origin of this gene family predating the divergence of the Metazoa. Furthermore, an expansion of this gene family correlating with the diversification of cell types can be observed (Presnell et al., 2015).

This diverse family of *Klf* proteins with 17 known mammalian members, forms a complex network of transcriptional control to mediate several developmental processes like genesis of blood vessels, skeletal muscles, skin, bones as well as neuronal tissue as well as the homeostasis of a multitude of different organ systems (Dang et al., 2000; Eaton et al., 2008; Brey et al., 2009; McConnell and Yang, 2010).

Several members of the Klf family such as *klf2*, 4 and 5 have been shown to be involved in the maintenance of pluripotency in embryonic stem cells of mice. Furthermore, reprogramming capacities regarding the conversion of epiblast-derived stem cells into pluripotent stem cells have been shown (Jiang et al., 2008; Jeon et al., 2016). *Klf4* which was shown to be a tumour related factor by repressing p53, is one of the Yamanaka factors used to reprogram somatic cells into induced pluripotent stem cells (Li et al., 2005; Rowland et al., 2005; Takahashi and Yamanaka, 2006).

A phylogenetic analysis of the *Klf* and other Zinkfinger proteins (Zfp) shows that Klf-spotty, which was investigated in this study clusters closely to the base of the *Klf* gene family (Denner, 2017 unpublished).

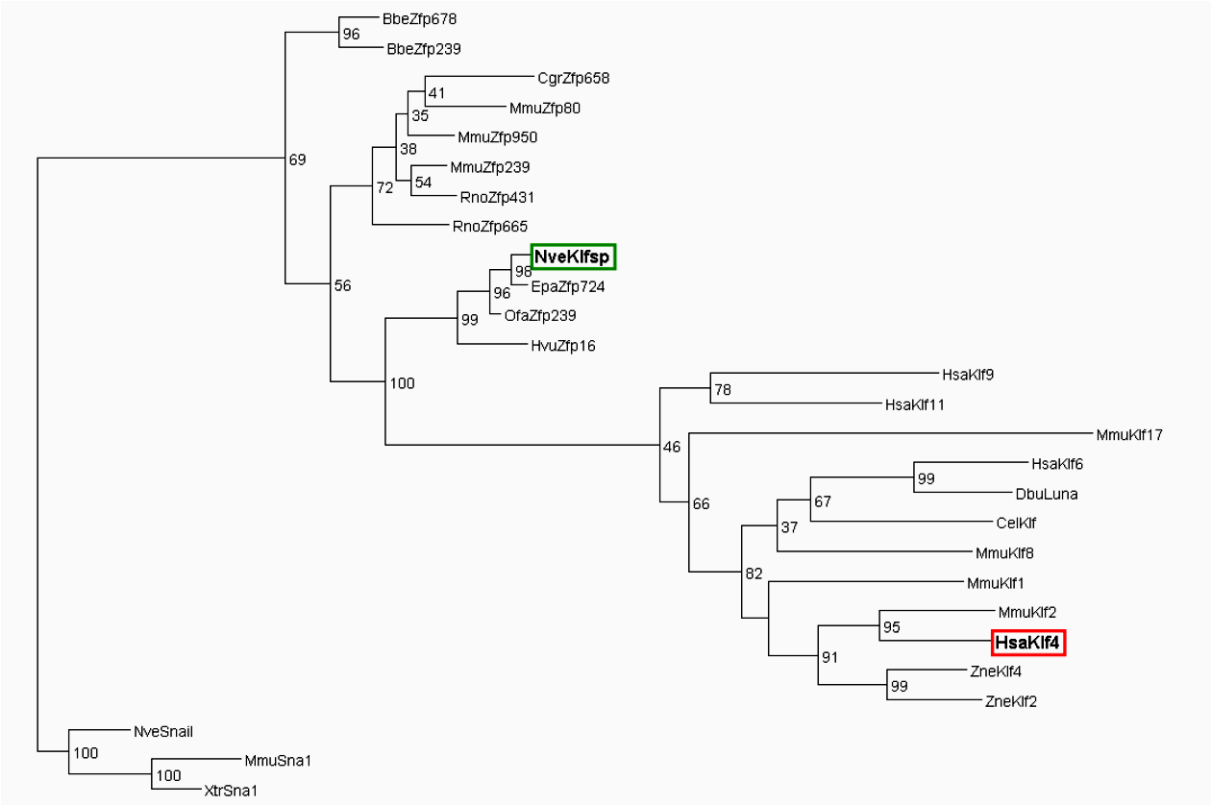

Figure 7: Klf and Zfp phylogenetic tree. *Red: Homo sapiens Klf4, Green: Nematostella vectensis Klf-spotty. Indicated at the nodes are bootstrap values from 1000 replicates. Nve: Nematostella vectensis, Has: Homo sapiens, Mmu: Mus musculus, Hvu: Hydra vulgaris, Ofa: Orbicella faveolata, Epa: Exiptasia pallida, Bbe: Branchiostoma belcheri, Rno: Rattus norvegicus, Cgr: Cricetulus griseus, Cel: Ceanorabditis elegans, Dbu: Drosohpila buschkii, Zne: Zootermopsis nevadensis. Image (Denner, 2017 unpublished)*

The expression pattern of Klf-spotty in the early developmental stages of *N. vectensis* investigated via *in situ* hybridization showed an abundant single cell expression throughout the gastrula and in primary polyp stage mainly in the tentacles and the ectoderm of the body column. Single cells could also be found in the pharynx (Denner, 2017 unpublished).

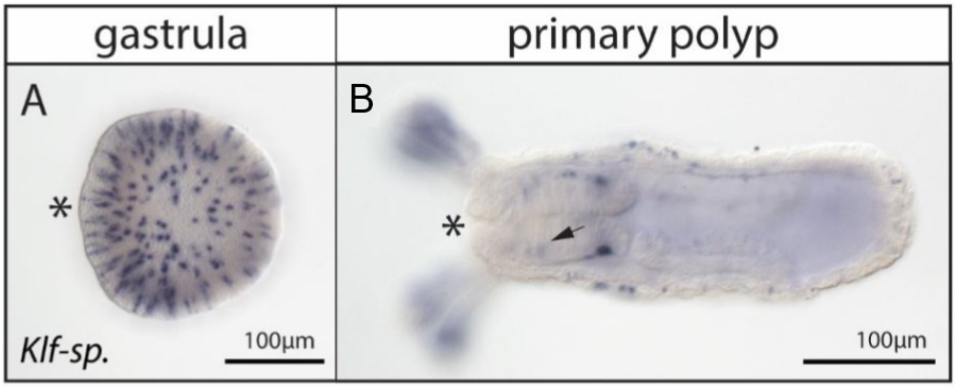

Figure 8: In situ hybridization of Klf-spotty. *A: gastrula stage, B: primary polyp, arrow indicates single cells in pharynx; oral pole indicated by * Image modified: (Denner, 2017 unpublished)*

The expression pattern of transgenic animals carrying a Klf-spotty-promoter::mOrange2 construct resemble the observed higher density of expressing cells in the tentacles (**Figure**

9A). Higher resolution images reveal that the cells expressing Klf-spotty are exclusively postmitotic differentiating nematocytes identified by the elongated shape containing a capsule and short protrusions (**Figure 9B**) (Technau et al., 2015; Ries, 2019 unpublished).

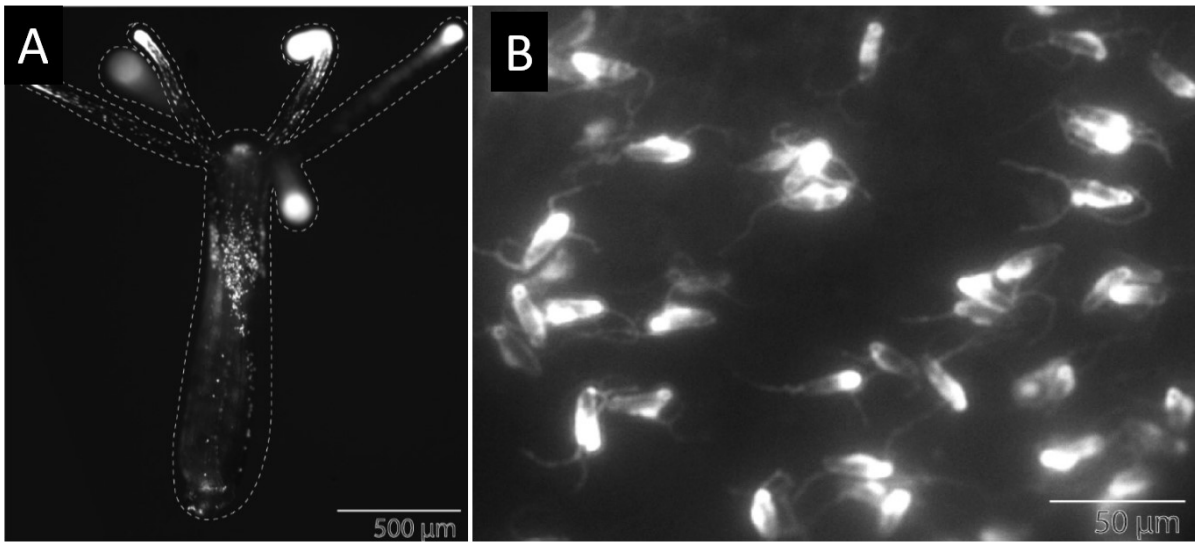

Figure 9: Klf-spotty transgenic reporter line. (A) *F0 transgenic 14d old polyp showing mOrange2 signal in the ectoderm with higher intensity at the tentacles.* ***(B)*** *High resolution image of the mOrange2 expressing cells in the ectoderm. Image modified: (Ries, 2019 unpublished)*

1.6. The Sox-Family

The first member of the highly diverse Sox gene family, which is specific to metazoans, was discovered as crucial player in mammalian male sex determination. Meanwhile the number of Sox genes known in vertebrates has climbed to about 20, which can be categorized in the eight *SoxA-H* groups by their protein structure. All Sox gene family members contain the high-mobility-group box (HMG-box) domain. This is a highly conserved DNA binding domain which can be traced back to choanoflagellates. The developmental regulatory functions throughout the family are highly divergent. Members of the same subgroup often share the same functions and are expressed in the same developmental tissue. *SoxB1* group proteins e.g. are found in the CNS whereas *SoxF* proteins can be found in the vascular system (Jager et al., 2011; Kamachi and Kondoh, 2013).

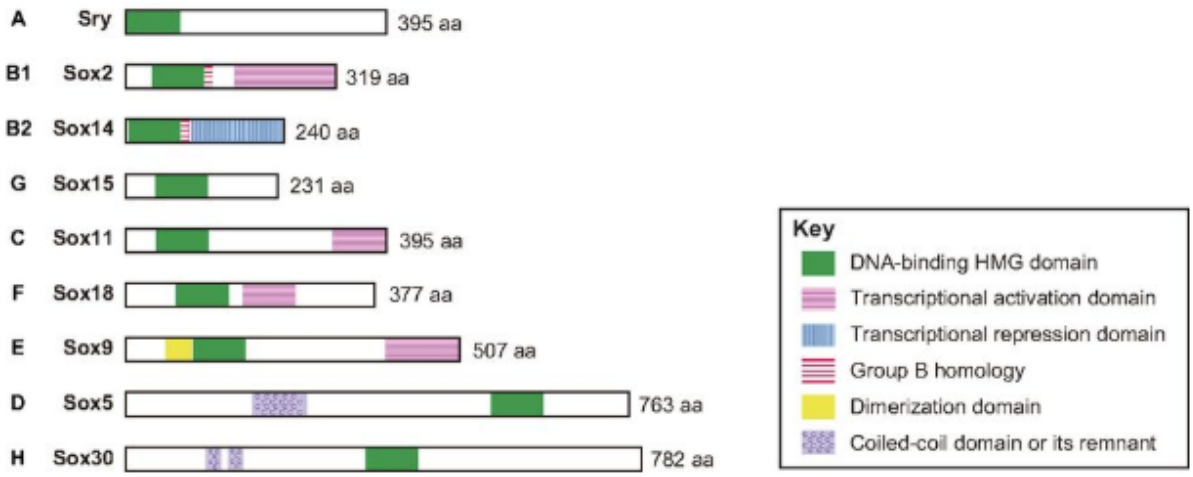

Figure 10: Structure of Sox Proteins. *Domainstructures of nine representative mouse Sox proteins showing the high conservation of structures among the family members. Image modified (Kamachi and Kondoh, 2013)*

One of the two Sox genes subject to this study is the gene *Nve2102* annotated as *NvSox21* (former *NvSox2*) and is member of a *SoxB*-related but Cnidarian-specific group (Jager et al., 2011; Denner, 2017 unpublished).

In vertebrates, the *SoxB* proteins are expressed in the neural plate throughout life in all neural progenitor/stem cells. This expression supports the survival of these cells since massive neural tissue degeneration was observed upon SoxB1 depletion in mice. Additionally, it maintains the progenitor/ stem cell state of these cells by preventing their transition to a more advanced developmental stage marked by *SoxC* expression (Kamachi and Kondoh, 2013).

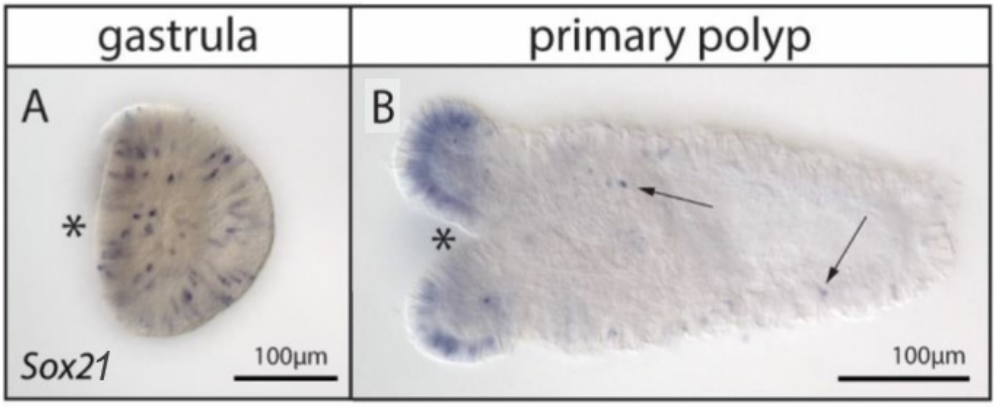

Figure 11: Expression pattern of NvSox21*. In situ hybridization of gastrula (A) and primary polyp (B) stage. ; oral pole indicated by * Image modified (Denner, 2017 unpublished)*

The expression pattern of *Sox21* in the early developmental stages of *N. vectensis* investigated via *in situ* hybridization demonstrated a single cell expression throughout the gastrula which accumulates in the tentacles during the primary polyp stage with only few expressing cells in the ectoderm of the body **(Figure 11)** (Denner, 2017 unpublished).

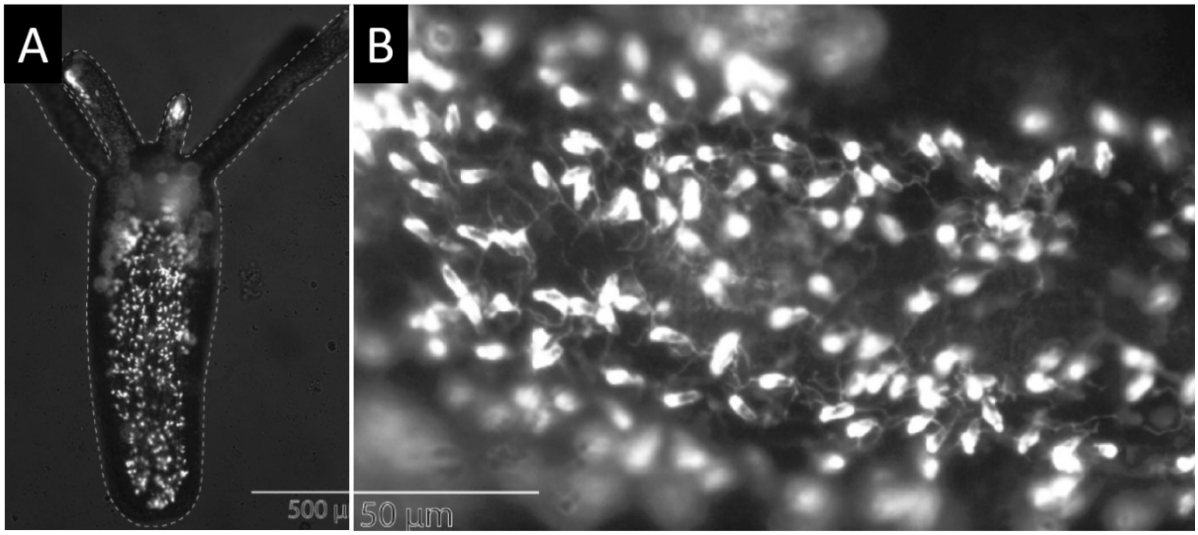

Figure 12: Sox21 transgenic reporter line. (A) *F0 transgenic 14d old polyp showing mOrange2 signal in the ectoderm with higher intensity at the tentacles.* ***(B)*** *High resolution image of the mOrange2 expressing cells in the ectoderm. Image modified: (Ries, 2019 unpublished)*

In *Nematostella vectensis* Nv*SoxB(2)* has been shown to be essential for the differentiation of neurons and cnidocytes via knock-down experiments. Transgenic reporter lines have shown an expression of *SoxB(2)* in a population of cells that give rise to ganglion cells, sensory neurons and cnidocytes therefore representing a population of neural progenitor cells (Richards and Rentzsch, 2014; Kelava et al., 2015).

The *SoxC* proteins can be mostly detected in post-mitotic differentiating neurons in mice. Experiments indicated that misexpression induces the expression of neuronal proteins, whereas the deletion of *SoxC* proteins in the spinal cord of embryonic mice causes a significant decrease in differentiated neurons (Bergsland et al., 2011). Single cell experiments demonstrated that the *Hydra SoxC* (*HvSoxC*) is almost exclusively active during the differentiation of cells undergoing nematogenesis, neurogenesis or early gland cell differentiation (Siebert et al., 2019). The expression pattern of the gene *NvSoxC* could so far only be visualized in ISH of the gastrula stage, showing an ectodermal single cell pattern.

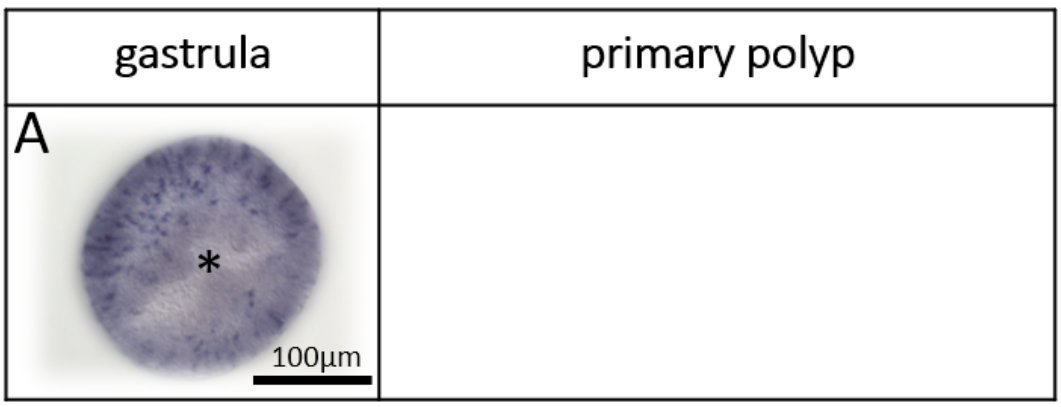

Figure 13: Expression pattern of NvSoxC. *(A) In situ hybridization of gastrula; oral pole indicated by * Image courtesy of Nina Znidaric*

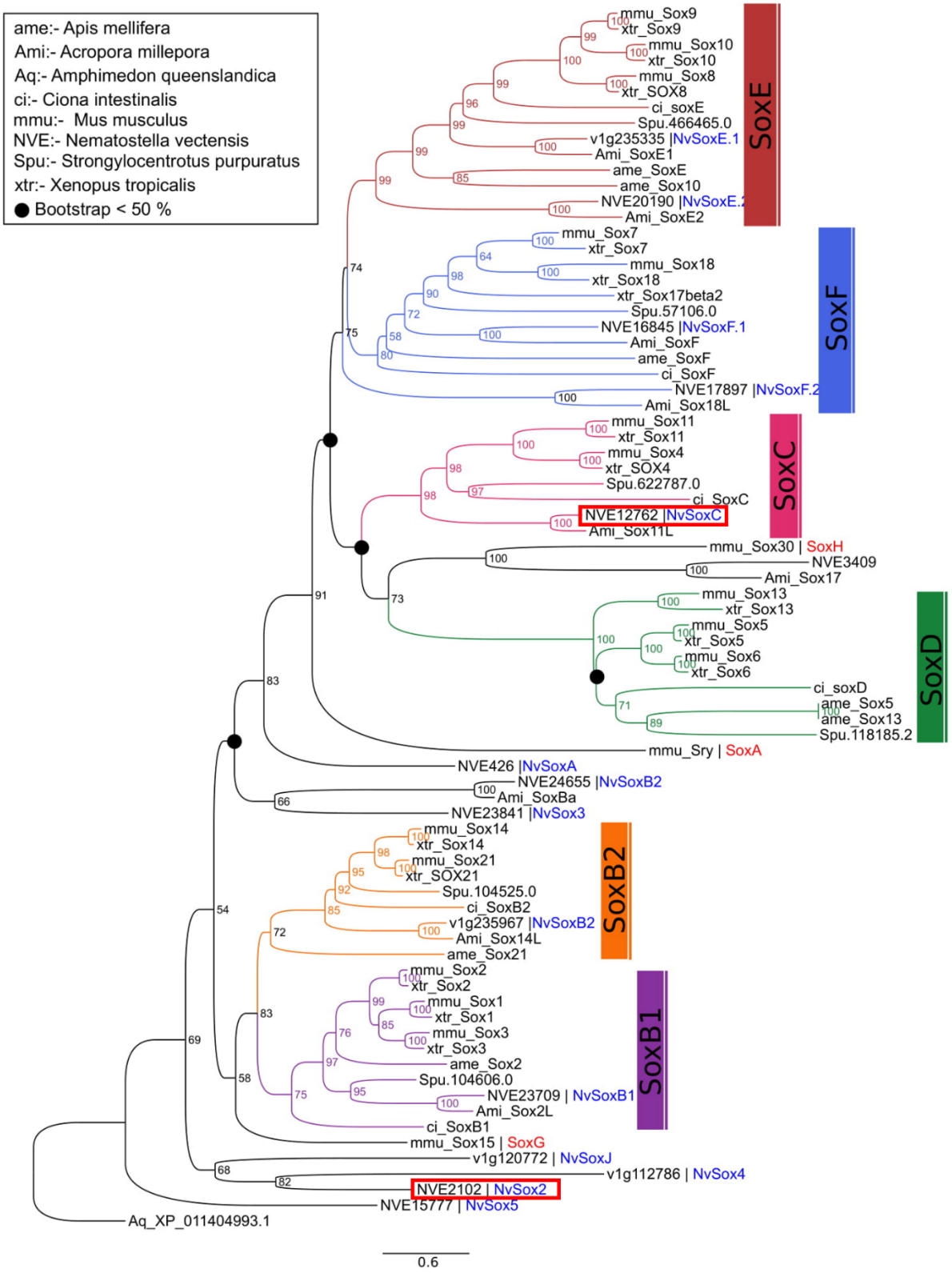

Figure 14 Sox family phylogenetic tree*: Indicated at the nodes are bootstrap values from 1000 replicates. Highlighted in red are the two Sox genes subject to this study; NVE12762/NvSoxC in the SoxC group and NVE2102/NvSox2 in a SoxB related group. Modified figure courtesy of Rohit Dnyansagar (unpublished).*

1.7. Transgenesis in *Nematostella* vectensis

The vector used in this study was published by Renfer et al. in 2010, where it was used in combination with the promoter of the *striated myosin heavy chain* (*MHC-st*) gene (Renfer et al., 2010).

