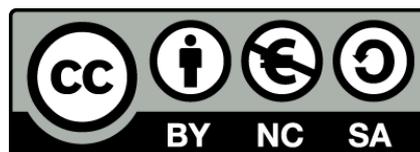# Autonomous and non-autonomous regulation on planarian growth and regeneration: *Smed-bls*, canonical Wnt signalling and Fox family

Eudald Pascual Carreras

Salid y disfrutad.

Johan Cruyff

UNIVERSITAT DE BARCELONA

Facultat de Biologia

Departament de Genètica, Microbiologia i Estadística

Programa de Doctorat de Genètica

2019

Tesis Doctoral presentada per

Eudald Pascual Carreras

amb el nom

Autonomous and non-autonomous regulation of planarian growth and regeneration: *Smed-bls*, canonical Wnt signalling and Fox family.

Par optar al títol de doctor per la Universitat de Barcelona

Tesi doctoral realitzada al Departament de Genètica, Microbiologia i Estadistica sota la direcció de la Doctora Teresa Adell i el Doctor Emili Saló.

Codirectora	Director i Tutor	Autor

Dr. Teresa Adell	Dr. Emili Saló	Eudald Pascual

Barcelona, 30 de Setembre de 2019

Abstract

Development requires an increment of cell growth and cell number, concomitant to a tightly control of cell differentiation. Thanks to cell communication, cells can be spatiotemporal patterned to acquire the required fate. Planarians are a unique model to study developmental processes due to their ability to regenerate and modulate their body size according to the nutrient availability. This body plasticity is based on the presence of pluripotent adult stem cells (neoblasts) and the continuous activation of the intercellular communication mechanisms. This active regulation of stem cells fate make them perfect models to study processes as growth, patterning, differentiation cell proliferation or cell death. In this thesis we have studied different molecular mechanisms that control planarian growth and pattern.

We have described a novel gene family, *blitzschnell* (*bls*), formed by de novo and taxonomically restricted genes, which control cell number trough the regulation of cell proliferation and cell death. Nutrient intake controls its expression suggesting that *bls* family have evolved in planarians as a mechanism by which to restrict cell number in nutrient-fluctuating environments. During growth and regeneration, planarians are not only able regulate their body and organ size accordingly but they also maintain a proper pattern. This regulation is mediated by different signalling centres that specify different regions along the 3 body axes (AP, DV and ML). Particularly, after an amputation, the anterior and the posterior planarian tips behave as organizers (signalling centre), specifying the fate of each planarian pole. The anterior organizer is defined by *notum* (a Wnt inhibitor) and the posterior by *wnt1* expression. The inhibition of any of those elements leads to a shift in polarity. During the first hours of regeneration both *notum* and *wnt1* are expressed in both poles, and it's around 36 hours that their expression becomes restricted to their respective tip. To decipher the molecular interactions that restrict the expression of *wnt1* to the posterior tip and confer the organizing activity we used genome wide approaches. ATAC-seq and RNA-seq analysis of regenerating wild-type and *wnt1* (RNAi) planarians allowed the identification of specific Cis-Regulatory Elements (CREs) of posterior regeneration. We found that already at 12 hours of regeneration the accessible CREs in posterior and anterior blastemas have essentially changed, indicating that specific posterior chromatin changes induced by amputation occur much earlier than the formation of

the organizers. Furthermore, we have identified specific transcription factors (TF) of the Otx and Fox families, which are enriched in posterior CREs. Particularly, *pitx* and *foxG* regulates wnt1+ cells and are essential for the specification of the posterior cells.

TFs regulate patterning events and developmental specification, particularly the Fox Family exerts crucial roles defining cell types of all germ cell layers or regulating cell cycle. Before this Thesis, poorly was known about the Fox family in *Schmidtea mediterranea* (*Smed*) neither in the Lophocotrozoan clade. In this study we have identified 27 Fox genes in *Smed*, classified in 13 families: A, At, C, D, E, G, L1t, QD, J1, N2/3, Nt, O and P. We have performed an extensive phylogenetic study of the family to understand the evolution of the Fox family in this clade. Furthermore, we have studied the sequence, expression and function of several planarian Fox genes.

Overall, we studied different molecular mechanisms that regulate planarian growth and regeneration, and that provide novel data concerning development and evolution.

Index

GLOSSARY OF FIGURES

INTRODUCTION

RESULTS

without changing cell size.
Figure R1.17: Cell death and cell proliferation dynamics after feeding.
Figure R1.18: *bls2*, *bls3* and *bls5* are down-regulated by food ingestion.
Figure R1.19: *bls2/3/5* (RNAi) animals show an increase in proliferation after feeding.
Figure R2.1: RNAi soaking protocol in *wnt1* (RNAi) animals.
Figure R2.2: *wnt1* expression during regeneration in wild type and irradiated animals.
Figure R2.3: Hierarchical clustering of RNA-seq libraries.
Figure R2.4: RNA-seq of *wnt1* (RNAi) and control animals reveal differentially expressed genes at different regenerating time points.
Figure R2.5: *wnt1* (RNAi) during posterior regeneration generates different expression pattern profiles.
Figure R2.6: *wnt1* (RNAi) produces genetic changes during regeneration.
Figure R2.7: *wnt1* (RNAi) affects genes involved in different stages of the regenerative respone.
Figure R2.8: RNAi of *wnt1* affects posterior gene expression patterns.
Figure R2.9: *wnt1* (RNAi) and *βcat1* (RNAi) RNA-seq present similar up and down-regulated genes.
Figure R2.10: Putative cWNT target genes display different regenerative expression patterns.
Figure R3.1: ATAC-seq of anterior and posterior wounds show specific accessible chromatin regions.
Figure R3.2: Accessible chromatin landscape after amputation at 12 hR.
Figure R3.3: ChIP-seq of anterior and posterior blastemas reveals specific enhancers.
Figure R3.4: Specific chromatin regions are open during anterior and posterior regeneration.
Figure R3.5: *wnt1* and *notum* (RNAi) change the genomic landscape.
Figure R3.6: *wnt1* inhibition changes transcription factor motif accessibility.
Figure R3.7: *pitx* (RNAi) animals lack *wnt1* and present a tailless phenotype.
Figure R3.8: *foxG* is expressed in a subset of muscle cells and coespress with *wnt1*.
Figure R3.9: *foxG* (RNAi) animals show a tailless phenotype.
Figure R3.10: *foxG* (RNAi) animals lavk *wnt1* and posterior markers expression.
Figure R4.1: *foxK* genes are expressed in the nervous system.
Figure R4.2: RNAi of *foxk* genes generates anterior and posterior defects.
Figure R4.3: *foxK1-2.1* (RNAi) animals show a tailless phenotype.
Figure R4.4: *foxK1-2.1* (RNAi) animals show differences in posterior markers expression.
Figure R4.5: *foxK1-2.1 (*RNAi) animals show defects in the nervous system.
Figure R4.6: FOXK1-2.1 and FOXK1-2.2 are similar at aminoacidic level
Figure R4.7: *foxK1-2.1* and *foxk1-2.2* double (RNAi) animals show a tailless-like phenotype and an increment of *wnt1* expression.
Figure R4.8: *foxK1-2.1* and *foxk1-2.2* (RNAi) animals show a stronger phenotype.
Figure R4.9: *Smed-foxK* genes could interact with *Smed-dvl* genes.
Figure R4.10: *foxK1-2.1* and *wnt1* (RNAi) animals show an increment of tailless phenotype.
Figure R5.1: The ML phylogenetic tree reveals a gene and family losses and some gene duplications in *Schmidtea mediterranea*.
Figure R5.2: Distribution of Fox homologs in Metazoan clade.
Figure R5.3: Distribution of Fox homologs in Metazoan species.
Figure R5.4: Fox family evolution in Lophotrochozoan clade.
Figure R5.5: Distribution of Fox homologs in Platyhelminthes clade.
Figure R5.6: Distribution of Fox homologs in Platyhelminthes species.
Figure R5.7: Summary of *fox* genes in *Schmidtea mediterranea*.
Figure R5.8: Evolution of FoxQ2 and FoxQD families.
Figure R5.9: Fox are not clustered in planarian genome.
Figure R5.10 : Domains of Fox proteins in *Schmidtea mediterranea*.
Figure R5.11: *fox* genes present different expression patterns.
Figure R5.12: *Semd-foxN* genes are coexpressed with *Smed-ror* and *Smed-roboC*.
Figure R5.13: Inhibition of *foxN* genes generate animal with bad anterior regeneration.
Figure R5.14: *foxN* (RNAi) animals show small brains and mistargeting of the optic chiasm.

DISCUSSION

GLOSSARY OF ABBREVIATIONS

Acessible Chromatin Regions	ACR
Amino acids	Aa
Anterior	Ant
Anteroposterior	AP
Adult stem cells	ASC
Assay for Transposase-Accessible Chromatin - sequencing	ATAC-seq
blitzschnell	bls
Bone Morphogenetic Protein	BMP
bais pairs	bp
Body-Wall Muscle	BWM
Coiledcoil domain	CC
Chromatin Immunoprecipitation - sequencing	ChIP-seq
Centimiter	cm
clonogenic Neoblast	cNeoblast
Central Nervius System	CNS
Core Promoters	CP
5'—C—phosphate—G—3'	CpG
Cis Regulatory Elements	CRE
canonical WNT pathway	cWNT
4 ',6-diamidino-2-fenilindol	DAPI
DNA Binding Domain	DBD
Distal promoters	Dis
Desxaribonucleic acid	DNA
days of regeneration dR	dR
double stranded RNA	dsRN
Dorsoventral	DV
Dishevelled	DVL
Dorsoventral muscle	DVM
emerging Anterior enhancers	eAnt
Extracellular matrix	ECM
Epidermal Growth Factor	EGF
emerging Posterior enhancers	ePost
Extracellular signal-regulated kinase	ERK
Embryonic stem cells	ESC
fold change	fc
Flase discovery rate	FDR
forkhead associated domain	FHA
Fluorescent in situ hybridization	FISH
forkhead domain	FKD
forkhead gene	fox

follistatin	fst
Frizzled 4	Fz4
Green fluorescent protein	GFP
Gene Ontology	GO
Gene regulatory networks	GRN
Groucho/transducin-like enhancer	Groucho/TLE
Growth Zone	GZ
27th lysine residue of the histone H3 protein	H3K27ac
Hedgehog	hh
High Mobility Group	HMG
hours of regeneration	hR
increasing Anterior ehancers	iAnt
Intestinal muscle	IM
increasing Posterior ehancers	iPost
internal dtructural disorders	ISD
c-Jun N-terminal kinase	JNK
lipoprotein receptor-related protein	LRP
milimiters	mm
square milimiters	mm2
messenger RNA	mRNA
mammalian target of rapamycin	mTOR
myogenic Differentiation	myoD
Neoblast	NB
non-canonical WNT signalling	ncWNT
Nuclear Localization Signal	NLS
Homeobox Orthodenticle protein	Otx
pValue adjusted	padj
positional control genes	PCG
phospho-histone 3	PH3
Positional information	PI
Paired-like homeodomain transcription factor	pitx
Posterior	Post
Proximal promoters	Pro
quantitative PCR	qPCR
regenerating Anterior enhancers	rAnt
RNA interference	RNAi
regenerating Posterior enhancers	rPost
Stem cells	SC
Single Cell Sequencing	SCS
Signal peptide domain	SP
Transcription factor	TF
Transforming growth factor beta	TFG-β

Tumor necrosis factor alpha	TNF α
Taxonomically Restricted Gene	TRG
teashirt	tsh
transcription starting site	TSS
Terminal deoxynucleodityl transferase dUTP nick end labeling	TUNEL
Ventral Nerve Chords	VNC
Whole mount in situ hybridization	WISH
Wingless-related integration site	WNT
wild type	WT
beta catenina 1	β-cat1
histone H2B	h2b
mammalian target of rapamycin	mTOR
Proprotein convertase 2	pc2
P-element Induced WImpy	piwi
teashirt	tsh
AKR mouse strain that develops thymoma	AKT
c-Jun N-terminal kinases	JNK

GLOOSARY OF SPECIES

Amphimedon queenslandica	*Amq*
Branchiostoma lanceolatum	*Bla*
Bothrioplana semperi	*Bose*
Capitella teleta	*Cap*
Catenulia	*Cate*
Crassostrea gigas	*Cgi*
Dugesia japonica	*Djap*
Dendrocoelum lacteum	*Dla*
Drosophila melanogaster	*Dme*
Echinoccocus multiocularis	*Emu*
Geocentrophora applanta	*Geap*
Gyrodactylus salaris	*Gsa*
Helobdella robusta	*Hbo*
Homo sapiens	*Hsa*
Leptoplana linguna	*Lli*
Intoshia linei	*Ili*
Lingula anatine	*Lan*
Lottia gigantean	*Lgi*
Mesostoma lingua	*Meli*
Macrostomum lignano	*Mli*
Monocelis sp	*Mosp*
Nematostella vectensis	*Nvec*
Octopus bimaculoides	*Obi*
Ptychodera flava	*Plf*
Polycelis nigra	*Pni*
Polycelis tenius	*Pte*
Planaria torva	*Pto*
Suberites domuncula	*Sdo*
Saccoglossus kowalevskii	*Sko*
Schistosoma mansoni	*Sman*
Schmidtea mediterranea (asexual strain)	*Smed*
Schmidtea mediterranea (sexual strain)	*Smes*
Schmidtea polychroa	*Spol*
Strongylocentrotus purpuratus	*Spu*
Tribolium castaneum	*Tca*
Taenia solium	*Tso*

INTRODUCTION

Introduction

1. Developmental biology

Developmental biology studies the processes that govern animal development, including embryonic development, growth, sexual maturation, regeneration and aging (1). Different model organisms such as nematodes *Caenorhabditis elegans* and *Nematostella vectensis*, the fruit fly *Drosophila melanogaster*, the zebrafish (*Dania rerio*) and the frog *Xenopus laevis* (2) have used to study developmental processes.

The definitive body size of an organism is reached by increasing either cell number or cell size. When organisms achieve their final body size, sexual maturation starts. This process consists of a complex transformation from a sexual immature individual into a sexual mature one, capable of reproduction (3).

During adulthood metazoan species have programs that protect these from physiological dysfunctions and allow tissue maintenance. In mammals, this capacity is used to restore particular tissues, being known as tissue repair (4). When missing structures, more complex processes to restore them, such as developmental programs, are required. This process is referred as tissue regeneration. Both tissue repair and regeneration affect different tissues and require cell replacement on a large-scale (5).

In the last decades, the impairment of many developmental mechanisms has been related to human diseases (6,7), increasing the scientific community's attention.

1.1. Regulation of developmental mechanisms

As mentioned previously, the regulation of the developmental mechanisms is crucial to support healthy organisms, from the molecular, cellular and tissue level (8). From a cellular perspective the mechanisms that regulate development can be classified as autonomous or non-autonomous. Autonomous mechanisms refer to the cellular behaviour that depends on its own genetic expression, while non-autonomous mechanism refers to the cellular behaviours that do not rely on the function of its own genes or proteins, but on signals received from other cells or extracellular components (1).

1.2. Cell-cell communication

To allow cell interactions and the regulation of developmental processes as tissue repair or apoptosis, cells need to be coordinated. Cells perceive and correctly respond to surrounding environment, using a variety of signal molecules that are secreted or expressed on the membrane surface. Cell communication can be classified as: mechanical, defined by forces exerted on the cell and the forces produced by the cell (9); or biochemical, with signals being molecules such as proteins, lipids, ions and gases. Biochemical communication can be categorized based on the cell-cell communication distance (Figure I1.1): 1) Intracrine signals are produced by a cell and stay within. 2) Autocrine signals are released in the extracellular environment by signalling cell, affecting the same secreting cell. 3) Juxtacrine, secreted signals target adjacent cells or interact with extracellular matrix (10,11). Juxtacrine signals are transmitted trough cell membranes via protein or lipid. Tight junctions, gap junctions, desmosomes and cell adhesion are examples of juxtacrine signalling. (12). 4) Paracrine factors (signalling molecules) target cells in the vicinity of the emitting cell. And 5) endocrine signals

target distant cells, producing hormones that travel through the blood to reach all parts of the body.
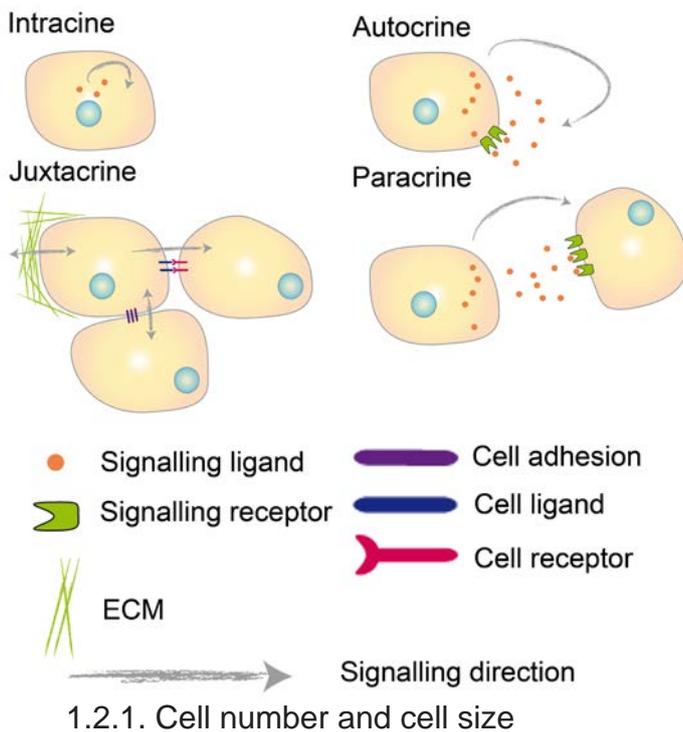
Figure I1.1: Cell communication. Cells are able to communicate with each other. Cell non-autonomous mechanisms implies extrinsic cues. These mechanisms can be paracrine when cells secret ligands that travel and bind receptors in surrounding cells. Autocrine, when ligands act at the cell source. And finally, cell-cell physical contact can occurs. Autonomous cell mechanisms are the ones generated by the cell itself.

1.2.1. Cell number and cell size

Embryonic development implies increases in cell size or cell number until reaching a definitive body size. Changes in cell size have been described in specific organs, e.g. liver and imaginal discs in *Drosophila* (13,14). However, regulation of cell number, achieved by modulating the balance between cell death and cell proliferation, is the main mechanism by which animals reach their definitive body size (15). In general, the main signalling pathways through to control growth, regulate cell proliferation and cell death in response to the nutritional environment. Studies in multiple species have identified the same key signalling pathways that appear to regulate body size. There are the JNK pathway, the Hippo pathway, and the insulin/Akt/TOR signalling network. The JNK signalling pathway controls cell death and proliferation, mainly in response to cellular stress (16). The Hippo signalling pathway regulates proliferation, apoptosis, and cell differentiation in response to mechanical stimuli (17,18). Genetic perturbation of both of these pathways (JNK or Hippo) leads to overgrowths or organ size changes. Nevertheless, these processes do not affect the overall body size (19,20). In contrast, activation of the insulin/Akt/TOR signalling network leads to increases in body size in animals as distant as *Drosophila* and mice (21). The insulin/Akt/TOR signalling is the most conserved molecular mechanism that relates nutrient intake, cell proliferation and cell growth. It can sense energy and amino acid levels, and as a consequence modulate different transcription factors (22,23).

1.2.2. Tissue patterning and organizers

During embryonic development, signalling factors (morphogens) (24) instruc and pattern tissues around them, therby generating positional information (PI). Morphogens can be grouped into four major families on the basis of their structure (25). These families are the fibroblast growth factor (FGF) family, the Hedgehog family, the Wingless (Wnt) family, and the TGF-β superfamily. Secreted proteins of those families were dicovered as "inducing factors" for classic embryologist experiments form last century (26–28). During development, morphogens source are known as organizers or signalling centres (25). Examples are the Spemann

organizer (29), Hensen's node in amniotes (30,31), the notochord (32), the zone of polarizing activity of the limb bud (33), and the mid-hindbrain boundary (34). Thanks to secreted molecules, organizers are able to instruct the surrounding cells (and tissue) to change their fate and/or pattern (Figure I1.2). It needs to be mentioned that Ethel Brown reported this same organizer behaviour 15 years before Spemann and Mangold published their results. Brown defined that in adult *Hydra* organisms, the tip head also presented organizer features (35).

Members of the TGF-β superfamily regulate some of the most important interactions in development, including: gastrulation, axis symmetry of the body, organ morphogenesis, and tissue homeostasis in adults. When TGF-β ligands bind to either Type I or Type II receptors (36), members of the SMAD family (TF) are activated. The bone morphogenetic protein (BMP) is a family part of the TGF-β superfamily, which plays a role in regulating bone formation, cell division, apoptosis, cell migration, and differentiation (37). BMP ligands also specify the anterior/posterior axis, induce growth, and regulate homeostasis (38). Spemann and Mangold firstly describe organizer activity in a *Xenopus* embryo. During the last century was revealed that this organizer was expressing inhibitors of BMP (39,40).

