Single-cell RNA Sequencing Analysis and Experimental Designs to Investigate the Role of Nitric Oxide Signaling in Progenitor Cell Survival in *Ciona*

An Undergraduate Research Thesis

By

Alexandra Gurgis

In Partial Fulfillment
of the Requirements for the Research Option
B.S. Neuroscience

Georgia Institute of Technology

May 2021

Table of Contents

Abstract

The tunicate Ciona is a marine chordate whose biphasic life cycle, relatively simple transcriptome, and ability to be genetically manipulated via electroporation makes it an ideal model for studying phenomena of neurodevelopment and regeneration. During the metamorphosis of larvae to adults, nearly the entire central nervous system is eliminated and rebuilt, with the notable exception of the "Neck", a compartment of quiescent neural progenitor cells. The mechanism by which these cells are spared from the wave of programmed cell death that occurs around them is not currently understood, but could reveal important principles of cell cycle regulation. In this work, I re-examine single-cell RNA sequencing data to further characterize differential gene expression in the Neck, with a particular emphasis on the nitric oxide signaling pathway as a potential suppressor of apoptosis. I also provide experimental constructs for future investigations of genes I believe to be relevant to understanding Neck cell survival, including designs for in situ hybridization probes and GFP reporters to verify gene expression in vivo, sgRNA primers for CRISPR/Cas9 knockouts, and peakshift primers to verify the efficacy of sgRNAs.

Introduction

The tunicate, specifically *Ciona robusta (intestinalis type A),* hereafter referred to as *Ciona,* provides an ideal model organism for studying developmental neurobiology. It is small, with a compact and well-characterized genome, can be easily genetically manipulated using DNA transfection via electroporation and other recently developed molecular techniques like CRISPR/Cas9, and has a rapid developmental timeline that makes it easy to track the stages of individual animals. Despite the relative simplicity of their nervous system, which has only 231 neurons, all of whose connections have been fully mapped out, its organization and development is remarkably similar to that of humans and other vertebrates. In fact, tunicates are the invertebrate most closely related to us [1] and are thought to have the most compact genome that still contain all the necessary components to create a chordate body. These factors combine to make *Ciona* a unique and highly accessible model with which to study mechanisms underlying neurodevelopmental processes that are conserved in more complex vertebrate nervous systems like our own.

Ciona have two distinct stages in their life cycle, changing from free-swimming larvae into stationary, sessile adults that remain attached to a substrate. During the metamorphosis, the larval nervous system undergoes massive changes, with nearly all of the peripheral nervous system being degraded, and the central nervous system (CNS) being replaced by a new adult nervous system[6]. While most of the larval neural tissue is eliminated, one exception is a population of neural progenitor cells referred to as the "Neck" region. This compartment of undifferentiated progenitor cells lies immediately posterior to the neurons of the posterior larval brain and anterior to the motor ganglion, and has been shown to give rise to the branchial motor neurons of the adult tunicate, thought to be homologous to the branchiomeric motor neurons of the vertebrate hindbrain[2]. During the larval stages, the neck cells are in a state of inactivity, or quiescence, during which they do not differentiate or divide. They remain undifferentiated until metamorphosis, during which they are able to survive the wave of programmed cell death that eliminates the adjacent differentiated larval neurons, and begin to proliferate[3]. The molecular mechanisms that allow them to survive are not completely understood, however, and further elucidating this would be extremely valuable to our understanding of the neurodevelopment of *Ciona*, as well programmed cell death and progenitor cell maintenance as a whole.

From prior studies, single-cell RNA sequencing (scRNAseq) data from *Ciona* is already available, from which it is possible to isolate and characterize the neck cells[4]. In initial exploration of this data, a number of were observed to appear expressed preferentially in Neck cells as compared to the surrounding, differentiated neural tissue.The purpose of this work is to further analyze differential gene expression in the Neck, to explore the potential roles of these genes in vivo during the Neck cell development, and to create experimental designs that can be used to elucidate the molecular mechanisms of Neck survival during metamorphosis. To do this, the scRNA-seq data was analyzed in greater depth for underlying gene expression differences between differentiated larval neurons and undifferentiated Neck cells. Furthermore, mRNA in situ hybridization probes, GFP reporter transgenes, and single-chain guide RNAs (sgRNAs) for CRISPR/Cas9 knockouts were designed for *in vivo* interrogation of gene function in the Neck.

Literature Review

During metamorphosis, the majority of neurons in the central nervous system are eliminated via apoptosis. Yet, during this crucial phase of destruction and restructuring, the Neck cells are spared[3]. The mechanism by which these cells survive apoptotic waves is not known, but scRNAseq data revealed that several genes known to be involved in the nitric oxide (NO) pathway appeared to be upregulated in the Neck. NO is a produced by the enzyme nitric oxide synthase (NOS), and is known to have protective function via the repression of apoptosis, especially at low levels[5,6]. It has previously been shown that increasing NO levels delays tail regression in *Ciona intestinalis*, which normally occurs through a apoptotic wave during the onset of metamorphosis[7] by preventing apoptosis[8]. This same study appears to show NOS expression in the Neck as well[8]. In addition to potential NOS expression, preliminary analysis of available scRNAseq data[4] has shown that Neck cells seem to express several related genes. This includes Nos1ap, an nNOS adaptor protein[9], and Slc15a1/2 and Slc6a14, which encode transporters thought to bring arginine-containing peptides into the cell, valuable since arginine is the immediate precursor to NO synthesis[10,11]. Finally GCH, which encodes for an enzyme produces another NOS co-factor, Tetrahydrobiopterin (BH4)[12], is known to be expressed in cells directly around the Neck[13], further suggesting a particular role for NO signaling in the Neck.

The idea that NO signaling is essential to metamorphosis for biphasic life is extensively discussed by Bishop and Brandhorst[14], who postulate that rising nitrogen levels in the atmosphere is tied to the evolution of biphasic life and the sharp increase in organismal diversity during the Cambrian explosion. They purport that as NO signaling is the first shared inhibitory pathway between phylogenetically disparate biphasic life forms, including mollusks, ascidians, and echinoderms, it may be the common factor allowing for metamorphosis and the protection of "set-aside cells" for the future growth[14]. Inhibition of NOS in the gastropod *Ilyanassa obsoleta*, solitary ascidians *Boltenia villosa* and *Cnemidocarpa finmarkiensis*, and sea urchin, *Lytechinus pictus*, promoted metamorphosis, while overexpression delayed its initiation[14]. NO is also noted to diffuse readily across cell membranes and be highly reactive with a short half-life, making it both a responsive, dynamic signaling system as well as a highly localized one[14]. NO regulation is a promising lead in addressing questions about the neck region, especially since the concept of "side-aside" cells is strongly reminiscent of how the Neck remains quiescent, undifferentiated, and unaffected by the destruction of surrounding cells, only to develop and proliferate after metamorphosis.

One interesting mechanism by which NO can have anti-apoptotic effects is by the inactivation of caspases that promote apoptosis[5,6,15]. All caspase proteases have a cysteine residue in the enzyme catalytic site required for function, and which has a thiol that can be *S*-nitrosylated. This reaction effectively inactivates caspase activity, thus suppressing apoptosis as has been noted in several different cell types, including in primary cultures of human adhesion fibroblasts[5, 15]. S-nitrosylation is the covalent attachment of NO to a cysteine thiol to create a **S**-nitrosothiol (SNO); however, this rarely occurs with a neutral NO at physiological pH, instead happening readily with highly reactive NO products formed when NO loses an electron. This often occurs with a transition metal acting as an electron acceptor to oxidize NO, forming nitrosonium (NO^+) which can rapidly react with a nearby thiol groups to S-nitrosylate the protein. This reaction is largely dependent on the presence of an electron acceptor, like O_2 or

Fe_3^+, to which NO gives up an electron before S-nitrosylation can occur. The requirement of the presence of electron acceptors is also supported by evidence that S-nitrosylation occurs effectively in iron-rich hepatocytes, but not in iron-poor MCF adenocarcinoma cells or RAW264.7 cells unless preloaded with iron[16, 17]. Proteins can also transfer -NO groups to thiol groups on nearby proteins in reactions known as transnitrosylation, with the ultimate effect of creating a NO-dependent reaction cascade capable of inhibiting protein activity, capable of regulating cell death and neuronal apoptosis[18]. Previous work shows Nos1ap, one of the genes believed to be highly expressed in the Neck, forms a ternary complex with a downstream target protein Dexras1 and nNOS, co-localizing them and thus facilitating the nitrosylation of Dextras by the NO produced by NOS[19]. It is possible that Nos1ap may play a similar bridging role for NOS and other proteins, including caspases. Nos1ap expression in the neck might therefore be an important component to a Neck-specific protective effect of NO.

The anti-apoptotic role of NO signaling via electron-acceptor dependent S-nitrosylation of caspases becomes increasingly intriguing when noting the upregulation of the gene coding for the vanadium-binding protein Vanabin4 in the neck, which has been observed in scRNA-seq data and confirmed in situ by the Stolfi lab. Vanadium (V) is a metal ion of group 5 that can exist in oxidation states II-V and is found primarily as the vanadyl, $V_{IV,}$ cation in sea water. While sea water typically has an average concentration of vanadyl ions around 30 nm[20], vanadium is found to be stored at much higher concentrations in tunicates, with 2,000 times higher fold concentrations in *Ciona* blood (an average of 0.06 mM)[21]. Despite this massive accumulation, the biological function for vanadium and vanabins remains elusive. The enriched expression of Vanabin4 in the neck and biochemical characteristics of vanabins suggest a possible role in facilitating NO S-nitrosylation reactions.

Vanabins are vanadium-binding metalloproteins associated with the transportation and accumulation of vanadium ions. They are structurally marked cysteine residues repeated at regular intervals; Vanabin4 has 18[21].The cysteine residues form disulfide bonds known to have redox activity in the reduction of vanadium [22, 23,24]. Vanabin2 specifically has been shown to catalyze the reduction of V_V to V_{IV} in the presence of glutathione reductase and glutathione or thioredoxin during redox cascades from NADPH[22], demonstrating both the role of vanabins as vanadate-reductases and the redox capability of vanabin cysteine residues. These characteristics make Vanabin4 a plausible player in protein S-nitrosylation by NO; it binds with vanadium, a transition metal that could act as an electron acceptor from NO, and has many cysteines with disulfide bonds and redox activity that could participate in trans-nitrosylation.

Given these known anti-apoptotic capabilities of NO signaling and our initial examination of scRNAseq data from the Neck, I hypothesize that the Neck is spared from the wave of apoptosis that eliminates the larval CNS by a NO-signaling pathway dependent upon the interaction of various related genes, including Nos, Nos1ap, and Vanabin4. Here I will present additional scRNAseq analysis for these genes of interest, as well as experimental constructs to be used in future experiments.