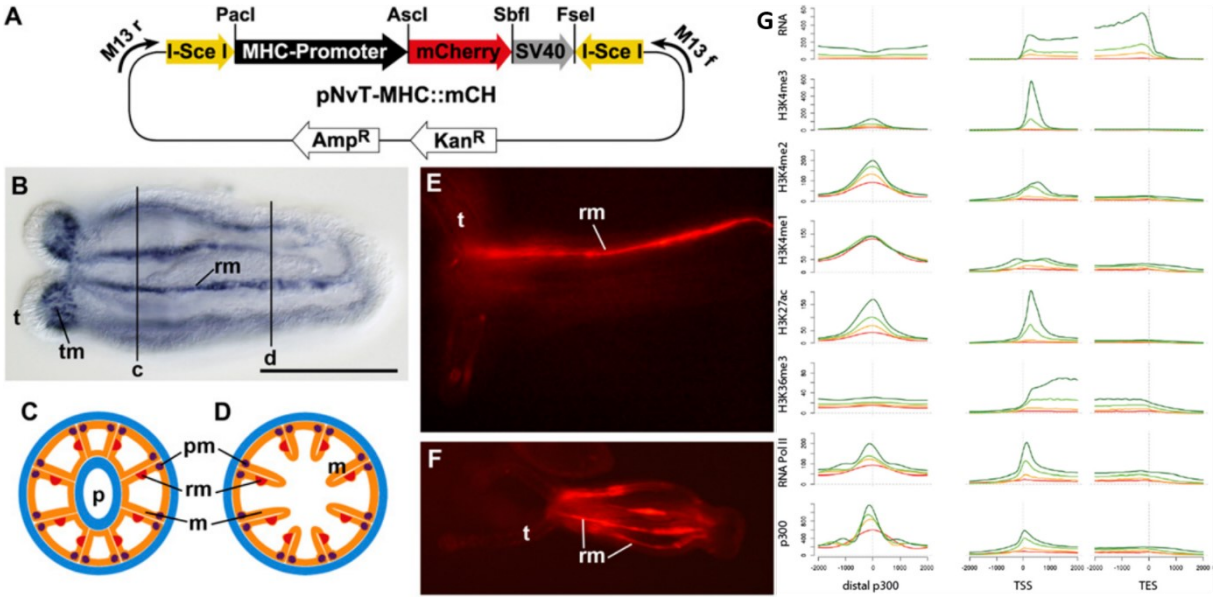

Figure 15: Schematic and specificity of the generation of gene specific fluorescence reporter lines. A: *Map of the vector containing the gene specific promoter site with a length of 1.6kb upstream the start codon. To generate a fragment for injection of the promoter, the fluorophore and the SV40 insulator are flanked by I-SceI meganuclease sites. B: in situ hybridization of a primary polyp targeting MHC. Retractor muscle (rm) and tentacle muscle (tm) are indicated. C,D: Drawing indicating the position of the muscles in sections of the pharynx (C) and gastric (D) region. E,F,: G0 mosaic transgenic animals with expression of the fluorophore in a single (E) or multiple (F) retractor muscles. G: Enrichment for epigenetic modifications related to the distal p300 site the transcription start site (TSS) and the transcription end site (TES) in respect to the FPKM ranging from <1.5 (red line), lowly expressed genes (orange line) medium expressed genes (green line) highly expressed genes FPKM >2 (dark green line). Images modified from A-F (Renfer et al., 2010) G from (Schwaiger et al., 2014)*

To clone the cis regulatory element of the *NvSoxC* gene, the epigenetic signatures upstream of the genes were analysed using the *Nematostella* genome browser. The histone modifications H3K4me2, H3K4me3 and H3K27ac which have been published as promotor specific regions, which were used to indicate a region of interest (Schwaiger et al., 2014). As shown in previous studies (Visel et al., 2009; May et al., 2011; Nègre et al., 2011; Shen et al., 2012) the binding of the p300 histone acetylase can be used to predict enhancer elements. Based on these parameters a 1-2.5kb genomic regions directly upstream of the *NvSoxC* coding sequence were used as promoter inserts.

Animals with successful integration of the transgenic construct into the genome determined by a visible fluorophore expression in the whole animal or in a mosaic pattern were observed during their development. Since these animals represent an F0 population it is most likely that only mosaic patterns of expression can be observed representing a subpopulation of cells deriving from cells which successfully integrated the construct.

1.8. The KAEDE Fluorophore

This fluorophore is based on the KAEDE protein cloned from the stony coral Trachyphyllia geoffroyi (Ando et al., 2002). This green (480nm) fluorophore (gKAEDE, visualised with cyan in the following pictures) undergoes self-cleavage upon illumination with UV light (350-410nm) and therefore switches irreversible to a red (540nm) emission (rKAEDE, visualised with magenta in the following pictures). This rather large shift increases the green-to-red ratio by 2 000-fold and is an irreversible process. (Ando et al., 2002; Dittrich et al., 2005). As Ando et al. state: *"…the red fluorescence emitted from the photoconverted Kaede is bright and stable for months without requiring rigorous anaerobic conditions*. *The red state is comparable to the green in terms of brightness and stability*". Unfortunately, this statement is neither aided by data nor by defined conditions.

1.9. Aims of the Project

To further investigate putative stem cell transcription factors this thesis explores the capacities of the KAEDE fluorophore for *in vivo* cell lineage tracing in the sea anemone *Nematostella vectensis.*

In situ hybridization (ISH) is a well- established technique to visualize protein expression (e.g. transcription factor). The mRNA of a transcribed gene is detected via an anti-sense probe and visualized by chemical reactions (Genikhovich and Technau, 2009a). This technique was and is heavily used to describe expression patterns of e.g. putative stem cell transcription factors, to reveal cell populations which express one or more of these transcription factors and are therefore putative stem cell populations (Denner, 2017 unpublished).

Two drawbacks of this technique are the fixation of the animals, which makes *in vivo* observations impossible, and the fact that this technique visualizes only cells currently expressing a certain mRNA and therefore does not allow to trace descendent cell populations where the gene of interest is not active anymore (**Figure 16**).

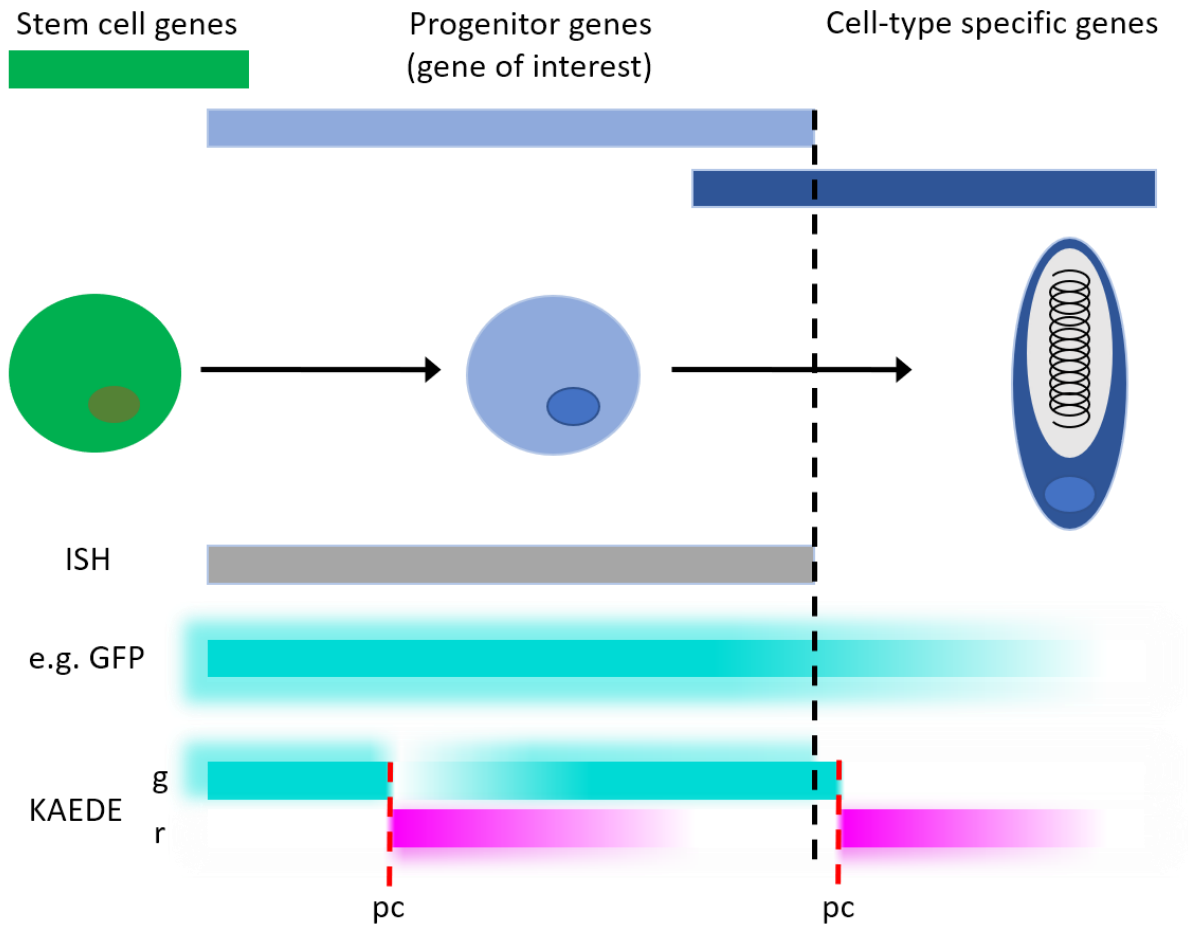

Figure 16: Comparison of different tools to investigate protein expression on the example of a progenitor gene. *A progenitor gene will be activated during a certain interval causing a differentiation from stem cells over progenitor cells toward a specific cell type. ISH will detect cells in which the gene of interest is currently expressed (on mRNA level) after the animal was fixed. Transgenic tools using fluorophores like GFP offer in vivo observations and the possibility of using immuno-histochemistry (IHC). Due to an unknown stability of the fluorophore a variety of cells with varying differentiation states might be detected although these are not expressing the gene of interest anymore. With the KAEDE-system the activity of the gene of interest can be monitored by photoconversion (pc) experiments which would result in either a regeneration of the gKAEDE signal upon activity of the gene promoter or no restoration if the promotor is inactive.*

Another useful tool to visualize protein expression is the use of transgenic fluorophores co-expressed under the same promoter as the gene of interest. The stability of these fluorophores allows the tracing of cells transiently expressing the gene of interest (**Figure 17**). Furthermore, fluorescent proteins (FP) are detectable *in vivo* and can be enhanced by antibody staining after fixation allowing detailed investigations at high magnifications with confocal microscopy (Renfer et al., 2010). The stability of these fluorescent proteins and the fact that descendants of former expressing cells can be observed hinders the identification of the progenitor cells that originally expressed the gene of interest (**Figure 17**).

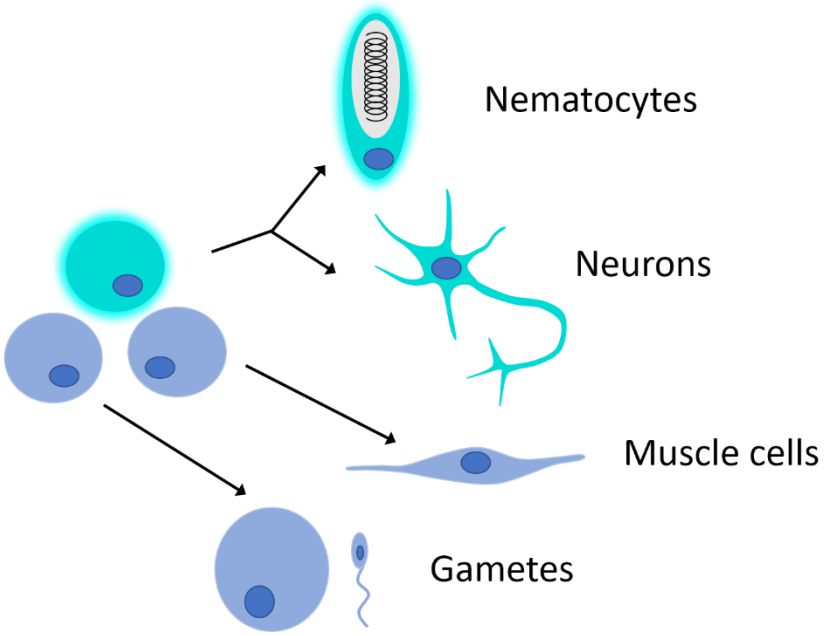

Figure 17: Cell lineage tracing using fluorescent proteins. *When a fluorescent protein (cyan) is coexpressed with a transcription factor specific for a precursor cell. Offspring of this cell will inherit the fluorescent protein even after the expression has been stopped.*

This is where the KAEDE fluorophore comes into play. As soon as the gKAEDE is detectable indicating an activity of the gene of interest photoconversion experiments can be performed *in vivo*. After photoconversion of gKAEDE only cells actively expressing the fluorophore would be able to restore the original gKAEDE signal. If the gene is turned off at a certain timepoint the gKAEDE signal would not be restored but only the rKAEDE signal would be detectable (**Figure 16**). The translucent nature of *Nematostella* facilitates this *in vivo* approach.

This study investigates the application of the KAEDE system in the model organism *Nematostella vectensis*. Transgenic reporter lines were created carrying the KAEDE fluorophore under the control of selected cis-regulatory elements. The lines generated with the *Klf-spotty* and the *NvSox21* reporter were used as proof of concept to validate the *KAEDE* transgenesis by comparing the observations with the already existing *mCherry* reporter lines of these promoters.

To investigate the function of *NvSoxC* in Nematostella and how conserved its function is related to *HvSoxC* and vertebrate *SoxC* the line was used to characterize the morphology of the detected cells, the activity of the TF in these cells for lineage tracing, and the migratory capabilities of these cells under amputation experiments and how well these properties can be visualised with the KAEDE system.

2. Material and Methods

2.1. *Nematostella* vectensis culture and spawning

Nematostella vectensis polyps were cultured at the animal facility of the Department of Molecular Evolution and Development of the University of Vienna. Female and male animals where kept separately at 18°C in darkness. Salinity of the *Nematostella* medium (NM) was 16‰ artificial sea water. The animals were fed five times a week with nauplius larvae of *Artemia salina*. The spawning of polyps was induced by a temperature rise to 25°C during a 10 hours light stimulus. After the stimulus the animals were brought back to 18°C under constant illumination and overnight-eggs were removed. After 2 hours most animals will have spawned and the eggs can be fertilized by transferring them into the male box with NM containing sperm for 30 min.. To remove the gelatinous package around the fertilized eggs they were treated with a 3% cysteine solution in NM at a pH 7.5 for 30min. The loose eggs were washed 5 times in NM and stored at 13°C to stall development for further applications. The embryos were raised at 21°C in darkness and fed with ground Artemia until polyp stage (Fritzenwanker and Technau, 2002; Genikhovich and Technau, 2009b).

2.2. Primer design and PCR

To generate transgenic lines of *Nematostella vectensis*, the promoter regions of the target genes were identified as described in (Schwaiger et al., 2014) using the genome generated in the Technau group (Fredman, Zimmermann, Rentzsch and Technau, unpublished). Primers were designed using the online tools Primer3 (Rozen and Skaletsky, 2000) and PCR Primer Stats to assure quality. The chosen primers were as well blasted via NCBI (Altschul et al., 1990) and JGI (Grigoriev et al., 2012) to assure uniqueness. After the primers were designed restriction sites for the enzymes PacI and AscI were added to the sequences preceded by 5 random bases at the 5' ends.

Table 1: Primer sequences for the cloning of promoters.

Gene Name	NVE	Direction	Primer sequence 5'-3'
NvSoxC Promoter	Nve12762	Forward	CAGTGTTAATTAACAAACATCAGTCTTAAACTC
		Reverse	GAGTCGGCGCGCCCACTAGAAACTAATTCAAAATG
KLFSpotty Promoter*	Nve81344	Forward	TTAATTAACGAAAGAAAACCAGTGGCCG
		Reverse	GGCGCGCCGATGGGTGTTGAGTATGGAG
Sox21 Promoter*	Nve2102	Forward	TTAATTAACTCCAAACTGGTCGGACTGC
		Reverse	GGCGCGCCGGTGCAACTGAAAATCAAGGA
Sox21 enhancer*	Nve2102	Forward	GGCCGGCCCAAACGTTTTATTTCATTGAGTCTAAAAAGCT
		Reverse	TAGGGATAACAGGGTAATGGCCGGCC

*by Andreas Denner

The PCR of the target region was performed following the standard Q5 PCR protocol.

Q5 polymerase PCR	
5x Q5 reaction buffer	5 µl
dNTPs (10mM each)	0,5 µl
Forward primer (10 µM)	1,25 µl
Reverse primer (10 µM)	1,25 µl
gDNA (1:10 dilution)	1 µl
Q5 High-Fidelity DNA Polymerase	0,25 µl
Water	15,75 µl
Total	25 µl

Q5 PCR program		
1	98 °C	30 sec
2	98 °C	10 sec
3	annealing T.	10 sec
4	72 °C	30s/kb
		Repeat step 2-4 40 times
5	72 °C	2 min
6	4 °C	Hold

To harvest the fragments a gel electrophoresis was performed with a 1% agarose gel using a voltage of 120V for 15min. Bands of target sequences were cut and eluted using a PeqGold elution kit.

2.3. Cloning

The pCRII-TOPO vectors used to generate transgenic animals carried either the mCherry fluorophore as used in (Renfer et al., 2010) or the KAEDE fluorophore (obtained from IST Kicheva Group) cloned in via the *Asc1, Fse1* restriction sites.

The amplicons generated were digested by *AscI* and *PacI* to generate sticky ends. The fragments were ligated into the pCRII-TOPO vector using the following ligation mix at 37° for >15min.

pCRII ligation mix	
5 µl	2x Ligation buffer
1 µl	Insert
0.5 µl	pCRII vector
3 µl	Water
0.5 µl	T4 DNA ligase
10 µl	Total

The following promoter-fluorophore pairs were generated this way:

NvSoxCprom::Kaede
NvSoxCprom::mCherry
Sox21prom::Kaede::enhancer
KLF-spottyprom::Kaede

To transform the vectors into chemically competent E. coli bacteria, 2x50 µl of the bacteria were slowly thawed on ice. For each aliquot 5 µl of ligation mix was added and mixed by gently flicking and incubated on ice for 30 min. A heat shock was performed for 60 sec at 42°C after which the cells were incubated on ice for 1 min. After adding 250 µl of SOC the bacteria were incubated for 30min at 37°C. 100 µl were spread on prewarmed LB-Ampicillin plates and incubated at 37°C over-night.