The mechanisms described above help to regulate developmental processes that generate differences between tissues and boundaries, also referred to as spatiotemporal barriers. These new spatiotemporal barriers are crucial to create new cell fates and tissues specific function and identity
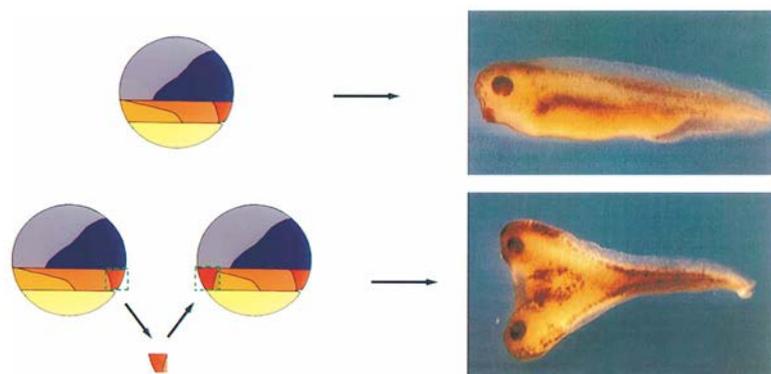
Figure I1.2: Organizer defines the identity during embryonic development. Experiment realized by Spemann and Mangold. Dorsal lip of an embryo is grafted into a ventral region of host embryo, resulting in the new axis formation. Adapted from 426.

1.3. Regeneration

Regeneration is the process by which, after an injury, tissue is able to produce new cells and regrow missing part (41). During this process, old tissue must be remodelled to adequate to the new one. The ability to regenerate cells, tissues, appendages or even the entire body is widely spread all over the animal kingdom (Figure I1.3).

1.3.1. Mechanisms of regeneration

Not all organisms have the same efficiency or potency to regenerate missing structures. Adult humans have a low regenerative capacity, just some organs and tissues have it, such as skin or liver. Liver regeneration is considered a compensatory hypertrophy of the pre-exiting tissue, function and mass recovery is accomplished by hepatocytes proliferation. In such process the exact morphology is not regained. By contrast, other vertebrates such as axolotls are able to regenerate an entire limb, retina and tails. In axolotl, after limb amputation, cells closely located to the wound start to dedifferentiate forming a proliferative blastema. Then, these cells differentiate accomplishing the required patterning and growth, of the new appendices (41). Zebrafish is another example that is able to regenerate fins, heart and spinal cord. In zebrafish, axonal regeneration after spinal cord injury is mediated by the proliferation

of glial cells, that migrate to the damage zone (42). Organisms like cnidarians as *Hydra* are able to regenerate the whole body after any amputation due to their highly proliferative stem cell population. In this type of regeneration the pre-existing tissue is remodelled in order to reshape the new body proportion (41). Summarizing, the regeneration capacity depends on different mechanisms: dedifferentiation, local proliferation or global proliferation (accompanied by extensive tissue remodelling).
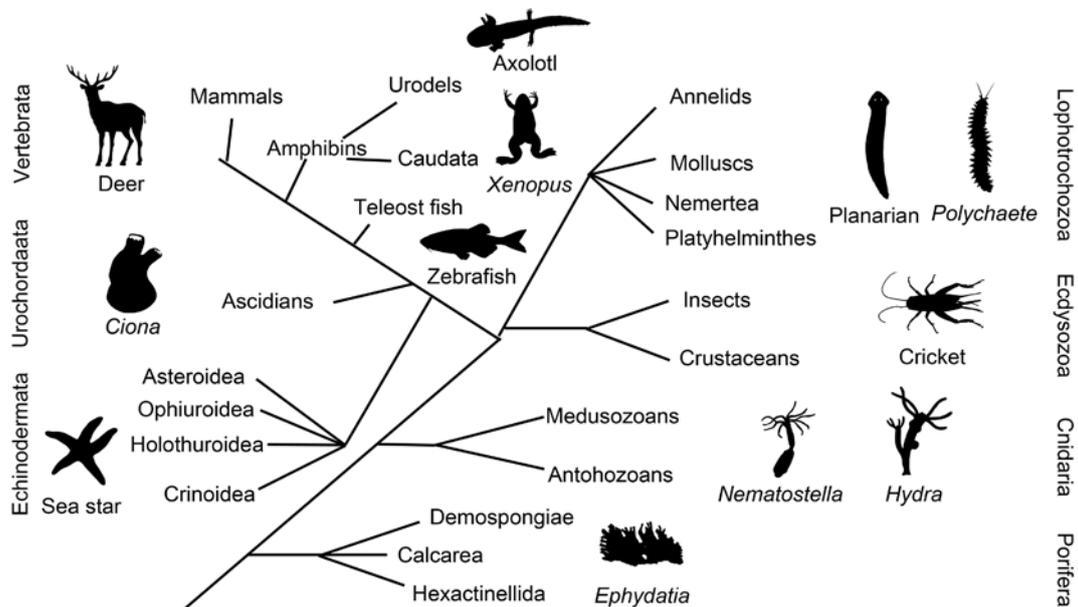
Figure I1.3: Species with high regenerative potential in Metazoan, and their classification in a phylogenetic tree. Animal silhouettes from Phylopic (www.phylopic.org). Adapted from (4)

It could be considered that regeneration results from the reactivation of embryonic development processes in adulthood (4,43). Galliot and Glia propose the term *development continuum*, which hypothesize that the genetic circuitries supporting development and regeneration should be highly similar, although not identical. Thus, regeneration might re-deploy developmental mechanisms. In newt limbs, regeneration was demonstrated to recapitulate limb development, with Hh pathway exerting a crucial role (44). In urodels as well as *Xenopus* tadpoles, genetic programs that regulate development and regeneration of limbs share similarities (45,46). Thanks to new sequencing approaches underlying processes supported the *development continuum* idea. In *Xenopus* it was confirmed that developing and regenerating limbs employ FGF signalling (47). Haberman et al. sequenced embryos and day-6 regenerating tail blastemas, identifying that developmental genes related with cell proliferation, cell differentiation and cell-cell communication are shared between these (45). Results from RNA-seq of differentiated tissues and regenerating limbs from an axolotl also demonstrated that some developmental pathways were re-deployed (48). Transcriptomic analysis corroborates that lizard tail regrowth involves the activation of conserved developmental pathways (49). In zebrafish it has been reported that adult caudal and larval fin underlie the same genetic profiles and share pathways, such as WNT and FGF pathways (50).

Although, it has been broadly demonstrated that mechanisms involved during development are also required for regeneration in different organisms. Other reaserch groups also demonstrated the presence of genes specifically activated during development. During motor nerve development in zebrafish; Notch signalling is crucial for its development but not in regeneration (51). Another example is Hoxc10L in axolotls, which is not expressed during forelimb development, but is expressed during forelimb regeneration (52). Additionally, other

research groups have reported genes only required for regeneration in zebrafish, where *sox2* is required for hair cell survival and regeneration (53).

1.3.2. Adult stem cells

Development, homeostasis and regeneration require a cell production to generate an organisms, restore cells for turn-over and regrow missing structures (5). The source of the new regenerated cells of some adult organisms are stem cells (SCs). For example *Hydra* can regenerate a whole body based on SCs. To enable this cell reproduction capacity, two features have been described in adult stem cells (ASCs). Firstly, stem cells divide maintaining themselves. This is referred to as self-renewal. Secondly, stem cells can differentiate all cell types to the tissue where are displayed. The ability to generate daughter cells that can undergo differentiation into different cell types is called pluripotency, multipotency or unipotency, depending on the amount of cell types the SC can differentiate into. Combination of self-renewal and potency define cell's stemness capacity (54).

SCs can divide symmetrically when their numbers need to be expanded, as it is happening during embryonic development or tissue regeneration. The way in which SCs balance self-renewal with the production of daughter cells is unique for each tissue. The microenvironment that regulates SCs is known as stem cell niche. The niche is crucial in supplying essential growth factors and adhesion anchors. Additionally, a niche also provides cues to asymmetric divisions necessary after the loss of stem cells or differentiated cells. The niche is formed by the vicinity cells and the extracellular matrix (ECM) (55), which generates ligands such as Notch, WNT and TFG-β, and a metabolic state that is indispensable to create and maintain the niche (56). An example are crypts in the intestinal epithelium which present a stem niche allowing the entire replacement in few days (Figure I1.4).
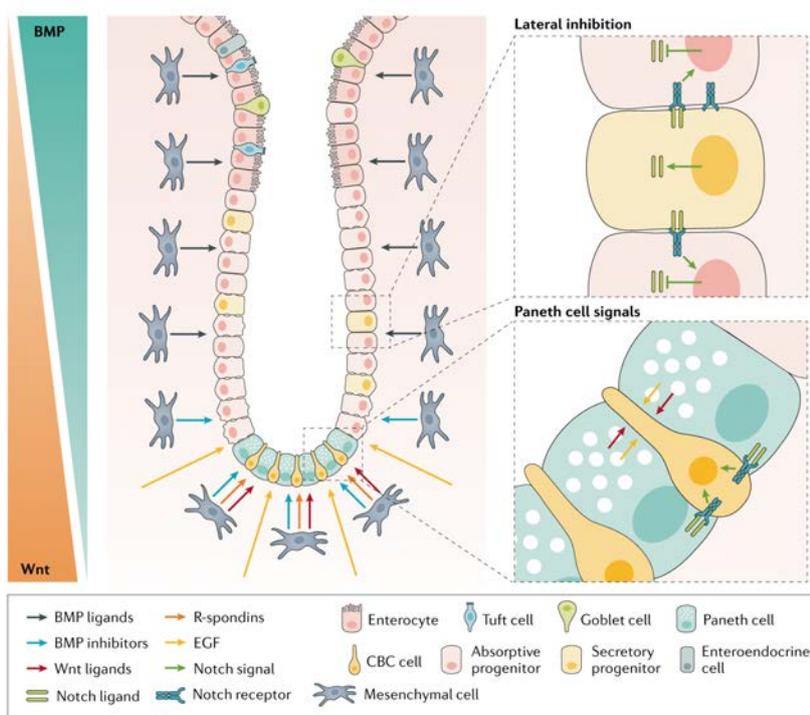
Figure I1.4: Main regulatory signals and their sources in the intestinal crypt. ASC situated in the intestinal crypts, referred to as crypt base columnar (CBC) cells. CBC cells are controlled by a surrounded niche. Paneth cells and surrounding mesenchymal cells secrete important signalling molecules such as WNT, EGF, TFG-β and Notch After an intestinal injury, WNT signalling seems to control stemness of ISCs and also participates in the dedifferentiation of late intestinal progenitors. Adapted from (427)

2. Planarian: a model organism to study regeneration and cell communication

Planarians are an ideal model to study regeneration and cell-cell communication. These organims are able to regenerate any missing body structure after any amputation. Planarian plasticity is sustained by a population of pluripotent adult SCs, known as neoblasts, which can differentiate into any planarian cell type. After an amputation, different processes such as cell proliferation, cell death, cell differentiation and patterning need to be regulated (57,58). Cell communication is crucial to properly integrate newly build cells in space and time. In planarian the following signalling pathways have been described as essential for cell communication: WNT (59–61), TFG-β (62,63) or EGF (64–66). Moreover, other pathways that sense environmental changes, like cell-cell contact or nutrient state are also essential in planarians to trigger regeneration. These are the JNK (67), mTOR (68–70), Insulin (71), AKT (72), PTEN (73), Hippo (74,75) pathways.

After any injury, planarians are able to close the wound and regenerate the missing part, meaning that two parts will regenerate all the missing structures, and will reshape the pre-existing tissue making tissues and organs smaller accordingly to the new size of the planarian. This body flexibility is also observed during their normal homeostasis, allowing planrian's to grow and degrow (shrink) depending on food availability (Figure I2.1). Such plastic capacities are sustained by the presence of neoblasts (76–79) as well as the continuous activation of cell communication mechanism. Regeneration and remodelling processes taken place during few days giving the opportunity to study key autonomous and non-autonomous mechanisms of regeneration and growth. Particularly, organ regeneration e.g. of brain, gut, eyes or epidermis can be carefully analyzed.

Their tremendous plasticity makes planarians a perfected candidate to study the regeneration process and the networks that regulate it. In this thesis, I used planarian as a model system to understand how cell communication mechanisms regulate regeneration and growth.
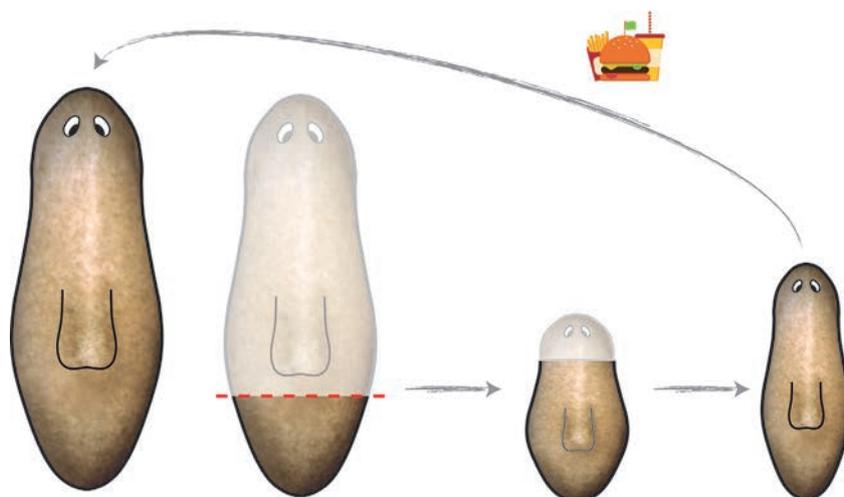
Figure I2.1: Planarian regeneration and growth. After an amputation, planarian fragments regenerate missing body parts and will result in a smaller animal. Small proportionate animals can eat and grow toward the original size.

2.1. *Schmidtea mediterranea*

Planarians are Lophotrocozoans, which together with Ecdyzozoa form the Protostomia clade. They belong to the Phylum Platyhelminthes, the Tricladida order and the Dugesiidae Family (80). Planarians are free-living flatworms found in many habitats; freshwater, marine and terrestrial. According to the habitats and transcriptomic data Riutort et al. proposed three taxonomic groups (suborders) within the Tricladida: Continenticola (freshwater planarians), Terricola (land planarians), and Maricola (marine planarians) (81). *Schmidtea meditteranea* is a Dugesiidae and thus belongs to Continenticola (Figure I2.2).
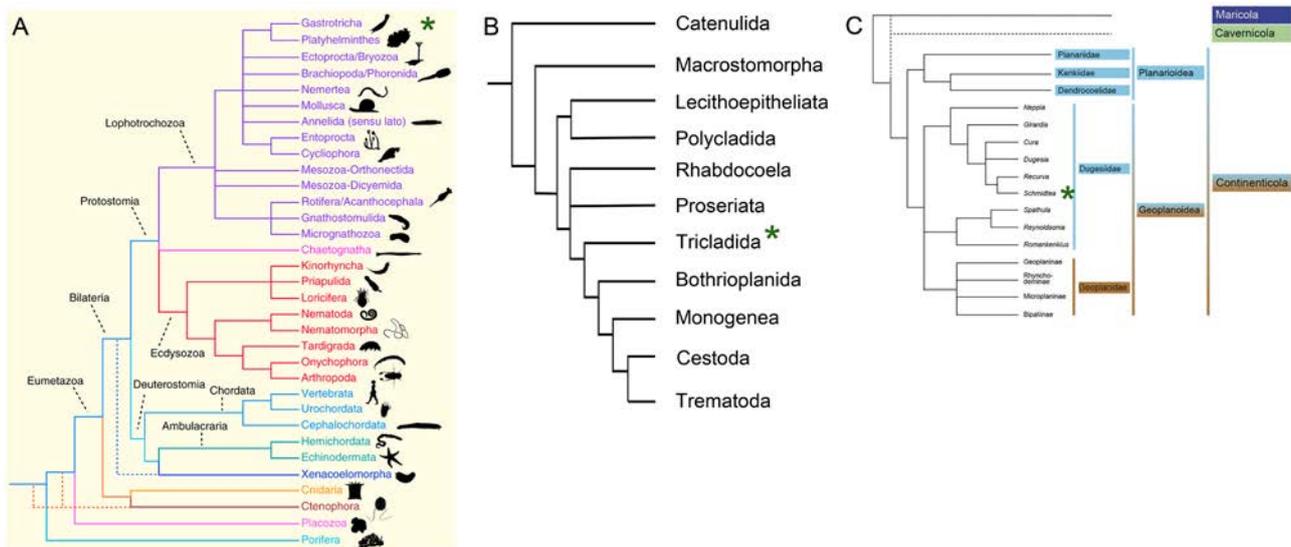
Figure I2.2: *Schmidtea mediterranea* evolutionary position. A phylogenetic tree of **(A)** the metazoan clade. Adapted from (428). **(B)** Platyhelminthes clade. Adapted from (80). **(C)** Tricladida Order. Color codes: light blue, freshwater; deep blue, marine; green, cavernicolan; brown, terrestrial. Adapted from (12). In each phylogenetic tree, *Schmidtea mediterranea* is indicated with an asterisk.

Planarians present a bilateral symmetry with two main axis, the anteroposterior (AP) and the dorsoventral (DV). Planarians are acoelomates; presenting a mass of cells among organs named parenchyma (82–85). They are triploblastic organisms with complex a tissue and organ structure. The central nervous system (CNS) is formed by two anterior lobes and two ventral nerve cords. Two eye spots are connected by an optic chiasm (86). Muscle fibers are distributed in four layers under the epidermis (87). The excretory system is composed by widely distributed protonephridia tubes (85). Food is taken from a pharynx which evaginates through a mouth, which also functions as an anus to excrete leftovers (57). The digestive system is formed by a blind gut with one anterior branch and two posteriors, which join altogether in the esophagus. The absence of a circulatory, skeletal and respiratory system produce the flat adaptation that allows the 0_2 diffusion trough the epidermis. Planarian species can be sexual or asexual. Sexual species present both female and male reproductive systems: ovaries and testis (57). They are hermaphrodite, and after a cross fertilization, they laie polyembrionic eggs called cocoons (88). Asexual planarians reproduce by fissioning their tails.

Nowadays, the most common planarian to address regenerative questions is *Schmidtea mediterranea* (*Smed*). Interestingly, *Smed* presents sexual and asexual strains. The asexual one is the most commonly used for regenerative studies. The features that make *Smed* a successful model are: 1) its fast regeneration capacity (less than 10 days), 2) is easy and cheap maintenance in the laboratory, 3) a clonal population can be obtained, reducing genetic variability, 4) availability of the genome, allowing to resolve epigenetic, gene regulatory or evolutionary questions (84); 5) the availability of a number of transcriptomes from different

regenerating time points and from different genetic backgrounds (90,91); and 6) well established molecular techniques, such as RNAi experiments (knockdown) (92), immunohistochemistry and westernblot (59,93,94), ISH (95), FISH (100) and FACS (96). Furthermore, very recently, sequencing at a single cell level has been successfully applied to planarians, offering a deeper understanding of the different cell types and unprecedented possibilities for understanding the cellular and molecular base of planarian plasticity.

2.2. Planarian anatomy

Since the publication of the planarian cell type atlas by Fincher et al. (98) and Plass et al. (99), the knowledge of planarian anatomy has substantially improved. In the next sections, I will be described planarian anatomy integrating those new findings with histological and cellular data. Plass et al.'s study will be taken as a guidance to describe cell types, and Fincher et al.'s results study will be used when extra information is provided. It should be mentioned that both studies facilitate a webpage to visualize the anatomy of the planarian: https://digiworm.wi.mit.edu and https://shiny.mdc-berlin.de/psca/.

2.2.1. Neural

The planarian CNS is formed by a bilobed arch-shaped cephalic ganglia connected by a single anterior commissure (100). It can be structurally divided into a central spongy region (neuropil) and lateral branches that project towards the periphery of the head. A pair of ventral nerve cords (VNC) runs from the head to the tip of the tail and is interconnected by transverse commissures (Figure I2.3). Neurons are also found innervating other tissues such as pharynx, subepidermal regions and the intestine (101). The brain is comprised by a myriad of different neuron types and glia (102,103) (Figure I2.4B). Single cell sequencing (SCS) reveal 8 major neuron types: ChAT 1, ChAT 2, cav-1, spp-11. npp-18, GABA, otf1 and otf2; resembling the previously identified neuron types: serotonergic (104–107), dopaminergic (100,105,108), glutamatergic (105,109), octopaminergic (110,111), gabaergic (105,111), cholinergic (105).
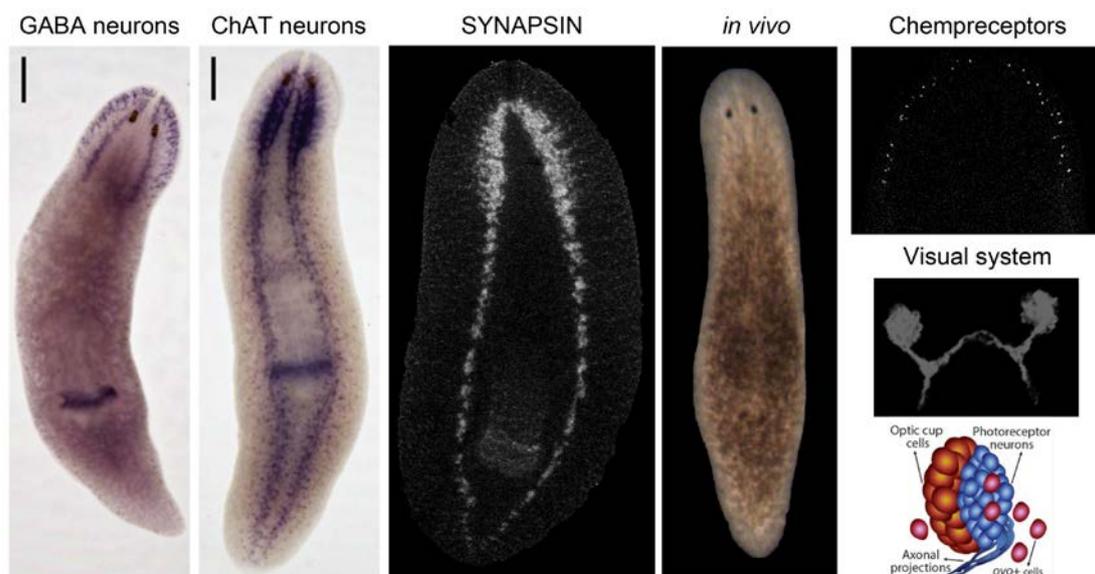
Figure I2.3: Planarian neural tissues. Immunostaining using an antibody against SYNAPSIN (3C11) showing the CNS: cephalic ganglia and two nerve cords. The pharynx plexus innervation can also be observed. in situ hybridization using twp neural markers show planarian neural diversity. Adapted from (99). Fluorescent in situ hybridization using an chemoreceptor marker showing its head location. The visual system is visualized by immunostaining using an antibody against ARRESTIN (VC1), observing the photoreceptors and the optic chiasm. Schematic cartoon of right eye, the disposition of the optic and photoreceptors cells is observed. Progenitor cells (*ovo*+) are also added. Adapted from (112)

Neural cells include sensorial cells as: photoreceptors, mechanoreceptors and chemorecep-tors. Those cells are mostly located in head area. The visual system is formed by rhabdome-ric photoreceptor neurons, pigment cells, and a pigmented optic cup structure. Rhabdomeric cells connect via axon tracts to the brain, and between them thought forming an optic chiasm (112). Pigment cells are distributed in a cub shape around the photoreceptors to protect them (113). Chemoreceptors allow planarians to localize food in the media.