Methods and Materials

scRNAseq Data Analysis

Re-processed raw scRNAseq data[4] from *Ciona* embryos in the Mid tail bud II stage, roughly 10 hours post-fertilization (hpf) was analyzed using the Seurat v3 package in R[25,26] to identify and recluster cell types based on unique gene expression markers. Prior to this work, replicates were integrated and pre-processing and clustering were performed using the **SCtransform** and **FindMarker** functions[27] to identify Neck and neural cell clusters. This data was re-analyzed to look specifically for differential gene expression in the neck relative to the brain. The **FeaturePlot** function was used to visualize feature expression in low-dimensional space, and to use colors to map out the relative expression levels for each gene of interest in the Neck and brain clusters (Fig. 1). Additionally, **ViolinPlot** and **RidgePlot** functions were applied to the expression distributions within the clusters, allowing for the heterogeneity of the Neck cluster to be further examined and the potential for additional sub-populations noted.

In situ Probe Design

Based on the expression trends in the re-analyzed scRNAseq data, several genes were chosen for additional investigation, and probes for in situ hybridization were designed. Probes were designed using the GHOST database[28,29] to access target gene sequencing and to find the open reading frames for different splicing variants. Then, 500 bp were selected and attached to a T7 promoter sequence to create a probe template. All probes used the following standard T7 promoter: CCCTATAGTGAGTCGTATTA. These sequences were then custom-synthesized by Twist Bioscience (South San Francisco, CA) for later *in vitro* transcription.

Design of GFP Reporter Constructs

GFP reporter constructs were made using potential cis-regulatory elements in introns attached to the basal promoter bpFOG fused to a GFP coding sequence, Unc-76::GFP. Potential cis-regulatory elements for genes of interest were identified using the WashU EpiGenome browser tunicate database, and looking for regions of untranslated DNA and intron that showed high degrees of conservation between the C. robusta genome and the C. savignyi genome. These elements were then added to a template with bpFOG and GFP and using AscI and XhoI restriction enzyme sites.

sgRNA oligos for CRISPR/Cas9

SgRNA designs were made using the predictive algorithms available in the CRISPOR[30] portal to maximize the likelihood of success. I used the genome *Ciona* robusta (formerly: C. intestinalis type S)- sea vase ascidian- GHOST Joined, KH2008, KH2013 Genes + SNPs: Ghost SNPs genome and the setting 20bp-NGG – Sp Cas9, SpCas9-HF1, eSpCas9 1.1 Exons from the target genes were inputted, and the resulting sgRNA targets were screened to exclude those that overlapped, as well as those with predicted single nucleotide polymorphisms or other inefficiencies. Targets were selected based on best predicted MIT specificity (>96) and Doench

‘16 scores (>50). The pre-designed primers for direct PCR in *C. intestinalis* from each selected target was used to generate sgRNA oligos. These constructs will be outscored to Twist Biosciences for synthesis and can later be used with CRSPR/Cas9 technology to analyze the effects of knocking out these genes on the development of *Ciona* Neck cells.

Peakshift Primer Design

Peakshift primers for each target site, as chosen by the sgRNA oligo designs above, were created using the BLASTN program[28] to search the genome for sequence uniqueness. The Multiple Primer Analyzer from ThermoFisher Scientific[31] was used to screen out primers with the potential to form self-dimers or cross-primer dimers as an additional quality control measure.

Results and Discussion

scRNAseq Analysis

Revisiting published scRNAseq data[4] processed and clustered using the methods described above, I compared the expression of 27 genes in the Neck as compared to the brain of *Ciona* embryos in the Mid tailbud II stage of development; results are listed in Table 1 and visualized in Figures 1 and 2. This list included several known Neck markers[32,2], Pax2/5/8.a, Phox2a/b, Hox1, FGF9/16/20, and Eph.c. as confirmation of the Neck cluster identity and provide a reference for other expression values As expected from previous in situ images, the Neck marker Pax2/5/8.a showed the largest fold change in expression between the two clusters (2.091272886), and the other previously identified Neck markers were also quite distinctVanabin4 was also greatly enriched (1.157771677) (Table 1, Fig. 2). NOS expression was confirmed to be slightly enriched (0.425761332), which consistent with our hypothesized protective role of NO signaling and previous work indicating the antiapoptotic effects of NO at low concentrations[5,6]. This analysis also revealed increased transcripts of Nos1ap, Slc15a1/2, and Slc6a14, further supporting the idea that NO signaling plays an important role in some aspect development unique to Neck cells versus surrounding tissues (Table 1, Fig. 1).

Several genes were notably under-expressed in the Neck compared to the differentiated brain neurons: Notch, CDK16/17, Mam1, Tuba, GBE.a, Crls1, and Gys. The lower expression of Tuba, which encodes tubulin proteins involved in the organization and formation of microtubules, is not surprisingly, as it plays an essential role in the development of axons[33] needed in the differentiated neurons but not in the Neck. It is particularly interesting to note of genes known to play important but diverse roles in development such as CDK16/17 and Notch and its transcriptional coactivator Mam1. Additionally, the lower expression of GBE.and Gys, both of which are involved in glycogen production and storage, suggests an altered metabolic state in the Neck as compared to rest of the brain. It is possible that the repression of these genes may be involved in the maintenance of the quiescent state of Neck cells during larval development, and would be interesting targets for the further study.

ANISEED ID	KH ID	*Name*	Top BLAST hits in human	Cluster 0 AveLog(FC)
Cirobu.g00013823	KH.S1363.2	***Pax2/5/8.a***	KH.S1363.2	2.091272886
Cirobu.g00003527	KH.C14.119	***Phox2a/b***	KH.C14.119	1.712462034
Cirobu.g00014681	KH.S555.1	***Htr7***	HTR1A; HTR5A; HTR7	1.709794864
Cirobu.g00011844	KH.L171.16	***Hox1***	HOXA1; HOXB1; HOXD1	1.460475549
Cirobu.g00010924	KH.L124.20	***Slc15a1/2***	SLC15A1; SLC15A2; SLC15A4	1.333518001
Cirobu.g00006092	KH.C3.88	***Vanabin4***	KH.C3.88	1.157771677
Cirobu.g00010142	KH.C9.582	***Nos1ap***	GULP1; LDLRAP1; Nos1ap	1.115298764
Cirobu.g00003715	KH.C14.291	***Dach***	DACH1; DACH2; SKIL	1.086857811
Cirobu.g00002001	KH.C11.189	***Abcb6***	ABCB6; ABCB7; ABCB8	1.042615438
Cirobu.g00002116	KH.C11.296	***Rerg***	RASL11A; RASL11B; RERG	0.99904173
Cirobu.g00002046	KH.C11.23	***Gephyrin***	GPHN	0.964671784
Cirobu.g00007857	KH.C6.44	***Kcnk***	KCNK10; KCNK2; KCNK5	0.950724909
Cirobu.g00004295	KH.C2.125	***FGF9/16/20***	FGF16; FGF20; FGF9	0.738473981
Cirobu.g00005712	KH.C3.526	***RasGRP***	RASGRP1; RASGRP2; RASGRP3	0.704826225
Cirobu.g00011572	KH.L154.32	***Gnaq***	GNA11; GNA14; GNAQ	0.702617478
Cirobu.g00009095	KH.C8.459	***Tesk***	LIMK1; TESK1; TESK2	0.692959659
Cirobu.g00008427	KH.C7.568	***Eph.c***	EPHA4; EPHA7; EPHB2	0.643789892
Cirobu.g00000141	KH.C1.1124	***Slc6a14***	SLC6A14; SLC6A5; SLC6A9	0.638526579
Cirobu.g00003129	KH.C12.562	***Noggin***	NOG	0.637940683
Cirobu.g00014898	KH.S674.1	***NOS***	NOS1; NOS2; NOS3	0.425761332
Cirobu.g00009697	KH.C9.176	***Notch***	NOTCH1; NOTCH2; NOTCH3	-0.717073035
Cirobu.g00007557	KH.C5.65	***Cdk16/17***	CDK16; CDK17; CDK18	-0.786535035
Cirobu.g00004335	KH.C2.162	***Maml***	MAML1; MAML3	-1.359294366
Cirobu.g00009570	KH.C8.892	***Tuba***	TUBA1C; TUBA3C; TUBA3D	-1.515928582
Cirobu.g00007750	KH.C6.224	***GBE.a***	GBE1	-1.927575092
Cirobu.g00002581	KH.C11.724	***Crls1***	CRLS1	-1.988920599
Cirobu.g00002477	KH.C11.628	***Gys***	GYS1; GYS2	-2.279430559

Table 1. List of selected genes, identified by unique gene ID (Cirobu.gxxxxxxxx), KyotoHoya (KH) gene model ID, name given in this study, and Top BLAST hit in humans. Single-cell RNAseq data from [4] using Seurat shows Average Log(FC), the log2 foldchange of gene expression difference between subcluster 0 (Neck) and subcluster 1 (brain). Positive values indicate increased expression in the Neck.