Supermedium (SOC)
25g peptone
15g east extract
5g NaCl
Water to 1 l
Add 15g agar for plates

2.4. Colony check PCR and plasmid extraction

Per plate 10 colonies were selected for a colony PCR to check for the insert of the amplicon.

Taq polymerase colony PCR	
0.8 µl	Taq polymerase
0.1 µl	M13 forward 10 µM
0.1 µl	M13 reverse 10 µM
0.1 µl	dNTPs (10mM each)
1 µl	Green Dream Buffer (10x)
8.62 µl	Water
10 µl	Total

Taq PCR program		
1	96 °C	2 min
2	96 °C	5 sec
3	55	10 sec
4	72 °C	1 min/kb
		Repeat step 2-4 25 times
5	72 °C	Time step 4 x 3
6	4 °C	Hold

The reaction was transferred on an agarose gel as above. Colonies with the appropriate insert size were used to set up a liquid culture overnight. The plasmid was harvested via miniprep with the innuPREP Plasmid Mini Kit Analytik jena). The plasmids were sent to Microsynth for Sanger sequencing and the results were analysed using ApE (ApE- A plasmid Editor, 2020).

2.5. Microinjection

To generate transgenic animals, fertilized eggs were prepared as described above and injected with a plasmid-meganuclease mix.

Transgenesis injection Mix	
1 µl	IScel buffer 10x
2 µl	Alexa 488
0.3 µl	IScel enzyme
20 ng/ µl	Plasmid
To 10 µl	Water

The microinjection setup was as follows:

- Two micromanipulators (Narishige)
- Inverted epifluorescence microscope (Nikon TS100F)
- Femtojet pump (Eppendorf)
- Flaming/Brown Micropipette puller (Model P-87 Sutter instrument CO)
- Holding capillaries 100 µm (Eppendorf VacuTip FCH)
- Water cooling system (Technau Lab)

2.6. Imaging and photoconversion of KAEDE

The Kaede fluorescence protein used to generated the transgenic animals has a switchable chromophore that can be switched from an emission of green (518nm) to red (580nm) light via exposure to UV light (350-400 nm) as stated in (Ando et al., 2002).

To perform this photoconversion (pc) *in-vivo,* the animals were relaxed in 1.5% MgCl2 in NM for at least 30min and placed on a 0.5% agarose gel in 1.5% MgCl2 in NM. The right orientation was determined under a Nikon SMZ18 fluorescence stereomicroscope and corrected via carefully pipetting the animal up and down. For animals with a bigger body diameter a little groove was cut in the gel beforehand. Due to the soft nature of the gel a coverslip could be put on the animals without squishing them.

Photoconversion and imaging were performed with a Nikon Eclipse 80i epifluorescence microscope. A conversion of the fluorophore was reached after 30sec of exposure to UV light under usage of a DAPI filter block. Before and after the conversion the animals were imaged using conventional filter blocks suitable for GFP and RFP observation. Images were processed with Fiji (ImageJ, 2019), Adobe Lightroom (Adobe Photoshop Lightroom, 2020) and Helicon Focus (Helicon Focus - Helicon Soft, 2020). *NvSoxC*.prom.::KAEDE animals with significant transgenic cell populations were selected an named a01-a16.

2.7. Amputation experiments

Some animals were used for amputation experiments. Their aboral part was further identified as “f” (for foot) to the animal name (e.g. a08f). Animals selected for amputation experiments were relaxed as described above (2.6) and observed under the stereomicroscope. The amputation of the foot was performed with a scalpel blade 3-4mm from the aboral end. Afterwards the animal was brought into NM and washed several times. The foot was treated as described above (2.6)

2.8. Image processing

KAEDE fluorescent images were converted into monochrome images and adjusted in brightness and contrast using Fiji - ImageJ (Schindelin et al., 2012). To visualise the overlay of the two fluorophores the images were merged using the mbf plugin for ImageJ (Tony Collins,

2013). When a high background was given due to autofluorescence the background was subtracted using a rolling ball radius of 50 pixels before the images were merged.

3. Results

3.1. Transgenic expression in early stages

3.1.1. *Klf-spotty* highlights nematocyte populations

To further investigate the origin and fate of *Klf-spotty* (Nv16303) expressing cells a transgenic line was established using the KLF-spotty promoter region of 6794 bp upstream of the start codon fused with the KAEDE fluorophore. This construct was digested from the vector using I-Scel Meganuclease and injected into 1.930 fertilized eggs.

Table 2: Survival and transgenesis rate Klf-spotty.promoter::KAEDE
**the animals were lost after the observation at 9dpf due to too long incubation in $MgCl_2$*

	Klf-spotty.prom.::KAEDE	
	n	%
Injected zygotes	1 930	100
24-h survival	820	42
24-h expression	169	9
9-d survival	0*	0*
9-d expression	0*	0*

The gastrula of animals injected with the *Klf-spotty.*promoter::KAEDE construct shows a distinct single cell expression pattern (**Figure 18A**) reminiscent of the pattern observed by *in-situ* hybridisation (**Figure 8A)** keeping in mind that these animals are only showing mosaic pattern expression.

In the primary polyp stage (**Figure 18B**) the transgenic animals show expression of the fluorophore in cells which resemble the classic features of nematocytes with a dimmer part of the cell inhabiting the nematocyst and fine, short cell protrusions, potentially functioning as a triggering sensor **Figure 18C** (Technau et al., 2015; Rentzsch et al., 2017). These cells are distributed along the ectoderm of the body column and an increased expression can be observed in the tentacles which is also reminiscent of the pattern in (**Figure 8B*)***.

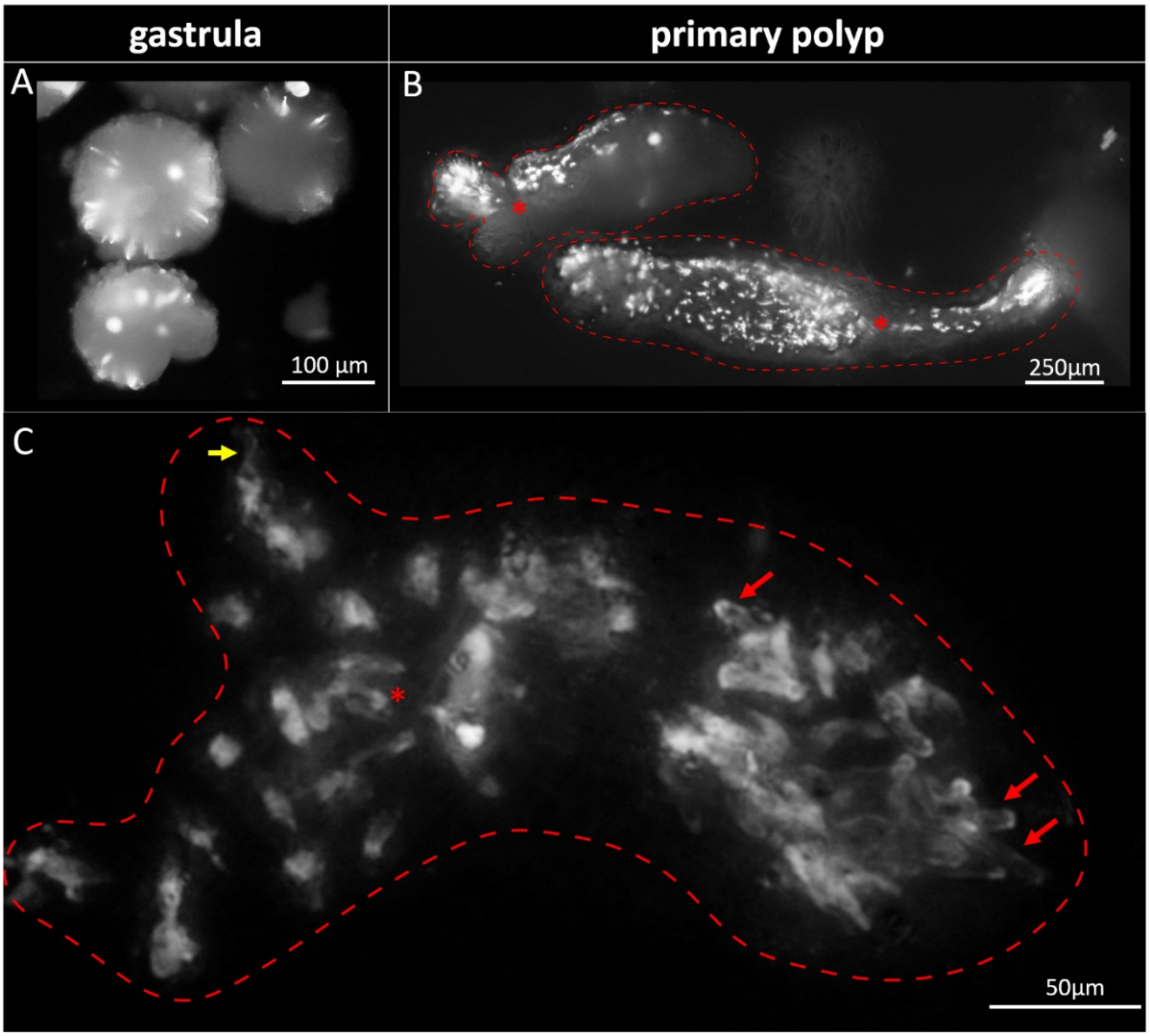

Figure 18: Klf-sp.prom.::KAEDE animals F0. *The transgenic animals resemble the salt and pepper expression observed in in situs at gastrula stage (A). In primary polyps the fluorophore is expressed in cells along the body column and higher concentrated towards the tentacle tips. A higher magnification reveals the morphology of the expressing cells as elongated structures (red arrows) with a dimmer spindle part and a small, faint sensilla (yellow arrow) (C). Some well visible examples are pointed out. Oral pole indicated by **

3.1.2. *NvSox21* highlights nematocytes and dim putative neurons

To further investigate the origin and fate of *NvSox21* (Nv2102) expressing cells a transgenic line was established using the *NvSox21* promoter region of 3050 bp upstream of the start codon fused with the KAEDE fluorophore followed by a 2782 bp enhancer sequence. This construct was digested from the vector using I-Scel Meganuclease and injected into 1.900 fertilized eggs.

Table 3: Survival and transgenesis rate NvSox21.prom.::KAEDE

	NvSox21::KAEDE	
	n	%
Injected zygotes	1 900	100
24-h survival	769	40
24-h expression	501	26
9-d survival	327	17
9-d expression	87	5

The KAEDE expression in the gastrula stage in transgenic animals resembles the salt and pepper pattern observed in the *in situ* (**Figure 11**). Nine days post fertilisation on the other hand the expression pattern shows a high number of cells in the ectoderm which resemble the features of nematocytes with a well distinguishable capsule and a short protrusion (**Figure 19**) (Technau et al., 2015).

In addition, some smaller and weaker cells are visible at higher magnifications (**Figure 19C, blue arrow**) which seem to have long cell protrusions. Sc data show an expression of Sox21 in neuronal progenitors (**Figure 6**) suggesting that these cells are of neuronal nature.

In addition to the short protrusions long, straight structures are visible, which seem to stretch from one nematocyte to another. Increased expression in tentacles can be observed in the transgenic animals as well.

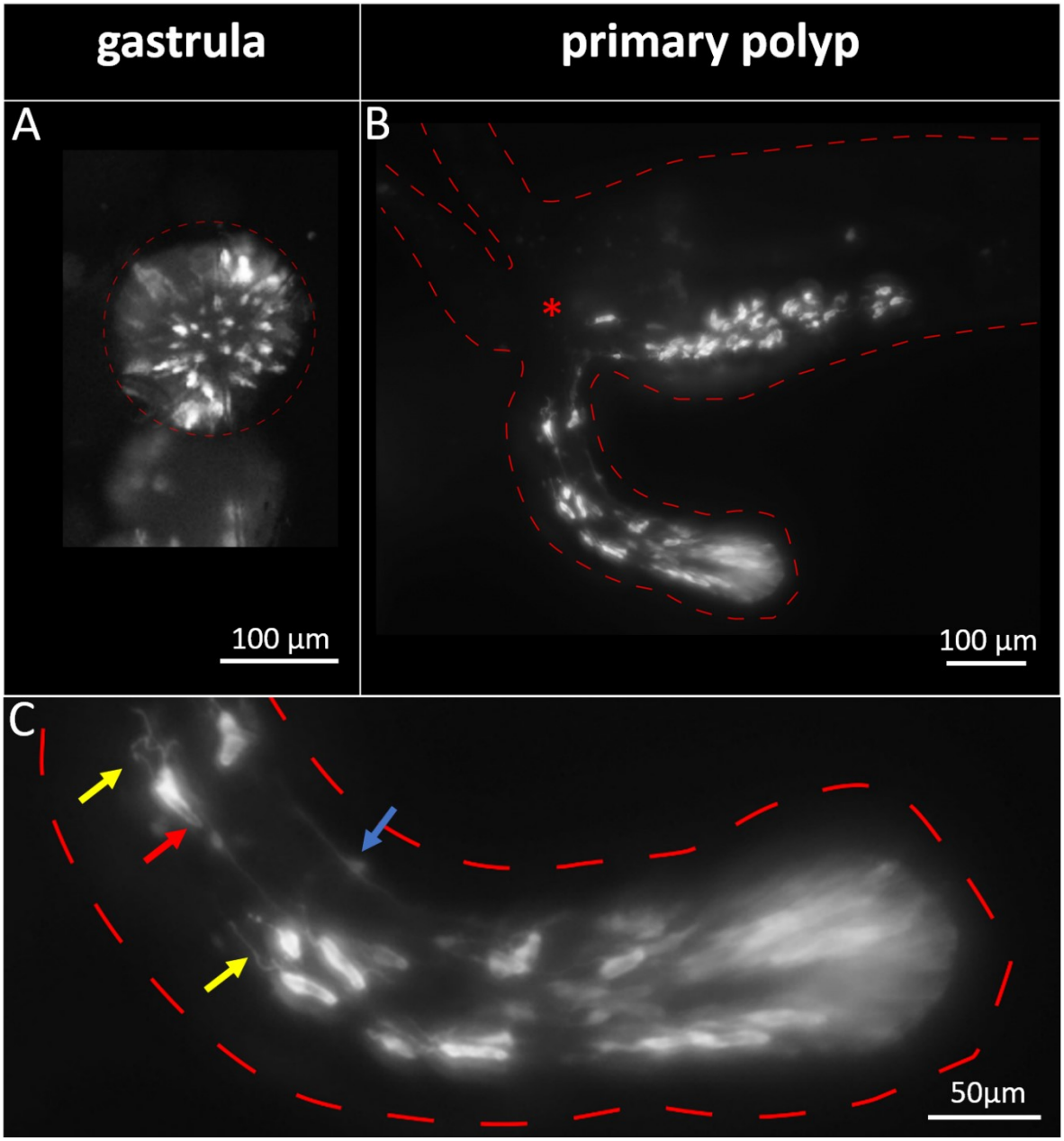

Figure 19: NvSox21.prom.::KAEDE animals F0. *The transgenic animals resemble the salt and pepper expression observed in in situs at gastrula stage (A). In primary polyps the fluorophore is expressed in cells along the body column and tentacles, with an increased signal at the tentacle tips. A higher magnification reveals the morphology of the expressing cells as elongated structures (red arrows) with a dimmer spindle part and a small, faint sensilla (yellow arrow). Weak cells with long protrusions (blue arrow) (C). Some well visible examples are pointed out. Oral pole indicated by **

3.1.3. NvSoxC highlights nematocyte and neuron populations

To further investigate the origin and fate of *NvSoxC* (Nv12762) expressing cells a transgenic line was established using the *NvSoxC* promoter region of 3603 bp upstream of the start codon fused with the KAEDE fluorophore. Furthermore, the same sequence was cloned into a mCherry plasmid. These constructs were digested from the vector using *I-SceI* Meganuclease and injected into 3.371 and 2.226 fertilized eggs respectively.

Table 4: Survival and transgenesis rate NvSox21.prom.::KAEDE and NvSox21.prom.::mCherry

	NvSoxCprom.::KAEDE		NvSoxCprom.::mCherry	
	n	%	n	%
Injected zygotes	3 371	100	2 226	100
24-h survival	573	17	589	26
24-h expression	470	14	261	12
9-d survival	363	11	216	10
9-d expression	105	3	63	3

The KAEDE expression pattern shown in the gastrula stage of the transgenic animals shows a single cell pattern. Contrary to the observations of the previous transgenic animals the single cells seem to be much bigger and do not resemble the previously observed cylindrical shape (**Figure 20A**). In the 9 day primary polyp stage a single cell pattern can be observed in the ectoderm of the body column as well as expression of the Fluorophore in the tentacles (**Figure 20B**). Contrary to the previous transgenics the cells expressing KAEDE seem to resemble the known nematocyte morphology of an elongated shape with a visible capsule. Cells with a round to triangle shape showing long cell protrusions in opposite directions were observed as well (**Figure 20C**).

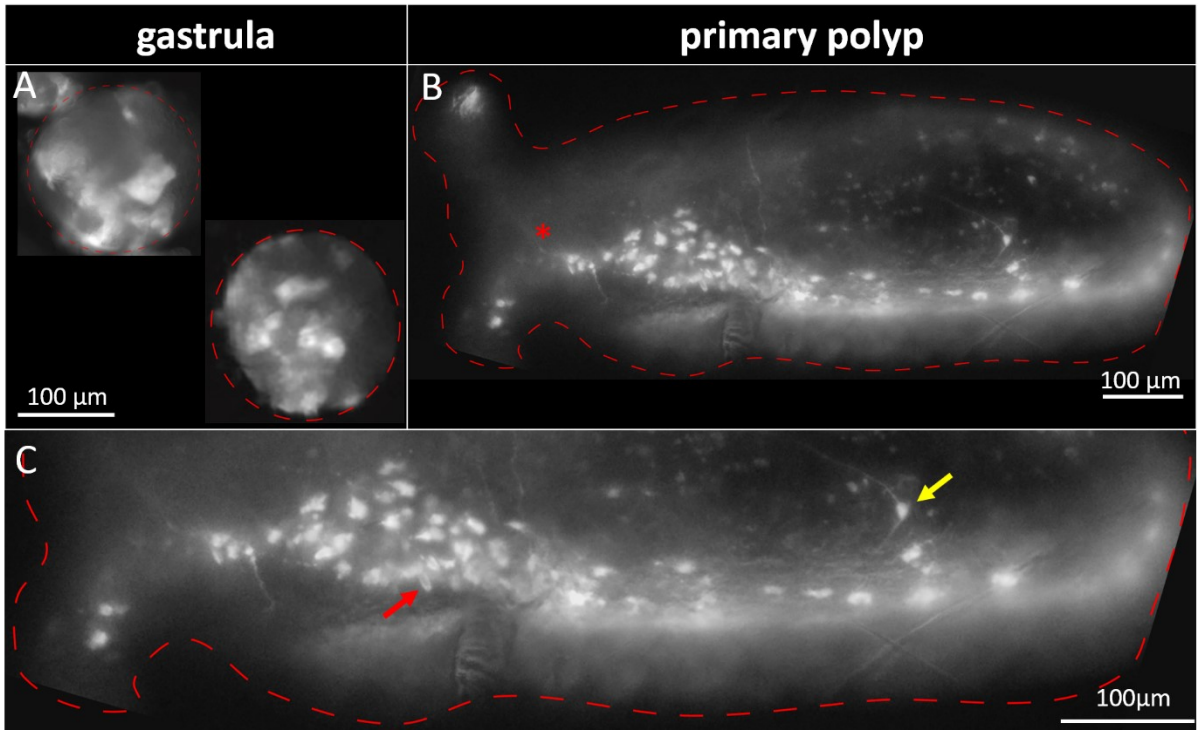

Figure 20: NvSoxC.prom.::KAEDE animals F0. *The transgenic animals show a single cell signal with cells in various shapes and bigger than observed in the two previous transgenic animals in gastrula stage (A). In primary polyps the fluorophore is expressed in cells along the body column and tentacles (B). A higher magnification reveals the morphology of the expressing cells as elongated structures (red arrow) with a dimmer spindle part. Other cells show a more rounded to triangle shape with long protrusions expanding (yellow arrow) (C). Some well visible examples are pointed out. Oral pole indicated by **

3.2. KAEDE on the *in vivo* test bench

In adult animals the KAEDE expressing cells are well visible with fluorescent microscopes. It was observed that upon periods of intense observation the fluorophore underwent photoconversion caused by the excitation light itself (**Figure 21**).