2.2.2. Epidermis

Planarian epidermis is a cell monolayer composed by ciliated and non-ciliated cell types deposited in a basal lamina (114). The epidermis produces mucus protecting the planarian against external insults. Epidermal ventral ciliated cells are used for gliding locomotion. In-terestingly, three different epidermal cell populations have been identified according to their localization: dorsal, ventral and DV boundary (99) (Figure I2.4).
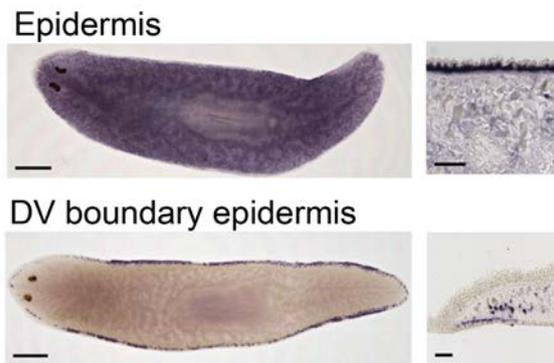
Figure I2.4: Planarian epidermis. *in situ* hybridiza-tion in whole mount and tissue sections, using epi-dermis and dorsoventral boundary markers. Scale bars: 500 μm for whole mount *in situ* hybridizations, 100 μm for in situ on sections. Adapted from (99)

2.2.3. Intestine

A highly branched blind gut distributes nutrients and connects to a muscular pharynx located centrally (66) (Figure I2.5A). The pharynx evaginates through a ventral opening that functions as a mouth and anus. The planarian gastrodermis is a monostratified epithelium composed of two cell types, absorptive phagocytes and secretory goblet cells. These are surrounded by an enteric muscular plexus (116–118) (Figure I2.5B). The function of phagocytes and goblet cells is to release enzymes in the lumen that facilitate nutrient digestion. Recently, it has been described that both intestinal cell types have different populations along the mediolateral axis (119). Fincher et al. described a third cell type defined as the outer intestine cell layer.
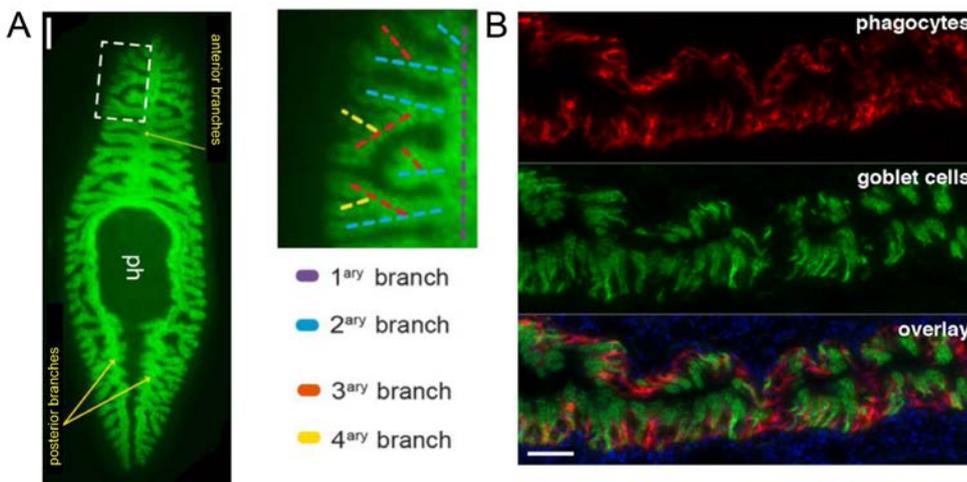
Figure I2.5: Planarian di-gestive system. (A) Fluo-rescent *in situ* hybridization using an intestine marker. Yellow arrows point to an-terior and posterior bran-ches respectively. ph is pharynx. Gut is ramified into secondary, tertiary and quaternary branches. Adapted from (66). **(B)** Double fluorescent *in situ* hybridization on sections using phagocytes and go-blet cells specific ribopro-bes. Their overlay shows not colocalization. Scale bars: 250 μm in A and 100 μm in B. Adapted from (99).

2.2.4. Phrotonephridia

The planarian excretory and osmoregulatory system consists of branched epithelial tubules (protonephridia) distributed throughout the entire body plan (120–122) (Figure I2.6), below the muscular plexus. Protonephridia are comprised of flame cells for filtering fluids, proximal and distal tubule cells, and a collecting duct.
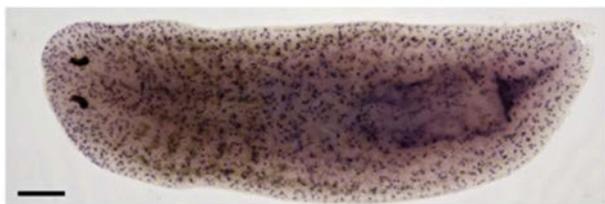
Figure I2.6: Planarian excretory system. *in situ* hybridization in whole mount using protonephridia marker observing its broad distribution. Scale bar is 100 μm. Adapted from (103)

2.2.5. Muscle

The body-wall musculature (BWM) contains circular, longitudinal, diagonal and longitudinal fiber, constituting the subepidermins in the presented order, from outside to inside (87) (Figure I2.7A, B). Structure has been described as a non-body wall muscle fiber that connects the dorsal and ventral planarian part (123) (Figure I2.7A). Altogether, this muscular system acts as a structural system for planarians. Different organs like the pharynx and the digestive are stabilized thanks to a musculature plexus (124) (Figure I2.7C, 7D). In addition to the mechanical function, all muscle fibers are the source of what are called positional control genes (PCGs). PCGs are secreted proteins that act as morphogens, instructing the fate of the receiving cells. PCGs are expressed in a well defined manner along the 3 planarian body axis (AP, DV and ML axis) and thus provide positional information during planarian homeostasis and regeneration. (125,126). Recently, is has been reported that muscle cells are also the major source of the extracellular matrix (ECM) and connective tissue. In particular they express a glycoprotein HMCN-1, which helps to maintain parenchymal cell localization. All the elements secreted by muscle cells are grouped and known as matrisome (127).
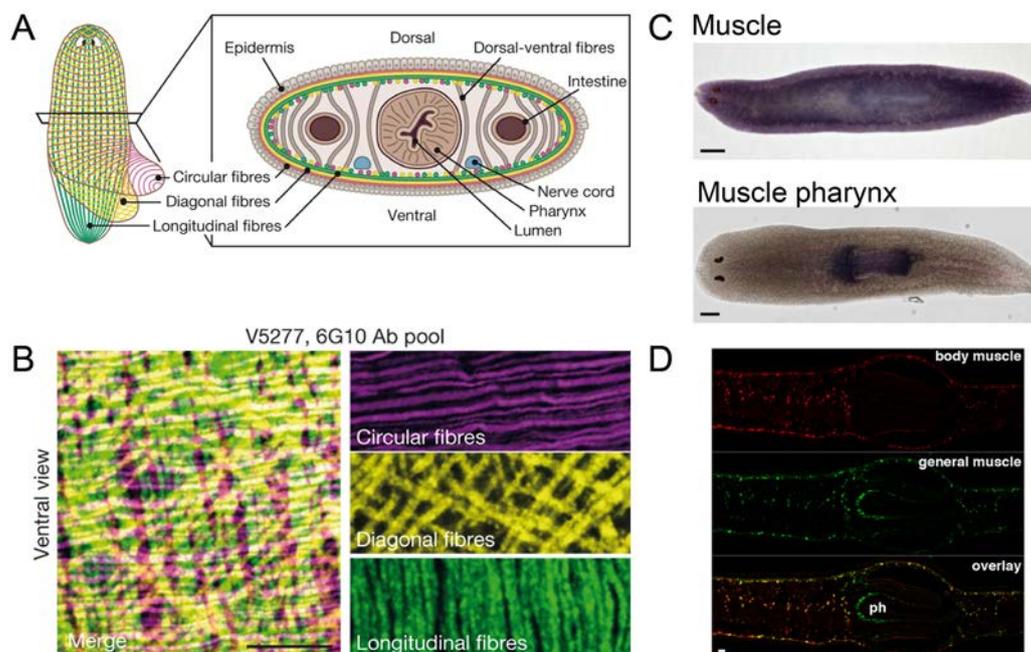
Figure I2.7: Planarian muscular system. (A) Diagram showing body wall muscle layers (circular, diagonal and longitudinal). Their disposition in transversal sections, are also indicating dorsal-ventral fibres. **(B)** Immunostaining using antibodies against muscle cells showing body wall muscle layers. Adapted from (184) **(C)** *in situ* hybridization in whole mount using muscle and muscle pharynx markers. **(D)** Double fluorescent *in situ* hybridization on sections using body muscle and general muscle riboprobes. Their overlay shows a colocalization with an exemption in the pharynx. Scale bars are 10 μm in (A), 250 μm in (B) and 100 μm in (C). Adapted from (99)

2.2.6. Pharynx

The planarian pharynx is a cylindrical muscle highly innervated to allow precise movement (85). It is located in an epithelial cavity and connects to the intestine via the esophagus (Figure I2.8). SCS data confirmed that the pharynx is composed by specific muscular cell type, and neural cells.
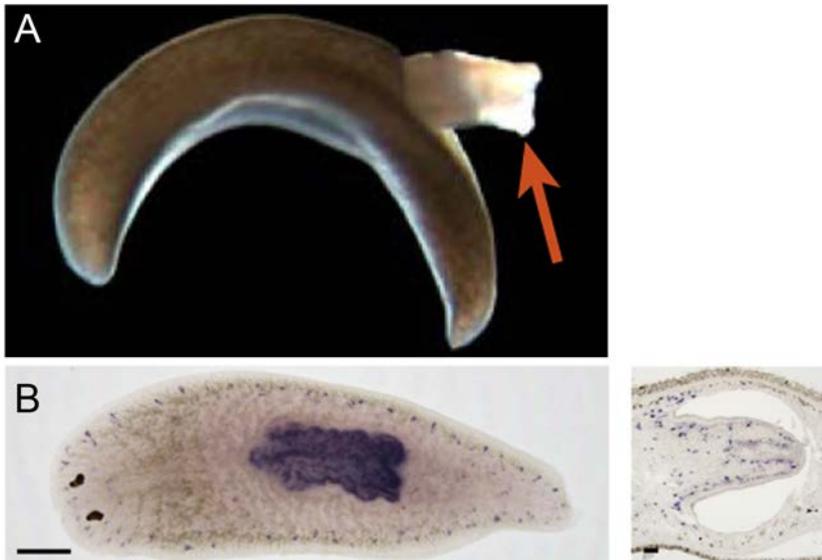
Figure I2.8: Planarian pharynx. (A) *in vivo* image of planarian with evaginated pharynx (red arrow). Adapted from (8). **(B)** *in situ* hybridization in whole mount using pharynx marker. Scale bar is 100 μm. Adapted from (99).

2.2.7. Parenchyma

The parenchymal tissue surrounds the internal organs (Figure I2.9A). It is a compartment composed by heterogeneous cellpopulations: phagocyte cells, glands cells, secretory cells, pigment cells and also glia cells (neural) (Figure I2.9B). The neoblasts and cellular progenitors for all cell types are also located in the parenchyma (76). Fincher et al. propose an extra parenchymal category: cathepsin+ cells, including cell types such as glia and pigment cells.
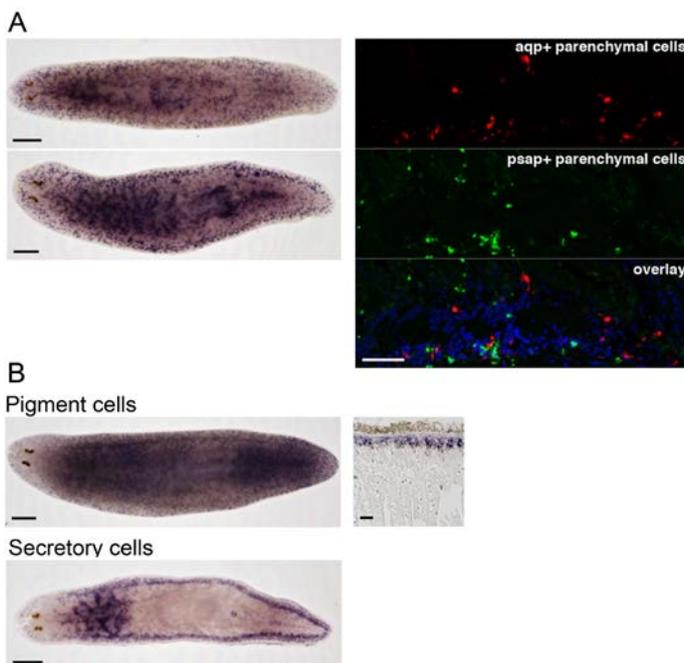
Figure I2.9: Planarian parenchyma. (A) *in situ* hybridization in whole mount and fluorescent *in situ* hybridization in sections using two parenchymal markers. Their overlay shows not colocalization. **(B)** *in situ* hybridization in whole mount and in sections using pigment and secretory cells markers, two populations of parenchyma. Scale bars in (A) are 250 μm for whole mount *in situ* hybridizations and 100 μm fluorescent *in situ* hybridization in sections; and in (B) are 500 μm for whole mount *in situ* hybridizations, 100 μm for in situ on sections. Adapted from (99)

2.3. Planarian plasticity

Planarians are able to regenerate from small body fragments (Figure I2.1). They are able to regenerate after any type of amputation, being either transversal or sagittal. After amputation, new cells derive from neoblasts and will form the missing tissues, but at the same time the pre-existing tissue remodels, to give raise a proportioned smaller animal. (58,128). Planarian plasticity is also evident during their normal homeostasis, since the individuals adjust their body proportions according to nutrient availability (Figure I2.1).

2.3.1. Neoblasts

Neoblasts (NB) are planarian pluripotent stem cells, which are spread in the parenchyma tissue. With an exception of the pharynx, inside the digestive system and in front of the photoreceptors (76,79) (Figure I2.10A, 10B). Morphologically, NBs are small cells (5-8 µm) with a huge nucleus:cytoplasm ratio (Figure I2.10C). They present basophilic cytoplasm, rich in free ribosomes, and few mitochondria and chromatoid bodies (129). Depending on the planarian size, NBs represent between 20-30% of the total cell number (76). NBs, like all SCs, are irradiation sensitive (Figure I2.10A). They present a unique genetic profile, which allows us to detect them by using WISH (or FISH) with specific riboprobes against genes as *piwi* (130) or *h2b* (131). Additionally, since NBs cyclin cell type these can also be detected by immunohistoquemistry with anti-PH3 antibody (136) or by BrdU incorporation (76).
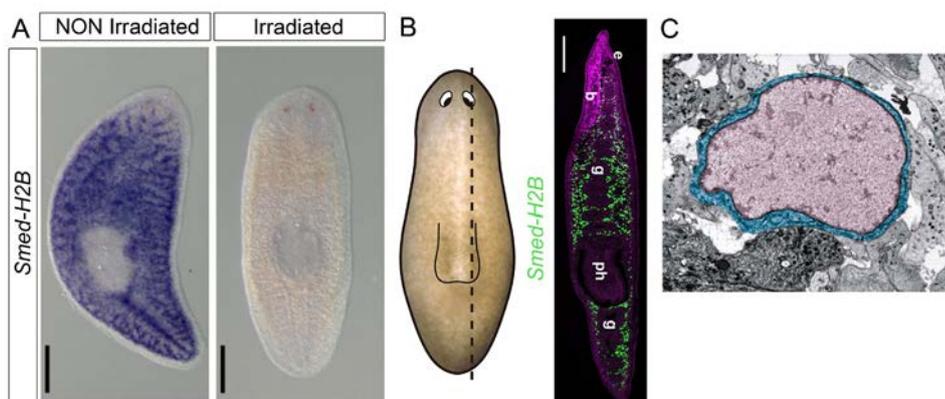
Figure I2.10: Neoblasts, planarian adult stem cells. (A) Whole mount *in situ* hybridization using a neoblast marker (*Smed-h2b*). Its expression disappears after irradiation. **(B)** Fluorescent *in situ* hybridization on sections using *Smed-h2b*. Schematic illustration shows where the section was approximately done. b, brain, g, gut; e, eye; ph, pharynx, adapted from (131). **(C)** Electron micrograph of planarian neoblast, its diameter is 8 µm. Nucleus is coloured in purple and cytosol in blue. Adapted from (58). Scale bars: 500 µm.

Wagner et al. (77) and recently Zeng et al. (133) demonstrated the pluripotency of NBs. Using irradiated planarian and single cell transplantation, a population of neoblasts that can form colonies in irradiated *in vivo* planarian. This population is known as clonogenic NB (cNB). This cell type is distributed all over the planarian body, presenting a specific genetic profile and certain surface molecules which allow investigators to isolate them by FACS (133). A single cNB is able to rescue a lethally irradiated planarian (77,133).

Mitotically active neoblasts are closely associated with the intestine in uninjured animals (79,132) and after feeding, neoblast proliferation increases dramatically (79,134). In addition, it has described that intestine-enriched transcripts encode regulators of metabolite processing, as well as putatively secreted proteins, suggesting an influence on neoblast dynamics (119). These results suggest a close communication between intestinal cells and neoblasts; acting as a neoblast "niche". No further results have been related intestine and nutrient intake with neither neoblast cycling activity nor gene expression.

2.3.2. Regeneration

Planarians are known for their regenerative capacity since the XIX century, when Morgan and Child started to describe regeneration and the underlying mechanisms (128,135). Even though both scientists left the field, others remained, fascinated by planarians, as commented by Dalyell who claimed "planarians can almost be called immortal under the edge of the knife" (136). Nowadays, it is known that after an injury, planarians regenerate based on adult pluripotent stem cells widely distributed in their body, and the continuous activity and coordination of cell-to-cell communication signalling pathways. Additionally, tight transcriptional regulation and cell differentiation are required (for successful regeneration).

2.4. Regeneration process in *Schmidtea mediterranea*

As commented above, the planarian regeneration process is based on neoblast proliferation and communication via different pathways. Those processes are widely studied, and the next chapter discusses how planarian regeneration is accomplished.

2.4.1. Regenerative stages

Just after amputation, musculature around the cut site rapidly contracts to minimize the surface of the wound and allow contact between dorsal and ventral epidermis. This process occurs within 30 minutes (after amputation) and is crucial to wound healing; allowing the confrontation of D and V epidermis. It is proposed that this confrontation will form an organizer enabling regeneration (137). Any type of regeneration, involving or not tissue loss triggers changes in planarian genetic profile. Wenemoser et al. described that during the first hours after amputation stress response genes are up-regulated in epidermis, muscle and neoblasts. Early transient genes start to be expressed 30 minutes post amputation until six to twelve hours after (Figure I2.11).
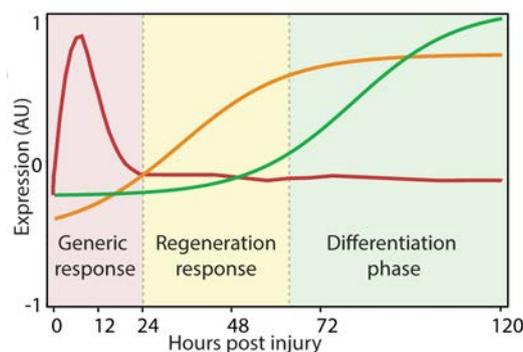
Figure I2.11: Model for planarian wound response and initiation of regeneration. A temporal model of planarian regeneration. Every injury triggers a prototypical generic response (red). If regeneration is not required following the injury, the response will decline. Otherwise, the expression of an injury-specific response emerges (yellow). These responses involve patterning molecules and neoblast-associated fate specialization genes. About 3 days following the injury, expression of differentiated tissue markers appears in association with the emergence of the newly regenerated structures (green). Adapted from (91).