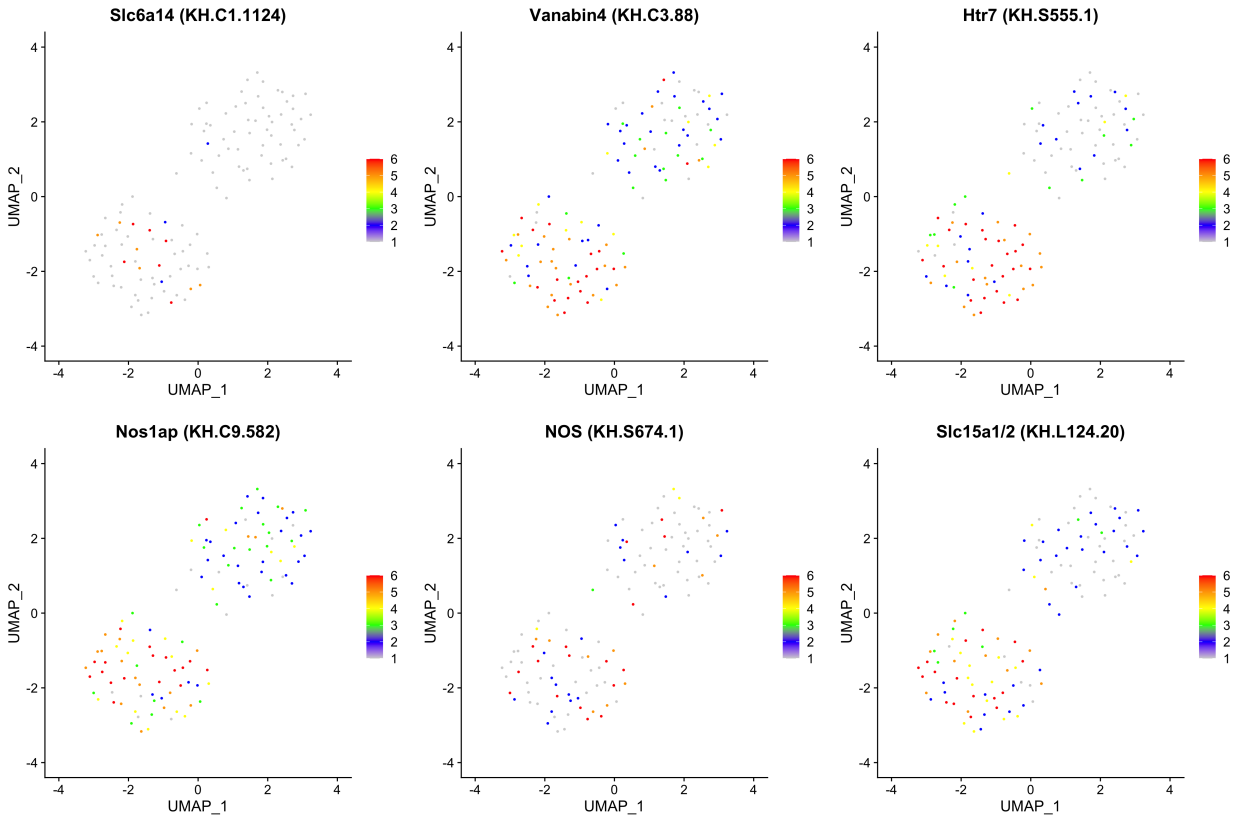

Fig. 1. Single-cell RNAseq differential gene expression "maps" of selected genes related to the nitric oxide signaling pathway and upregulated in in the Neck (bottom left) as compared to the brain (top right) subclusters. Genes are labelled by the common names used in this text as well as KyotoHoya (KH) gene model ID. Data reanalyzed from [4]. Cells are colored by relative gene expression, normalized as maximum (red) to minimum (grey) values for each gene, as indicated by color scale at top.

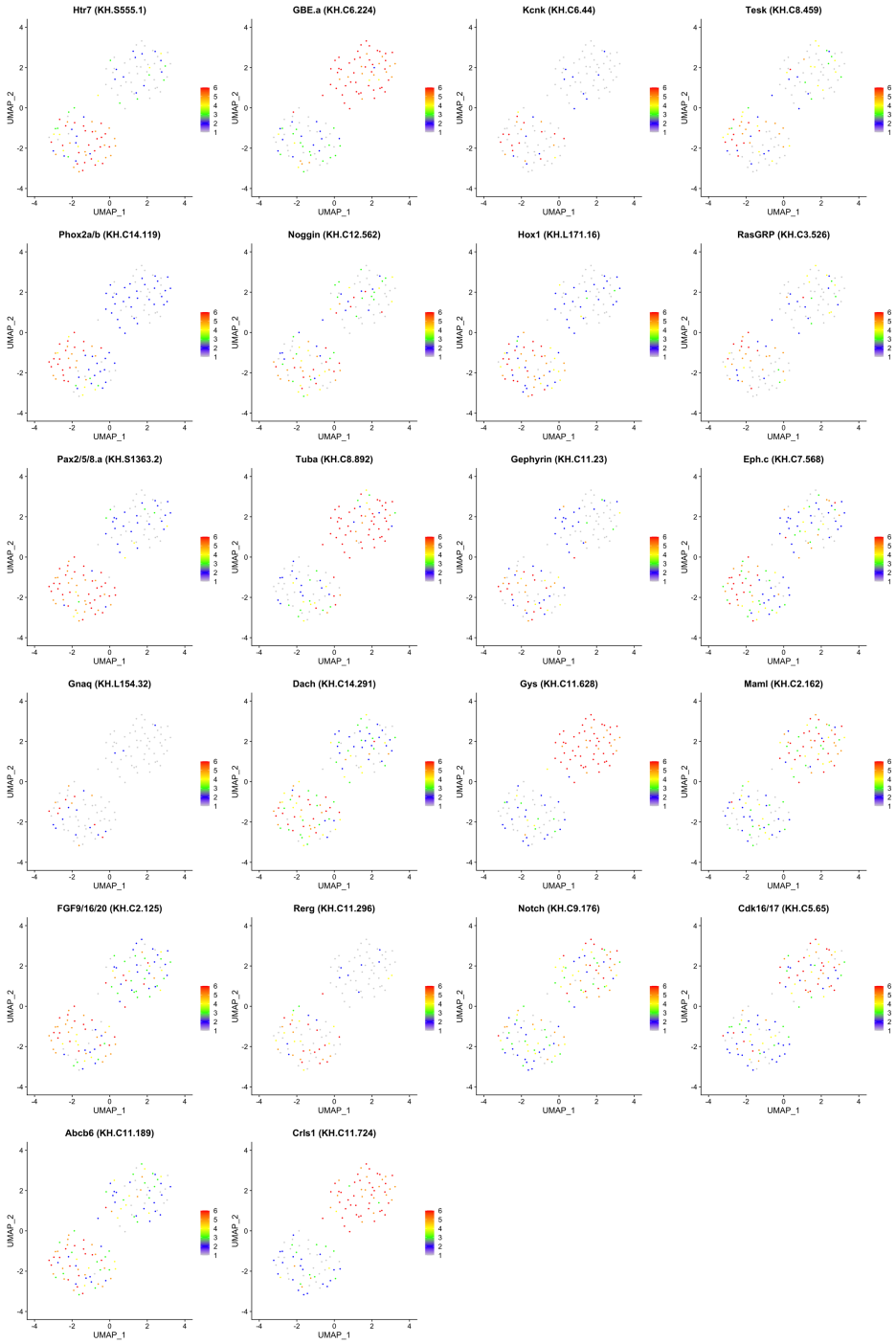

Fig. 2. Single-cell RNAseq differential gene expression "maps" of additional selected genes in the Neck (bottom left) and brain (top right) subclusters. Genes are labelled by the common names used in this text as well as KyotoHoya (KH) gene model ID. Data reanalyzed from [4]. Cells are colored by relative gene expression, normalized as maximum (red) to minimum (grey) values for each gene, as indicated by color scale at top. Neck cells show increased expression of several known markers[32,2], Pax2/5/8.a, Phox2a/b, Hox1, FGF9/16/20, and Eph.c and a notably decreased expression of Notch, CDK16/17, Mam1, Tuba, GBE.a, Crls1, and Gys.

When examining the Neck cluster expression mapping, it was noted that for some genes, there appeared to be a grouping distribution of expression within the Neck cluster itself. This grouping, seen as two distinct regions in the top left and bottom right is most prominent in the expression map for Phox2a/b, as annotated in Fig. 3 below, but is also appears to align with expression patterns visible in Eph.c, RasGRP, Kcnk and Tesk. Additional analysis of these genes' expression in the Neck cluster 0 may, therefore, be able to reveal that it can be further divided into additional subclusters delineating only parts of the Neck. Additionally, I created two types of plots to visualize the distributions of gene expression, which can be seen in figures 4 and 5 below. In these representations, expression in the Neck does shown more patched or clumped distribution patterns, especially visible with RasGRP and Kcnk. This could be due real stochasticity in expression or simply be a reflection of "gene dropout" due to limited coverage in sequencing or other technical artifacts.

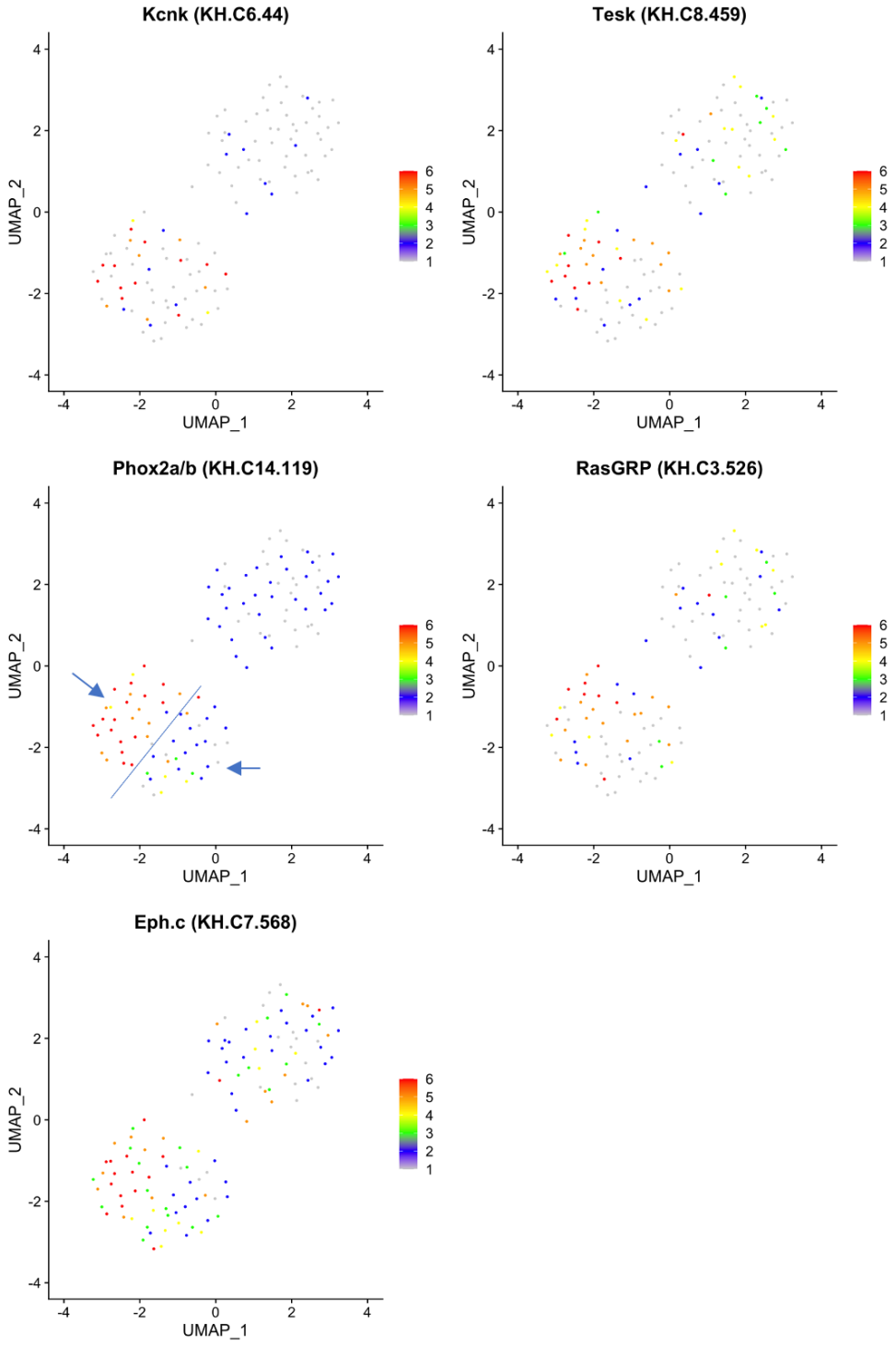

Fig 3. Annotated single-cell RNAseq differential gene expression "maps" to show the potential of an additional subcluster division within the Neck cluster. Blue arrows and line on Phox2a/b indicate the potential dividing of the Neck cluster into smaller subsets. Genes are labelled by the common names used in this text as well as KyotoHoya (KH) gene model ID. Data reanalyzed from [4]. Cells are colored by relative gene expression, normalized as maximum (red) to minimum (grey) values for each gene, as indicated by color scale at top.