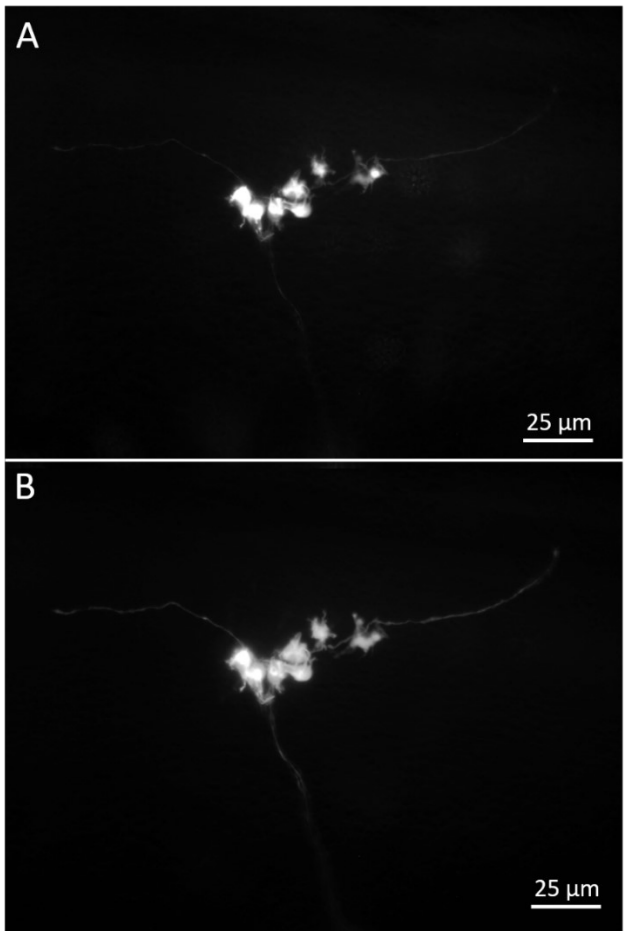

Figure 21: KAEDE expressing cells in animal a06 after an extended observation. *After intensive observation of gKAEDE positive cells the fluorophore is partly photoconverted by the excitation light* ***(A)*** *Green channel****. (B)*** *Red channel****.***

To determine how efficient the photoconversion of KAEDE *in vivo* is and how long the signal could be detected a number of simple experiments were performed using the former generated *NvSoxC*.prom.::KAEDE animals.

The animal selected for these experiments (called a12) expressed KAEDE in a stripe like cell population in the outer body wall. Two spots were selected for the photoconversion. Before pc cells expressing gKAEDE show a weak signal in the red channel (**Figure 22C**). Immediately after photoconversion, the green signal is strongly diminished whereas the red signal is considerably brighter in the selected spots.

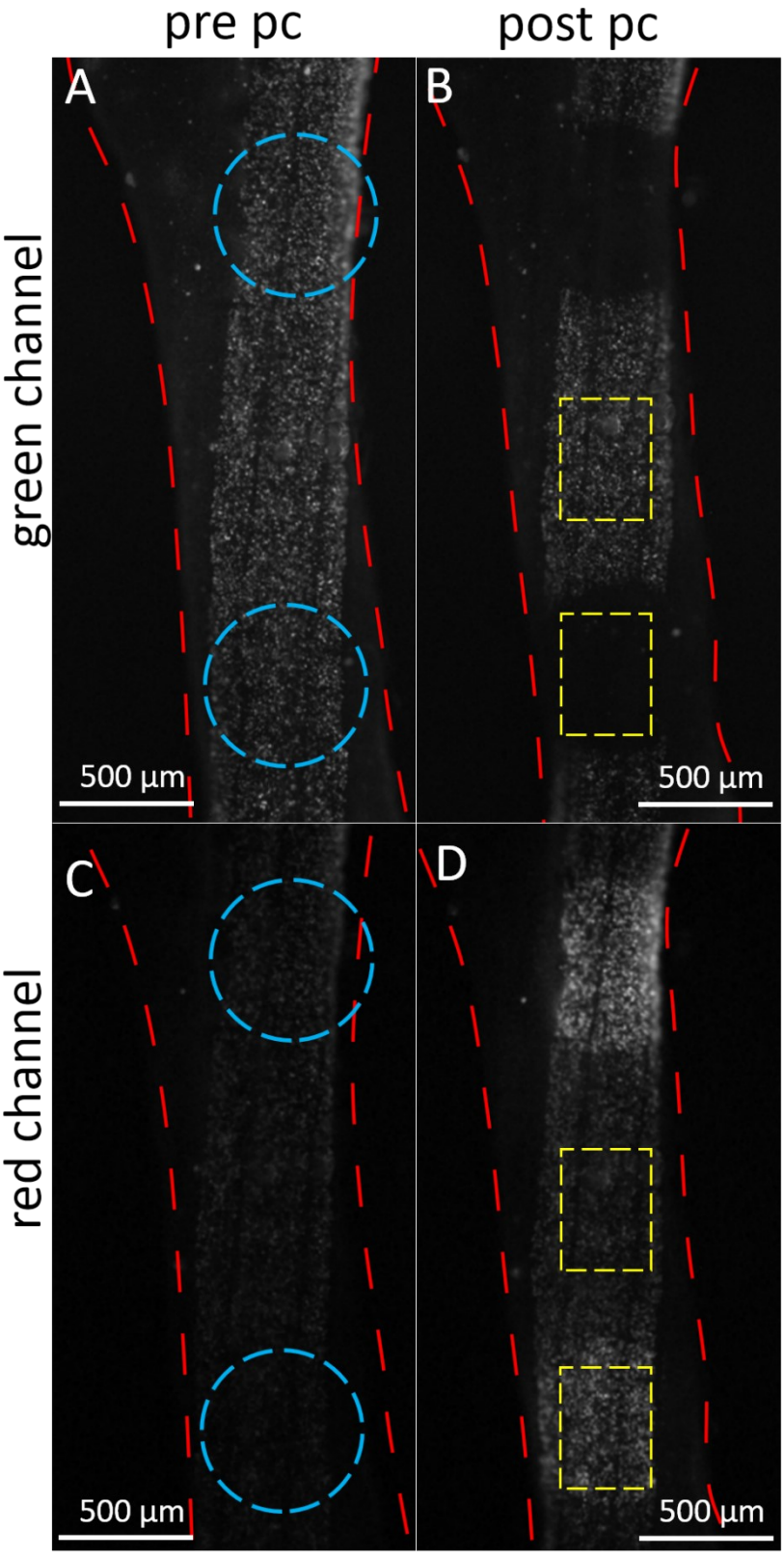

Figure 22: partial body column of the KAEDE transgenic animal a12 imaged in the green (A,B) and red (C,D) channel before (A,C) and after (B,D) photoconversion*. In this partial body-column of animal a12 expressing the KAEDE fluorophore in a certain cell population along the body wall, two spots were selected for photoconversion (pc) (blue circles).* ***(A,C)*** *Images taken before the photoconversion . A very low but detectable signal in the red channel is already visible pre-conversion. After the conversion (**B,D**) the green signal is obliterated at the selected spots (**B**) whereas the red channel shows a strong increase in fluorescence post-conversion (**D**); Mean brightness measurements were performed to evaluate the change in green-to-red ratio (yellow rectangles).*

Average brightness measurements of selected areas (yellow rectangles **Figure 22B,D**) show that the green-to-red ratio (R_{gr}) changed from R_{gr}= 1,55 pre conversion to R_{gr}= 0,05 post conversion and therefore shows a ~32-fold change. This change, although far from the stated 2,000-fold observed by (Ando et al., 2002) in cell cultures, is sufficient for further studies.

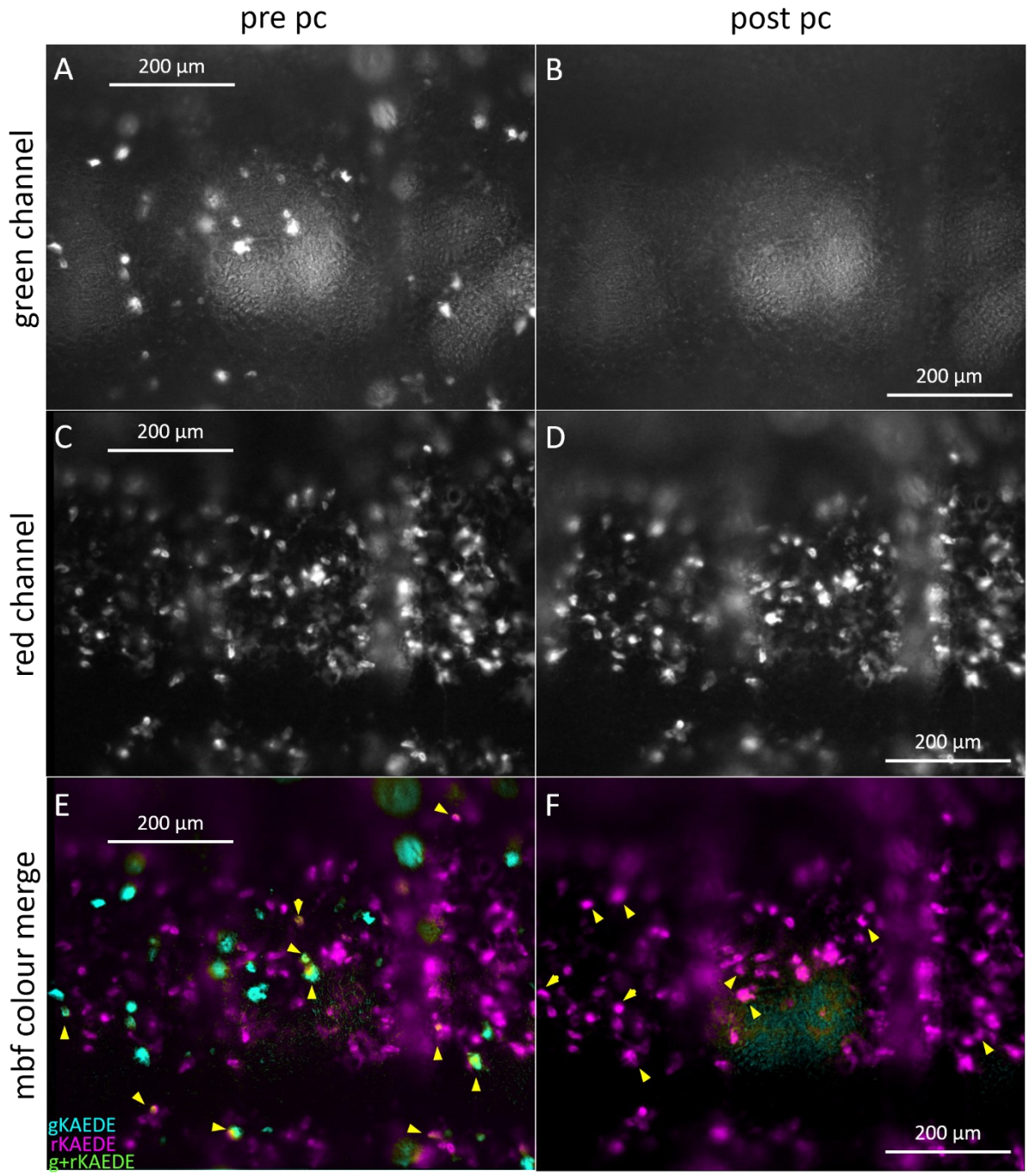

Figure 23: KAEDE transgenic animal a02 7days post pc. *(A,C) segment of cells after being photoconverted 7 days ago. Some cells replenish gKAEDE. (E) colour merge using the mbf plugin for ImageJ (Tony Collins, 2013) gKAEDE= Cyan, rKAEDE=Magenta, g+rKAEDE (overlap)= Green. Cells which replenished gKAEDE while still containing rKAEDE are indicated by yellow arrow heads. Some cells show only gKAEDE and no rKAEDE signal (cyan).(B) Same segment post photoconversion shows no gKAEDE. (D) photoconverted cells from (A) are now visible in the red channel in addition to the cells from (C). (F) Colour merge shows all cells have been pc, newly photoconverted cells are indicated by yellow arrowheads.*

To investigate the stability of the fluorophore and the behaviour of cells illuminated with UV light, a patch of cells was reimaged 7 days after photoconversion (**Figure 23A,C,E**). The observation of this patch in the red channel (**Figure 23C**) shows a clear rKAEDE signal above background concluding a sufficient stability of the cleaved fluorophore within the cytosol. Observations in the green channel (**Figure 23A**) reveal that some cells continue to express gKAEDE (red arrows **Figure 23E**). In addition to photoconverted cells replenishing gKAEDE, cells which were not expressing gKAEDE upon photoconversion are now present as visible, cyan cells in (**Figure 23E**). To proof that the cells expressing gKAEDE (**Figure 23A,E**) are not artefacts, the same patch of cells was photoconverted again (**Figure 23B,D,F**). All cells visible before this photoconversion (**Figure 23A**) were not visible in the green channel afterwards (**Figure 23B**) but appear in the red channel (**Figure 23D**). These cells stand out with an especially bright signal in the red channel (**Figure 23D**) and are pointed out with yellow arrowheads in (**Figure 23F**).

These findings suggest that cells not replenishing their gKAEDE are cells in which the promoter is no longer active. The expressed gKAEDE remains stable for a considerable long time beyond the suppression of the promoter and after photoconversion as rKAEDE.

3.2.1. Pretty but disruptive autofluorescence of *Nematostella vectensis*

Bright fluorescence signals in green and red which could be mistaken for KAEDE signal come from endogenous fluorescent proteins of *Nematostella vectensis*. So far seven putative fluorescent proteins have been predicted in the *Nematostella vectensis* genome and investigated. The fact that these proteins are related to the widely used type of fluorescent proteins like GFP, RPF, mCherry etc. and therefore heavily overlap in their spectral properties, denies the possibility of using certain filters to distinguish KAEDE from autofluorescent signals (Ikmi and Gibson, 2010). The presence of autofluorescence can vary heavily between individual animals and shows predominantly green and red emission. Typical sources of strong autofluorescence in red and green are the in the inner pharynx region (**Figure 24A,C**) and a striped pattern along the tentacles in green autofluorescence (**Figure 24B**). Some animals also show fluorescent clusters of various shapes and size along the body wall (**Figure 24D,F**) or strong green autofluorescence along the mesenteries (**Figure 24E**). The position and appearance of the autofluorescence might be different enough from the KAEDE expressing cell population to perform experiments and distinguish KAEDE signals (red arrowheads **Figure**

24C) from autofluorescence. Especially around the pharynx region the autofluorescence will present a troubling amount of background often outshining the KAEDE signal (**Figure 31**).

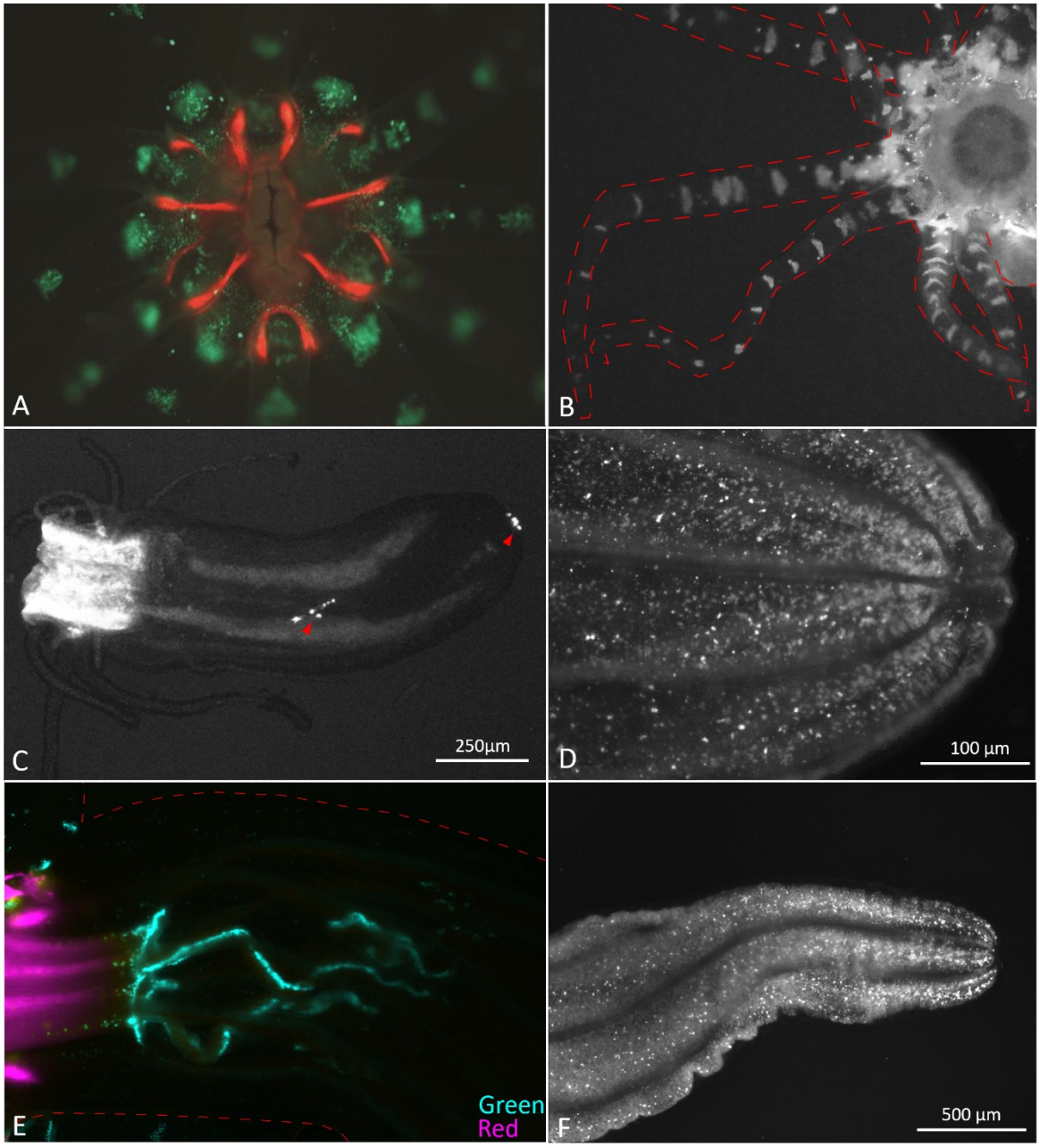

Figure 24: Autofluorescence of Nematostella vectensis KAEDE transgenic animals. *(A) green and red autofluorescence in of the pharynx region of a wild type animal. (B) This wild type shows abundant autofluorescence in the oral region as well as a well a typical stripe pattern along the tentacles. (C) a07f in red channel showing strong autofluorescence in the regenerating pharynx and small cell clusters with rKAEDE signal pointed out by red arrowhead. (E) Colour merge of Green and Red autofluorescence signals of a07. Red signal predominantly in Pharynx region while in green signals originate from mesenteries. (D,F) a16 and a15 respectively show autofluorescence along the body column in ununiform granules.(A) Image courtesy of Adrien Demilly*

3.2.2. *In vivo* imaging reveals neuronal cell types and features

The different transgenic animals are showing mosaic patterns with a variety of cell types. The most abundant cell type over all the selected animals were nematocytes distributed along the outer body column and tentacles (**Figure 25B and 14C4,D4,E4**). The strongly elongated shape with a short sensilla as well as the capsule within these cells are a clear indicator that they are nematocytes (Technau et al., 2015). Activity of the *NvSoxC* promoter in nematocytes matches the single cell data (**Figure 6**). As observed in the polyp stage of *NvSoxC*.prom.::KAEDE transgenic animals (**Figure 20**) there are cells displaying the KAEDE signal, which do not resemble the typical nematocyte shape but possess long cell extrusions which span up to 0.86 mm, more than a quarter of the body circumference (**Figure 25A, red arrowhead**). The single cell transcriptome data suggests an expression of the *NvSoxC* in neurons and their precursors which implies that these cells might be neuronal (**Figure 6**). The protrusions can be assigned different shapes like circling back to the cell of origin with ramifications (**Figure 25C1,F1,G1**) or in a strongly directional manner (**Figure 25A,C,D,E,H**). Especially later seemed to either run in pairs or have a certain inner structure which causes this appearance (**Figure 25C2,E2,H2**). Two of these protrusions seem to end at or be in contact with protrusions of other cells (**Figure 23C3**)

Another noticeable different cell type appears in a round shape with a dim structure in the middle (**Figure 25F5,G5**). These cells might be immature nematocytes.