Early transient genes are activated in the proximity of the wound in a translation-independent manner (91,138), and encode for signalling and transcription factors. Wurtzel et al. named those genes as generic wound response genes, being *runt1*, *jun-1*, *fos-1* or *egrl1* examples of it (91). Then a second response wave starts close to the wound site. It starts during the first six hours and involves translation-dependent genes, expressed in muscle, epidermis and neoblasts (Figure I2.13). These genes are involved in differentiation, proliferation and patterning. Examples are *wnt1*, *notum* and *follistatin* (*fst*), all of which are crucial for regeneration and patterning (139,140). The expression of the early response genes occurs after any amputation, independently if there is tissue loss or not. However, the expression of the wave genes occurs only after amputations that remove tissue and require the formation of new patterned tissue. Finally, the activation of a third genetic wave occurs in the so-called differentiation phase. (Figure I2.11).

It has been demonstrated that after an injury, extracellular signal-regulated kinase (ERK) phosphorylation is triggered within minutes. Pharmacologic inhibition of ERK, blocks wound-induce genes such as *wnt1*, *notum* and *runt1*, impairing regeneration (141). Recently, it was published that as fast as ERK is expressed, the wound also induces reactive oxygen species (ROS) expression after an injury, putting forward early signals in the induction of planarian regeneration (142).

2.4.2. Regeneration is stem cell dependent

The planarian regeneration process happens due to the proliferation of neoblasts. After injury the proliferative ability of neoblasts is activated to provide a new source of cells to rebuild the missing tissue. This proliferative response is coordinated starting 6 hours after an injury; a body-wide proliferative peak is observed (Figure I2.12). Neoblasts involved in this first mitotic event are in G2 phase (76,132). A second mitotic peak occurs at 48 hR. This is not general but located next to the wound region (Figure I2.12). This second peak only takes place when the damage involves tissue loss. Simple poking or small incisions just trigger the first mitotic response (132,143,144). Even though the blastema continues growing, mitotic cells are restricted to the pre-existing tissue close to the wound (post-blastema) (76). It was proposed that the neoblast accumulation during the second wave is due to the wound itself, which triggers migratory response to attract new cells (145,146).

In the undifferentiated tissue (blastema) cells differentiate to form the missing tissues and organs. Structures such as the eyes are easy to visualize, only four days after amputation. Animals will completely restore the missing tissue within seven to ten days after amputation (57).
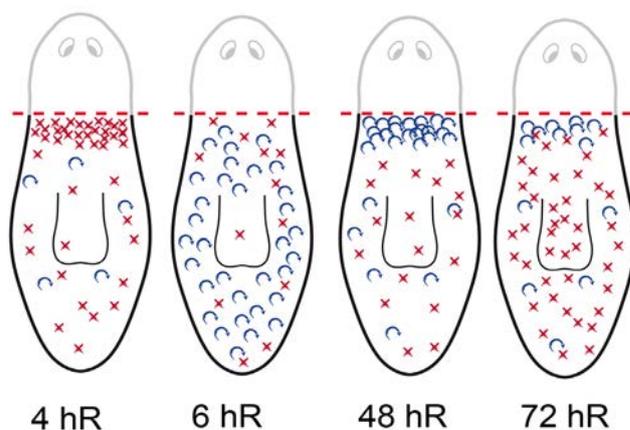
Figure I2.12: Cell proliferation and cell death regeneration. Schematic cartoons illustrating cell death disposition (red cross) and cell proliferation (blue semicircle) after an anterior amputation. At 4 hR, a burst of cell death is localized close to the wound. At 6 hR, general mitotic events occur broadly over the planarian body, followed by a second mitotic response localized close to the wound at 48 hR. Finally, at 72 hR a genereal cell death event is observed all over the planarian body.

2.4.3. Cell death remodellenig during regeneration

Planarian regeneration is a global process which involves the production of new cells at the amputation site, as well as remodelling of the pre-existing tissue to adjust organism proportions. Cell death and autophagy modulates the remodelling process eliminating differentiated cells. The apoptotic response also presents two events associated with injury. A burst of apoptosis (TUNEL+ cells) occurs proximal to a wound site four hours after any injury (Figure I2.12) (147), and could be associated with the triggering of the early regenerative response. The second event occurs at 72 hR and it is general (Figure I2.12), being associated with the remodeling of pre-existing tissue, thus only occurring when tissue was lost (151). Autophagy is also activated very early during regeneration, with activity observed one day post-amputation, mainly in the post-blastema region, spreading gradually to all existing tissues as remodelling processes occur (148).

Silencing several signalling pathways, such as the JNK, TOR or Hippo pathways, results in the formation of smaller blastemas in which cell proliferation and apoptosis are impaired. JNK is required for the early apoptosis response, since it allows G2-M transition of neoblasts entering mitosis and the activation of early response genes (67). Akt signalling mediates cell death activation during tissue repair (72), TOR hyper-activation gives rise to larger blaste-mas, although they remain undifferentiated (69), and Hippo hyper-activation enhances the wound response, promoting the expansion of cell populations (74).

2.4.4. Axis establishment

A successful regenerative process does not only require the regeneration of the precise number of cells but also their proper pattern. Planarians have been an excellent model to understand patterning mechanism. In adult intact planarians, different molecule patterns, position and identify the three main axes: anteroposterior (AP), dorsoventral (DV) and medio-lateral (ML) (Figure I2.13). These axial coordinates are provided by positional control genes (PCGs). PCG expression is restricted to muscular cells that are peripheral to the neoblasts. SCS data also confirms that muscles form different regions along the AP and DV axes re-gionally express PCGs (98,1273,126). PCGs have the prominent role in harbouring positional information. Indeed, two or more PCGs could be in overlapping spatial domains and expres-sed together in the same cells to a substantial degree. This leads to suggest that expression patterns of PCGs are reminiscent of patterning gene expression in animal embryos (such as *Drosophila*) but are present in adults (125). Inhibiting PCGs, leads to animals showing a different gradient of phenotypes, presenting different number, shape or size of organs, such as eyes, brain, pharynx or mouth. PCGs encodes for secreted proteins, that have been functionally studied specify the 3 planarian body axes (125,126,149). The AP axis is defined by the cWNT pathway (Figure I2.13A). The inhibition of certain elements such as *wnt1* and *wnt11-2* leads to regeneration of animals with posterior loss identity (150–155). Meanwhile the inhibition of *notum* (Wnt inhibitor) produces planarians with anterior loss identity (156). BMP signalling is another prominent model that involves adult positional information, which regulates the (DV) axis (Figure I2.13B). The inhibition of BMP components such as *bmp4*, *smad1*, *smad4* or *tolloid* leads to regeneration and homeostatic phenotypes affecting the DV axis (157–159). *bmp* inhibition results in progressive ventralization, with ventral tissue and ciliated epidermis, appearing dorsally. *admp* (157) and a variety of *noggin* (*nog*) and *noggin-like* (*nlg*) genes (62,157) inhibition produces the opposite phenotype, animals dorsalize their ventral side. *bmp4* is expressed dorsally, in a medial-to-lateral gradient. The ML axis is de-fined by the combination of two molecular pathways: 1) the non canonical Wnt (ncWNT) or the βcatenin-independent pathway. ncWNT ligands (WNT5) act through other receptors and do not activate βCATENIN. In planarian, this is related axon guidance. 2) On the other hand, SLIT molecules which bind their receptor Robo are repelling axons. *wnt5* is expressed laterally and inhibits the medially expressed *slit*. *wnt5* and *slit* (RNAi) expression results in medial-lateral patterning abnormalities (150,155,160–162) (Figure I2.13C).
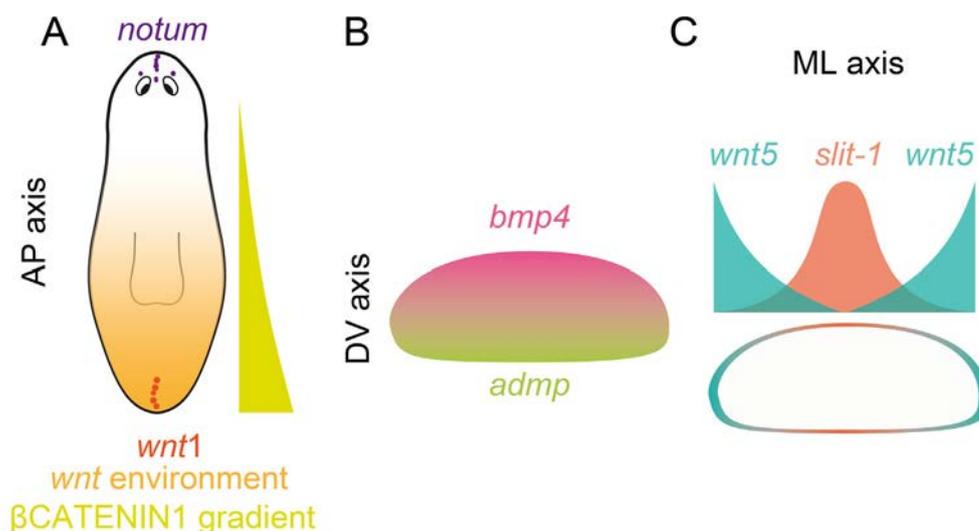
Figure I2.13: Positional control genes mediate the positional information in planarians.
Schematic depicting currently known primary patterning signals and their deployment along the indicated cardinal body axes: **(A)** anterior-posterior, **(B)** dorso-ventral and **(C)** medio-lateral. Schematic signalling gradients are hypothetical extrapolations from the expression patterns of the respective genes. Adapted from (429).

After amputation PCGs are used to pattern, position and identify the new structures. Specifically pre-existing tissue close to the wound contributes to pattern the missing tissue. The tissue will provide guidance to form the new anterior pole. Oderberg et al. carefully describes this process. As a first step, anterior pole progenitors form at anterior-facing wounds. The second step starts when *slit*, from the pre-existing tissue (ML), specifies the anterior pole progenitors. Those pole progenitors are dispersed in the ML from the dorsal to the ventral part. This distribution allows the promotion of the new pole in the midline blastema, and connects the pre-existing ML with the new one, which is forming in the blastema (Figure I2.15A). In the final step, scattered pole progenitors fuse to the pre-existing DV median plane, from where they will grow and pattern the AP axis (161) (Figure I2.15B). In that moment three orthogonal axes will be formed. With those observations, it is corroborated that the pre-existing tissue close to the wound patterns and promotes the formation of the new anterior pole. Even that the architecture of anterior formation is well studied, posterior pole is poorly investigated.
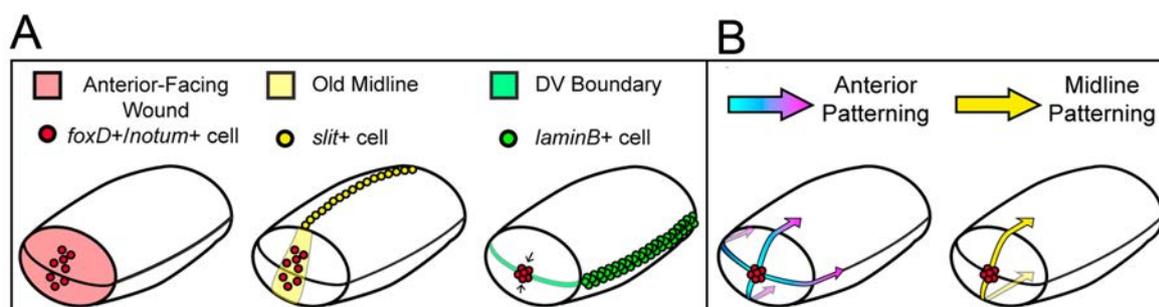
Figure I2.14: Landmarks in existing tissue at wounds are utilized to generate pattern in regenerating tissue. (A) Schematic illustration showing anterior pole formation, which relies on three landmarks at wounds in order to integrate the pattern of the new and pre-existing tissues: an anterior-facing wound, the prior midline, and the boundary between the dorsal and ventral sides of the animal. **(B)** Once the anterior pole is formed, it acts to help pattern the AP and ML axes of the regenerating head. Adapted from (161).

2.4.5. AP axis establishment and WNT signalling pathway

After transversal amputation, two signalling centre (pole) types are formed: the anterior (A) is located in the tip of A-facing wounds (or blastemas) and drives the regeneration of lost A structures; and the posterior (P), which is located in the tip of P-facing wounds and drives the regeneration of lost P structures. Each group of cells appears between 12 and 24 hours after amputation, and is formed by muscle cells. Cell appearance is not dependent of neoblast proliferation, making them known as independent SCs (Figure I2.15). Wound induced expression stops within two–three days after injury, and a new group of cells provided by neoblast, is established in each pole independently (125,126,150,151,154,155,156,163,164). A pole expresses *notum* (Wnt inhibitor) and requires different TFs for its specification, for instance: *foxD* (165,166), *zic* (166,167), *prep* (168), *pbx* (169,170), and *pitx* (105,106). Blocking anterior pole formation causes smaller head blastemas and midline collapsing (165,169). The P pole expresses *wnt1* and requires *islet* (105,163), *pitx* (105,106), *teashirt* (*tsh*) (171,172) and *pbx* to be expressed (Figure I2.15). Their inhibition produces animals without tail formation ("tailless").

WNT1 belongs to the WNT ligands, a family of glycoproteins that are secreted to modulate different developmental steps. When wnt ligands bind with their receptors Frizzled (Fz) captures Dishevelled (Dvl) in the membrane. This implies the destruction of the βCATENIN (βCAT) destruction complex. As a consequence, βCAT will be free to enter to the nucleus and bind the TCF/LET transcription factor, which has a DNA binding site for HMG. This pathway is known as canonical WNT (cWNT) or βcatenin-dependent pathway. The cWNT pathway is one of the most studied pathways in planarians, since its implication has been related with posterior identity (60,150,151,164,173). This function defining identity has been
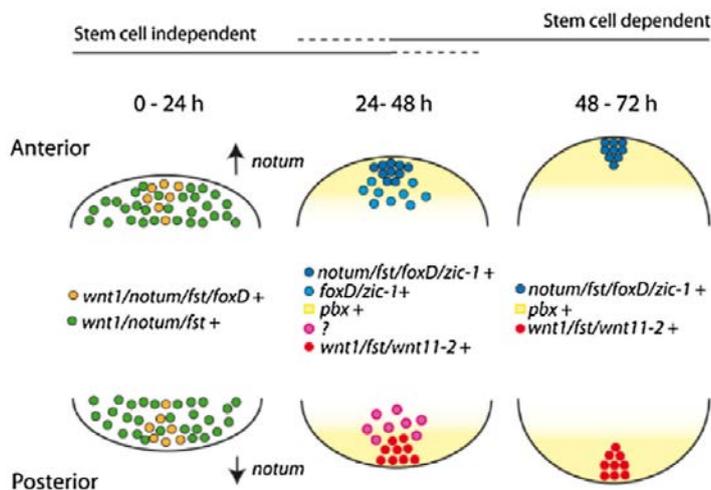
Figure I2.15: Gens involved in anterior and posterior organizers determination. During the first 12 hours after amputation *notum*, *wnt1* and *foxD* are expressed in isolated muscular cells. The differential expression of notum during this early stage triggers the inhibition or activation of the cWNT signal and thus the establishment of the A or the P program in each wound, respectively. From 24 hR (after amputation) new muscular cells expressing *notum*, *foxD*, *zic-1* and *fst* appear in the most A tip. They arise from precursor cells that express *foxD* and *zic-1*. In the P tip new muscular cells expressing *wnt1* and *fst* apper. From 48 h to 72 h the A and the P organizers are definitively formed by muscular cells expressing *notum*, *foxD*, *zic-1* and *fst* in the A, and *wnt1* and *wnt11-2*, in the P organizer. From 24 hR the expression of *pbx* is activated in every wound with a diffusive pattern, which does not correspond to muscular cells. Adapted from (305).

reported to be evolutionary conserved, since cWNT's presence in most of the studied clades. In planarians, first evidence of its role was studying the phenotype of *βcatenin-1* (*βcat1*) (RNAi), which led to dramatic regeneration: heads were regenerated in place of tails, resulting in "two-headed" animals (61,164,173). Similarly, inhibition of cWNT components as *wnt1* (150,151), *evi/wentless* (150), *dishevelled* (174), or *tsh* (171,171) can result in two-headed animals. On the other hand, up-regulating cWNT signalling, through *APC* (173) or *notum* (156) (RNAi) results in the regeneration of tails, resulting in two-tailed animals

Certain Wnt ligands, such as *wnt11-5*, *wnt11-1*, *wnt11-2*, and *wnt1* are regionally expressed along the AP axis posterior (also known as posterior wnts) (Figure I2.13A). And their inhibitions produce different degrees of posterior identity loss and trunk duplications, with animals developing two pharynges rather than one (126,154,175). Conversely, Wnt inhibitors such as *sFRP-1*, *sFRP-2*, and *notum* are expressed in the anterior pole (155,156,164). Two studies demonstrated a posterior-to-anterior gradient of βCATENIN-1 protein levels (59,149) (Figure I2.13A), indicating that constitutive regional expression of Wnt ligands in the posterior pole and Wnt inhibitors in the anterior pole control regionalization of the planarian AP axis.

wnt1 confers posterior identity since inhibition produces animals with the tailless phenotype. Animals which are able to close the wound but lose the posterior morphology. Since they miss posterior organizer, develop without tail extension and VNC is fusioned in U shape. Stronger phenotypes could be observed depending on the degree of inhibiting, regenerating planarians are able to generate a head instead of a tail, called two-headed phenotype. (60,150,151). *notum* (RNAi) animals regenerate tails instead of head, suggesting that *notum* promotes head (and anterior identity) (156). Both genes are generically expressed at each wound site during the first hours of regeneration. The *wnt1* and *notum* expression cease, and in the later stages of the regenerating process, a second expression of each gene (stem cell dependent) is focus in anterior or posterior, respectively. Hence, what is regulating *wnt1* just in the posterior? Transcriptomic data revealed that *notum* was the only gene early differentially expressed at the A facing wound compared to the P pole (95), assuming that the decision to generate a tail or a head is taken within this period of time. Regenerating tips seem to be a transient structure since their properties are lost in adult organism, since the inhibition of *notum* or *wnt1* produces changes in the head or tail shape, but never produces a head-to-tail or tail-to-head transformation (60,176). Thus, adult planarians do not have active organizers.

The Hedgehog pathway helps cWNT signalling. Hedgehog (hh) ligand bind its receptor Patched (ptc). This interaction nulls the effect of the receptor Smoothness (Smo). Smo will initiate the signalling that triggers GLI (TF) activation, and their target genes (177). In planarians *ptc* inhibitions lead to regenerate two-tailed animals. Whereas, *hh*, *smo* and *gli* inhibition produce tailless planarians; the combination of *hh* with *smo* or *gli* produces a bi-headed phenotype (177,178). Leading to suggest that the Hh pathway regulates cWNT. Moreover, it is supported that *hh* regulated *wnt1* expression (178).

2.4.6. From neoblasts to organs

Neoblasts are a heterogenic population formed by truly pluripotent stem cells (cNeoblasts) as well as lineage-committed progenitors that can differentiate into all cell types present in the tissues (77,133,165,180,181). Neoblasts and their progenitor populations are characterized by the co-expression of *smedwi-1* (*piwi1*). Zeng et al. sort high content *piwi1*+ cells. After cell analyses, it was proposed that 11 Nb major types cover 3 germ layers: epidermis, neural and protonephridia (ectoderm), muscular and pharynx (mesoderm) and gut (endoderm) (Figure I2.16). Previous studies revealed the presence of different NBs populations: gamma (γ)-class, zeta (ζ)-class and nu (v)-class (99). γ-class is related with gut linage; ζ-class with epidermal lineage and v-class with the neural one (180,182) (Figure I2.16).

Committed progenitors express *piwi1* and transcription factors specific to the different cell lineages, suggesting that these are characteristic of committed cell progenitors still capable of proliferating. A homologue of the MEX3 RNA-binding protein is proposed to play a general role in cell differentiation, since silencing of its expression results in expansion of the stem cell compartment in parallel with a decrease in the number of lineage-restricted progenitors (183). Independently and thanks to SCS data, Plass et al. and Fincher et al. could determine almost all neoblast trajectories, describing progenitors for all cell types and the intermediated state that they could have (Figure I2.16). Performing cell type atlas, Plass et al. characterized that some differentiated cell types rise from the same intermediate progenitors, suggesting that even if specialized NB are avaliable, different cell types could come from the same progenitor (99). During the last decades, markers for different tissues, cell populations and their progenitors were identified using homologous genes previously published in other organism, or based on a genetic screening. SCS data helped to discriminate whether a gene was expressed in some tissue without properly knowing the specificity of that gene to that tissue. Thanks to SCS data this problem is almost solved. Particularly, experimental results provide the planarian community with genes specifically expressed in neoblast populations, committed progenitors and differentiated cells. For instance, each muscle fibre has a specific transcription factor related to it: *myoD* is present in longitudinal muscle dibre, *nkx1* with circular fiber (184) and *foxF-1* is present in non-body wall fiber (123).