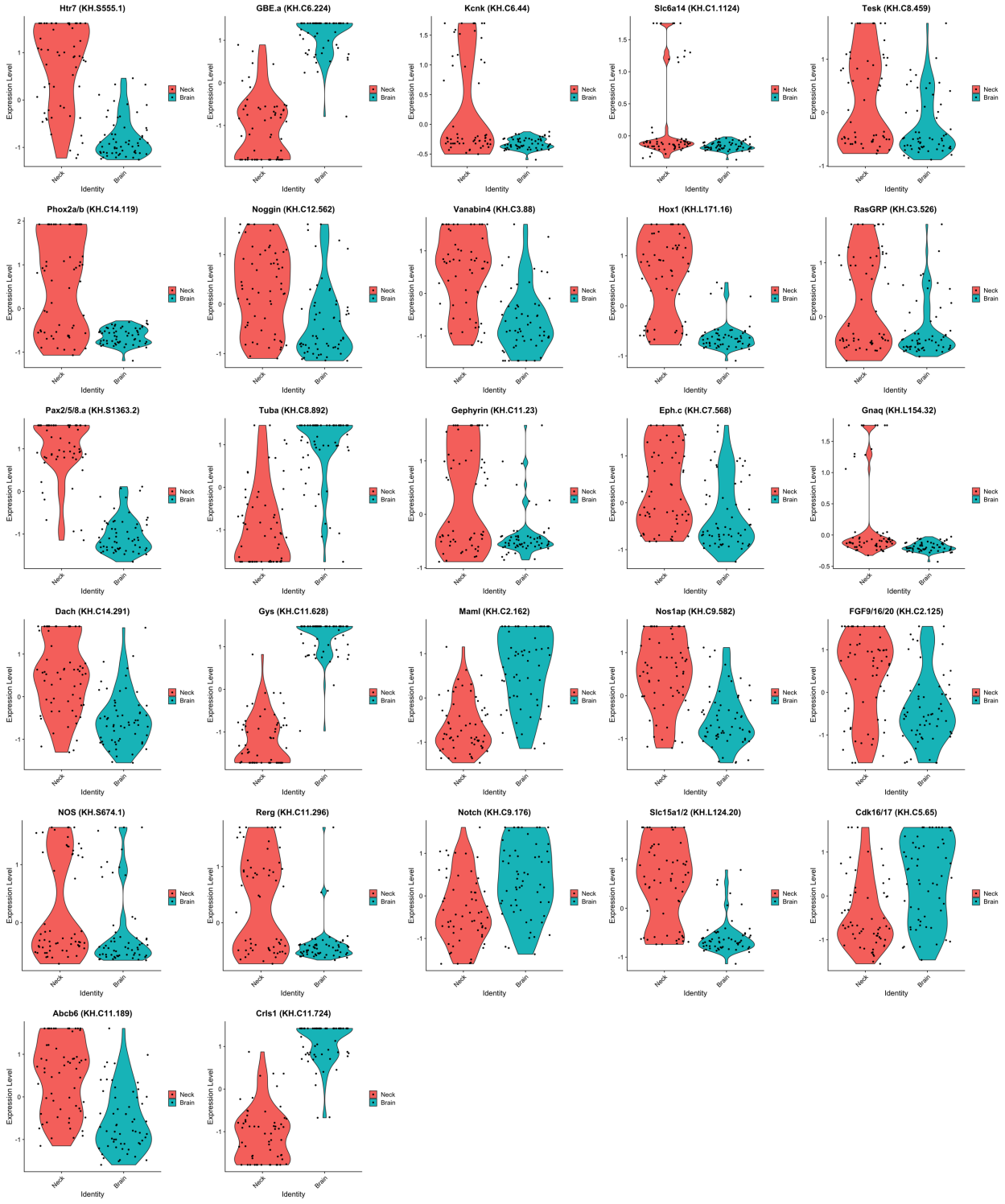

Fig 4. Violin plots to characterize gene expression distributions within the clusters. . Genes are labelled by the common names used in this text as well as KyotoHoya (KH) gene model ID. Data reanalyzed from [4]. X-axis width represents the number of cells within that expressed the gene

at that level, while y-axis position indicates the relative expression level. Blue is brain cluster 1 , red is Neck cluster 0.

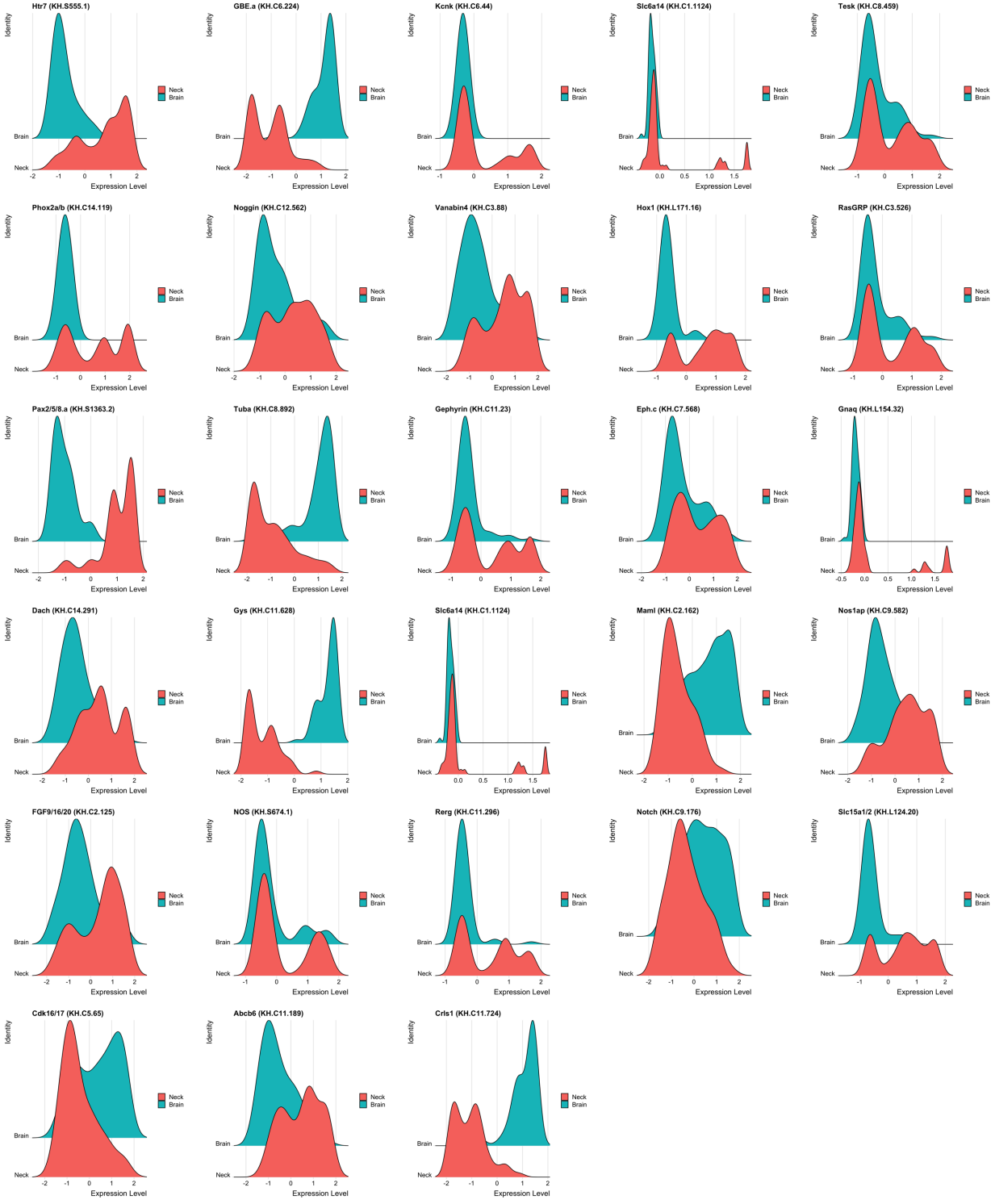

Fig 5. Ridge plots to characterize gene expression distributions within the clusters . . Genes are labelled by the common names used in this text as well as KyotoHoya (KH) gene model ID. Data reanalyzed from [4]. X-axis position indicates the relative expression level, while y-axis height

the number of cells within that expressed the gene at that level. Blue is brain cluster 1 , red is Neck cluster 0.

Experimental Designs

For sgRNA and peakshift primer designs, I followed the approaches previously laid out for the design and evaluation of guide RNAs for CRISPR/Cas9-mediated mutagenesis in *Ciona* specifically[34]. Using the methods outlined earlier, I designed sgRNA primers (Table 2) to be used for CRISPR/Cas9 mutagenesis and peakshift primers (Table 3) to verify their efficacy[34], for several of the genes shown in the scRNAseq data, as well as a few additional targets of interest. These constructs will be valuable to quickly determine the necessity of each gene in Neck survival. These will be made into oligos and ordered from Twist Biosciences, and then assembled into a plasmid vector for cloning and amplification in *E. Coli* cells, and finally electroporated into newly fertilized *Ciona* embryos.