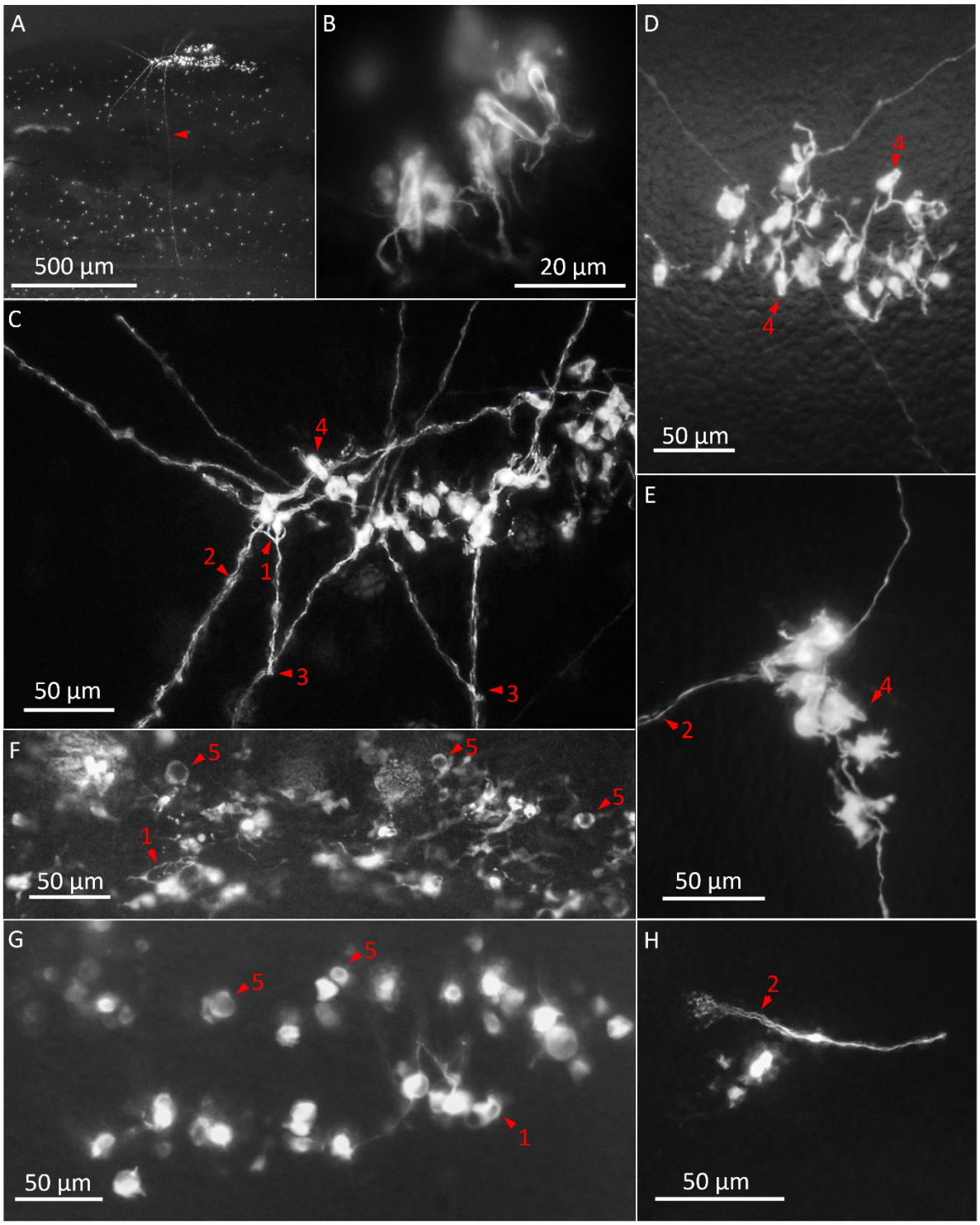

Figure 25. morphology of KAEDE expressing cells in NvSoxC.prom.::KAEDE transgenic animal a06 (A-E,G,H) and a02 (F). *(A) long cell protrusion spanning roughly on quarter of the circumference of the juvenile (red arrowhead) animal a06. (B) typical nematocytes. (C,D,E,G,H) high magnifications of a06. (F) high magnification image of animal a02. 1: small curvy cell protrusions; 2: Long rather straight cell protrusions seemingly running in pairs; 3: possible contact of cell protrusions; 4: nematocytes; 5: round shaped cells with dim centre (round-hole-cells).*

3.2.3. Sequential pc in lineage tracing

To establish a method for lineage tracing, cells possessing a gKAEDE signal were tested for their capabilities of restoring this signal upon photoconversion. Three transgenic animals with gKAEDE signals in nematocytes were chosen for this experiment. From the three animals a12 (**Figure 26A-C**), a08f (**Figure 26D-F**), a07 (**Figure 26G-I**), certain areas containing nematocytes with gKAEDE signal were photoconverted and observed again after 7 days. As observable in the amputation experiments this time span is long enough to build up a genuine nematocyte population (**Figure 33**).

Towards the centre of the photoconverted cell population of animal a12 (**blue circle Figure 26A-C**) it is clearly visible that cells did not restore the gKAEDE signal, which would result in a visible colocalization in **Figure 26C**. There are on the other hand some cells visible, which have a gKAEDE signal, suggesting that these cells arose after the photoconversion. Towards the edges of the photoconverted area some cells seem to show colocalization of gKAEDE and rKAEDE, although this is most likely caused by the decrease in intensity of the UV light towards these edges and therefore only a partial photoconversion of the fluorophore.

Animal a08f (**Figure 26D-F**) again shows a mix of cells with either gKAEDE or rKAEDE signal but contrary to a12, some of the cells seem to give a colocalised signal therefore representing photoconverted cells, which have restored the gKAEDE signal.

The cell population observed in animal a07 (**Figure 26G-I**) does not show new cells with gKAEDE signal and a few cells with colocalised signals.

These findings show that nematocytes mostly do not express *NvSoxC* but merely inherited the KAEDE fluorophore from precursor cells with an activated *NvSoxC* gene. In fact, 23,1% ±1,3% of mature nematocytes observed in **Figure 26D-I** were found to restore their gKAEDE signal.

Some of the nematocytes seem to be capable of restoring gKAEDE and therefore need to have an active *NvSoxC* promoter. Since only a few nematocytes show this property it might indicate that these cells were photoconverted in a state where they were not completely matured and therefore had a still active *NvSoxC* promoter.

An even more striking observation, however, is the emergence of cells with a clean gKAEDE signal which means that these cells did either not express gKAEDE prior to the photoconversion or that these cells were located outside photoconverted area.

Especially in animals a12 (**Figure 26A-C**) and a08f (**Figure 26D-F**) where the photoconverted area was surrounded by a high number of gKAEDE signalling cells this might be a hint towards

migratory capabilities of nematocytes and the visible but unidentified other cells shown in **Figure 26F**. The fact that the photoconverted cell population in a07 (**Figure 26G-I**) was a segregated pool of transgenic cells and does not show any new gKAEDE signal would support this hypothesis.

The assumption that these gKAEDE signals come from newly emerged cells would indicate that the *NvSoxC* is not expressed in true stem cells but rather gets activated along the way towards the determined cell fate.

This hypothesis is supported by observations in a07 (**Figure 27A-C**) and a06f (**Figure 27G-I**) where another patch of segregated cells was photoconverted and showed the emerging cells with a clean gKAEDE signal that could not be identified more closely.

In general observations of a07 (**Figure 27A-C**) and a06f (**Figure 27D-F**) show that the neuron-like cells observed in these transgenic animals seem to be more likely to have an active expression of the *NvSoxC* gene therefore showing strong restoration of gKAEDE signal. This observation could also be made in another cell population of a06f (**Figure 27G-I**) where the same cell population seen in (**Figure 25F,G**) showed that some of the cells with the round-hole phenotype (**Figure 25F5,G5**) restored gKAEDE as well.

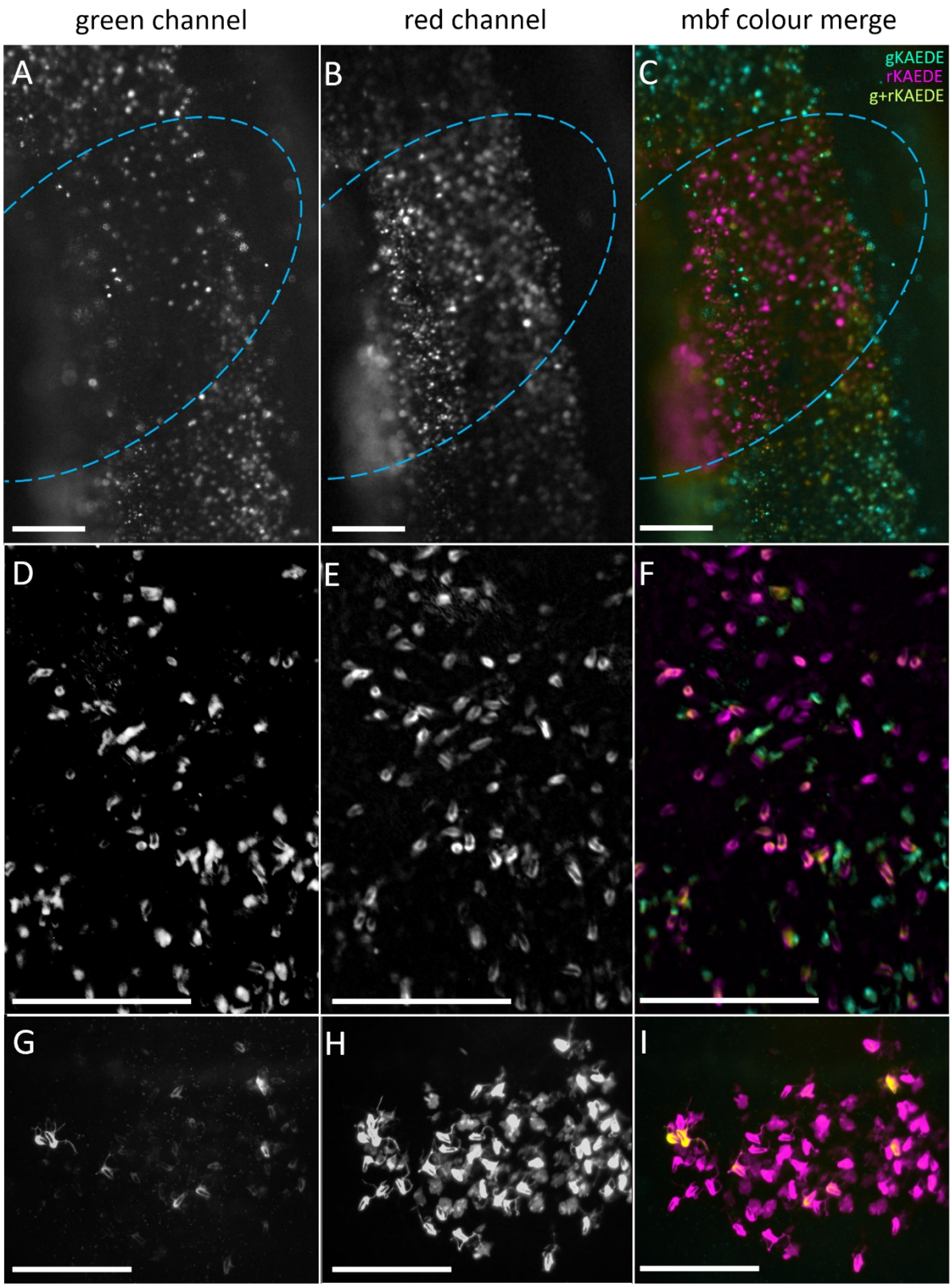

Figure 26: Photoconversion experiment of animals a12 (A-C), a08f (D-F) and a07 (G-I) imaged 7 days after photoconversion. *A-C: the blue circle shows the area that was photoconverted 7 days prior to taking the pictures. Upon the rKAEDE signals some cells show new gKAEDE signal. D-F: A high number of cells show gKAEDE signal. Some nematocytes partially restored the gKAEDE signal. Both converted areas are surrounded by non-converted cells. G-I: This patch of photoconverted cells secluded of other cells. Only a few nematocytes restored the gKAEDE signal but no cells show new gKAEDE signal. Scale Bars, 250µm*

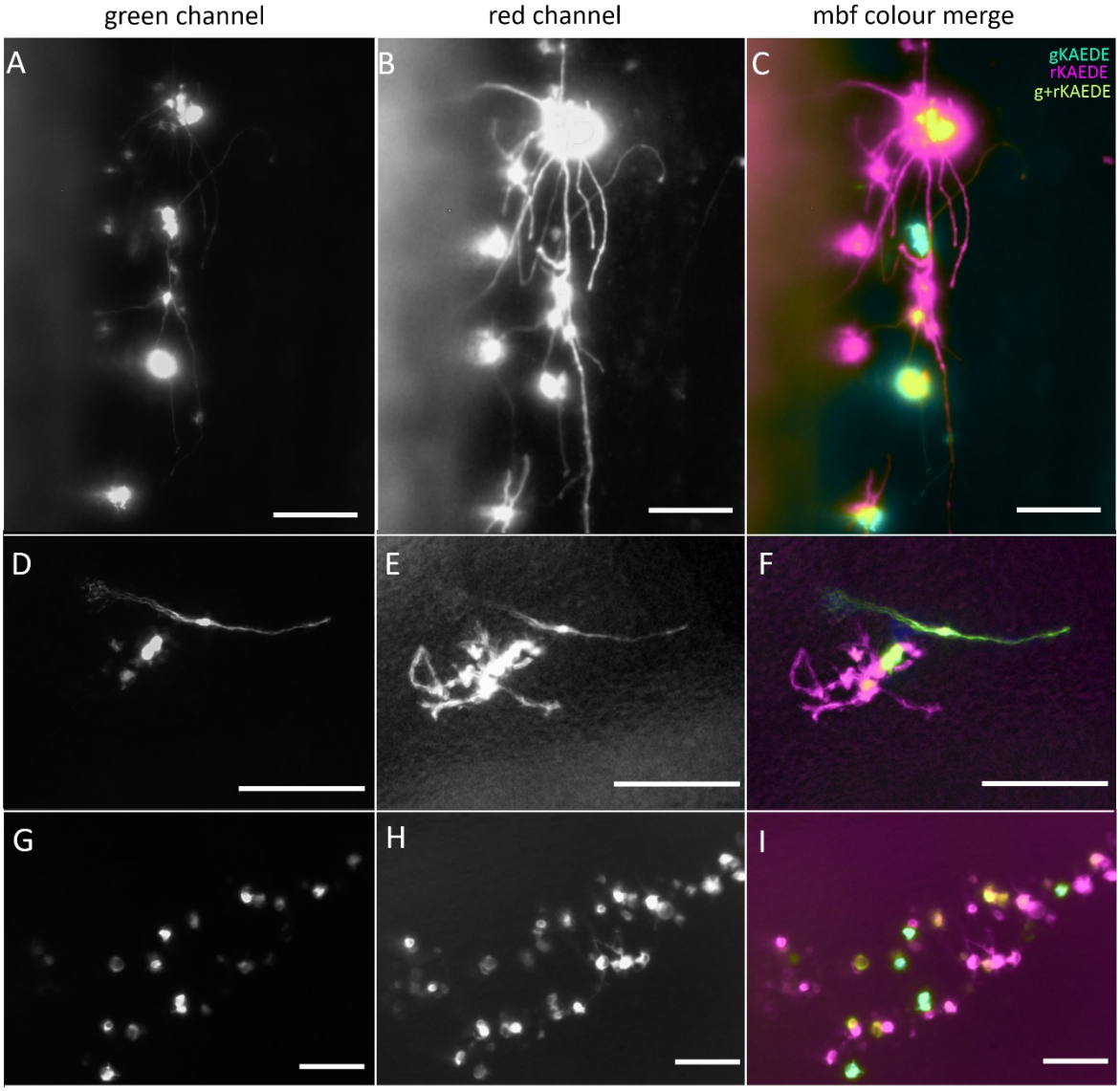

Figure 27: Photoconversion experiment of neuron-like and round-hole cells imaged 7 days post photoconversion animals a07 (A-F) and a06f (D-I). *A-F: A higher portion of this secluded neuron like cells show strong gKAEDE activity. (A-C) one cell is expressing the gKAEDE fluorophore genuinely. (G-I) This secluded cell patch shows neuron like and round-hole cells. Some cells restored the gKAEDE signal and a high number of cells show native gKAEDE expression. Scale bars, 50 µm*

The amputated foot of animal a12 further referred to as a12f shows an extremely high activity of the *NvSoxC* promoter in the regenerating head 7 days post amputation (pa). Most of these cells can be identified as nematocytes.

3.2.4. Cell migration, and homeostasis during regeneration

A preliminary approach to investigate the migration behaviour of the *NvSoxC* expressing cells was to photoconvert a small patch of cells and after a certain time observe if cells with a rKAEDE signal could be found outside the initially photoconverted area.

In addition, a second experiment was designed involving the photoconversion of a proximal part of a tentacle with a certain distance to an amputation site. After regeneration of the tentacle it was monitored if cells with a rKAEDE signal could be found outside the formerly converted area or even in the newly built tentacle (**Figure 28**).

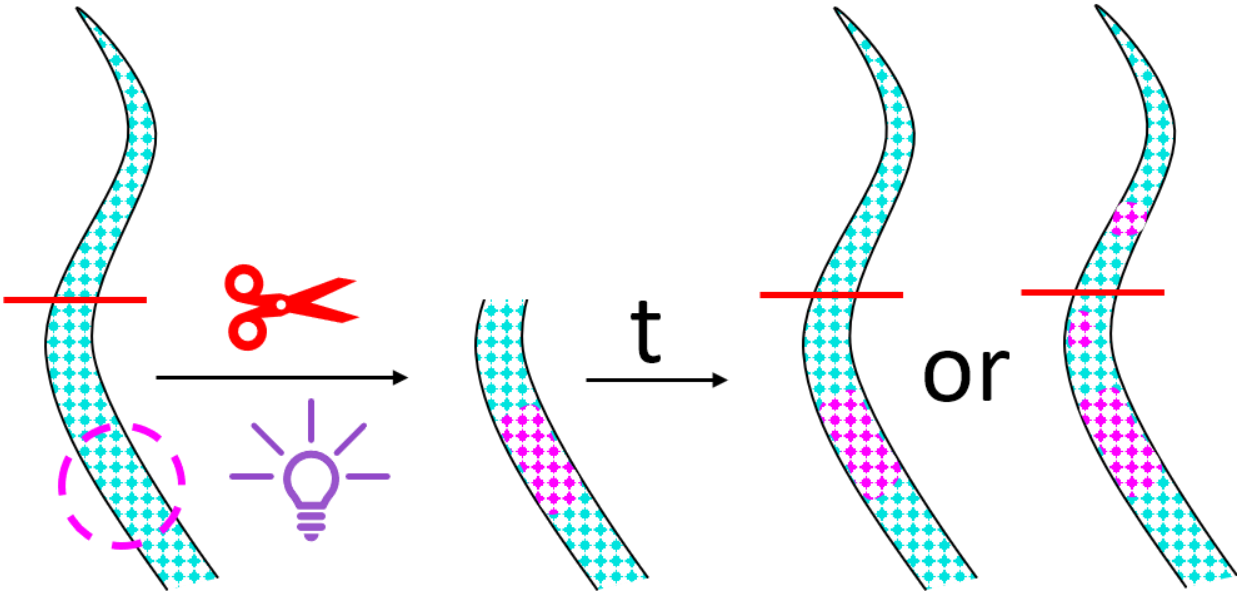

Figure 28: Initial design for amputation experiments. *A tentacle with transgenic cells is photoconverted at a certain spot proximal to the amputation site. After regeneration of the Tentacle cells which migrated will be visible outside the photoconverted site.*

Unfortunately, a regular occurring problem was the difficulty of orienting the animal in a position which would recover the same image section as it was taken at the start of the experiments. Mainly the soft and motile nature of the animal exacerbated this task resulting in repeated correction of the animal's position causing injuries or triggering an aversive contraction of the animal despite the relaxation with $MgCl_2$. Efforts to keep the animals in a constantly relaxed state mounted as described turned out to be deficient or lethal.

The experimental design was adapted so that the foot of the animals would be amputated with a length of 3-4mm, then photoconversion of the desired area was executed and reimaged after 7 days (**Figure 29**). Changing the orientation of a very small animal was more manageable

and a new head was formed within 7 days and regenerate to a functional anatomical state containing tentacles within 14 days (**Figure 31**).

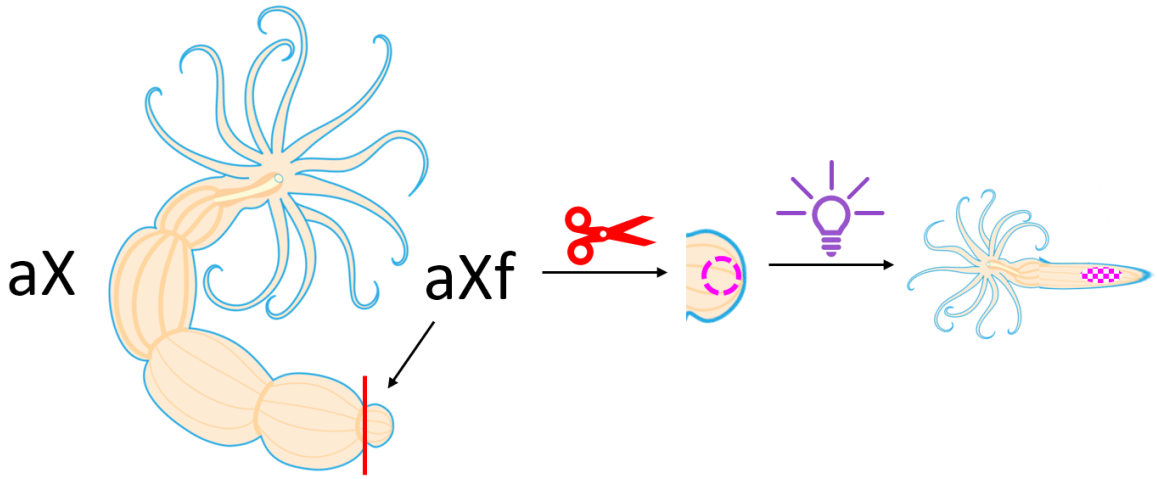

Figure 29: Adapted experimental approach. *3-4mm short pieces were amputated from the foot of the animals. These pieces got the postfix f. When this experiment was repeated with another amputate of the adult animal another postfix with a running number was added. These amputates underwent proportion regulation of shape upon regeneration distorting the photoconverted area. (Nematostella drawing by Emmanuel Haillot).*

Several observations using this approach showed an axial elongation and radial contraction of the remaining fragment (proportion regulation of shape). Due to this change in body shape no reliable statements could be made about the cell positions and therefore their migration behaviour. This change in body shape can be seen in the foot of animal a12 henceforth referred to as a12f (**Figure 31**).