Specialized Nbs and their progenitors are broadly distributed. During regeneration and homeostasis, different tissues exert a role with organizer or self-organizer activity to properly target these. Below, different examples will be commented. Eye progenitors are specified in the pre-pharyngeal region and migrate to precise, predictable locations in the head. The position in *de novo* organ formation has been defined as the progenitor target zone (TZ). Eyes are maintained in a specific region called the targetable zone (TAZ) (Figure I2.17), defined as the zone where progenitors are able of going to maintain the organ (162). These findings indicate that it is the combination between self-organization and extrinsic cues (TAZ) that determines the destination of migrating regenerative progenitors.
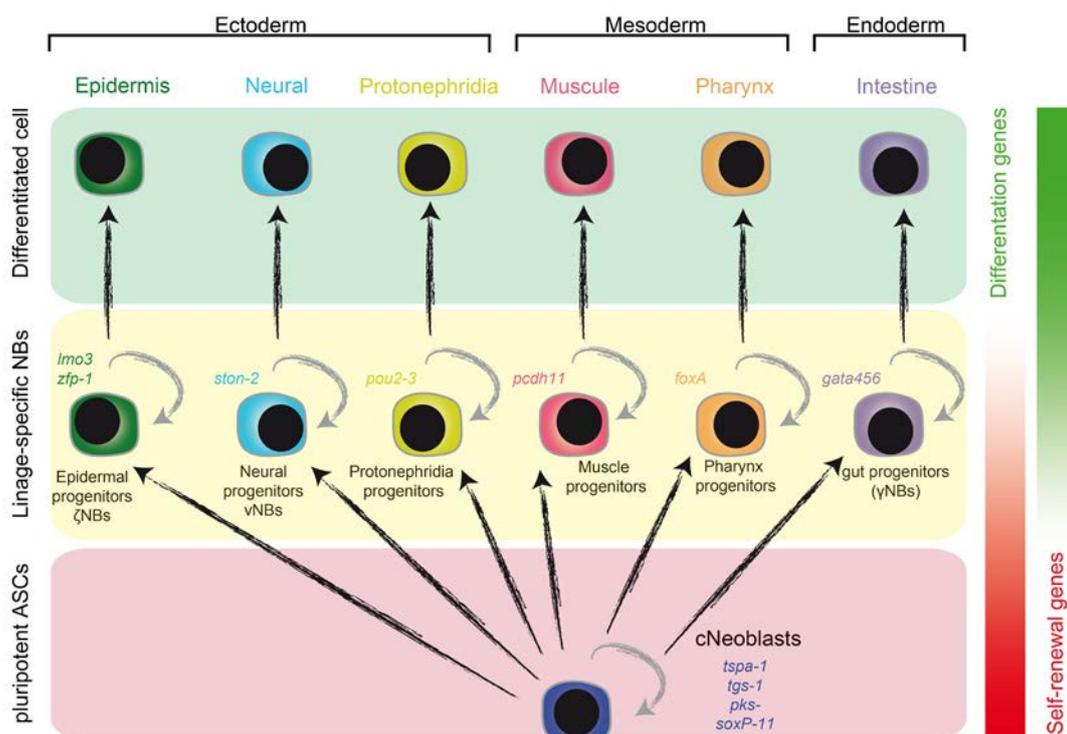
Figure I2.16: Model of stem cell hierarchies. cNeoblasts give rise to all linage-specific neoblasts. ζ-neoblasts to epidermis, v-neoblasts to neural tissue, protonephridia progenitors to excretory system, muscle progenitors to muscle cells, pharynx progenitors to pharynx and γ-neoblasts to intestine. Linage-specific neoblasts will proliferate and cell will be committed to each tissue: ectoderm with epidermal, neural and protonhephridic tissue; mesoderm with muscle and pharynx; and endoderm with intestine. At last one gene representing each neoblast population was added. Adapted from (133, 182).
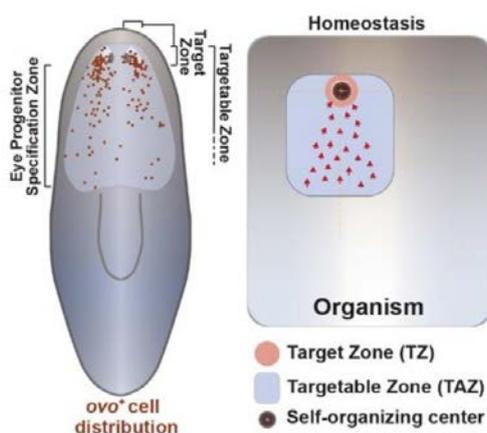
Figure I2.17: Eye progenitor specification. eye progenitor cells (*ovo*+) are displayed in a targetable zone (TAZ) and eye acts self-organization tissue, in an area called Target Zone (TZ). Red arrows indicate migratory progenitors and their directions. Adapted from (162).

2.5. Body and organ size regulation

In most animals adult body size is determined by growth during embryonic and juvenile stages, while the adult stage consists of tissue renewal. However, long-living species such as planarians change their body size according to nutrient availability during their entire live retaining their body proportions (Figure I1.1). These alterations in planarian body size are mediated by changes in cell number (185,186) resulting from modulation of the balance between cell proliferation and apoptosis (Figure I1.18). Thus, the ratio of proliferation to apoptosis decreases in starvation conditions and increases in times of nutritional abundance (147,187). Furthermore, the balance between the SC population and all types of differentiated cells relies on robust signalling mechanisms that allow continuous adjustment of cell proliferation, cell death, and cell differentiation.
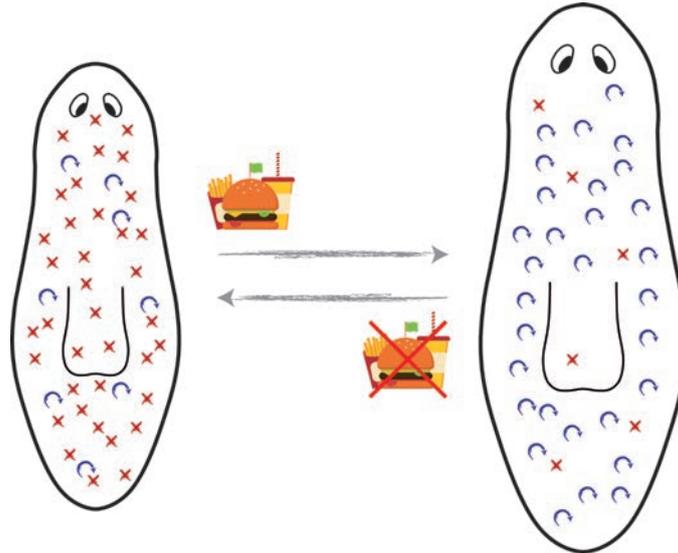
Figure I2.18: Cell proliferation and cell death growth and degrowth. Schematic cartoons illustrating cell death disposition (red cross) and cell proliferation (blue semicircle) during growth and shrinking periods. After feeding, neoblast proliferate all over the body and cell death is reduced, resulting in the increment of cell number and body size. Conversely, in starved conditions cell proliferation is reduced and cell death increased <cross the whole planarian body, resulting in the decrease of the total cell number and body size.

Mechanisms that allow planarian regeneration also participate in organ adjustment to reshape pre-existing structures. Mechanisms that control planarian body size and growth remain to be fully elucidated. The insulin/mTOR pathway is the only pathway demonstrated to control planarian body size. Inhibition of insulin-like peptides or TOR attenuates cell proliferation, prevents planarian growth after feeding, and accelerates shrinking during starvation (68,71). Hyper-activation of mTOR using *PTEN* or *smg-1* (RNAi) does not give rise to larger organisms but does promote over-proliferation and outgrowth formation (69,73). In planarians, JNK is required for organ remodelling through the induction of apoptotic cell death (67). Moreover, hippo inhibition increases mitosis, inhibits apoptosis, and promotes dedifferentiation, leading to the formation of overgrowths but not to changes in body size or cell number (75).

Not only the whole body but also organs must maintain proper size and proportions. Planarian brain remodels according to the planarian body size. The mechanism that controls the size of the brain appears to be a *wnt11-6*/*notum* signal. *notum* is expressed in anterior brain neurons and promotes brain growth. On the contrary, *wnt11-6* is expressed in the posterior brain neurons and acts as an inhibitor of brain growth (150,188). This Wnt/notum negative-feedback loop regulates brain:body proportions through control of neoblast differentiation (189).

3. Genomic landscape

During animal development, a single cell can give rise to a multitude of different cell types that contain the same genome but show unique morphologies and functions. This cell variety emerge due to the particular gene expression of each cell, which will define each and very single cells identity (190,191). Genes are encoded on the DNA, which can wrap to chromatin. Chromatin is a macromolecule formed by DNA and nucleosomes, which are octamers containing a pair of all known histone proteins: H2A, H2B, H3 and H4. Each nucleosome is wrapped with 146 base pairs (bp) of genomic DNA and separated by 20-50 bp. This first level of chromatin organization is called euchromati. In euchromatin regions, genes can be transcribed, it is hence an active gene transcription DNA region. Thanks to folding proteins, the level of chromatin compaction can increase. The packed chromatin state is known as heterochromatin (192–194) and does no allow for gene transcription. Thus the chromatin state is crucial to determine if DNA is accessible for the transcriptional machinery. Changes between the distinct chromatin stages are crucial to dynamically regulate gene expression and define each cells genetic profile. Histones play a key role in this aspect, since they can be chemically modified, being: acetylated, methylated, phosphorylated and ubiquitinated (195,196). Those modifications are driven by internal cell mechanisms, and can influence nucleosome shifting, gene accessibility and transcription.

3.1. Epigenome

Gene transcription begins with the recruitment of RNA polymerase II (Pol II) and auxiliary factors to core promoters, short DNA sequences located on the genome around transcription start sites (TSSs). While core promoters are sufficient to recruit Pol II and drive basal levels of transcription (197,198), they also require cis-regulatory elements (CREs) or enhancers to be fully activated (194) (Figure I3.1A). Enhancer sequences on the genome contain short DNA motifs (specific nucleotide sequences) that act as binding sites for sequence-specific transcription factors (TF). Several such motifs could be identified in a unique enhancer, allowing cooperation between them (Figure I3.1B). Enhancers could be located upstream or downs-
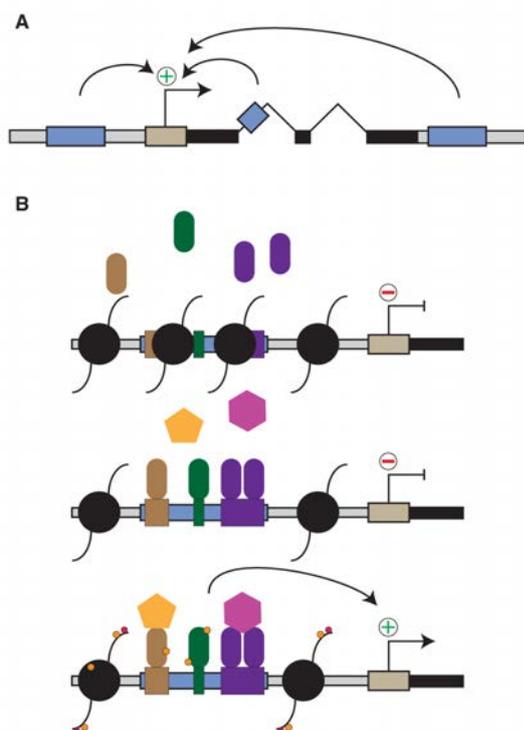
Figure I3.1 Enhancers: Elements of the epigenome. **(A)** Gene (black bars) transcription starts at the TSSs (straight arrow) within core promoter elements (light brown). Enhancers (blue boxes) are CREs and are often found in introns or distal intergenic regions both upstream and downstream of the gene. **(B)** Nucleosomes (black circles) bind to DNA and decrease accessibility to other proteins, such as TFs (coloured rods). Enhancers contain TF-binding motifs (coloured boxes), where TFs bind in competition with nucleosomes. TFs recruit transcriptional cofactors (coloured polygons) and activate gene expression. Cofactors often show catalytic activity and post-translationally modify TFs and histones (small coloured circles indicate such modifications). Adapted from (194)

tream from a gene position, and be classified proximal and distal depending on their relative position to the TSS (199). Enhancers function independently of the distance and orientation to their target genes, and can do so at large distances of several hundred kilobases or even megabases by looping (200).

Enhancers are able to recruit and activate Pol II at target gene promoters. Activity of the enhancers is determined by the recruitment co-activators and co-repressors. Genes can present different numbers of enhancers that modulate their spatiotemporal expression (192). Enhancer activity has been shown to correlate with certain properties of chromatin; active enhancers are typically devoid of nucleosomes promoting DNA accessibility by TFs. Nucleosomes in the vicinity of active enhancers contain histones with post-translational modifications, which modulate DNA accessibility due to 1) enhancing or weaken non-covalent interactions between DNA and histones, and 2) acting as a platform for recruitment of other proteins that particularly recognize these modifications (201). Typical active enhancer histone modifications are histone H3 lysine 4 monomethylation (H3K4me1) and trimethylation (H3K79me3), and H3K27 acetylation (H3K27ac) (202,203). Histone modifications are dynamically added and removed by chromatin-modifying enzymes in a highly regulated manner. An example is acetyltransferase p300/CBP which acetylates H3K27, recruiting chromatin remodelers and transcription regulators. H3K27ac also decreases nucleosome stability and chromatin de-compactation due to changes in net charge of histones (204). Another relevant modification is DNA methylation, which when it occurs at CpG islands of promoters is associated with silencing of genes. The regulation of gene expression by DNA methylation and histone protein modifications is known as epigenetics. The particular state of DNA methylation and histone modification in a particular cell is called the epigenome. Interestingly, during the last years different techniques have emerged to understand the epigenome, and particularly to detect CREs. The main experimental methods are based on chromatin-immunoprecipitation, high throughput sequencing (ChIP-seq), and chromatin accessibility techniques in conbination with high throughput sequencing, such as MNase-, FAIRE-, DNase- and ATAC-sequencing. ChIP-seq is based on *in vivo* crosslinking of DNA and a protein (193). Emerging complexes are pulled-down (of a cellular particle mix) by using a specific antibody and finally sequenced. This allows consequent sequencing. With this method two main approaches can be taken 1) Using antibodies identifying specific TFs and detecting their binding sites, or 2) using antibodies against histone modifications or co-factors such as p300 that would reveal active enhancers distribution across the genome. ATAC-seq (or transposase-accessible chromatin using sequencing) is the most common technique to assess the accessible chromatin distribution (Figure I3.2). It is based on a modified Tn5 transposase enzyme that thanks to tagmentation reaction, is able to clave and tag DNA with universal sequences (193).
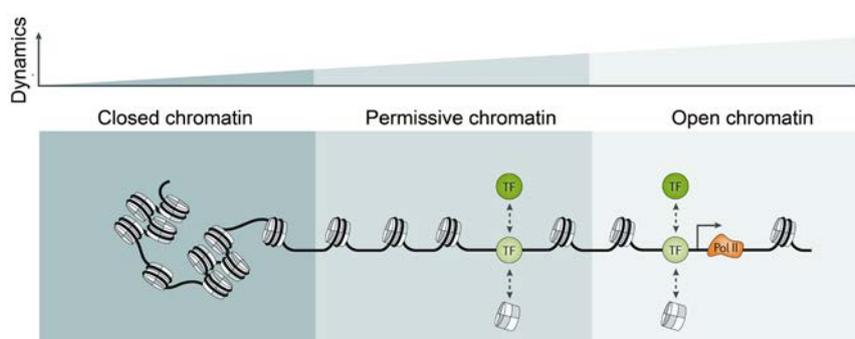
Figure I3.2: Chromatin accessibility. Accessibility state is a dynamic continuum transition across the genome. Closed chromatin is not permissive to transcription factors, as long as chromatins became permissive, TFs initiates sequence-specific accessibility remodelling and establish an open chromatin conformation and transcription activity (RNA polymerase II bonded to chromatin). Adapted from (430)

Understanding how enhancers regulate gene expression is an area of increasing interest because these are essential not only for developmental gene expression but also to understand evolution and human diseases.

3.1.1. Epigenetics in regeneration

Previously, it has been described how regeneration deploys mechanisms from embryonic development, such as the expression of developmental genes and the use of conserved signalling pathways. In the same way, some epigenetic changes have also been described to be recapitulated in regeneration, whereas others have been described to be specific for regeneration in different organs and organisms (205–207). Thus, specific genes important to trigger regeneration seem to be regulated by enhancers specifically active during regeneration (205) (Figure I3.3). For example, Kang et al. discovered that the zebrafish gene *leptin b* (*lepb*) is robustly induced during regeneration in both fin and cardiac tissues. Studying H3K27ac's genomic profile putative enhancers distally located to *lepb* could be identified. These were associated directly to regeneration-specific gene expression, being inactive during embryonic development (206,2107). Vizcaya-Molina et. al also reported, that in highly regenerative wing imaginal disc enhancers exclusively active after damage, co-opted enhancers from other tissues. A set of conserved transcription factors (TF) that control regeneration across metazoans can bind to those enhancers (208). Revealing TF function, Gehrke et al. found that whole body regeneration in acoel *Hofstenia miamia* is driven by a wound induced master regulatory TF (Egr) (209).

The existence of regeneration/injury-specific enhancers suggests that different enhancers control the transcription of a subset of genes expressed in both developmental and regenerative contexts. Identifying the essential motifs of regeneration-specific enhancers, their binding partners, and upstream regulators will uncover how injury signals are transformed to trigger regeneration programs.
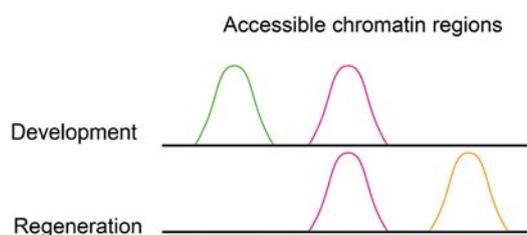
Figure I3.3: Developmental and regenerative enhancers. Schematic ATAC-seq peaks illustrating enhancers only open during development (green) or regeneration (orange), or development peaks de-ployed during regeneration.

3.2. Epigenetics in planarians

The epigenome of planarian cells is poorly studied. Early experiments in this model system have shown that the knockdown of orthologs of mammalian epigenetic regulators can lead to different SC defects and errors in lineage commitment of stem cell progeny, culminating in a loss of regenerative capacity. Particularly, it has been analyzed how DNA methyl-transferases (*mbd2/3*) are involved in the methylation of CpG islands. RNAi of *mbd2/3* resulted in a loss of certain differentiated cell lineages (epidermis, gut and pharynx) without reducing neoblats number (210,211). RNAi inhibition of different deacetylases such as *Smed-CHD4* (212), *Smed-HDAC1* (213–215) or nucleosome interaction proteins such as *RbAp48* (216,217), and gene *p66* (218), led to an abrogation of SC differentiation. It was also reported that a group of methylases are related with neoblast maintenance and differentiation of different cell lineages (219,221).

Polycomb proteins mediate gene silencing through post-transcriptional modification (222). Three planarian genes encoding homologs of Polycomb proteins have been identified. *Smed-ezh*, *Smed-suz12-1*, and *Smed-eed-1* were shown to be necessary for stem cell clonal expansion (223).

Interestingly, it has been published, that ChIP analysis in *mex-1* (RNAi), a gene related to differentiation in planarian, revealed changes in chromatin state (224). Although, it was not further investigated, I think this represents a good starting point to elucidate how chromatin changes after gene inhibition. Thus, published studies demonstrate that the planarians epigenome is regulated by evolutionary conserved mechanisms. However, until today the use of high throughput epigenome-sequencing techniques allowing to understand the epigenetic changes during planarian regeneration has not been reported. In the second chapter of this thesis I will present new data ferived from ATAC-seq and Chip-seq analysis of regenerating blastemas focusing on the epigenomic differences between anterior and posterior regenerating wounds.
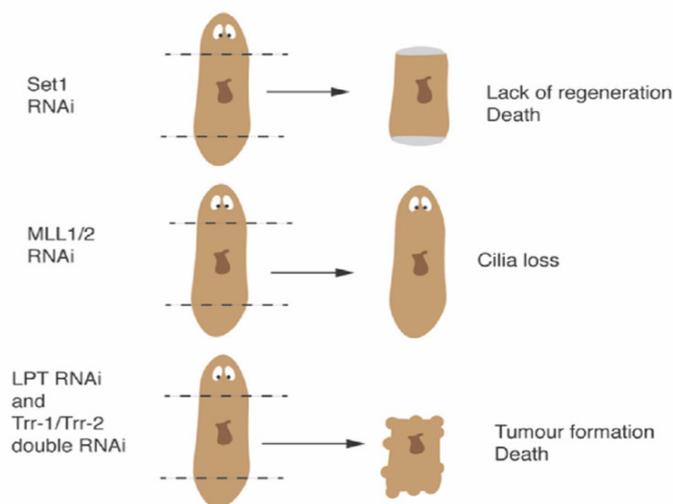
Figure I3.4: Epigenetics machinery regulate regeneration in planarians. Schematic illustration showing different phenotypes after amputation. *Set 1* knockdown animals presents a lack of regeneration and end up dying. *MLL1/2* (RNAi) organisms show mobility defects due to cilia loss. The inhibition of *LPT* in one hand, and *Trr-1* together with *Trr-2* on the order hand, leds to tumour formation and death. Adapted from 431.