Name	Target Sequence
vanabin.exon1.78fw.sgF	GTGGAAACCGCCATCATGG
vanabin.exon1.134fw.sgF	ACGTGCAAAGATGATTGCA
vanabin.exon1.145rv.sgF	GACATGTAGCTAAAGCGCA
vanabin.exon1.41fw.sgF	ATTGTACTTGTGCTTGCAT
Tesk.exon1.94rv.sgF	ATGATCTCATTACGACTGG
Tesk.exon1.156rv.sgF	AGATATATTCAAGCTTCGA
Tesk.exon1.212fw.sgF	CAAAATGCAGTCTCTACGC
Tesk.exon1.67rv.sgF	TTACTGGTTGTTGGGATTG
RasGRP.exon3.112fw.sgF	CAGTGTTGAAGCTCTAGTG
RasGRP.exon3.26fw.sgF	ACTGTCCAGTCGTCAACTG
RasGRP.exon3.35rv.sgF	TCGTCCAAACCAGGTACGC
MamI.exon2.33fw.sgF	CAAAACCGTCCATCGAGCA
MamI.exon3.92rv.sgF	GGCAATCCATTACCACCGG
MamI.exon3.63rv.sgF	GGGGTCACAACCATTCATG
MamI.exon3.39fw.sgF	GCAGCAACTTATGCAACAA
Abcb6.exon1.41fw.sgF	GGTATTAAAGACAGCGTCG
Abcb6.exon2.70rv.sgF	AGTTGAGCTTTGGCACACC
Abcb6.exon4.95fw.sgF	ATACTTATTGCAGGACGTG
Abcb6.exon4.25fw.sgF	GCTACTATGGCCTTACATG
Cdk16/17/18.exon1.65fw.sgF	AGAAGCCGCAGCAACTTGG
Cdk16/17/18.exon1.89rv.sgF	TCTTCCATAGTAAGGTGCT
Cdk16/17/18.exon2.87rv.sgF	TATTCCGGAGGGTGACAGT
Cdk16/17/18.exon2.157rv.sgF	GGAGGACCGAACGGGACAG
Nos1ap.exon1.64rv.sgF	AAGCTTCTTCATTATGAAG
Nos1ap.exon1.90rv.sgF	TGCCTTGAAATTGATCCCA
Nos1ap.exon1.1rv.sgF	TCCTCCTGGACATTCGTGA
Nos1ap.exon2.41fw.sgF	CCTCGGCCAACAGGACGAA
NotchReceptor.exon3.76fw.sgF	TATCGCATTGTAACAACGG

NotchReceptor.exon3.220fw.sgF	GCGAATGCGAAATACAAGA
NotchReceptor.exon3.106rv.sgF	TAAACACATTTGAACGGTG
NotchReceptor.exon3.144rv.sgF	AGCACTTGACAAGTATCAC
SLC15A1.exon1.53fw.sgF	GAAACAAAGATAATTGGTG
SLC15A1.exon3.160rv.sgF	TGACCGATGACGTATACGA
SLC15A1.exon3.55rv.sgF	GTGAAGGCGTGGAAGATGG
SLC15A1.exon3.132fw.sgF	AATAGCTGATTCTTATTGG
SLC6A14.exon5.86fw.sgF	GTAACCAGGCACAGTTCAG
SLC6A14.exon5.25rv.sgF	GGTGCATTCAGACCTACTC
SLC6A14.exon9.135fw.sgF	TTGTGTGTACGGTAAACTG
SLC6A14.exon9.198fw.sgF	GTCACATGGCACATCGACT
NOS.exon2.140fw.sgF	CCAACTGATAAACAATCAC
NOS.exon5.147fw.sgF	AATCTCAAAAATTGGGAGA
NOS.exon5.80rv.sgF	TCTGAAGCATTTCTGACAC
NOS.exon6.56fw.sgF	CTCATGCAACCAAGCAACA
GCH.exon1.5rv.sgF	GCTGCCATGACTGTATCAG
GCH.exon3.28fw.sgF	AATGAGGCAATATTTGACG

Table 2. sgRNA targets (N19) for direct PCR in *C. intestinalis* as designed using the CRISPOR portal[30].

Name	**Sequence**
GCH target 1 Forward	AAAATTTAGCGTTGTATCAAAAAAG
GCH target 1 Reverse	GCTTATAGTGCTAAAGCAGA
GCH target 2 Forward	CCACCCTATTGTTTATACATTTA
GCH target 2 Reverse	CGTACCACACAAACGGTATC
SLC6A14 target 1 Forward	TCGGCATTTACTACAACGT
SLC6A14 target 1 Reverse	CTTCCAATATTCTTCACTTGCG
SLC6A14 target 2 Forward	CAAGATGCAGAGGTATGAGA
SLC6A14 target 2 Reverse	GTCAAAAAGTTTGGTCCAGC
SLC6A14 target 3 Forward	CCAATTTAAAAGTGAGCATGTG
SLC6A14 target 3 Reverse	CATGGTCAGGATATTTTGCTA
SLC15A1/2 Target 1 Forward	GCTATTGTGAAGTAATAATCGGC
SLC15A1/2 Target 1 Reverse	GGATACGGTTATAGAAACCTGT
SLC15A1/2 Target 2 Forward	CTCATATTATAACGACTGACGT
SLC15A1/2 Target 2 Reverse	AATCGCTGCTATTGTCTT
SLC15A1/2 Target 3 Forward	AATGCTATCCATTTACTCAC
SLC15A1/2 Target 3 Reverse	GTCATAATTACGTGTGTGTC
Vanabin4 Target 1 Forward	GAGCACAAGGTTAATCCAA
Vanabin4 Target 1 Reverse	TTCGGCCGCTAAATATAAA

Vanabin4 Target 2 Forward	AGCGATTTACTTAGCTTGC
Vanabin4 Target 2 Reverse	CGGCAAACATTGAGTCAA
Vanabin4 Target 3 Forward	GACTGTAAGTAGTTGCTGTT
Vanabin4 Target 3 Reverse	TATATTTCTGCAATTACGTTTCA
Vanabin4 Target 4 Forward	GAAGCTATCAACCGATCAAA
Vanabin4 Target 4 Reverse	CTTTACAGCTTGGGAATCT
Nos1ap Target 1 Forward	CGATACACTACTCTGTGGTTA
Nos1ap Target 1 Reverse	TCCATTCAACCATCACATAG
Nos1ap Target 2 Forward	TTTACCATAAGCCATCACCT
Nos1ap Target 2 Reverse	TCCAATGTACTGTAACAAAAAG
Nos1ap Target 3 Forward	CCAATCAATAGCCGAGTG
Nos1ap Target 3 Reverse	CCCTTATTCTTCTCATGGC
Nos1ap Target 4 Forward	TGATACAAGAGTCCCACTT
Nos1ap Target 4 Reverse	TTTCTATTTCAGGAAATATGGTT
NOS Target 1 Forward	GAAAAAAGTACAAACCAGAAC
NOS Target 1 Reverse	ATACACGTGACATGCCAAA
NOS Target 2 Forward	TGTTAGAAATGATACGCACT
NOS Target 2 Reverse	CCAAGCAGCCATATACATATT
NOS target 3 Forward	CAAAACGCTGTTCGTTTG
NOS Target 3 Reverse	AGTAACGATGTCCCAAGC
NOS Target 4 Forward	ACGTGAACTATATACTAGAAACT
NOS Target 4 Reverse	CTTTCAAAAAGTAAAAGCAGTT

Table 3. Peakshift Primers, designed to be used with PCR and Sanger-sequencing to determine the efficacy of our CRISPR/Cas9 sgRNAs as previously discussed in [33].

Additionally, I designed probes for whole-mount mRNA in situ hybridization and GFP reporters for GCH, NOS, Nos1ap, Slc6a14, Vanabin4 and Slc15a1/2 to further examine the expression of these genes in the Neck. These genes were chosen for additional investigation due to their potential links in NO signaling as presented earlier. See the Additional Materials for complete designs.

Conclusion/Future Directions

Due to COVID-19 lab closures, I have been unable to validate the designs achieved above experimentally, nor carry out experiments to test the proposed NO and Vanabin-dependence of Neck cell survival. Moving forward, experiments can apply the different construct designs experimentally to begin to investigate this hypothesis.

First, the in situ probes and GFP reporters should be tested to confirm the scRNAseq data and further visualize gene expression across developmental stages in vivo. Specifically, it will be important to verify NOS expression in Neck cells. To test the function of NO pathway genes identified in our scRNAseq analysis (NOS, Nos1ap, Slc15a1/2, Slc6a14*)*, future experiments

could knock them out using CRISPR/Cas9 with the sgRNA designs presented in Table 2 and *Pax2/5/8.a>Cas9* to target them specifically in the Neck. Assaying resulting NO levels would allow one to determine if local NO presence production is directly dependent on gene expression in the neck, while monitoring apoptosis and Neck cell survival under each knockout condition would test the protective activity of each. If, as I predict, the expression of these genes is an important contributor to Neck cell survival, their knockout should result in both decreased NO levels and increased apoptosis in Neck cells. Additionally, to test the potential role of Vanabin4 in S-nitrosylation and caspase inhibition in this pathway, a Vanabin4 knockout could be made using the sgRNA designed, and assays performed for apoptosis and Neck cell survival as well as nitrosylation, caspase activity and NO levels. If Vanabin4 is indeed acting to facilitate the S-nitrosylation of Caspase-3 by NO, knocking out this gene would not affect NO production, but would result in a decrease of nitrosylation and an increase of caspase activity and apoptosis.

If these further experiments demonstrate anti-apoptotic role for NO signaling independent of caspase activity levels, other aspects of the nitric oxide signaling pathway may be valuable to investigate. Examining the expression of soluble guanylate cyclase genes (GUCYs), which encode for an enzyme that makes cGMP, one of the primary effectors for NO signaling, as well as well as phosphodiesterases that inhibit cGMP formation would help to determine if NO may be acting via a different mechanism than S-nitrosylation of caspases. It would also be interesting to further investigate what function Vanabin4 might be playing in the Neck instead.

This work provides a compelling motivation for additionally researching the role of NO signaling in Neck cell survival from the scRNAseq expression patterns noted, as well as tools for further investigations into the role of NO signaling in the Neck to build upon. Beyond this, the scRNAseq gene expression data presented can be used to continue characterizing this unique population of progenitor cells, and to identify other potential genetic regulators of Neck cell development and survival. Understanding the genetic regulators and cell signaling systems underlying Neck cell development in *Ciona* will enhance our understanding of progenitor cell maintenance, cell cycle regulation, and apoptotic survival as a whole.