The effect of this change in body shape on a small patch of cells can be seen in the amputated foot of animal a07 furthermore referred to as a07f in (**Figure 30A,B oral is right**). Over the period of 7 days this cluster of cells elongated in shape. This elongation could indicate migration behaviour of these cells. The fact that the whole amputates underwent proportion regulation of shape is also a possible explanation for this observation. Not just morphological changes can influence the appearance of such a cell cluster but also the short-term contractions of the animal. The pictures (**Figure 30 C, D**) were only taken several seconds apart while the animal was contracting. Nevertheless, the appearance of the cell cluster as well as the orientation of some of the cells, which are manly nematocytes, changed significantly.

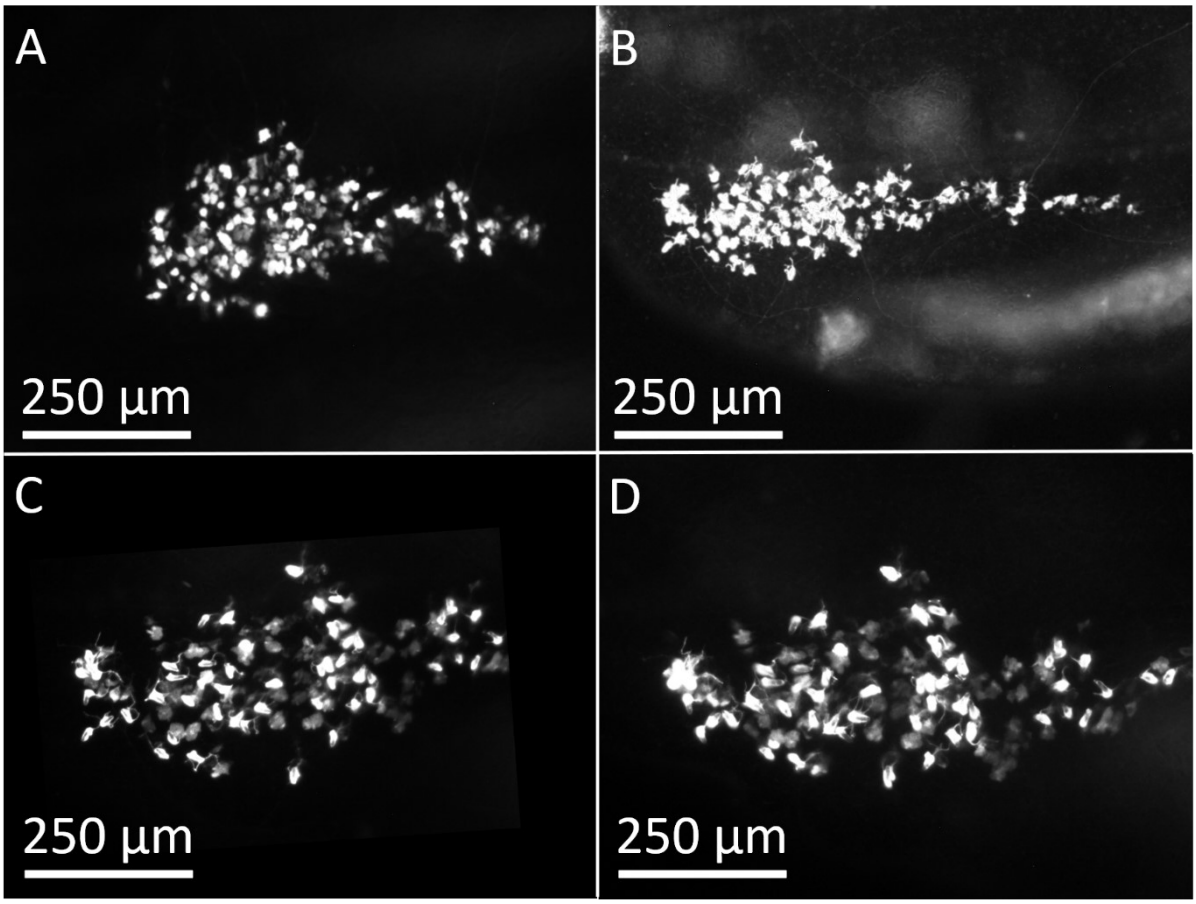

Figure 30: a07f over different time courses. (A,B) *The same path of transgenic cells after 7 days**. (C,D)** the same patch of cells after a few seconds. All images taken in red channel*

The animal a16f possesses a population of transgenic cells in an elongated patch along the body axis. This patch of cells mainly consists of nematocytes. The patch was photoconverted after amputation (**Figure 31A**). Over the course of 14 days post amputation this patch elongates together with the whole animal. Overall, the diameter of the animal diminished by half.

Unfortunately, the red autofluorescence of the newly built pharynx particularly impairs the observations of cells near the oral end (**Figure 31B,C**). gKAEDE signal which arose in some cells on day 7 and 14 were photoconverted again (**Figure 31B,C**).

The observation of a16f shows that over the timespan of regeneration the overall signal of the cell population is decreasing. This decrease in signal is on one hand a matter of decreasing cell number as it can be seen in the upper half of the cell patch but also a matter of degradation of rKAEDE as cells in the lower half of the patch seem to dim out over time.

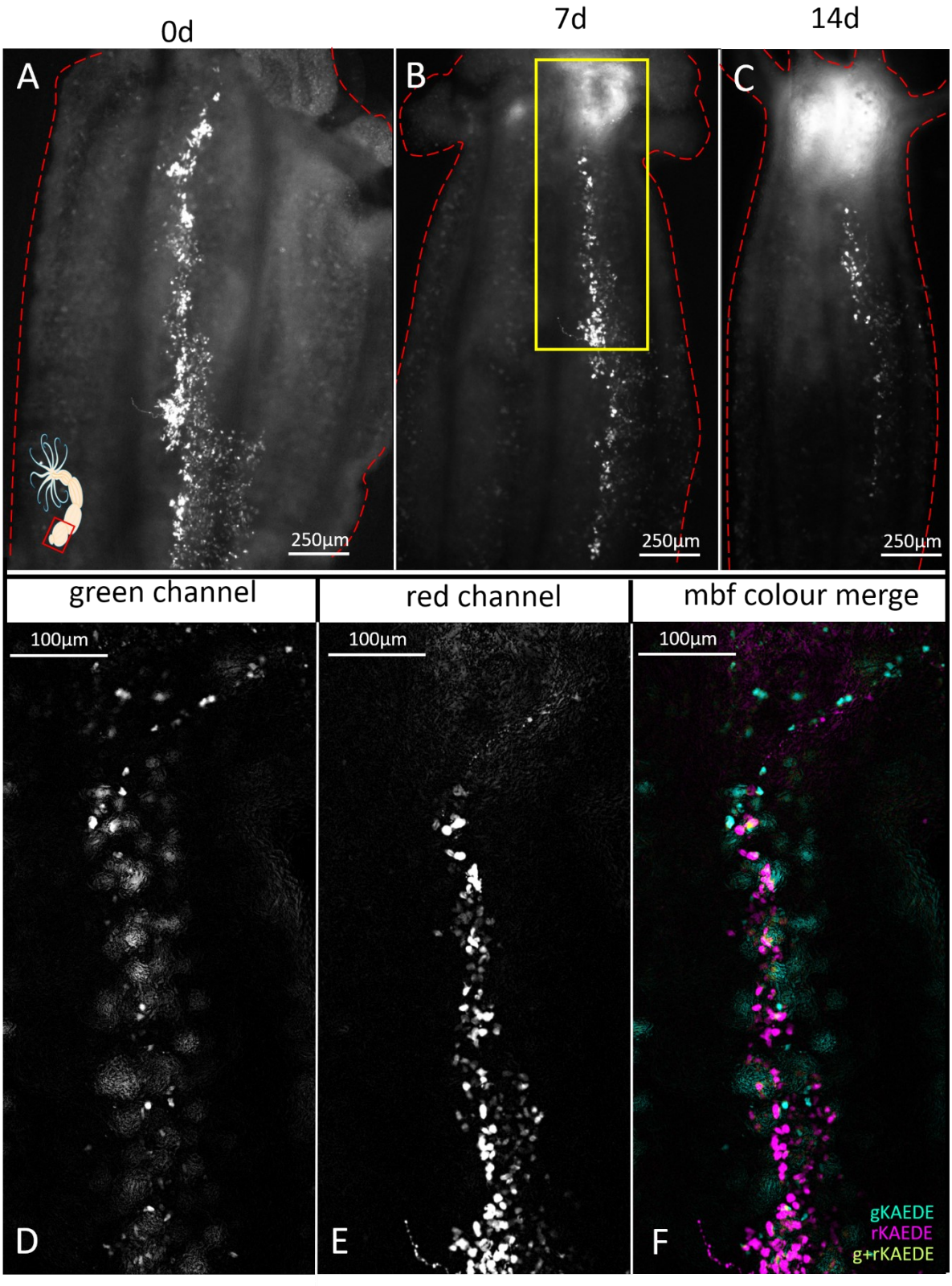

Figure 31: Amputation experiment of a16. (A-C) *The amputate a16f 0d, 7d and 14d post amputation in the red channel. Yellow box indicates detail of* ***(D,E,F):*** *A photoconverted cell patch after 7 days of regeneration. Clean gKAEDE signalling cells are mainly visible at oral side of the patch, at the regeneration site. Autofluorescence background was removed. Oral is up.*

Closer observations were only made 7d post amputation since the autofluorescence of the pharynx grew too bright after 14 days. Higher magnification of the upper part of the pharynx

show that most of the cells showing a gKAEDE signal after 7 days are located at the regenerating pharynx (**Figure 31D-F**). The morphology of these cells cannot be assigned to a neuronal or nematocyte category. These images also show that cells exhibiting an rKAEDE signal are not migrating towards the regenerating pharynx (**Figure 31E**).

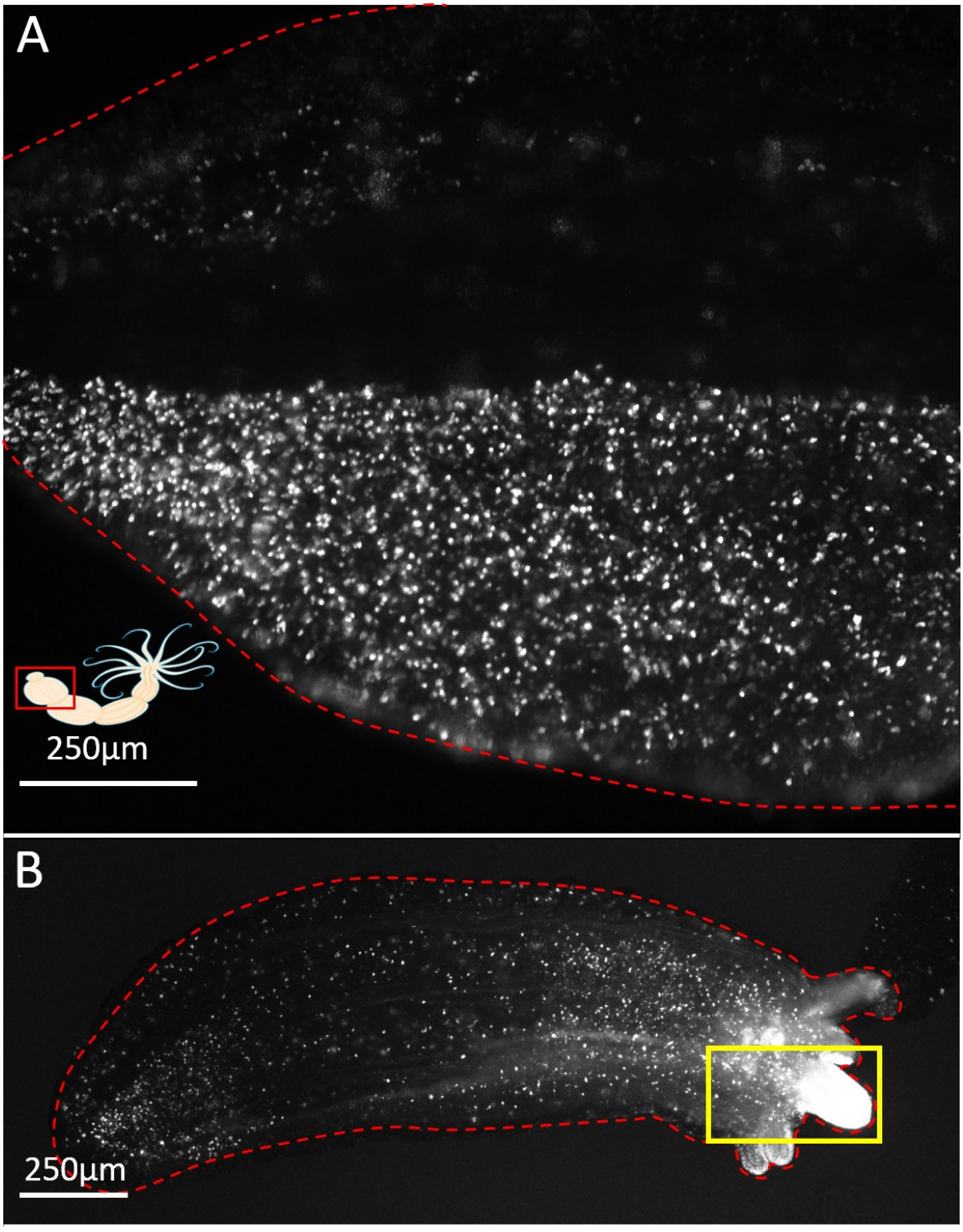

Figure 32: Animal a12 and a12f. (A) *Animal a12 shows a very sharp-edged stripe mosaic transgenesis. (B) Animal a12f 7 days post amputation. The tentacles regenerated at the transgenic stripe show higher gKAEDE signal then those to the side. Yellow box indicates location of detail (**Figure 33**) (A, B) taken in green channel. Oral is right.*

Animal a12 (**Figure 32A**) shows a particularly sharp-edged stripe of transgenic cells. The amputation experiment performed with this animal (**Figure 32B**) reflects this expression by displaying a bright signal in one tentacle and a weaker signal in adjacent ones. This bright signal originates mainly from nematocytes, which are newly differentiated in the tentacles as they are regenerating 7 days post amputation. Another interesting observation is that although the animal a12 shows a sharp-edged transgenic stripe, in the amputate a12f this sharp edge is not recognisable anymore.

A closer look at a12f shows mature nematocytes with their specific morphology including the nematocyst. These cells display a rKAEDE signal from the initial photoconversion on the day of amputation (**Figure 33B**). Besides the strong signal of the new differentiating nematocytes at the tentacle tip, there are a few particularly bright cells which do not have the round or elongated morphology of nematocytes. These cells have no rKAEDE signal indicating that they arose after the amputation (**Figure 33A,C** yellow arrowheads).

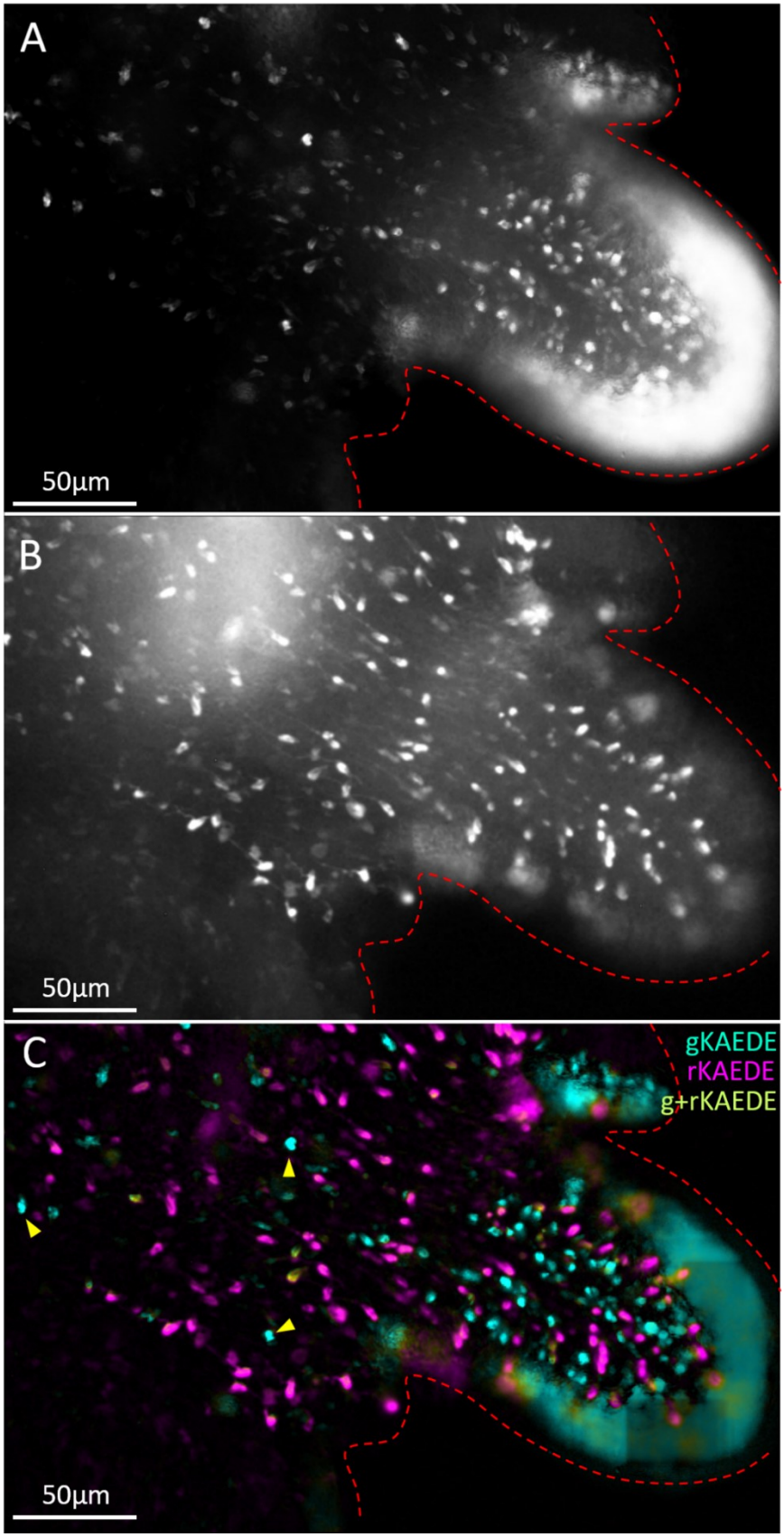

Figure 33: Animal a12f 7 days post amputation. *(A) Green Channel. The majority of the gKAEDE signalling cells is located in the regenerating tentacle and emitted by nematocytes. Some cells further aboral show a distinct morphology different from nematovcytes (yellow arrowheads (C). (B) Red Channel. Some nematocytes which were photoconverted before the amputation can be found in the regenerating tentacle. (C) Overlay.*

4. Discussion

4.1. The KAEDE-system in *Nematostella vectensis*

4.1.1. High transgenic efficiency of the transgenesis

Animals injected with the *NvSoxC* constructs show an overall lower survival and expression rate than previously reported (Renfer et al., 2010). The biggest divergence is caused by the low survival rate after 24h which might be due to the naïve nature of the experimenter rather than a cytotoxic effect of the injection. The success rates of the KAEDE construct and the *mCherry* construct under the *NvSoxC* promoter show similar survival and expression rates. Neglecting the high mortality rate of the zygotes and setting the expression rate in relation to the survival rate the *NvSoxC*ptom.::KAEDE construct and the *NvSoxC*prom.::mCherry construct show a rate of expression of 82% and 44% after 24h and 29% each after 9 days respectively. This is comparable with the 40% after 24h and 22% after 9 days of the *NvMyHC*prom::mCherry construct. Therefore, one could argue, that a more sophisticated execution of the injection would lead to overall similar success rate as conducted by Renfer et al.

Table 5: Comparison of efficiency of I-SceI-mediated transgenesis of NvSoxCprom.::KAEDE, NvSoxCprom.::mCherry and NvMyHCprom::mCherry (Renfer et al., 2010)

	NvSoxCprom.::KAEDE	NvSoxCprom.::mCherry	NvMyHCprom.::mCherry
	%	%	%
24-h survival of injected	17	26	90
24-h survivors expressing	82	44	40
9-d survival of injected	11	10	33
9-d survivors expressing	29	29	22

4.1.2. Handling of transgenic animals and visibility of KAEDE

Animals with transgenic cells expressing the KAEDE construct were clearly distinguishable from the background apart of the sites of autofluorescence (pharynx and tentacles). The bright gKAEDE signal could be observed with standard filter cubes suitable for detection of Alexa488 or GFP. For rKAEDE detection the mCherry filter cube was appropriate. The UV photoconversion could be performed with a standard DAPI filter cube. These setups are common in labs working with fluorophores and therefore the establishment of the KAEDE system does not compel to invest in further equipment. Nevertheless, specialized filters could enhance the observation of the decaying rKAEDE signal since mCherry and rKAEDE show a distinct difference in emission and excitation spectra (Dittrich et al., 2005).