OBJECTIVES

Objectives

To further understand the molecular mechanisms that regulate planarian growth and regeneration, the main objectives of this thesis were:

1. Study of mechanisms involved in the control of cell number in *Schmidtea mediterranea*.

    1.1. Characterization of new gene family *Blitzschnell* (*bls*): identification and classification of *bls* genes.

    1.2. Investigation of *bls* genes role in regulating cell proliferation and cell death during growth, degrowth and regeneration.

    1.3. Study of the *bls* gene expression regulation.

2. Study how the cWNT pathway affects the planarian epigenome during posterior regeneration

    2.1. Identification of target genes of *wnt1* related with posterior identity.

    2.2. Characterization of the epigenomic landscape during anterior and posterior regeneration.

    2.3. Identification of TFs related with the posterior organizer formation and/or function.

3. Characterization of the Fox family of transcription factors in *Schmidtea mediterranea.*

    3.1. Identification all Fox genes in *Schmidtea mediterranea* and their classification accordingl to their phylogeny.

    3.2. Study of Fox Family in *Schmidtea mediterranea* at genomic, sequence and functional level.

RESULTS

Results

4. Chapter I. Planarian size depends on *Blitzschnell*, a novel gene family that controls cell number through balancing cell proliferation and cell death

Planarians are able to change their body size, growing and shrinking accordingly to their nutrient status. This huge plasticity is based in the context-dependent control of the total number of cells, which is regulated by the ratio between cell proliferation and cell death. This chapter will focus in the study of a novel gene which regulates the balance of both processes and as a consequence planarian cell number and size.

4.1. *Blitzschnell* is a new gene family organized in two clusters of tandem repeats in *Schmidtea mediterranea*

We performed an RNAi screen to find candidate genes involved in planarian eye regeneration, and identified an unknown gene whose inhibition resulted in faster regeneration of the eyes after head amputation. We named this gene *Blitzschnell* (*bls*), which means "quick as a flash" in German. Surprisingly, upon attempting to identify the genomic locus of this gene in *Schmidtea mediterranea*, we found that *bls* belongs to a gene family composed of 15 members distributed on 2 distinct genomic scaffolds (Figure R1.1A).

Although all *bls* sequences shared more than 70% of identity (Figure R1.1B), a phylogenetic analysis using the nucleotide sequence, allowed us to classify *bls* genes into five subfamilies (Figure R1.1C). Four of them (*bls1*, *bls2*, *bls4* and *bls5*) contained two putative genes apparently originated by duplications (named a and b). Subfamily *bls3* contained 7 *bls* sequences (named *bls3a-g*). These were also apparently derived from recent successive tandem duplications, as suggested by their genomic organization (Figure R1.1A) and near identical DNA sequence (Annex I; Figure R1.1B). One band of the expected size was successfully amplified using primers spanning the junction between *bls3a* and *bls3b*, confirming the existence of at least 2 repeats (Figure R1.1A, 1D). Interestingly, in the repeated genomic block harbouring *bls3* members and in the vicinity of other *bls* genes we identified complete or fragmentary transposon related genes, such as Reverse Transcriptase, RNAse H and Integrase (Figure R1.1A).

Figure R1.1: Bls family is composed by 11 genes and 4 pseudogenes. (A) Cartoon illustrating the genomic organization of *bls* family members. *bls1*, *2* and *3* subfamilies are found in scaffold 54 and *bls 4* and *5* subfamilies in Scaffold 49. The primers used to amplify the junction of the first *bls3* repeats are indicated in blue. Transposon elements are indicated with squares. Scale bar indicates base pairs. **(B)** Table comparing the identity among all the *bls* members in *Schmidtea mediterranea*. **(C)** Phylogenetic analysis of all members of *Smed bls* family using nucleotide sequences. They group into 5 subfamilies. Scale indicates expected nucleotide substitution per site. **(D)** PCR analysis using primers flanking the junction of the first *bls3* repeat (between *bls3a* and *b*) showing the expected 1 Kb band.

% Identity Nt	bls1a	bls1b	bls2a	bls2b	bls3a	bls3b	bls3c	bls3d	bls3e	bls3f	bls3g	bls4a	bls4b	bls5a	bls5b
bls1a	100	97.95	78.35	78.16	74.15	73.35	73.11	73.35	73.15	73.35	72.95	75.67	75.67	78.5	78.5
bls1b	x	100	77.12	76.94	74	73.15	73.29	73.15	72.96	73.15	72.42	75	75	77.78	77.78
bls2a	x	x	100	99.79	82.23	84.65	83.5	84.24	84.44	84.65	81.02	70.27	70.27	87.87	87.45
bls2b	x	x	x	100	85.03	84.82	84.62	84.41	84.62	84.82	81.84	70.87	70.87	88.05	87.84
bls3a	x	x	x	x	100	99.58	99.58	99.37	99.37	99.58	99.58	82.92	82.92	81.52	81.31
bls3b	x	x	x	x	x	100	99.59	99.38	99.38	99.59	99.59	82.52	82.52	81.2	81
bls3c	x	x	x	x	x	x	100	99.38	99.38	99.59	99.59	82.78	82.78	80	79.8
bls3d	x	x	x	x	x	x	x	100	99.18	99.38	99.38	82.89	82.89	80.96	80.76
bls3e	x	x	x	x	x	x	x	x	100	99.38	99.79	82.48	82.48	81.2	81
bls3f	x	x	x	x	x	x	x	x	x	100	99.59	82.48	82.48	81.2	81
bls3g	x	x	x	x	x	x	x	x	x	x	100	83.03	83.03	78.65	78.8
bls4a	x	x	x	x	x	x	x	x	x	x	x	100	100	72.53	72.9
bls4b	x	x	x	x	x	x	x	x	x	x	x	x	100	72.67	72.9
bls5a	x	x	x	x	x	x	x	x	x	x	x	x	x	100	99.8
bls5b	x	x	x	x	x	x	x	x	x	x	x	x	x	x	100

By mapping reads from the transcriptome of intact planarians (75) against the genome of *S. mediterranea* (99), we detected transcripts for subfamilies *bls2*, *bls3*, and *bls5*, but not for subfamilies *bls1* or *bls4* (Figure R1.2A). Furthermore, the predicted open reading frame (ORF) for *bls2*, *bls3*, and *bls5* encoded peptides containing an N-terminal signal peptide (SP), suggesting that they could be secreted, and a highly conserved C-terminal coiled-coil (CC) domain (Figure R1.3). Non-detectable transcription, together with a much shorter ORF, strongly suggests that subfamilies *bls1* and *bls4* are made up of pseudogenes (Ψ).

Taken together these data demonstrate that *bls* is a new gene family consisting of 11 genes and 4 pseudogenes. The genes encode very similar peptides that may be released into the extracellular space, as suggested by the presence of a signal peptide.
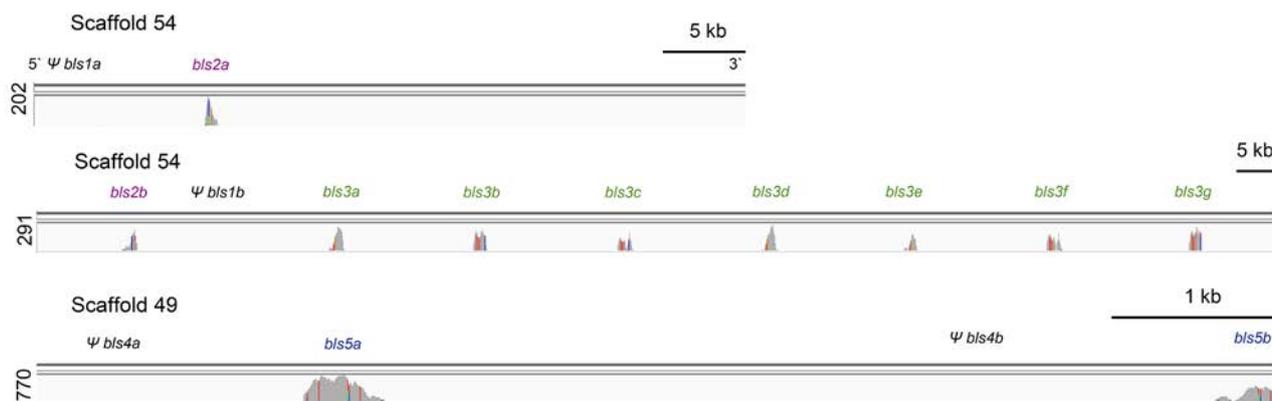
Figure R1.2: Genomic features of *bls* subfamilies in *Smed*. Transcriptomic reads mapping in the three non-consecutive parts of the two scaffolds where *bls* genes are located. Peaks represent reads accumulation. The absence of peaks is considered a lack of expression. Lateral number represents the highest summit in each track.

Species	Protein	SP	CC	Molecular Weight (Kd)
Smed	BLS2A	Y	Y	16,89
Smed	BLS2B	Y	Y	16,6
Smed	BLS3A	Y	Y	17,38
Smed	BLS3B	Y	Y	17,43
Smed	BLS3C	Y	N	11,87
Smed	BLS3D	N	Y	16,12
Smed	BLS3E	Y	N	4,87
Smed	BLS3F	Y	N	11,9
Smed	BLS3G	Y	Y	19,54
Smed	BLS5A	Y	N	16,72
Smed	BLS5B	Y	Y	11,37

Figure R1.3: Poteomic features of *bls* subfamilies in *Smed*. Sequence analysis of the BLS proteins: presence of the signal peptide (SP) and coiled coild (CC); their molecular weight, and the sequences. Schematic illustration of BLS protein domains: signal Peptide (SP) in red and Coiled Coil (CC) in blue. *Schmidtea mediterranea* (*Smed*).

4.2. The *bls* family is taxonomically restricted to the Tricladida order

A BLAST search using *S. mediterranea bls* sequences against non-redundant transcriptomic and proteomic databases of all species (NCBI) produced no significant results. More specific BLAST searches against genomic and transcriptomic datasets for Platyhelminth species (80) (NCBI and Planmine) indicated that homologs of the *bls* family are only found in species of the order Tricladida (planarians) (Figure R1.4A): *Schmidtea polychroa* (*Spol*), *Dugesia japonica* (*Djap*), and the sexual *S. mediterranea* strain (*Smes*) (Figure R1.4B). Although genomic databases are only available for a few Lophotrochozoa species, this result suggests that the *bls* family is taxonomically restricted to order Tricladida. Interestingly, a BLAST search of the available transcriptomic databases for Tricladida species returned more than one hit for those species (Figure R1.4C, 1.5), with a high degree of similarity at the nucleotide level (Annex I). Phylogenetic analysis performed with amino acid sequences revealed that the *bls5* subfamily was present in all Tricladida species studied, the *bls3* subfamily was present in *Smed*, *Spol*, and *Djap*, and the *bls2* subfamily was present only in *Smed* (Figure R1.4C, 4D). However, it should be borne in mind that transcriptomic databases for Tricladida species other than *Smed* are incomplete. These findings suggest that the *bls* family is taxonomically restricted to Tricladida.
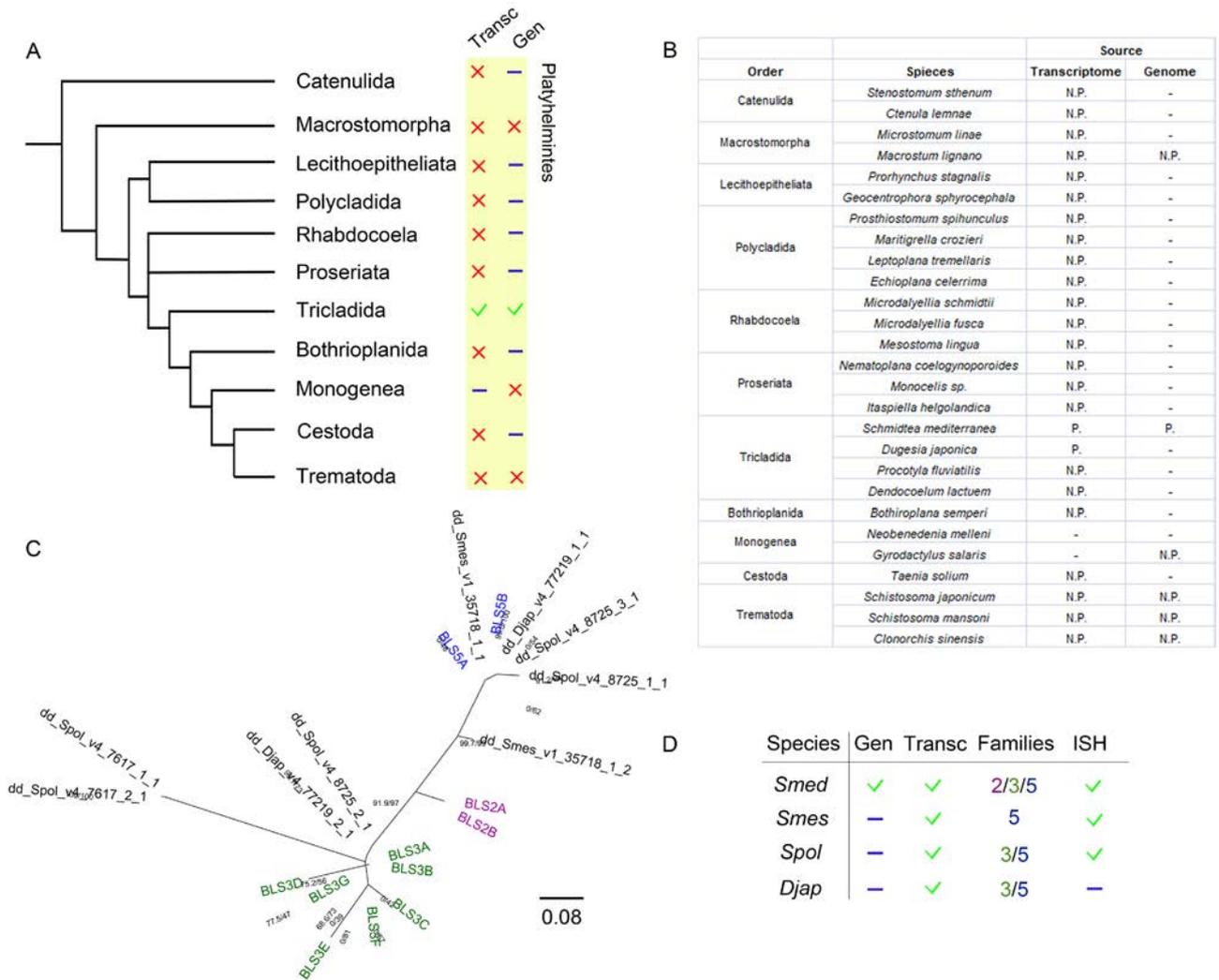
Figure R1.4: Genomic features of *bls* subfamilies in Tricladida species. (A) Presence of *bls* homologs in the transcriptomic (Transc) and genomic (Gen) available databases from different Platyhelminth species. Green check means that presence of some homolog, red cross indicates that no homologs have been identified and blue line indicates no available data. **(B)** All used species in the study, belonging to different Platyhelminthes Orders. (P.) indicates *bls* presence, (N.P.) indicates non presence and (-) indicates no availability data. **(C)** Phylogenetic tree of the *bls* homologs in the Tricladia Order using amino acidic sequences. *bls5* subfamily is present in all species, *bls3* was not found in *Smes*, and *bls2* was only found in *Smed*. Scale indicates expected amino acidic substitution per site. **(D)** *bls* homologs found in the available genome (Gen) and transcriptomes (Transc) of planarian species. ISH expression detection is indicated. Green check indicates presence; blue line indicates no available data. ISH, *in situ* hybridization, *Smed*, *Schmidthea mediterranea* (asexual strain); *Smes*, *Schmidthea mediterranea* (sexual strain); *Spol*, *Schmidthea polychroa*; *Djap*, *Dugesia japonica*.

Species	SP	CC	Planmine Id	Molecular Weight (Kd)
Djap	N	Y	dd_Djap_v4_77219_2_1	12,77
Djap	Y	Y	dd_Djap_v4_77219_1_1	16,65
Spol	N	Y	dd_Spol_v4_8725_2_1	15,97
Spol	Y	Y	dd_Spol_v4_7617_2_1	17,59
Spol	Y	Y	dd_Spol_v4_7617_1_1	17,58
Spol	N	Y	dd_Spol_v4_8725_3_1	7,15
Spol	N	Y	dd_Spol_v4_8725_1_1	10
Smes	N	Y	dd_Smes_v1_35718_1_1	15,07
Smes	N	Y	dd_Smes_v1_35718_1_2	15,27

Figure R1.5: Proteomic features of *bls* subfamilies. Amino acid sequence analysis of Tricladida BLS protein homologs: presence of the signal peptide (SP) and coiled coild; their molecular weight, and the sequences. The Planmine Id for each one is included. *Djap, Dugesia japonica. Spol, Schmidtea polychroa. Smes, Schmidtea mediterranea* sexual strain.

4.3. Subfamilies *bls2*, *bls3* and *bls5* are expressed in secretory cells

Although the 3 transcribed *bls* subfamilies (*bls2*, *bls3*, and *bls5*) shared a high percentage of sequence identity at nucleotide level (Figure R1.1B), we designed riboprobes spanning different gene regions (Figure R1.6A; Annexe I, IV) to specifically detect genes from each subfamily. Whole-mount *in situ* hybridization (WISH) with each riboprobe revealed the same pattern of expression and labelled specific dorsal-prepharyngeal cells (Figure R1.6A, 7A). Double fluorescence *in situ* hybridization (FISH) revealed coexpression of genes from the 3 families in the same cells although showing not identical subcellular localization (Figure R1.6B), confirming riboprobe specificity.
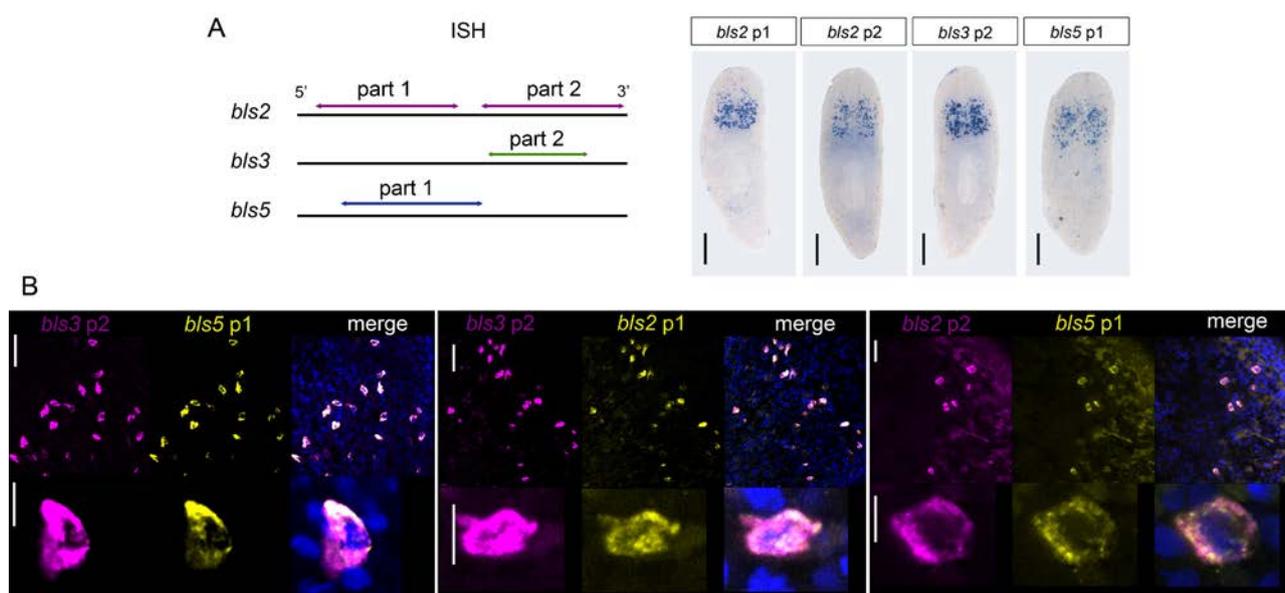
Figure R1.6: *bls2*, *bls3* and *bls5* are expressed in the same cell type in intact animals. (A) Scheme indicating the riboprobes designed for each gene family. WISH with the different riboprobes in intact animals showing the similar expression pattern. **(B)** Double FISH combining all specific riboprobes. Each panel represents each gene combination. A magnification is also shown. All riboprobes colocalize in most of the cells, and magnifications demonstrate that riboprobes present a different cellular distribution. Images from (B) correspond to confocal images. Scale bars: (A) are 500 μm. In (B) are 50 μm and 10 μm in magnifications.

The *bls3* riboprobe revealed that *bls*+ cells were located dorsally and in the marginal cells throughout the body (Figure R1.7A). These *bls*+ cells were differentiated, since they were insensitive to irradiation (Figure R1.7B), and corresponded to secretory cells, since they coexpressed *dd4277*, a secretory and parenchymal cell marker (98,99) (Figure R1.7C). *bls3* was not expressed in blastemas during regeneration, but re-established its expression pattern according to the remodelling of the fragment in question (Figure R1.7D). Interestingly, WISH in sexual *S. mediterranea* (*Smes*) and *S. polychroa* (*Spol*) revealed the same expression pattern as observed for *Smed* (Figure R1.7E), supporting a conserved function among Tricladida species (Figure R1.4D).