References

1. Delsuc, F.; Brinkmann, H.; Chourrout, D.; Philippe, H., Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* **2006,** *439* (7079), 965-8.
2. Dufour, H. D.; Chettouh, Z.; Deyts, C.; de Rosa, R.; Goridis, C.; Joly, J.-S.; Brunet, J.-F., Precraniate origin of cranial motoneurons. *Proceedings of the National Academy of Sciences* **2006,** *103* (23), 8727-8732.
3. Tarallo, R.; Sordino, P., Time course of programmed cell death in Ciona intestinalis in relation to mitotic activity and MAPK signaling. *Developmental Dynamics* **2004,** *230* (2), 251-262.
4. Cao, C.; Lemaire, L. A.; Wang, W.; Yoon, P. H.; Choi, Y. A.; Parsons, L. R.; Matese, J. C.; Wang, W.; Levine, M.; Chen, K., Comprehensive single-cell transcriptome lineages of a proto-vertebrate. *Nature* **2019,** *571* (7765), 349-354.
5. Kim, Y.-M.; Bombeck, C. A.; Billiar, T. R., Nitric Oxide as a Bifunctional Regulator of Apoptosis. *Circulation Research* **1999,** *84* (3), 253-256.
6. Chung, H.-T.; Pae, H.-O.; Choi, B.-M.; Billiar, T. R.; Kim, Y.-M., Nitric Oxide as a Bioregulator of Apoptosis. *Biochemical and Biophysical Research Communications* **2001,** *282* (5), 1075-1079.
7. Karaiskou, A.; Swalla, B. J.; Sasakura, Y.; Chambon, J.-P., Metamorphosis in solitary ascidians. *genesis* **2015,** *53* (1), 34-47.
8. Comes, S.; Locascio, A.; Silvestre, F.; d'Ischia, M.; Russo, G. L.; Tosti, E.; Branno, M.; Palumbo, A., Regulatory roles of nitric oxide during larval development and metamorphosis in Ciona intestinalis. *Developmental Biology* **2007,** *306* (2), 772-784.
9. Li, L.-L.; Ginet, V.; Liu, X.; Vergun, O.; Tuittila, M.; Mathieu, M.; Bonny, C.; Puyal, J.; Truttmann, A. C.; Courtney, M. J., The nNOS-p38MAPK Pathway Is Mediated by NOS1AP during Neuronal Death. *The Journal of Neuroscience* **2013,** *33* (19), 8185.
10. Yang, X. D.; Ma, J. Y.; Barger, M. W.; Ma, J. K., Transport and utilization of arginine and arginine-containing peptides by rat alveolar macrophages. *Pharm Res* **2002,** *19* (6), 825-31.
11. Stevens, B. R., Amino Acid Transport by Epithelial Membranes. *Epithelial Transport Physiology* **2009**, 353-378.
12. Alp, N. J.; Channon, K. M., Regulation of Endothelial Nitric Oxide Synthase by Tetrahydrobiopterin in Vascular Disease. *Arteriosclerosis, Thrombosis, and Vascular Biology* **2004,** *24* (3), 413-420.
13. Razy-Krajka, F.; Brown, E. R.; Horie, T.; Callebert, J.; Sasakura, Y.; Joly, J.-S.; Kusakabe, T. G.; Vernier, P., Monoaminergic modulation of photoreception in ascidian: evidence for a proto-hypothalamo-retinal territory. *BMC Biology* **2012,** *10* (1), 45.
14. Bishop, C. D.; Brandhorst, B. P., On nitric oxide signaling, metamorphosis, and the evolution of biphasic life cycles. *Evol Dev* **2003,** *5* (5), 542-50.
15. Jiang, Z. L.; Fletcher, N. M.; Diamond, M. P.; Abu-Soud, H. M.; Saed, G. M., S-nitrosylation of caspase-3 is the mechanism by which adhesion fibroblasts manifest lower apoptosis. *Wound Repair Regen* **2009,** *17* (2), 224-229.
16. Kim, Y.-M.; Kim, T.-H.; Seol, D.-W.; Talanian, R. V.; Billiar, T. R., Nitric Oxide Suppression of Apoptosis Occurs in Association with an Inhibition of Bcl-2 Cleavage and Cytochrome cRelease*. *Journal of Biological Chemistry* **1998,** *273* (47), 31437-31441.

17. Kim, Y.-M.; Chung, H.-T.; Simmons, R. L.; Billiar, T. R., Cellular Non-heme Iron Content Is a Determinant of Nitric Oxide-mediated Apoptosis, Necrosis, and Caspase Inhibition*. *Journal of Biological Chemistry* **2000,** *275* (15), 10954-10961.
18. Nakamura, T.; Lipton, S. A., Emerging role of protein-protein transnitrosylation in cell signaling pathways. *Antioxid Redox Signal* **2013,** *18* (3), 239-49.
19. Fang, M.; Jaffrey, S. R.; Sawa, A.; Ye, K.; Luo, X.; Snyder, S. H., Dexras1: a G protein specifically coupled to neuronal nitric oxide synthase via CAPON. *Neuron* **2000,** *28* (1), 183-93.
20. Chatterjee, P. B.; Crans, D. C., 3.13 Vanadium Biochemistry. 2013; Vol. 3, pp 323-342.
21. Trivedi, S.; Ueki, T.; Yamaguchi, N.; Michibata, H., Novel vanadium-binding proteins (vanabins) identified in cDNA libraries and the genome of the ascidian Ciona intestinalis. *Biochim Biophys Acta* **2003,** *1630* (2-3), 64-70.
22. Kawakami, N.; Ueki, T.; Amata, Y.; Kanamori, K.; Matsuo, K.; Gekko, K.; Michibata, H., A novel vanadium reductase, Vanabin2, forms a possible cascade involved in electron transfer. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2009,** *1794* (4), 674-679.
23. Ueki, T.; Adi, T. K., Mechanism of vanadium accumulation and possible function of vanadium in underwater adhesion in ascidians. *AIP Conference Proceedings* **2019,** *2120* (1), 020001.
24. Ueki, T.; Uwagaki, M.; Yamamoto, S.; Michibata, H., Participation of thioredoxin in the V(V)-reduction reaction by Vanabin2. *Biochim Biophys Acta* **2014,** *1840* (11), 3238-45.
25. Satija, R.; Farrell, J. A.; Gennert, D.; Schier, A. F.; Regev, A., Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology* **2015,** *33* (5), 495-502.
26. Stuart, T.; Butler, A.; Hoffman, P.; Hafemeister, C.; Papalexi, E.; Mauck, W. M., 3rd; Hao, Y.; Stoeckius, M.; Smibert, P.; Satija, R., Comprehensive Integration of Single-Cell Data. *Cell* **2019,** *177* (7), 1888-1902.e21.
27. Hafemeister, C.; Satija, R., Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *bioRxiv* **2019**, 576827.
28. Satou, Y.; Kawashima, T.; Shoguchi, E.; Nakayama, A.; Satoh, N., An integrated database of the ascidian, Ciona intestinalis: towards functional genomics. *Zoolog Sci* **2005,** *22* (8), 837-43.
29. Satou, Y.; Mineta, K.; Ogasawara, M.; Sasakura, Y.; Shoguchi, E.; Ueno, K.; Yamada, L.; Matsumoto, J.; Wasserscheid, J.; Dewar, K.; Wiley, G. B.; Macmil, S. L.; Roe, B. A.; Zeller, R. W.; Hastings, K. E. M.; Lemaire, P.; Lindquist, E.; Endo, T.; Hotta, K.; Inaba, K., Improved genome assembly and evidence-based global gene model set for the chordate Ciona intestinalis: new insight into intron and operon populations. *Genome Biology* **2008,** *9* (10), R152.
30. Concordet, J.-P.; Haeussler, M., CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Research* **2018,** *46* (W1), W242-W245.
31. Multiple Primer Analyzer. https://www.thermofisher.com/us/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/multiple-primer-analyzer.html.
32. Imai, K. S.; Stolfi, A.; Levine, M.; Satou, Y., Gene regulatory networks underlying the compartmentalization of the Ciona central nervous system. *Development* **2009,** *136* (2), 285-93.

33. Aiken, J.; Buscaglia, G.; Bates, E. A.; Moore, J. K., The α-Tubulin gene TUBA1A in Brain Development: A Key Ingredient in the Neuronal Isotype Blend. *Journal of Developmental Biology* **2017,** *5* (3), 8.
34. Gandhi, S.; Haeussler, M.; Razy-Krajka, F.; Christiaen, L.; Stolfi, A., Evaluation and rational design of guide RNAs for efficient CRISPR/Cas9-mediated mutagenesis in Ciona. *Dev Biol* **2017,** *425* (1), 8-20.

Additional Materials

In Situ Probe Sequences:

GCH:

GGATACATTCCAAACAAAAGGGTGTTAGGAATTAGCAAGCTTGCAAGAATCGTGGA
AATGTACAGTCGTCGATTGCAGGTTCAGGAGAGGCTAACAAAACAAATAGCGTCTG
CACTCGTTGAGGTTATTGAACCTTCCGGGGTGGCTGTTGTTATTGAAGCATCTCATAT
GTGTATGGTGATGCGAGGTGTTCAGAAACCAAGAGCAACTACAATGACTAGCAGTA
TGTTTGGCGTTTTCCGAGACGATCCGAAAACTCGCGAAGAATTCCTAACCTTGCTCA
ACAAAACCTAAACTGTTTCAAACTCCTAAGAAATTTTCATTGGCTTTTAAGATGAAT
GAATGTAACTCGCTTCATCCTTGCGTGGCCGGAAAACGACCTTATCATATAGGTTTA
ACACCTCGATCGCGTGCCAGCTTACGAGTTGCCATGTATGTTACTTTGTGGGTTAAT
ATGTTTCTATTTTCATGTCATATTGATTCTAAAACAATTTTACTTTCCCTATAGTGAGT
CGTATTA

NOS:

ATCATGATCACCTCTATAAACAAGAAGTTGCGGCTGCCATGAAAGAAGGAGCCATA
TCTCATGCTTACACTGCATTTTCAAGGGAAAAGGGTAAACCAAAGCGATATGTCCAA
AATGTGATTATGAGTGAGATTCCAGATTTTGTTTATGACATTGTTCATGAAAAGAAA
GGCCATTTTTATGTGTGTGGTGATGTCACAATGGCAACAGAAGTAAACAAAACACTT
CTAAAAATCCTTTGTATAAAAGGAGAGATGACAAAAGAAAAAGCTCAGGAGTTTAT
CAACTCAATGAAGACGGAAGAGAGGTATCACGAAGATATATTTGGAGTCACGCTTC
GTACTTATGAAGTCAACAGAAGTGTGAGACTCAAGGGAAAGGAAAAGAGTGAAAG
CCTGTCGCATAAGTGAATCATACGTTTTTGAGAGATTACCACTTGTATTAAAACTTAT
TTTGAAACCCTTCTTATTAGCAAGCACTTTAAGTTTTGCTTTTATAAGCCCTATAGTG
AGTCGTATTA

Nos1ap:

CAAACAGTGATCCGCTAGAGGCATCCGTGGTTGACAGAACCCCAGTTTTATCACCAG
TGACGTCACCTACCTCGCTTAGCAGTGGTACGTATAACGATCCTGCAAACTCTCCTG
CACGGGATTTAAACTGTGCGAAGTCATTACCGTTCAGTGGAAGCGAGCCGACTTCTA
TTTCGTCAAGTGACTCAGGCGTTATGTCAAGTTCAACAAGTGGAAACATCCAGGACA
TTCCAAACGTAGTACATCATACGTGACTTTCATTAGCGTACAATCTATAGCGGACTG
AATATACAGCACGGGTATAGTACTGTGAAAGATCAGTGTGAAACAGACGCTAGACG
GTGCCCGTAATTTATTTCTTTTGAGGTACCACATTCGTTATTAATCTAGTGTCCGAGT
ACCTCTGTATCTCACTTGGAAAGTTGGCTGTTCCTATTGATTTTAAAGTTCGTATTGT
TGTTAGAGGTAGATATAGTCTACATTTTGCATTCCTTCCGTTTCCCTATAGTGAGTCG
TATTA