The photo-convertibility of KAEDE was accidentally discovered, when cells expressing KAEDE were left in daylight (Dittrich et al., 2005). Unfortunately, the spawning process requires the animals to be exposed to bright light for several hours. The impact of this induction on the stability of gKAEDE in the adult animals is yet to be determined.

Nevertheless, to avoid unnecessary exposure of the fluorophore, light exposures should be carefully planned and short (**Figure 21**). The KAEDE tetrameric protein diffuses freely in the cytoplasm of the cell advancing into long axons spanning at least 0.86 mm and thin sensilla of nematocytes. This observation matches previous findings in a rat hippocampal primary culture where the rKAEDE diffused into axons within 3 min (Dittrich et al., 2005).

In general, the autofluorescence of *Nematostella* causes heavy restrictions on the possible observations, especially regeneration experiments involving the observation of the pharynx or tentacles.

The body plan of *Nematostella* displays only a few morphological landmarks (pharynx, siphonoglyph) or radially symmetrically arranged structures (tentacles, mesentery) makes it more difficult to follow the possible migration of cells. The morphogenetic shape changes during regeneration add to this problem as mentioned in by other groups (Havrilak et al., 2019). Therefore, orienting animals in reproducible positions after several days with or without amputation to compare images taken in certain intervals poses a major challenge. Attempts to keep the animals in a defined position over several days with the help of $MgCl_2$ were lethal for the animals.

Not only the reproducibility over a longer time period but also capturing images in a short interval (few seconds) is a problem. *Nematostella* shows strong avoidance behaviour to short

wavelength light. In unrelaxed animals this behaviour is already visible when exposed to the excitation light for GFP. The aversive reaction manifests as a contraction of the whole animal with retraction of the tentacles. Exposure to DAPI excitation (UV) light triggered this behaviour even quicker. Relaxing animals with $MgCl_2$ results in a slower and delayed reaction. Unfortunately, high magnifications still revealed residual movements of the animal especially after photoconversion with UV light which leads to blurry images and as a main issue non congruent multi-channel images which must be aligned in post processing.

This study shows, that the gKAEDE fluorophore is stable enough to be visible in cells after the expression was repressed. The photoconverted rKAEDE is stable enough to be well detected after 7 days of photoconversion without noticeable drop in brightness. After 14 days a reduction in brightness is visible. The rKAEDE is stable enough to be detected after 3-4 weeks at lower levels which means that photoconversion experiments cannot be performed in the same animals in short intervals. This observation was one of the reasons to change the layout of the experiments to the photoconversion taking place after the amputation so the straylight of the microscope during photoconversion in one part of the animal would not photoconvert the remaining cells. This way, several amputation experiments of an aboral amputate could be performed right after the aboral foot of the original animal was regenerated.

4.1.3. The observations meet the expectation

Anticipated observations made with the KAEDE system in combination with a transcription factor X could be as follows.

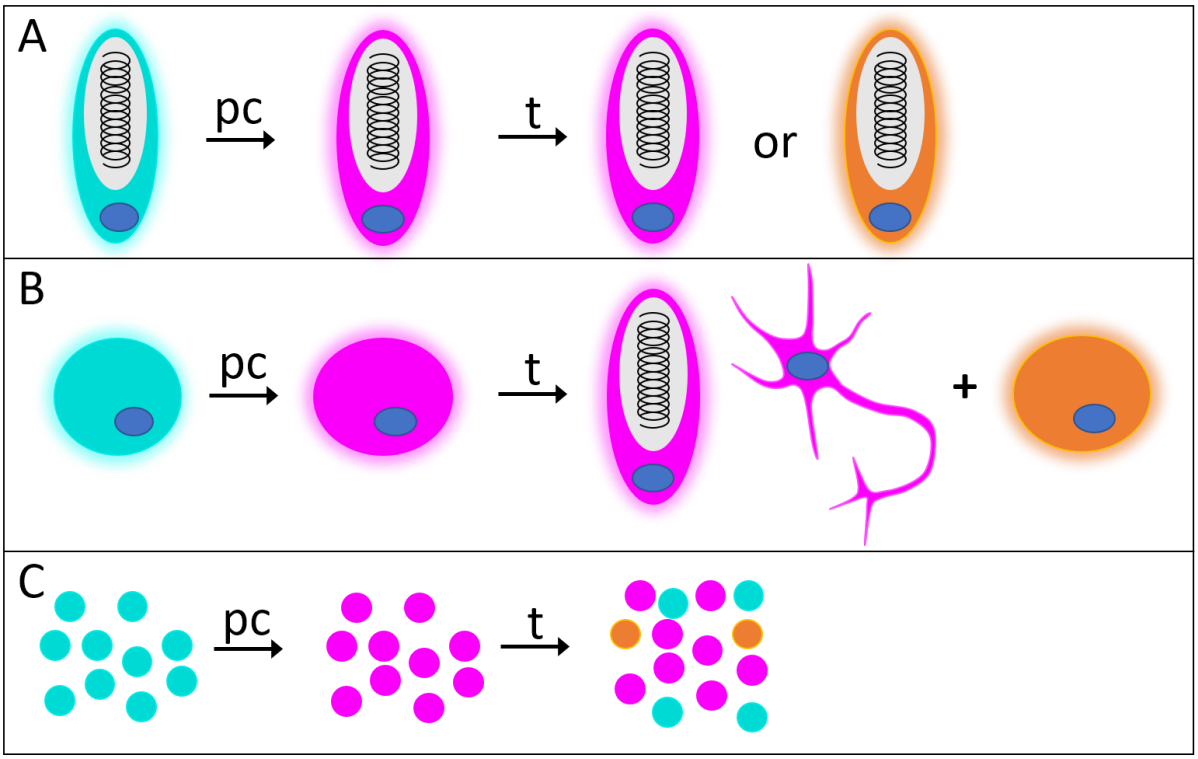

Figure 34: Possible scenarios for observations made with the KAEDE fluorophore of photoconverted cells after a given time. gKAEDE=cyan, rKAEDE=magenta, overlay=orange. *(A) Nematocytes might not restore the gKAEDE signal indicating an inactive promoter while on the other hand restoring the signal will lead to an overlay and indicate a promoter which is still active. (B) Photoconversion of potential precursor cells will give offspring to cells inherited with the photoconverted fluorophore. Since these cells change their identity the promoter will be turned off and the cells maintain the a clean rKAEDE signal. The mother cell of this asymmetric division will continue to express the TF and therefore also express the gKAEDE signal leading to an colocalized signal. (C) When a whole cell population is photoconverted a mix of cells capable or not capable of restoring the gKAEDE signal will be visible. In addition, some cells might give a clean gKEDE signal indicating that these cells either started expressing the TF after the photoconversion or that they migrated into this area.*

Differentiated cells like nematocytes giving a gKAEDE signal might, after photoconversion, not restore the gKAEDE indicating that the former gKAEDE was inherited from precursor cells with an active TF X, or might be capable of restoring the signal suggesting activity of the TF questioning the assumed stem cell significance (**Figure 34A**). The observations made using the *SoxC* promoter show that most nematocytes maintain a rKAEDE signal whereas a few nematocytes which are still expressing *SoxC* show a rKAEDE as well as a gKAEDE signal (**Figure 26**).

If a gKAEDE signalling cell has stem cell features, after photoconversion, differentiated offspring originating from an asymmetric cell division of this cell will be inherited the photoconverted rKAEDE, while the other daughter cell regains the gKAEDE signal and

therefore indicates an activity of TF X in the stem cell or progenitor cell (**Figure 34B**). The performed experiments were not be able to support this hypothesis due to a lack of identification of possible stem cell or progenitors as well as technical constrains not allowing the photoconversion of single cells.

The same experiments done with whole cell populations might show a mixture of cells not restoring and restoring the gKAEDE fluorophore upon photoconversion but also new cells showing a clean gKAEDE signal (**Figure 34C**). The detection of clean gKAEDE cells might indicate that these cells arose after the photoconversion from KAEDE-negative precursors or from migratory cells settling in the photoconverted area. Experiments performed with the *NvSoxC* promoter resemble these observations (**Figure 23**).

These observations show the potential of the KAEDE system for lineage tracing.

4.2. The neuronal transcription factor NvSoxC

4.2.1. Lineage tracing with temporal information of KAEDE

The photoconversion experiments allow to postulate a time course of neuronal type cells developing in *Nematostella* under the control of the *NvSoxC* promoter.

The fact that a few cells show a native gKAEDE signal implies that these neuronal precursor cells (NPC) derive from a *NvSoxC* negative stem cell lineage. From these NPCs nematocytes develop. These juvenile nematocytes are still expressing *NvSoxC* and therefore show an overlay signal of the photoconverted rKAEDE as well as a restored gKAEDE. Mature nematocytes no longer express *NvSoxC*. Neurons deriving from the NPC show a similar behaviour. The observation of neurons still expressing *NvSoxC* although they show long axon structures makes it seem that a placement of these cells on a temporal axis between mature neurons and NPCs might be daring. But observations in *Hydra* support a developing nature of *SoxC* expressing cells (Siebert et al., 2019).

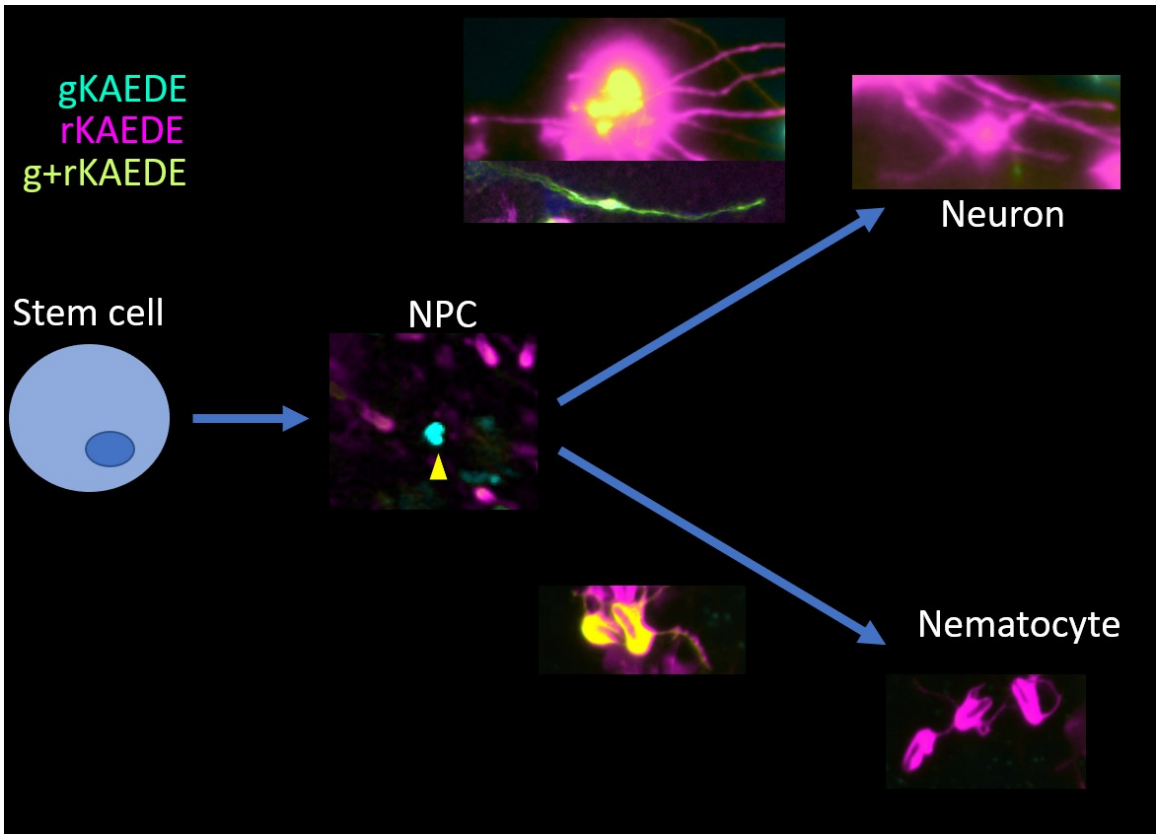

Figure 35: Schematic reconstruction of lineage tracing in neuronal cells. *The images taken during these photoconversion experiments can be aligned into a developmental time course from NPC to neuron or nematocytes. Immature cells would show a mixed signal whereas NPC and maturated cells would either show a native gKAEDE signal or a clean rKAEDE signal respectively. This hypothesis would require that the neurons restoring the gKAEDE signal are still immature despite their elaborated cell protrusions.*

The *Hydra* sc-data showed that *HvSoxC* was found to be almost exclusively active during the differentiation of cells undergoing nematogenesis and neurogenesis or early gland cell differentiation. *HvSoxC* cannot be found in i-cells or differentiated cells (Siebert et al., 2019). Regarding bilaterians, members of the *SoxC* genes were found to have a conserved function in neuronal differentiation in vertebrates by establishing a pan-neural protein expression. Although the existence of a *SoxC* gene has been reported in several species, not many functional or expression data were reported from nonvertebrate species. In the lophotrochozoan *Plarynereis, SoxC* is only expressed in the late stage of neurogenesis. In the ecdysozoan *Drosophila* a *SoxC* homolog exists, but it is ubiquitously expressed during early development and is apparently not involved in neurogenesis (Bergsland et al., 2006; Kerner et al., 2009).

In none of the observed animals, gland cells could be found. This might be due to the reason, that these cells are manly located in the septal filament in the endoderm which is harder to observe *in vivo* than the ectoderm is.

4.2.2. SC-data refine optical lineage tracing experiments

The latest, unpublished *Nematostella* sc-data by Cole and Steger can be used to augment the observations of the KAEDE experiments. Single cell transcriptomes of animals from different developmental stages were clustered using a nearest neighbour graph approach (Butler et al., 2018; Stuart et al., 2019). Cells with similar transcriptomes are grouped closer together than cells with less similar transcriptomes. For a better visibility the resulting tree was reduced to two dimensions using UMAP (**Figure 36A**). This type of analysis allows for pseudo time investigation in the trajectories of cell type development. Cell types with a high turnover rate like nematocytes will constantly be generated by precursor cells. Therefore, at any given moment an animal will have nematocytes at different stages of differentiation with transcriptomes that differ only slightly to each other therefore generating trajectories to mature nematocytes, which have a transcriptome less common with early nematocytes. Cells, which do not have a high turnover rate, like retractor muscle cells are less likely to show these trajectories since at a given moment there might not be differentiating muscle cells.

Another dimension of information is added by the colour clustering of the cells. These clusters are expressing a common gene set, which is not expressed in other cells. The trajectories plus the clusters taken together we can see that the nematocytes, neurons as well as gland cells seem to show a common origin as the trajectory of the neuronal-glandular progenitors connects these clusters. A subgroup of cells assigned to the neuronal-glandular progenitor cluster is expressing the proliferation marker *NvPCNA* giving a hint towards the origin of these cells from proliferative precursors.

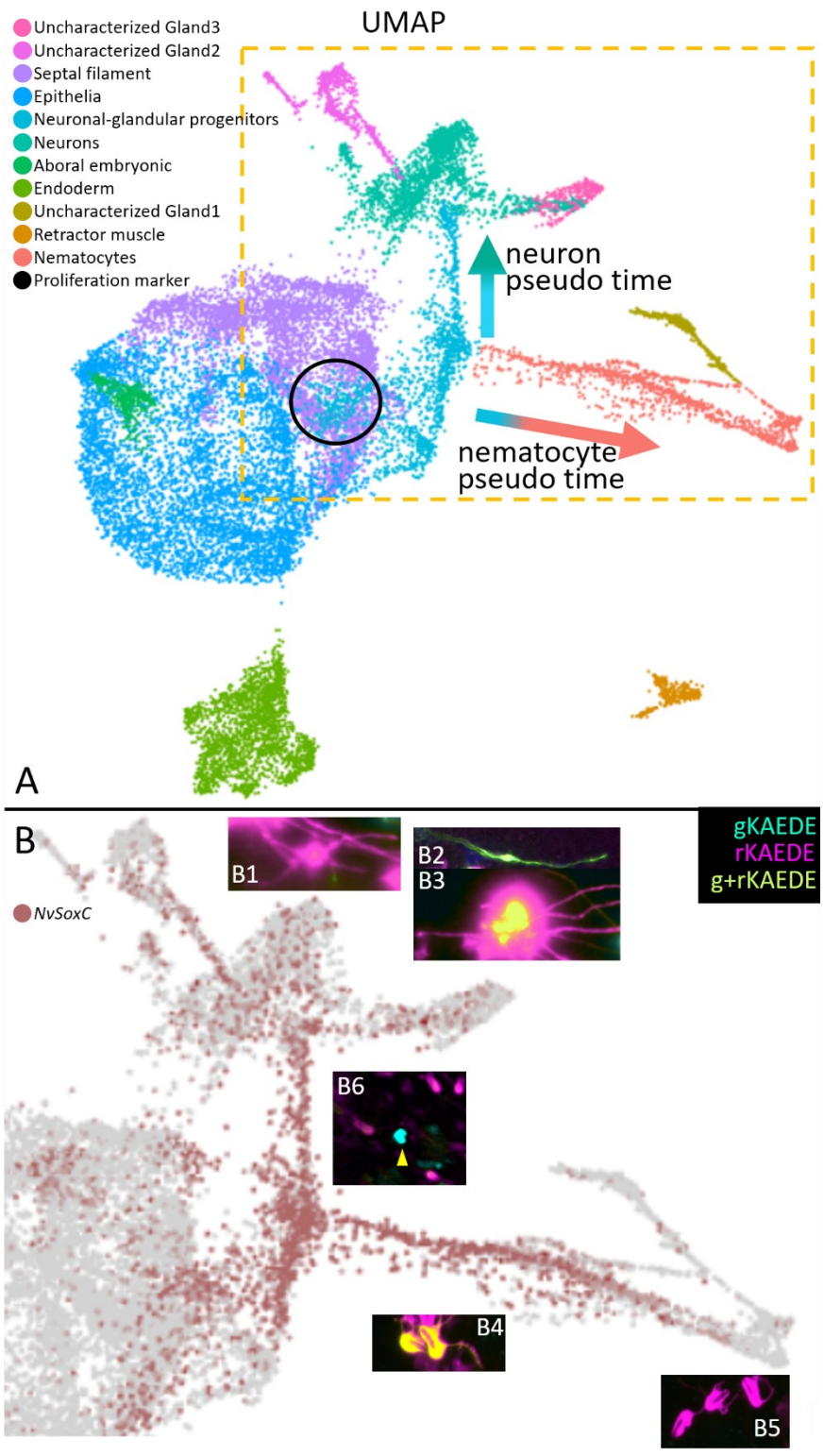

Figure 36 UMAP Single cell data. *(A) all single cell data. (B) NvSoxC expressing cells*

Viewing cells which show *NvSoxC* expression in these data, we can see, that *NvSoxC* is expressed along the trajectories towards the neuronal and gland clusters as well as the nematocyte and gland clusters (**Figure 36B**). Combining these sc-data with the findings using the KAEDE system the observed cells can be assigned to the clusters (**Figure 36B1-6**). The details in (**Figure 36B1-6**) show cells 7d post photoconversion.

The single cell data on the nematocyte trajectory shows that around 23% of nematocytes are expressing *NvSoxC*. This trajectory leads towards a population of nematocytes negative for *NvSoxC.* These two observations are consistent with my observations using the photoconversion in the KAEDE transgenic line. Nematocytes, that are restoring the gKAEDE signal therefore must have an activated Nv*SoxC* promoter (**Figure 36B4**). Thus, they can be assigned to the transition of the trajectory. On the other hand, nematocytes that do not restore the gKAEDE signal suggests a state with a repressed *NvSoxC* promoter at the distal end of the trajectory (**Figure 36B5**).

The second cell type which could be identified by their morphology are neuronal cells. In the single cell data, we can see that another trajectory is leading away from the main *NvSoxC* expressing cluster upwards towards neuronal and gland cell populations. *In vivo* neuronal cells of different phenotypes were found to either restore or not restore the gKAEDE signal (**Figure 36B1,2,3**). Looking at the neuronal population in the single cell data we see that these observations can be generally attributed to cells which are expressing *NvSoxC*, or which do not express *NvSoxC* any longer. The observed neurons expressing *NvSoxC* therefore do not exclusively have to be on the developmental trajectory since also within the cluster of differentiated neuron there are cells showing *NvSoxC* transcripts in the single cell data (**Figure 36B**). If *NvSoxC* is a marker for neuronal differentiation, an activity of this gene in differentiated neurons might imply that these cells might still have progenitor abilities like transdifferentiating upon recruitment by regenerative processes. Data show a high plasticity of the neuronal network in *Nematostella* depending on animal size changes (amputations) as well as nutritional changes (starvation) (Havrilak et al., 2019) .

The third cell type/ cell state which was found in regenerating animals (**Figure 31D-F, Figure 33C**) could not be assigned to neurons or nematocytes by morphology since they lack cell protrusions capsules. These small round cells are characterized by a clean gKAEDE signal, thus, must have turned on the transgene newly after the photoconversion form a *NvSoxC*-negative state. Therefore, it is tempting to assign these cells to a population of a yet unknown stem cell cluster.