Taken together, our data indicate that genes from subfamilies *bls2*, *bls3*, and *bls5* are expressed in a specific subpopulation of secretory prepharyngeal cells in planarians.
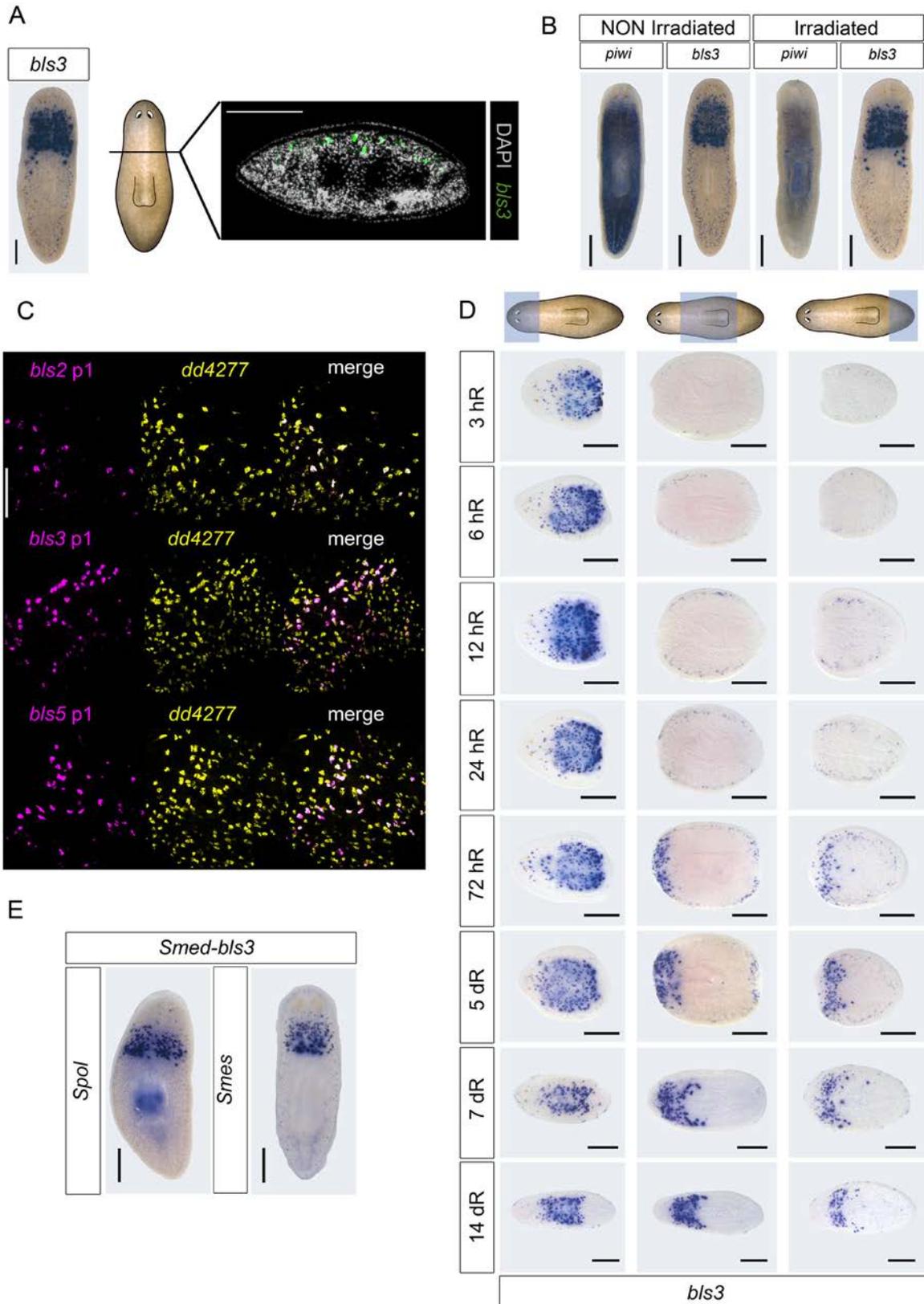
Figure R1.7: *bls* expression pattern in intact and regenerating animals. (A) *bls3* WISH (blue) in whole mount and FISH (green) in a transversal section. Nuclei are stained with DAPI. **(B)** WISH of *bls3* and *piwi* in non-irradiated and irradiated animals. After irradiation the neoblast marker (*piwi*) expression decreases but not *bls3*, corroborating that *bls* gene family is expressed in differentiated cells. **(C)** *bls2*, *bls3* and *bls5* co-expression with *dd4277*. **(D)** WISH of *bls3* during regeneration at different time points. From 3hR to 24hR no *bls3* expression is observed in the blastemas. From 72hR to 14dR the new expression and redistribution of *bls3* is observed. **(E)** WISH of *bls3* in *Schmidtea polychroa* (*Spol*) and *Schmidtea mediterranea* sexual strain (*Smes*), showing the same expression pattern than in *Smed*. Scale bars: (C), (D) and F are 500 µm. In B are 50 µm and 10 µm in magnifications. Scale bars: 200 µm in (A). In (B), (D) and (F) are 500 µm.

4.4. *bls* inhibition promotes faster regeneration

Because *bls2*, *bls3*, and *bls5* genes share a high percentage of identity (Figure R1.1B), specific inhibition of any of these genes using RNAi was technically impossible. Furthermore, the high level of shared identity and cellular colocalization (Figure R1.6B) suggested that at least some paralogs may perform similar functions. For this reason we designed double-stranded RNAs (dsRNAs) corresponding to a highly conserved region in order to inhibit genes of each of the 3 subfamilies (Figure R1.8A). qPCR analysis using primers specific to each subfamily (Figure R1.8B, Annex IV) showed that expression levels of each of the 3 subfamilies were down-regulated after RNAi. Sequencing of the fragments amplified by each qPCR corroborated inhibition of the genes of each of the 3 subfamilies (see Materials and Methods; Figure R1.8C). These animals are referred to henceforth as *bls2/3/5* (RNAi) animals.
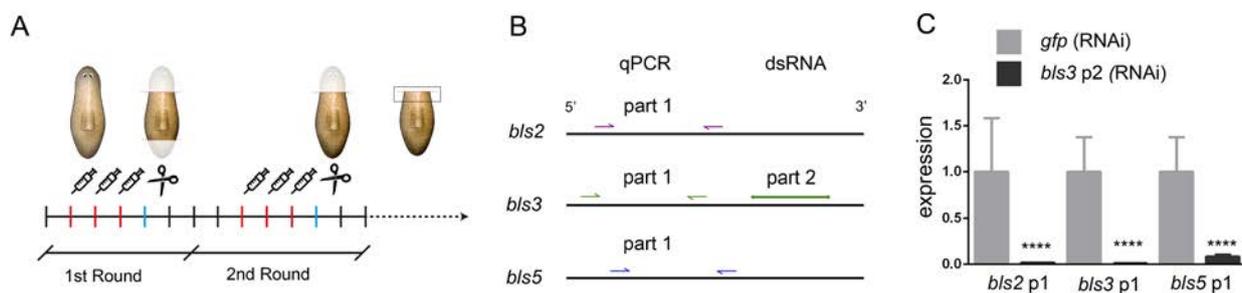
Figure R1.8: dsRNA of *bls3* inhibits all *bls* genes during regeneration. (A) Cartoon illustrating the protocol of RNAi inhibition during planarians regeneration. One week starved planarians were injected 3 consecutive days and amputated the following day. The following 3 days, planarians were let to regenerate (this is what is called one round of inhibition). Planarians were amputated anterior and posteriorly, and a second round of inhibition was performed only with trunk fragments. At the second round trunks were amputated just anteriorly. **(B)** Scheme indicating the fragments used for RNAi and qPCR analysis. (C) qRT-PCR analysis quantifying *bls2*, *bls3* and *bls5* expression after *bls3* inhibition at 3dR, demonstrating that all three subfamilies were down-regulated after injection of *bls3* dsRNA. Relative expression is plotted as $2^{-\Delta\Delta CT}$ values. Data are plotted as mean and error bars represent s.e.m. (****$P<0.0001$).

RNAi of *bls2/3/5* confirmed our initial observation of faster regeneration after head amputation in planarians. We observed earlier differentiation of the eye spots (Figure R1.9A), and earlier differentiation of photoreceptor cells (identified by anti-arrestin immunostaining): after 3 days of regeneration (3dR) the optic chiasm was visible in most *bls2/3/5* (RNAi) animals but not in control animals (Figure R1.9B). In addition to the visual system, other anterior structures such as the brain branches and chemoreceptors regenerated faster than controls (Figure R1.9B), as evidenced by quantification of *gpas*+ (86) and *cintillo*+ (225) cells, respectively. Quantification of *pitx*+ cells (105,106) revealed an increase the number of differentiated neural cells in the blastema of *bls2/3/5* (RNAi) planarians as early as 18 hours of regeneration (hR) (Figure R1.9C). These results demonstrate that inhibition of *Smed-bls2/3/5* promotes faster regeneration.
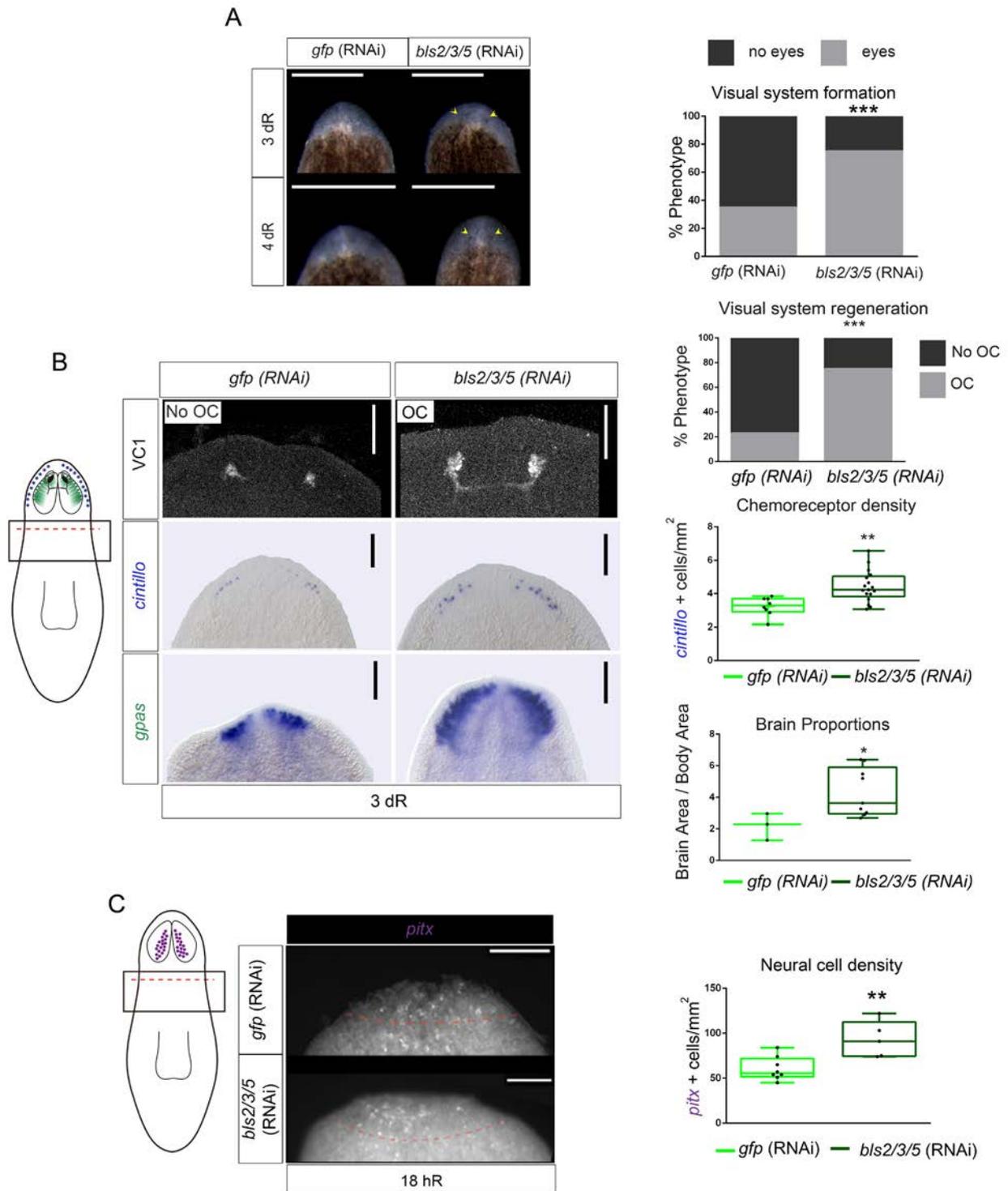
Figure R1.9: *bls2/3/5* (RNAi) animals regenerate faster. (A) *in vivo* images of planarians showing that in *bls2/3/5* (RNAi) animals regenerating eyes are more evident (yellow arrows) than in controls at 3 and 4dR (n of controls=23, n of RNAi=23, ***$P<0.001$). **(B)** Immunohistochemistry against arrestin (VC1), labelling the visual system, WISH of *cintillo* (chemoreceptors) and *gpas* (brain branches). Quantification of the appearance of the optic chiasm (n of controls=23, n of RNAi=23, ***$P<0.001$), *cintillo*+ cells (n of controls=8, n of RNAi=17, **$P<0.01$) and *gpas*+ area (n of controls=3, n of RNAi=9, *$P<0.05$) in *gfp* (RNAi) and *bls2/3/5* (RNAi) animals is shown. Illustration indicating where *gpas* and *cintillo* are expressed, and where arrestin (VC1) is detected, is shown. The amputation level and the area analyzed at 3dR are indicated (dashed red line and black square, respectively). (C) Representation of *pitx* expression in control animals. Red dashed line represents amputation level, and black square represents the area analyzed at 18 hR. FISH of *pitx* shows an increase in *pitx*+ cells in the blastema (dashed red line limits it) at 18 hR. Quantification of *pitx*+ cells / mm² is showed (n of controls=8, n of RNAi=5, **$P<0.01$). Scale bars: 250 μm in (A), 100 μm in (B) and (C).

4.5. *bls* attenuates cell proliferation and promotes cell death after injury

To understand the mechanism by which *Smed-bls2/3/5* (RNAi) promotes faster regeneration, we analysed the proliferative and apoptotic responses triggered by amputation. In planarians, amputation triggers a general proliferative response, which peaks at 6 hR, and a local response that peaks at 48 hR. Quantification of mitotic cells using an anti-phospho-histone 3 (PH3) antibody (132) revealed an increase in the mitotic response at both 6 hR and 48 hR in *bls2/3/5* (RNAi) versus control animals (Figure R1.10A, 10B). The apoptotic response after amputation consists of 2 apoptotic peaks: one at 4 hR, which occurs close to the wound, and a second at 3 days of regeneration (dR), which is generalized (147). Using a TUNEL assay (Figure R1.10C) and by quantifying caspase-3 enzymatic activity (Figure R1.10D) we demonstrated lower rates of apoptosis in *bls2/3/5* (RNAi) versus control planarians at both time points.
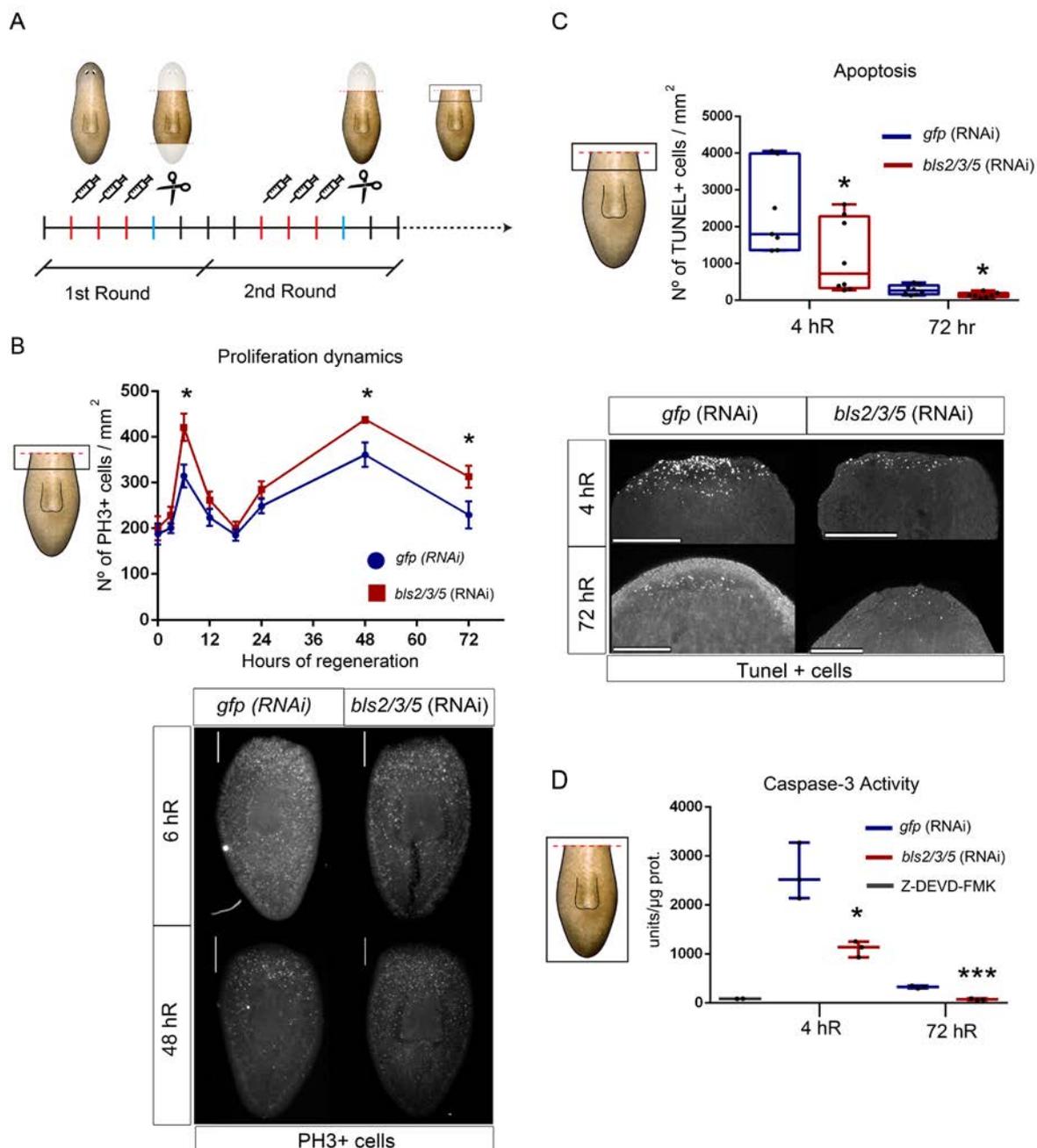
Figure R1.10: *bls2/3/5* (RNAi) animals show an increase in proliferation and a decrease of apoptosis during anterior regeneration. (A) Cartoon illustrating the protocol of RNAi inhibition during planarians regeneration. One week starved planarians were injected 3 consecutive days and amputated the following day. The following 3 days, planarians were let to regenerate (this is what is called one round of inhibition). Planarians were amputated anterior and posteriorly, and a second round of inhibition was performed only with trunk fragments. At the second round trunks were amputated just anteriorly. **(B)** Quantification of PH3+ cells at different regeneration time points (n of controls >5, n of RNAi>5, *P<0.05) Anti- PH3 immunostaining of *gfp* (RNAi) and *bls2/3/5* (RNAi) animals are showed bellow. **(C)** Quantification of TUNEL+ cells in *bls2/3/5* (RNAi) and controls (n of controls>7, n of RNAi>7, *P<0.05). TUNEL images are showed bellow. Images correspond to Z projections. **(D)** Quantification of caspase-3 activity in *bls2/3/5* (RNAi) animals and controls (n of controls=3, n of RNAi=3, *P<0.05, ***P<0.001). Each biological replicate represents 5 animals. Illustrations show the amputation plane and the area analyzed (dashed red line and black square, respectively). Scale bars: 500 μm in (B) and 200 μm in (C).
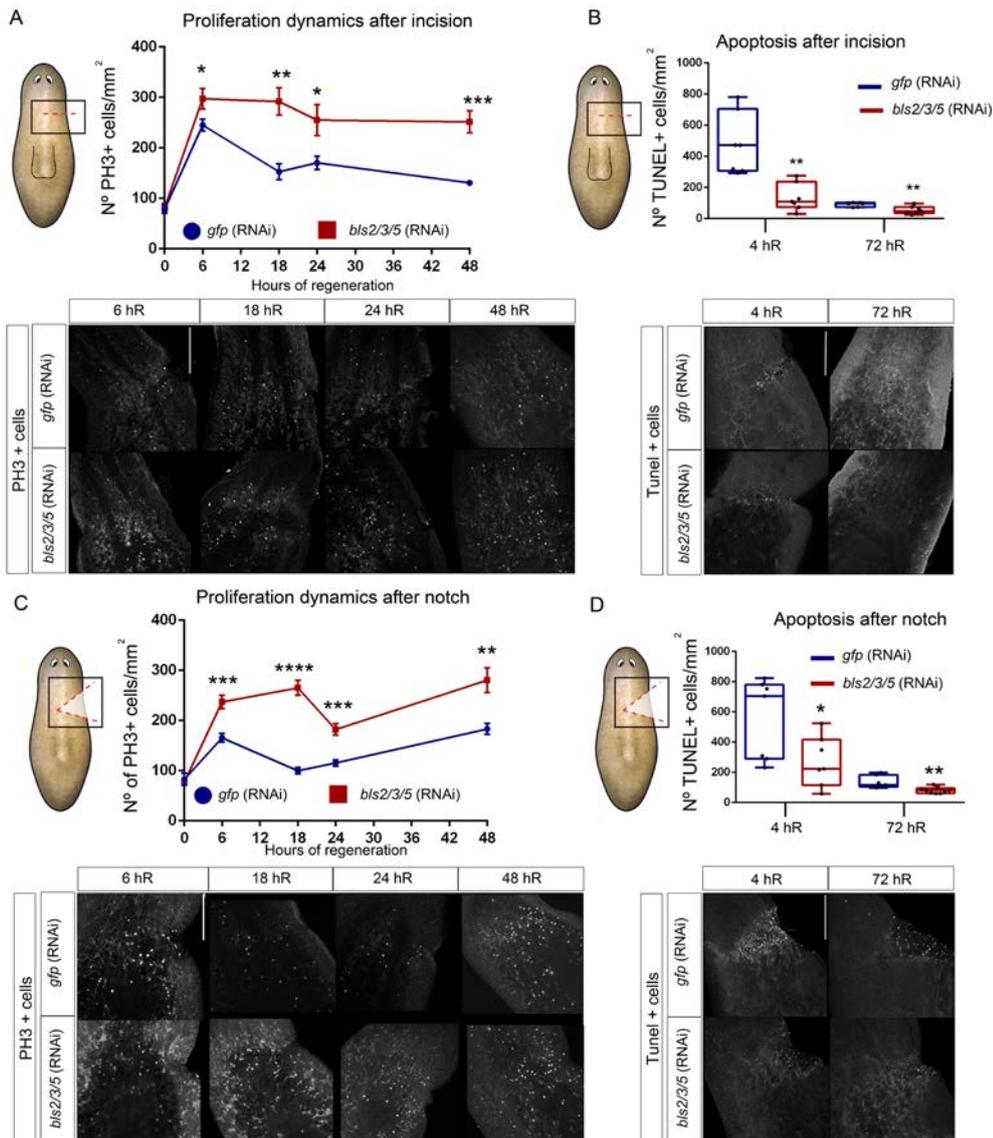
Figure R1.11: *bls2/3/5* (RNAi) animals presented an increase in proliferation and a decrease of apoptosis after any injury. (A) Quantification of PH3+ cells at different time points after incision shows an increment of mitotic cells/mm^2 in *bls2/3/5* (RNAi) animals. Time points were 6 hR (n of controls=5, n of RNAi=7, *P<0.05), 18 hR (n of controls=6, n of RNAi=7, **P<0.01), 24 hR (n of controls=6, n of RNAi=8, *P<0.05) and 48 hR (n of controls=8, n of RNAi=8, ***P<0.001). anti-PH3 immunostaining images are shown below. **(B)** Quantification of TUNEL+ cells show a decrease of apoptotic cells in *bls2/3/5* (RNAi) animals after incision at 4 hR (n of controls=7, n of RNAi=7, **P<0.01) and 72 hR (n of controls=6, n of RNAi=7, **P<0.01). TUNEL images are shown below. **(C)** Quantification of PH3+ cells at different time points after notching shows an increment of mitotic cells/mm^2 in *bls2/3/5* (RNAi) animals at 6 hR (n of controls=8, n of RNAi=9, ***P<0.001), 18 hR (n of controls=8, n of RNAi=8, ****P<0.0001), 24 hR (n of controls=8, n of RNAi=9, ***P<0.001) and 48 hR (n of controls=8, n of RNAi=8, **P<0.01). anti-PH3 immunostaining images are shown below. **(D)** Quantification of TUNEL+ cells show a decrease of apoptotic cells in *bls2/3/5* (RNAi) animals after notching at 4 hR (n of controls=7, n of RNAi=7, *P<0.05) and 72 hR (n of controls=7, n of RNAi=9, **P<0.01). TUNEL images are shown below. All images correspond to Z projections. Illustrations show the amputation plane and the area analyzed (dashed red line and black square, respectively). Scale bars: 200 μm in all panels.