Slc6A14:

AGCATGTTGGTTCTTCATATCACCGTTGCTGTTAACGGCGATCTTAGTTTGGTCTCTT
GCGACGTTTGATACCAAATCTATCGGAGGTTATGAGTATCCAGAATTCGCGACAGTG
ATTGGTTGGCTCATCATTGCAACTGCTATCATTTTCGTTCCAATATTTGGCATCGTGG
CTATTGTGAAAAATAAAGGAAGTGTATTCAAGGCTTCGTCCCCTGCATCAGACTGGG
GACCTTACCTAAAAGAACATCGTGGTGTAAGGTATGCCAATATGGATGAACCGGAG
GAGGCATTAGCAATTGCTGAAGAAAGAAACCAGCATGAGAGCGAGAAAGGAAAGG
GCAATCCAGTGTATGCACATGACAACGGTGAAACTGAACAAACTAAGATGTAAACA
TGATGTGTAGTTTTACAATTCATGCTGTAAATAAACTTTTCACCTGATGGTGCTACAA
AGGGAATTTAAACTATACTGAATCATCAATTATATAGGTCTACAACCCTATAGTGAG
TCGTATTA

Slc15a1/2:

GATCACAGTAGGGGAGGTTTTCCTCTCCATTACTGGGCTCGAGTTCTCTTACACGCA
GGCCCCAGCCAGTATGAAGTCCGTAGTATCAAGTATATGGCTCCTCACCGTTTCGAT
GGGCAGCATTATCGTACTCGTCGTCGCCGAATCTCGTATGATTGACAGGCAGGCCCA
TGAATTTATTTTCTTCGCATGTTTGATCGGAGTCGCTTCCATCGTGTTTGCTTTCCTGG
CTCGCAGATACAAAGCATCGAAACCGCAAGAAATTCCAAGCAAGAAAGAAAAGAA
CGAAATCGAAAAGACAGAACTTTTAGAAGAAAAAACAGACTTTAGTGAATGAATTC
GTCGTAAAATATATTTTAATCCTGATTTTCTAATTTAGTAACAAACTAATTAATTCTC
TTTATATAAAACAGAATTCACAATCTTGTTTTGCGATAACTTAAACCCAATGCTTAA
ATTGTCATTAAACTATTGCCTTTGCCCCTATAGTGAGTCGTATTA

Vanabin4:

CTGAAAACGTGCAAAGATGATTGCACGGAAATGAAACCTTGCGCTTTAGCTACATGT
CCTTCTGTCTGCCACGCAACTCGCGAAGCGGCCGAAGGCAGTGGCGCTAATCGCTGC
ATGATCAGGTGCGGGTTGACTCAATGTTTGCCGAGATTCCCAAGCTGTAAAGCGTGC
GTCGCCCGTTGCGCTGCACCGGTTACCGCATGCAAGCGAAGCAGTTGTGCATCTGAG
TGCCCAGCCGGGATGACCATACCTGAGCACATGCAACTAAGTGGATGCGTTCGTTGC
ATGAAACGTAATTGCAGAAATATAATGAATGGGAACTAGTGGATACGACTAGATTC
TGACGATCACCCCCCAAAAAAATATACCCACTAACCTAAAATACTAACAATTTGTAA
GTGATCTGTCGTAGAGTGTGAAATATAACTCGTTAAAGAAAACTAATAAAATTGCAT
TCTAACATATTTGATATTTCGTATTGGAATTTACCAACAGTTAATCCCTATAGTGAGT
CGTATTA

GFP Reporters:

GCH:
GGCGCGCCTTTCAAATTCAAATATTTACACTGTACTTGAGAAAGGCGCCGACGGCAT
ATTGTTTCAACAAATTATAAATACTAGCTTTTTAAAAATATTGAAAGACGTTAAACC
AAAACTAAGTTACTTGTTTGTCTCGAGCAGCTGAAGCTTGCATGCCTGCAGGTCGAC
TCTAGAGGATCCGGCAAAGCTTCGTGTATTGTACCGGCCCATTGTCAATCATGCAAA
CTTGATATTATATTGACAAGAGAAGAAGGCAGTTTAAATTAAAACTCTAAAGTAGA
GAGACATTAATCTCAGCTGACAAGGCAGGTGGTCACAGTAAGTTCATTTAAATAGTT
GGCCAACAATAGCCTTTCCAAGAAAGTATTTTTGTTCCAGGTCTATACAAAAATAAC

ACACAACATGGCGGCCGCAACCATGGCGGATCTGCGAGTACCGGACATTCCGCTCG
CCTCGTGTGATGATGATGATATCGATAGTAATAAGAATTTGAGCAACCATTCATCAG
ACGAGAAACATCACTGCAACAGCAACAGCGACGAGGAACGTCTTCATGACGAGTTC
TCTGGATCCCTTGAGGACCTTGTCGGCAACTTTGACGAAAAAATTGCGGCATGCCTG
AAGGACCACGAGGTGACGACAGCGGATATTGCACCTGTGCAGATACGTACTCAAGA
GGAAGTTATGAATGAAAGCCAAACATGGTGGACATTAACCGGAAACTTTGGAAACA
TTCAACCTCTCGACTTTGGAACCTCTTCGATATGTAAAAAGATGGCCGCAGCTCTGG
ACAGTGATTCATTGAAAGACGACGCATCTACACGCCGAAGTATGACAAATTCCGAT
GATGAGGATCTTTTACGACAACAAATGGATGTTCATCAAATGATTGGACATCATCAT
GGATCTACGGATACTGGTGGTGAAACACCTCCACAGACTGCTGATCAAGTTATCGAA
GAAATTGATGAAATGTTACAGGTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGA
GCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCC
ACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACC
CTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACC
ACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCA
CGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTT
CAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACC
CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCT
GGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACA
AGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGC
AGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGT
GCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCA
ACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACT
CTCGGCATGGACGAGCTGTACAAGTAAGAATTCCAGCTGAGCGCCGGTCGCTACCA
TTACCAGTTGGTCTGGTGTCAAAAATAATAATAACCGGGCAGGCCATGTCTGCCCGT
ATTTCGCGTAAGGAAATCCATTATGTACTATTTAAAAAACACAAACTTTTGGATGTT
CGGTTTATTCTTTTTCTTTTACTTTTTTATCATGGGAGCCTACTTCCCGTTTTTCCCGA
TTTGGCTACATGACATCAACCATATCAGCAAAAGTGATACGGGTATTATTTTTGCCG
CTATTTCTCTGTTCTCGCTATTATTCCAACCGCTGTTTGGTCTGCTTTCTGACAAACTC
GGAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATT
TCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAA
TGTATCTTATCATGTCTGGATCGACAAAGTCAAAGCGGCCATCAGATCTGCCGGTCT
CCCTATAGTGAGTCGTATTAATTTCGATAAGCCAGGTTAACCTGCATTAATGAATCG
GCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCA
CTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGG
CGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCA
AAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA
TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGC
GAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTG
CGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGG
GAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCG
TTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCT
TATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGG
CAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAG
TTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGC
GCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA

CAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGA
AAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGG
AACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACC
TAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAA
ACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGT
CTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGG
AGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCG
GCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGG
TCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTA
AGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTG
GTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG
CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCG
ATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTG
CATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACT
CAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGT
CAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGA
AAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCG
ATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT
CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGAC
ACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAG
GGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATA
GGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATT
ATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTACGTATTAATTAA

Yellow Highlighted Region: AscI
Blue Highlighted Region: GCH cis-regulator element to test
Red Highlighted Region: XhoI
Magenta Highlighted Region: Basal promoter
Green Highlighted Region: GFP

Nos1ap:
GGCGCGCCTAAAAAAGAGGAAGTACTAACAGAACTGTTTTAATTAATAATGTAACT
TCTCCTTTTCACCCATTTTGTTTAATAAGCAAAAATGAATCACCTGTTCAATTATTTG
TTTTTCTGCTCGAGCAGCTGAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCG
GCAAAGCTTCGTGTATTGTACCGGCCCATTGTCAATCATGCAAACTTGATATTATATT
GACAAGAGAAGAAGGCAGTTTAAATTAAAACTCTAAAGTAGAGAGACATTAATCTC
AGCTGACAAGGCAGGTGGTCACAGTAAGTTCATTTAAATAGTTGGCCAACAATAGC
CTTTCCAAGAAAGTATTTTTGTTCCAGGTCTATACAAAAATAACACACAACATGGCG
GCCGCAACCATGGCGGATCTGCGAGTACCGGACATTCCGCTCGCCTCGTGTGATGAT
GATGATATCGATAGTAATAAGAATTTGAGCAACCATTCATCAGACGAGAAACATCA
CTGCAACAGCAACAGCGACGAGGAACGTCTTCATGACGAGTTCTCTGGATCCCTTGA
GGACCTTGTCGGCAACTTTGACGAAAAAATTGCGGCATGCCTGAAGGACCACGAGG
TGACGACAGCGGATATTGCACCTGTGCAGATACGTACTCAAGAGGAAGTTATGAAT
GAAAGCCAAACATGGTGGACATTAACCGGAAACTTTGGAAACATTCAACCTCTCGA
CTTTGGAACCTCTTCGATATGTAAAAAGATGGCCGCAGCTCTGGACAGTGATTCATT

GAAAGACGACGCATCTACACGCCGAAGTATGACAAATTCCGATGATGAGGATCTTT
TACGACAACAAATGGATGTTCATCAAATGATTGGACATCATCATGGATCTACGGATA
CTGGTGGTGAAACACCTCCACAGACTGCTGATCAAGTTATCGAAGAAATTGATGAA
ATGTTACAGGTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGG
GGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCG
TGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATC
TGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTAC
GGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAG
TCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGG
CAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCA
TCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTG
GAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGG
CATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCG
CCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGAC
AACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGA
TCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGA
GCTGTACAAGTAAGAATTCCAGCTGAGCGCCGGTCGCTACCATTACCAGTTGGTCTG
GTGTCAAAAATAATAATAACCGGGCAGGCCATGTCTGCCCGTATTTCGCGTAAGGA
AATCCATTATGTACTATTTAAAAAACACAAACTTTTGGATGTTCGGTTTATTCTTTTT
CTTTTACTTTTTTATCATGGGAGCCTACTTCCCGTTTTTCCCGATTTGGCTACATGACA
TCAACCATATCAGCAAAAGTGATACGGGTATTATTTTTGCCGCTATTTCTCTGTTCTC
GCTATTATTCCAACCGCTGTTTGGTCTGCTTTCTGACAAACTCGGAACTTGTTTATTG
CAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCA
TTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGT
CTGGATCGACAAAGTCAAAGCGGCCATCAGATCTGCCGGTCTCCCTATAGTGAGTCG
TATTAATTTCGATAAGCCAGGTTAACCTGCATTAATGAATCGGCCAACGCGCGGGGA
GAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTC
GGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATC
CACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAG
GCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCT
GACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACT
ATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGAC
CCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT
CAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGC
TGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGT
CTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA
CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGC
CTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAG
TTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTA
GCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAG
AAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTT
AAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT
AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTT
ACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCAT
AGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGG
CCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC

AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCG
CCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTA
ATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTT
TGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCC
CATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAA
GTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTC
ATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGA
GAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACC
GCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGA
AAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA
CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACA
GGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATAC
TCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAG
CGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATT
TCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTA
TAAAAATAGGCGTATCACGAGGCCCTTACGTATTAATTAA

Yellow Highlighted Region: Ascl
Blue Highlighted Region: Nos1ap cis-regulator element to test
Red Highlighted Region: XhoI
Magenta Highlighted Region: Basal promoter
Green Highlighted Region: GFP

Slc6A14:
GGCGCGCCACGAAAGCCACTGACACCTTCCTTCATGTTTTCGTCCCGATAATATAAC
GCTTCAGTGACGTCAGCACACGGGCGAATGTTGCGGGACTACACGGTTCGCCTGTTT
GCAAAAATTCAATGAGACACTGTTCAGTAACTGCTGTTTTGGGTGAATGCTTCGTTT
TGGGAAATAGCCTTTTTCTGCACATGCATGAGTGCGGAATTGGGGCGTTAACTATAG
ACGTGTCCCTCGAGCAGCTGAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCG
GCAAAGCTTCGTGTATTGTACCGGCCCATTGTCAATCATGCAAACTTGATATTATATT
GACAAGAGAAGAAGGCAGTTTAAATTAAAACTCTAAAGTAGAGAGACATTAATCTC
AGCTGACAAGGCAGGTGGTCACAGTAAGTTCATTTAAATAGTTGGCCAACAATAGC
CTTTCCAAGAAAGTATTTTGTTCCAGGTCTATACAAAAATAACACACAACATGGCG
GCCGCAACCATGGCGGATCTGCGAGTACCGGACATTCCGCTCGCCTCGTGTGATGAT
GATGATATCGATAGTAATAAGAATTTGAGCAACCATTCATCAGACGAGAAACATCA
CTGCAACAGCAACAGCGACGAGGAACGTCTTCATGACGAGTTCTCTGGATCCCTTGA
GGACCTTGTCGGCAACTTTGACGAAAAAATTGCGGCATGCCTGAAGGACCACGAGG
TGACGACAGCGGATATTGCACCTGTGCAGATACGTACTCAAGAGGAAGTTATGAAT
GAAAGCCAAACATGGTGGACATTAACCGGAAACTTTGGAAACATTCAACCTCTCGA
CTTTGGAACCTCTTCGATATGTAAAAAGATGGCCGCAGCTCTGGACAGTGATTCATT
GAAAGACGACGCATCTACACGCCGAAGTATGACAAATTCCGATGATGAGGATCTTT
TACGACAACAAATGGATGTTCATCAAATGATTGGACATCATCATGGATCTACGGATA
CTGGTGGTGAAACACCTCCACAGACTGCTGATCAAGTTATCGAAGAAATTGATGAA
ATGTTACAGGTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGG
GGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCG

TGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATC
TGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTAC
GGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAG
TCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGG
CAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCA
TCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTG
GAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGG
CATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCG
CCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGAC
AACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGA
TCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGA
GCTGTACAAGTAAGAATTCCAGCTGAGCGCCGGTCGCTACCATTACCAGTTGGTCTG
GTGTCAAAAATAATAATAACCGGGCAGGCCATGTCTGCCCGTATTTCGCGTAAGGA
AATCCATTATGTACTATTTAAAAAACACAAACTTTTGGATGTTCGGTTTATTCTTTTT
CTTTTACTTTTTTATCATGGGAGCCTACTTCCCGTTTTTCCCGATTTGGCTACATGACA
TCAACCATATCAGCAAAAGTGATACGGGTATTATTTTTGCCGCTATTTCTCTGTTCTC
GCTATTATTCCAACCGCTGTTTGGTCTGCTTTCTGACAAACTCGGAACTTGTTTATTG
CAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCA
TTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGT
CTGGATCGACAAAGTCAAAGCGGCCATCAGATCTGCCGGTCTCCCTATAGTGAGTCG
TATTAATTTCGATAAGCCAGGTTAACCTGCATTAATGAATCGGCCAACGCGCGGGGA
GAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTC
GGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATC
CACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAG
GCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCT
GACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACT
ATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGAC
CCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT
CAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGC
TGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGT
CTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA
CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGC
CTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAG
TTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTA
GCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAG
AAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTT
AAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT
AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTT
ACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCAT
AGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGG
CCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC
AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCG
CCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTA
ATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTT
TGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCC
CATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAA

GTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTC
ATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGA
GAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACC
GCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGA
AAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA
CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACA
GGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATAC
TCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAG
CGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATT
TCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTA
TAAAAATAGGCGTATCACGAGGCCCTTACGTATTAATTAA

Yellow Highlighted Region: AscI
Blue Highlighted Region: Slc6A14 cis-regulator element to test
Red Highlighted Region: XhoI
Magenta Highlighted Region: Basal promoter
Green Highlighted Region: GFP

Vanabin4:
GGCGCGCCAACAAAACCCAGCGGCGTCTGAGCGGTTCGAGCTCGAGCACCGCAGAC
TAAGATTCTGTGGGTATGTTTTCGACTCAATAACCCCCACCAGCTCAAAACAACAAT
ATTTTCCTCCTTTAATTCTCTTCAACGTTTCTTTCTTTTATTCCACCACCGACGTGAAG
AGAGAAAAAACGCAGGAGGTGGTTCGAACATTCTAGCAAACGCGGGTCAAACGAG
GCAGCAAGACTCGAGCAGCTGAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATC
CGGCAAAGCTTCGTGTATTGTACCGGCCCATTGTCAATCATGCAAACTTGATATTAT
ATTGACAAGAGAAGAAGGCAGTTTAAATTAAAACTCTAAAGTAGAGAGACATTAAT
CTCAGCTGACAAGGCAGGTGGTCACAGTAAGTTCATTTAAATAGTTGGCCAACAATA
GCCTTTCCAAGAAAGTATTTTTGTTCCAGGTCTATACAAAAATAACACACAACATGG
CGGCCGCAACCATGGCGGATCTGCGAGTACCGGACATTCCGCTCGCCTCGTGTGATG
ATGATGATATCGATAGTAATAAGAATTTGAGCAACCATTCATCAGACGAGAAACAT
CACTGCAACAGCAACAGCGACGAGGAACGTCTTCATGACGAGTTCTCTGGATCCCTT
GAGGACCTTGTCGGCAACTTTGACGAAAAAATTGCGGCATGCCTGAAGGACCACGA
GGTGACGACAGCGGATATTGCACCTGTGCAGATACGTACTCAAGAGGAAGTTATGA
ATGAAAGCCAAACATGGTGGACATTAACCGGAAACTTTGGAAACATTCAACCTCTC
GACTTTGGAACCTCTTCGATATGTAAAAAGATGGCCGCAGCTCTGGACAGTGATTCA
TTGAAAGACGACGCATCTACACGCCGAAGTATGACAAATTCCGATGATGAGGATCT
TTTACGACAACAAATGGATGTTCATCAAATGATTGGACATCATCATGGATCTACGGA
TACTGGTGGTGAAACACCTCCACAGACTGCTGATCAAGTTATCGAAGAAATTGATGA
AATGTTACAGGTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCG
GGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGC
GTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCAT
CTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTA
CGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAA
GTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACG
GCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGC

ATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCT
GGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACG
GCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTC
GCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGA
CAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCG
ATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACG
AGCTGTACAAGTAAGAATTCCAGCTGAGCGCCGGTCGCTACCATTACCAGTTGGTCT
GGTGTCAAAAATAATAATAACCGGGCAGGCCATGTCTGCCCGTATTTCGCGTAAGG
AAATCCATTATGTACTATTTAAAAAACACAAACTTTTGGATGTTCGGTTTATTCTTTT
TCTTTTACTTTTTTATCATGGGAGCCTACTTCCCGTTTTTCCCGATTTGGCTACATGAC
ATCAACCATATCAGCAAAAGTGATACGGGTATTATTTTTGCCGCTATTTCTCTGTTCT
CGCTATTATTCCAACCGCTGTTTGGTCTGCTTTCTGACAAACTCGGAACTTGTTTATT
GCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGC
ATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATG
TCTGGATCGACAAAGTCAAAGCGGCCATCAGATCTGCCGGTCTCCCTATAGTGAGTC
GTATTAATTTCGATAAGCCAGGTTAACCTGCATTAATGAATCGGCCAACGCGCGGGG
AGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCT
CGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTAT
CCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAA
GGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCC
TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC
TATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGA
CCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTC
TCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGG
CTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG
TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA
CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGC
CTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAG
TTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTA
GCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAG
AAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTT
AAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT
AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTT
ACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCAT
AGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGG
CCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC
AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCG
CCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTA
ATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTT
TGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCC
CATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAA
GTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTC
ATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGA
GAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACC
GCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGA
AAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA

CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACA
GGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATAC
TCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAG
CGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATT
TCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTA
TAAAAATAGGCGTATCACGAGGCCCTTACGTATTAATTAA

Yellow Highlighted Region: AscI
Blue Highlighted Region: Vanabin4 cis-regulatory element to test
Red Highlighted Region: XhoI
Magenta Highlighted Region: Basal promoter
Green Highlighted Region: GFP

Acknowledgements

I wish to express my sincere thanks to Dr. Alberto Stolfi, for providing me with the opportunities and guidance needed to complete this theis. I am also grateful to Dr. Liang Han for acting as my second reader, and sharing her valuable expertise and feedback in this process.

I also take this opportunity to express my extreme gratitude to all of the Stolfi lab, especially Florian Razy- Krajka and Susanne Gibboney for their help and support in starting my undergraduate research experience, and giving me such welcoming and patient support as I learned what it means to do research.

Finally, I'd like to acknowledge Ryan Longood for his continued encouragement, his computer science skills, and his willingness to teach and troubleshoot code with me.