In summary, the following hypothesis could be postulated. As the neuronal cell type developed, *NvSoxC* as a lineage specific transcription factor was responsible for the differentiation of neurons and remained active to maintain identity but still maintaining a certain degree of plasticity (Havrilak et al., 2019). Other neuronal cell types, like nematocytes,

are starting their differentiation form NPCs via *NvSoxC* expression but deviate from the trajectory upon activation of nematocyte specific transcription factors and repression of *NvSoxC*. This deviation takes place at a very early stage of the neuronal trajectory (**Figure 36B**). In later branching cnidarians like *Hydra* the role of *SoxC* was then further restricted to be a differentiation marker only while maintenance of neuronal identity and plasticity was assigned to "new", yet unknown genes.

4.2.3. Migration and homeostasis of *NvSoxC* expressing cells.

Investigating the migration of cells in *KAEDE* transgenic animals happened to be more challenging than expected. First, one needs to define morphological landmarks, which help visualize the animals under the same orientation several days after the photoconversion. Using the oral pole of the animals as a landmark is not feasible since the pharynx has high levels of autofluorescence. Using the aboral pole of the animal as a landmark might be feasible in F1 animals but the mosaic transgenic nature of F0 animals causes transgenic patches not necessarily to be close enough to the aboral pole. The difficulty of anesthetizing the polyp leads to movement of *Nematostella*. The soft and motile nature of the relaxed animal introduces the challenge of orienting the animal even when a very prominent patch of cells was used as landmark. Amputation experiments generating amputates of 3-4mm to fit in the field of view of the microscope, underwent proportion regulation of shape upon regeneration. Proportion regulation of shape. Furthermore within 7 days post amputation the regenerating pharynx established the autofluorescence common to it, impairing further observations.

The proportion regulation in the regenerating amputate is not just of morphologically visible nature. The nerve net itself was shown to scale with size in animals which were amputated or starved (Havrilak et al., 2019). These studies could explain the observations regarding the transgenic cell patch of animal a16f (**Figure 31**) which was reducing in cell number during regeneration.

However, the mosaic nature of the F0 transgenic line offers some insights with respect to the migration of *SoxC* cells: the fact that the transgenic cells mostly stay together in clearly defined cell clusters with rather sharp edged boundaries (**Figure 24C** and **Figure 30**) and (**Figure 22** and **Figure 32A**) pver the course of 7 days instead of getting scattered across the body column implies that there is only limited migration of transgenic NvSoxC cells and the stem cells of which they originated. Animal a12 (**Figure 22** and **Figure 32A**). which has the broadest KAEDE patch is interesting since its transgenic stripe runs along the oral-aboral axis with a rather

sharp edge. This defined border implies a strong control of the patterning in the directive axis via e.g. *Hox* genes as it is shown by (Technau and Genikhovich, 2018). This would be coherent with observations in Hydra where neuronal precursor cells showed only movement along the oral-aboral axis but little circumferential migration (Technau and Holstein, 1996; Hager and David, 1997).

4.3. Conclusion

The aim of this work was to establish and characterise a new tool for lineage tracing in the cnidarian model organism *Nematostella vectensis.* This was accomplished by generating transgenic animals carrying the KAEDE fluorophore. To drive the expression of the fluorophore the promoter of the putative NPC transcription factor *NvSoxC* was selected based on unpublished single cell transcriptome data. The hypothesis was that utilizing a photoconvertible fluorophore would add temporal resolution, which allows to generate lineage maps for the selected driver gene.

The survival and transgenesis rates of the new KAEDE system are comparable with established reporter lines or fluorophores. This indicates that the fluorophore does not exhibit a toxic effect to be expected from this new fluorophore.

The photoconversion experiments performed in mosaic transgenic animals have shown, that the signal is bright in the original and the converted state. Unfortunately, the fluorophore is susceptible to photoconversion upon exposure to intense light. The rKAEDE signal decayed barely after 7d but was significantly reduced 14d post photoconversion. A comparably high red background remains 4 weeks post photoconversion.

The replenishment of the gKAEDE expression in some cells and while the rKAEDE is kept in other cells after photoconversion suggests that a temporal resolution was achieved. Based on these data a hypothesis could be made on *NvSoxC* being active in early nematogenesis as well as in mature neurons and an unidentified putative NPC lineage. This hypothesis could be refined by data form sc-approaches.

Where common fluorophores struggle to add information about the current activity state of the promoter the KAEDE system can clearly show if a promoter is still active in a given photoconverted cell or not. This aspect adds valuable temporal information to the investigation of a promoter refining the characterisation of its expression.

From our data we can state that our observations are partially consistent with those of *Hydra* and other model systems depicting *NvSoxC* as a gene active during nematogenesis, which is considered a specialised neuronal cell type. In contrast to *HvSoxC*, *NvSoxC* is also expressed in differentiated neurons. Thus, *NvSoxC* appears to play a crucial role in the common progenitor of both neurons and nematocytes.

We hypothesise further that this could be to keep a certain level of plasticity to these cells since *Nematostella* as well as other cnidarians show high regenerative and plastic capacities of the nervous system. In *Hydra* this plasticity might be maintained by a gene other than *SoxC* placing *HvSoxC* in the differentiation cascade only.

The observation that *NvSoxB(2)* is crucial for the formation of neurons and cnidocytes (Richards and Rentzsch, 2014) and that our single cell data show a broader expression of *NvSoxC* including the cell types depending on *NvSoxB(2)* expression **(Figure 6)** suggests a hierarchical order of *NvSoxC* above NvSoxB(2) the differentiation cascade of the neuronal lineage.

4.4. Outlook

4.4.1. The KAEDE system, confocal imaging and ISH

To properly establish a transgenic line a F1 generation should be generated to achieve a stable *NvSoxC*.prom.::KAEDE reporter line. The expression of the fluorophore throughout the whole animal might be overwhelming in the number of visible cells. These animals could be fixed and stained with available anti-KAEDE antibodies to prepare confocal investigations for more detailed observations.

Confocal experiments with *NvSoxC*.prom.::KAEDE transgenic animals will allow photoconversion of single neurons and tracing of axons (Dittrich et al., 2005). Due to technical constraints this requires totally relaxed animals since the capturing of confocal images is much more time consuming and movement of the animal will have a devastating impact on the image quality. Light sheet confocal microscopy might be better suited due to shorter image capture times. Anyhow, confocal imaging would provide the advantage of adjusting the excitation and emission wavelength to optimized values for the KAEDE fluorophores. Additionally, the influence of the autofluorescence background would be reduced drastically.

Should confocal approaches not be feasible with *in vivo* animals, another option to remove of the autofluorescence background could be the use of endogenous green/red double knockout lines. Experiments involving CRISPR/Cas9 successfully generated so called black-out lines with considerably reduced autofluorescence (Ikmi et al., 2014; Denner, 2017 unpublished).

ISH on *NvSoxC.prom.::KAEDE* animals would help to further validate the temporal resolution of the KAEDE-System., creating a correlation between the *NvSoxC* mRNA as well as the gKAEDE expression. It is yet to be clarified if the gKAEDE and rKAEDE fluorophores are stable enough to maintain their fluorescent properties after the harsh treatment during the ISH process. Already established combined methods of IHC and ISH using anti-KAEDE antibodies could be an alternative.

4.4.2. Functional dissection of NvSoxC using KAEDE transgenic lines

We have proposed that cells with a certain KAEDE expression patterns can be assigned to certain populations on the sc-data. To validate this hypothesis the next step would be to use fluorescence activated cell sorting (FACS) in combination with single cell transcriptomics. Since these experiments would have a high number of gKAEDE, rKAEDE and g+rKAEDE positive cells coming from photoconversion experiments of F1 animals they would be able to augment existing data by adding fluorescent labels to certain *NvSoxC* expressing cell populations.

To shed light on how crucial *NvSoxC* expression is for the development of the neuronal and glandular cell populations experiments involving knock-down or knock-out methods in combination with ISH and EdU labelling could be performed. These experiments would also be able to elucidate the relationship of *NvSoxC* and *NvSoxB(2)* in the neuronal transcription factor network.

References

(2020). Adobe Photoshop Lightroom. https://www.adobe.com/at/products/photoshop-lightroom.html. 6.4.2020.

Alié, A., Leclère, L., Jager, M., Dayraud, C., Chang, P., Le Guyader, H., Quéinnec, E., and Manuel, M. (2011). Somatic stem cells express Piwi and Vasa genes in an adult ctenophore: Ancient association of "germline genes" with stemness. Developmental biology *350*, 183-197.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. Journal of Molecular Biology *215*, 403-410.

Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H., and and Miyawaki, A. (2002). An optical marker based on the UV-induced greento- red photoconversion of a fluorescent protein. PNAS Biochemistry *99*, 12651-12656.

(2020). ApE- A plasmid Editor. https://jorgensen.biology.utah.edu/wayned/ape/. 6.4.2020.

Bergsland, M., Ramsköld, D., Zaouter, C., Klum, S., Sandberg, R., and Muhr, J. (2011). Sequentially acting Sox transcription factors in neural lineage development. Genes & development *25*, 2453-2464.

Bergsland, M., Werme, M., Malewicz, M., Perlmann, T., and Muhr, J. (2006). The establishment of neuronal properties is controlled by Sox4 and Sox11. Genes & development *20*, 3475-3486.

Bosch, T.C.G. (2007). Why polyps regenerate and we don't: towards a cellular and molecular framework for Hydra regeneration. Developmental biology *303*, 421-433.

Bosch, T.C.G., Klimovich, A., Domazet-Lošo, T., Gründer, S., Holstein, T.W., Jékely, G., Miller, D.J., Murillo-Rincon, A.P., Rentzsch, F., and Richards, G.S., et al. (2017). Back to the Basics: Cnidarians Start to Fire. Trends Neurosci *40*, 92-105.

Brey, C.W., Nelder, M.P., Hailemariam, T., Gaugler, R., and Hashmi, S. (2009). Krüppel-like family of transcription factors: an emerging new frontier in fat biology. Int J Biol Sci *5*, 622-636.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nature biotechnology *36*, 411-420.

Dang, D.T., Pevsner, J., and Yang, V.W. (2000). The biology of the mammalian Krüppel-like family of transcription factors. Int J Biochem Cell Biol *32*, 1103-1121.

Denner, A. (2017 unpublished). Characterization of conserved stem cell genes in Nematostella vectensis; unpublished. Master's dissertation (Vienna).

Dittrich, P.S., Schäfer, S.P., and Schwille, P. (2005). Characterization of the photoconversion on reaction of the fluorescent protein Kaede on the single-molecule level. Biophys J *89*, 3446-3455.

Dunn, C.W., Hejnol, A., Matus, D.Q., Pang, K., Browne, W.E., Smith, S.A., Seaver, E., Rouse, G.W., Obst, M., and Edgecombe, G.D., et al. (2008). Broad phylogenomic sampling improves resolution of the animal tree of life. Nature *452*, 745-749.

Eaton, S.A., Funnell, A.P.W., Sue, N., Nicholas, H., Pearson, R.C.M., and Crossley, M. (2008). A network of Krüppel-like Factors (Klfs). Klf8 is repressed by Klf3 and activated by Klf1 in vivo. J Biol Chem *283*, 26937-26947.

Ferraro, F., Lo Celso, C., and Scadden, D. (2010). Adult stem cels and their niches. Adv Exp Med Biol *695*, 155-168.

Fritzenwanker, J.H., and Technau, U. (2002). Induction of gametogenesis in the basal cnidarian Nematostella vectensis(Anthozoa). Development genes and evolution *212*, 99-103.

Genikhovich, G., Kürn, U., Hemmrich, G., and Bosch, T.C.G. (2006). Discovery of genes expressed in Hydra embryogenesis. Developmental biology *289*, 466-481.

Genikhovich, G., and Technau, U. (2009a). In situ hybridization of starlet sea anemone (Nematostella vectensis) embryos, larvae, and polyps. Cold Spring Harbor protocols *2009*, pdb.prot5282.

Genikhovich, G., and Technau, U. (2009b). Induction of spawning in the starlet sea anemone Nematostella vectensis, in vitro fertilization of gametes, and dejellying of zygotes. Cold Spring Harbor protocols *2009*, pdb.prot5281.

Grigoriev, I.V., Nordberg, H., Shabalov, I., Aerts, A., Cantor, M., Goodstein, D., Kuo, A., Minovitsky, S., Nikitin, R., and Ohm, R.A., et al. (2012). The genome portal of the Department of Energy Joint Genome Institute. Nucleic acids research *40*, D26-32.

Hager, G., and David, C.N. (1997). Pattern of differentiated nerve cells in hydra is determined by precursor migration. Development (Cambridge, England) *124*, 569.

Havrilak, J.A., Al-Shaer, L., Akinci, N., Amiel, A., Röttinger, E., and Layden, M.J. (2019). Dynamics and variability in regenerative potential of neuronal subtypes in the Nematostella nerve net. bioRxiv, 685917.

(2020). Helicon Focus - Helicon Soft. https://www.heliconsoft.com/heliconsoft-products/helicon-focus/. 6.4.2020.

Ikmi, A., and Gibson, M.C. (2010). Identification and in vivo characterization of NvFP-7R, a developmentally regulated red fluorescent protein of Nematostella vectensis. PLOS ONE *5*, e11807-e11807.

Ikmi, A., McKinney, S.A., Delventhal, K.M., and Gibson, M.C. (2014). TALEN and CRISPR/Cas9-mediated genome editing in the early-branching metazoan Nematostella vectensis. Nature communications *5*, 5486.

(2019). ImageJ. https://imagej.net/Welcome. 6.4.2020.

Jager, M., Quéinnec, E., Le Guyader, H., and Manuel, M. (2011). Multiple Sox genes are expressed in stem cells or in differentiating neuro-sensory cells in the hydrozoan Clytia hemisphaerica. Evodevo *2*, 12.

Jeon, H., Waku, T., Azami, T., Le Khoa, T.P., Yanagisawa, J., Takahashi, S., and Ema, M. (2016). Comprehensive Identification of Krüppel-Like Factor Family Members Contributing to the Self-Renewal of Mouse Embryonic Stem Cells and Cellular Reprogramming. PLOS ONE *11*, e0150715.

Jiang, J., Chan, Y.-S., Loh, Y.-H., Cai, J., Tong, G.-Q., Lim, C.-A., Robson, P., Zhong, S., and Ng, H.-H. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. Nature Cell Biology *10*, 353-360.

Juliano, C.E., Reich, A., Liu, N., Götzfried, J., Zhong, M., Uman, S., Reenan, R.A., Wessel, G.M., Steele, R.E., and Lin, H. (2013). PIWI proteins and PIWI-interacting RNAs function in Hydra somatic stem cells. Proceedings of the National Academy of Sciences of the United States of America *111*, 337-342.

Kamachi, Y., and Kondoh, H. (2013). Sox proteins: regulators of cell fate specification and differentiation. Development (Cambridge, England) *140*, 4129.

Kelava, I., Rentzsch, F., and Technau, U. (2015). Evolution of eumetazoan nervous systems: insights from cnidarians. Philosophical transactions of the Royal Society of London. Series B, Biological sciences *370*.

Kerner, P., Simionato, E., Le Gouar, M., and Vervoort, M. (2009). Orthologs of key vertebrate neural genes are expressed during neurogenesis in the annelid Platynereis dumerilii. Evolution & development *11*, 513-524.

Kirillova, A.O., Kraus, Y.A., and Markov, A.V. (2018). Dissociation–Reaggregation Experiments in Cnidarians as a Model System for the Study of the Regulative Capacity of Metazoan Development. Biology Bulletin Reviews *8*, 1-11.

Layden, M.J., Rentzsch, F., and Röttinger, E. (2016). The rise of the starlet sea anemone Nematostella vectensis as a model system to investigate development and regeneration. WIREs Dev Biol *5*, 408-428.

Li, Y., McClintick, J., Zhong, L., Edenberg, H.J., Yoder, M.C., and Chan, R.J. (2005). Murine embryonic stem cell differentiation is promoted by SOCS-3 and inhibited by the zinc finger transcription factor Klf4. blood *105*, 635-637.

Liu, N., Lu, M., Tian, X., and Han, Z. (2007). Molecular mechanisms involved in self-renewal and pluripotency of embryonic stem cells. J. Cell. Physiol. *211*, 279-286.

Martínez, D.E. (1998). Mortality Patterns Suggest Lack of Senescence in Hydra. Experimental Gerontology *33*, 217-225.

May, D., Blow, M.J., Kaplan, T., McCulley, D.J., Jensen, B.C., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., and Wright, C., et al. (2011). Large-scale discovery of enhancers from human heart tissue. Nature genetics *44*, 89-93.

McConnell, B.B., and Yang, V.W. (2010). Mammalian Krüppel-like factors in health and diseases. Physiological reviews *90*, 1337-1381.

Nègre, N., Brown, C.D., Ma, L., Bristow, C.A., Miller, S.W., Wagner, U., Kheradpour, P., Eaton, M.L., Loriaux, P., and Sealfon, R., et al. (2011). A cis-regulatory map of the Drosophila genome. Nature *471*, 527-531.

Nishimiya-Fujisawa, C., and Kobayashi, S. (2012). Germline stem cells and sex determination in Hydra. The International journal of developmental biology *56*, 499-508.

Presnell, J.S., Schnitzler, C.E., and Browne, W.E. (2015). KLF/SP Transcription Factor Family Evolution: Expansion, Diversification, and Innovation in Eukaryotes. Genome biology and evolution *7*, 2289-2309.

Renfer, E., Amon-Hassenzahl, A., Steinmetz, P.R.H., and Technau, U. (2010). A muscle-specific transgenic reporter line of the sea anemone, Nematostella vectensis. Proceedings of the National Academy of Sciences of the United States of America *107*, 104-108.

Renfer, E., and Technau, U. (2017). Meganuclease-assisted generation of stable transgenics in the sea anemone Nematostella vectensis. Nature protocols *12*, 1844-1854.

Rentzsch, F., Layden, M., and Manuel, M. (2017). The cellular and molecular basis of cnidarian neurogenesis. WIREs Dev Biol *6*, e257.

Richards, G.S., and Rentzsch, F. (2014). Transgenic analysis of a SoxB gene reveals neural progenitor cells in the cnidarian Nematostella vectensis. Development (Cambridge, England) *141*, 4681-4689.

Richards, G.S., and Rentzsch, F. (2015). Regulation of Nematostella neural progenitors by SoxB, Notch and bHLH genes. Development (Cambridge, England) *142*, 3332-3342.

Ries, A. (2019 unpublished). Tracing the fate of conserved stem cell marker genes in the sea anemone Nematostella vectensis; unpublished. Master's dissertation (Vienna).

Rowland, B.D., Bernards, R., and Peeper, D.S. (2005). The KLF4 tumour suppressor is a transcriptional repressor of p53 that acts as a context-dependent oncogene. Nature Cell Biology *7*, 1074-1082.

Rozen, S., and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. Methods in molecular biology (Clifton, N.J.) *132*, 365-386.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., and Schmid, B., et al. (2012). Fiji - an Open Source platform for biological image analysis. Nature methods *9*.

Schwaiger, M., Schönauer, A., Rendeiro, A.F., Pribitzer, C., Schauer, A., Gilles, A.F., Schinko, J.B., Renfer, E., Fredman, D., and Technau, U. (2014). Evolutionary conservation of the eumetazoan gene regulatory landscape. Genome research *24*, 639-650.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., and Lobanenkov, V.V., et al. (2012). A map of the cis-regulatory sequences in the mouse genome. Nature *488*, 116-120.

Siebert, S., Farrell, J.A., Cazet, J.F., Abeykoon, Y., Primack, A.S., Schnitzler, C.E., and Juliano, C.E. (2019). Stem cell differentiation trajectories in Hydra resolved at single-cell resolution. Science (New York, N.Y.) *365*.

Steele, R.E. (2002). Developmental signaling in Hydra: what does it take to build a "simple" animal? Developmental biology *248*, 199-219.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888-1902.e21.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Technau, U., and Genikhovich, G. (2018). Evolution: Directives from Sea Anemone Hox Genes. Current biology : CB *28*, R1303-R1305.

Technau, U., Genikhovich, G., and Kraus, J.E.M. (2015). Cnidaria. In Evolutionary Developmental Biology of Invertebrates 1: Introduction, Non-Bilateria, Acoelomorpha, Xenoturbellida, Chaetognatha, A. Wanninger, ed. (Vienna: Springer Vienna), pp. 115–163.

Technau, U., and Holstein, T.W. (1996). Phenotypic maturation of neurons and continuous precursor migration in the formation of the peduncle nerve net in Hydra. Developmental biology *177*, 599-615.

Tony Collins (2013). MBF "ImageJ for Microscopy" Collection.

Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., and Chen, F., et al. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature *457*, 854-858.

Voog, J., and Jones, L.D. (2010). Stem cells and the niche: a dynamic duo. Cell Stem Cell *6*, 103-115.

Watanabe, H., van Hoang, T., Mättner, R., and Holstein, T.W. (2009). Immortality and the base of multicellular life: Lessons from cnidarian stem cells. Seminars in Cell & Developmental Biology *20*, 1114-1125.