Distinct molecular and cellular responses are induced during healing of amputated tissue, notches (which imply tissue loss), and incisions (in which no tissue is removed). While control of cell proliferation and cell death is required in all scenarios, incision gives rise to just the first proliferative and apoptotic peaks. To examine how general was the role of *bls2/3/5* in attenuating cell proliferation and promoting cell death, we analysed the response to notching and incision in *bls2/3/5* (RNAi) animals. In both situations, compared with controls RNAi animals showed an increase in the number of mitotic cells (Figure R1.11A, 11C) and a decrease in apoptosis (Figure R1.11B, 11D), indicating that *bls2/3/5* attenuates proliferation and promotes cell death.

These findings indicate that S*med-bls2/3/5* attenuates cell proliferation and promotes cell death after any injury type, regardless of whether tissue is removed.

4.6. Cells are more numerous but smaller in starved *bls* (RNAi) planaran, resulting in no overall change in body size

The pattern of *bls2*, *bls3*, and *bls5* expression in secretory cells in the prepharyngeal region and along the planarian margin suggests that these peptides may play a role in controlling cell proliferation and cell death, not only after injury but also during homeostasis, since planarians undergo continuous growth and degrowth according to nutrient availability. These changes in size are thought to be primarily due to modulation of cell number (185,189) through regulation of the balance between proliferation and apoptosis (147,226). In nutrient-poor environments planarians shrink by decreasing mitosis and increasing cell death. To determine whether *bls2/3/5* participates in regulating the proliferation/apoptosis equilibrium and body size during degrowth, we injected starved animals with *bls2/3/5* dsRNA for 3 weeks (Figure R1.12A, 12B). After 3 weeks of dsRNA injection, starved *bls2/3/5* (RNAi) animals showed increased mitosis (Figure R1.12C) and decreased apoptosis compared with controls (Figure R1.12D, 12E). While this alteration in the proliferation/apoptosis equilibrium did not give rise to larger animals (Figure R1.12F), total cell number was higher in RNAi-injected animals versus controls (Figure R1.12G). The fact that total cell number but not body size was increased in *bls2/3/5* (RNAi) animals implies a decrease in cell size. To examine changes in cell size we focused our analysis on the epidermis, since epidermal cells form a monolayer that can be easily imaged in 3 dimensions.
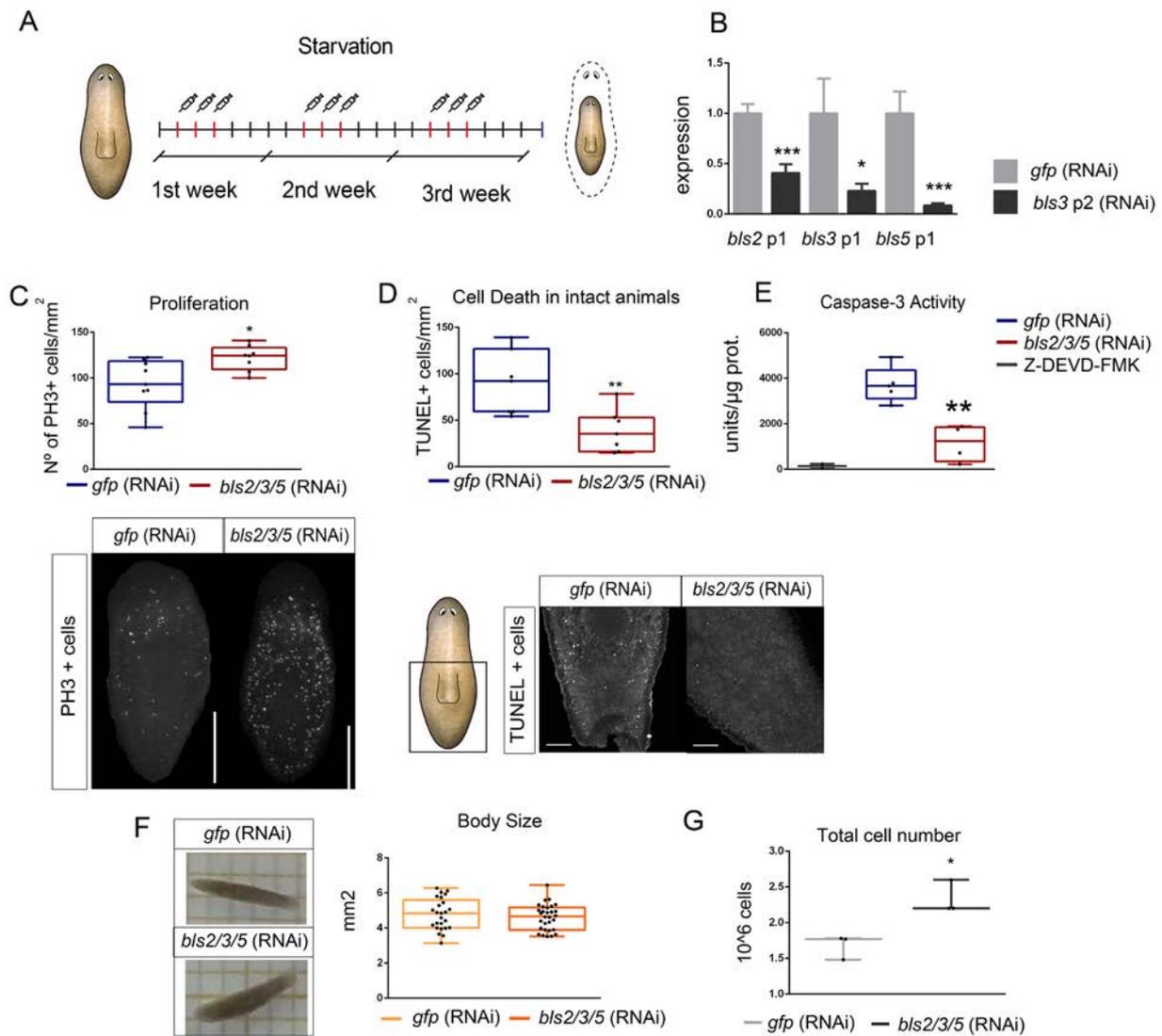
Figure R1.12: In starved conditions, *bls2/3/5* (RNAi) animals show an increase in proliferation and a decrease in cell death, which leads to cell number increment but no bigger animals. (A) Scheme of the RNAi procedure to inhibit *bls2/3/5* in starved planarians. Each week, animals were injected 3 consecutive days. At the end of the third week, animals were fixed and analyzed. **(B)** qRT-PCR analysis measuring *bls2*, *bls3* and *bls5* expression after 3 rounds of *bls3* p2 inhibition, demonstrated that all three genes are down-regulated. Relative expression is plotted as $2^{-\Delta\Delta CT}$ values. Data are plotted as mean and error bars represent s.e.m. (*$P<0.05$; ***$P<0.001$). **(C)** Quantification of PH3+ cells after 3 weeks of RNAi treatment (n of controls=9, n of RNAi=8, *$P<0.05$). anti-PH3 immunostaining images of *bls2/3/5* (RNAi) animals and controls. **(D)** Quantification of TUNEL+ cells after 3 weeks of RNAi treatment (n of controls=7, n of RNAi=7, **$P<0.01$). TUNEL assay images of *bls2/3/5* (RNAi) animals and controls, corresponding to the posterior region of the animals. **(E)** Quantification of caspase-3 activity in *bls2/3/5* (RNAi) animals and controls (n of controls=5, n of RNAi=4, **$P<0.01$). Each biological replicate represents 5 animals. **(F)** Area quantification of *in vivo* animals (n of controls=25, n of RNAi=30, n.s.). **(G)** Cell number quantification (n of controls=3, n of RNAi=3, *$P<0.05$). Each biological replicate represents 5 animals. Images from C correspond to Z projections. Scale bars: 500 µm in (C), 100 µm in (D) and one side of a square equate to 1 mm in (F).

Nuclear staining revealed a higher density of epithelial cells in *bls2/3/5* (RNAi) animals (Figure R1.13A, 13B), and quantification of mean epidermal cell area confirmed that this parameter was reduced in *bls2/3/5* (RNAi) animals as compared with controls (Figure R1.13B'). A decrease in mean epidermal cell area could be due not to a reduction in the total cell volume but to narrowing of the cells in *bls2/3/5* (RNAi) animals. To quantify epidermal cell volume we measured epidermal cell height (i.e., the mean distance from the apical to the basal margin of the cell) in animals immunostained with 6G10 antibody. We observed no differences in apical-basal distance in *bls2/3/5* (RNAi) animals with respect to controls (Figure R1.13C, 13C').
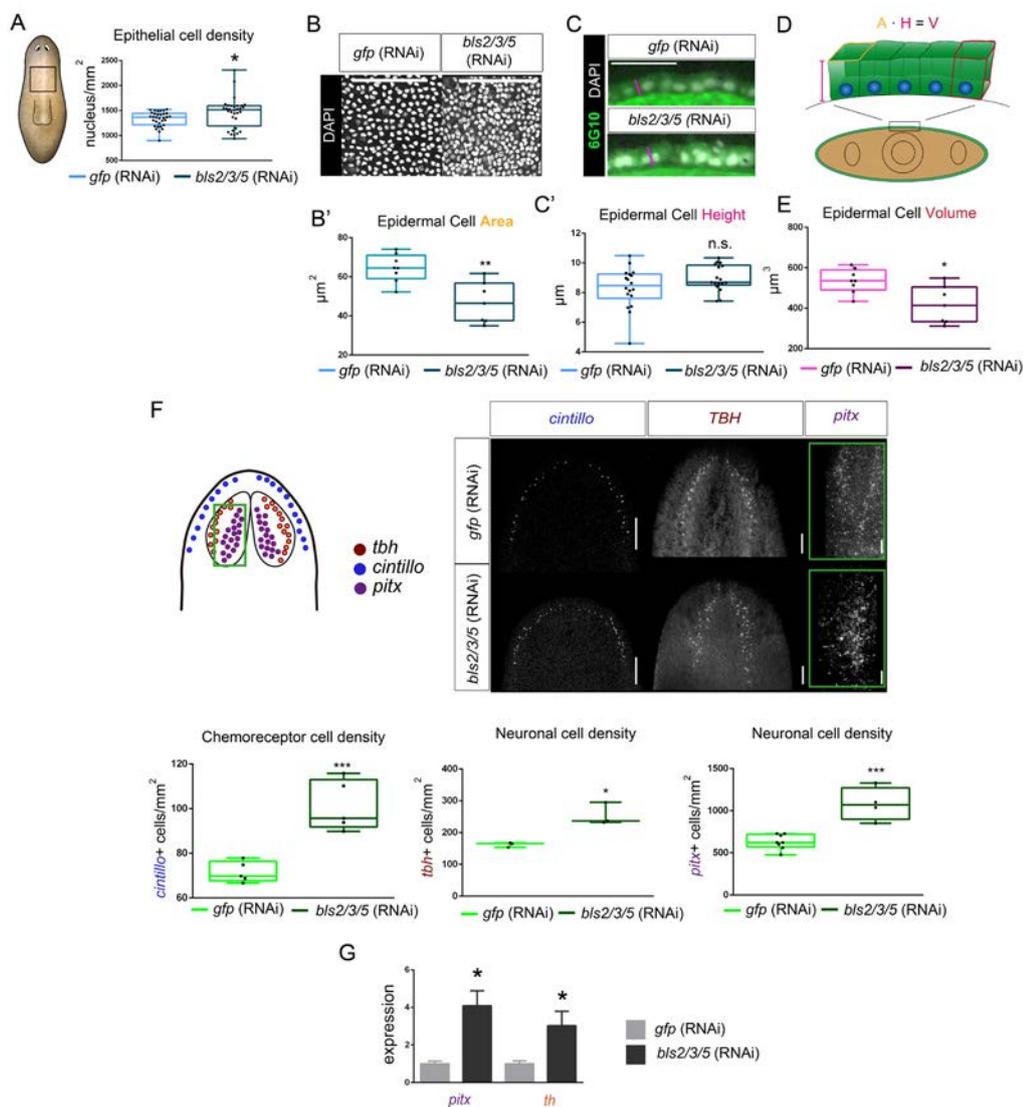
Figure R1.13: In starved conditions, *bls2/3/5* (RNAi) animals show cell number increment and cell size reduction. **(A)** Quantification of Nuclei of epidermal cells stained with DAPI per area (n of controls=15, n of RNAi=15, *P<0.05). The illustration indicates the area quantified with a black square. **(B)** DAPI staining of epithelial cells of the prepharyngeal region. **(B')** Quantification of the epidermal cell average Area (A) (n of controls=8, n of RNAi=7, **P<0.01). **(C)** Transversal sections of planarian epidermis immunostained with anti-6G10. The distance from the basal to the apical part of the cells measured (epidermal cells height, H) is indicated with a pink line. **(C')** Quantification of the H is shown (n of controls=18, n of RNAi=17, n.s.). (D) Schematic illustration of the measurements performed to quantify the V. **(E)** Quantification of the Epidermal Cell Volume (V) (n of controls=8, n of RNAi=7, *P<0.05). (F) Quantification of *cintillo*+ cells (n of controls=5, n of RNAi=5, ***P<0.001) *pitx*+ cells (of controls=8, n of RNAi=4, ***P<0.001) *tbh*+ cells (of controls=3, n of RNAi=3, *P<0.05). Confocal images of the expression of neural (*tbh* and *pitx*) and chemoreceptor (*cinillo*) markers in *bls2/3/5* (RNAi) animals and controls. The illustration shows the expression of *tbh*, *pitx* and *cintillo*, and the green square shows the area were *pitx* was quantified. (G) qRT-PCR quantification of the expression of neural markers (*pitx* and *th*). Relative expression is plotted as $2^{-\Delta\Delta CT}$ values. Data are plotted as mean and error bars represent s.e.m. (*P<0.05). Images from (B) and (F) correspond to Z projections. Scale bars: 20 µm in (B) and (C); 200 µm in (F).

Multiplication of mean cell area by mean cell height confirmed a decrease in epidermal cell volume in *bls2/3/5* (RNAi) animals versus controls (Figure R1.13D, 13E). Changes in specific neural populations were evaluated by confocal imaging (Figure R1.13F) and qPCR (Figure R1.13G). The density of serotoninergic (*pitx*+) (105,109), octapaminergic (*tbh*+) (214), *dopaminergic* (*th*+) (65) neurons and of chemoreceptors (*cintillo*+) (225) was increased in *bls2/3/5* (RNAi) animals.

Importantly, continuous inhibition of *bls2/3/5* (RNAi) for 4 weeks resulted in the formation of overgrowths composed of epidermal progenitor cells *(nb21+, agat1+)* (214) (Figure R1.14).

These data indicate that *bls2/3/5* promotes cell death and attenuates mitosis during periods of shrinkage. *bls2/3/5* inhibition in starved planarians prevents the necessary reduction in cell number. Because cell size is reduced in *bls2/3/5* (RNAi) versus control animals, the increase in cell number observed in the former does not translate to larger body size. However, the accumulation of cells following long term inhibition does lead to overgrowths.
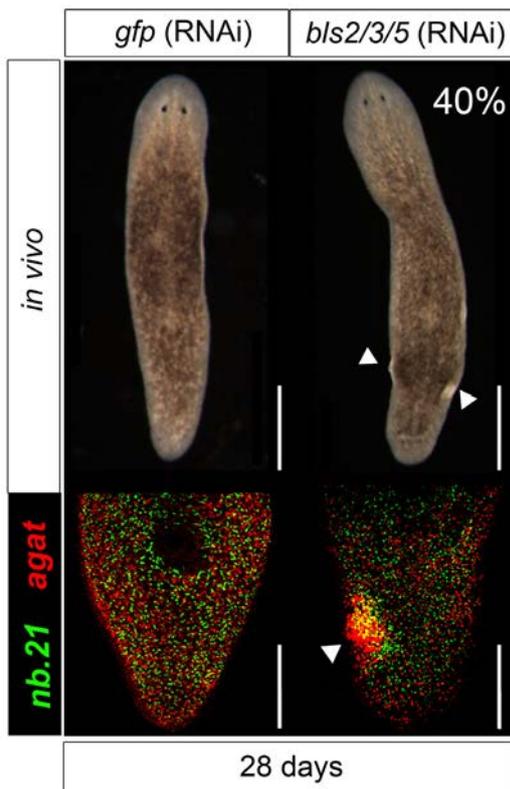
Figure R1.14: In long starved conditions, *bls2/3/5* (RNAi) animals generate overgrowths. Overgrowths observed after 4 weeks of *bls2/3/5* inhibition in starved animals and FISH with *agat* and *nb.21* riboprobes (n of controls=40, n of RNAi=35). White arrows point to over-growths. Images from (F) and K correspond to Z projections. Scale bars in *in vivo* imatges are: 500 μm and 200 μm in FISH and *in vivo* images, respectively.

4.7. *bls* (RNAi) in fed planarians results in increases in cell number and body size

Planarians grow in size in nutrient-rich environments. This growth is due to an increase in cell number resulting from an increase in the mitosis:apoptosis ratio (147,226). Our previous findings suggest that *bls2/3/5* inhibition in continuously fed planarians may lead to an increase in cell number and possibly also in body size. To test this hypothesis, planarians fed twice per week were injected with *bls2/3/5* dsRNA for 3 weeks (Figure R1.15A,15B).

Compared with controls, these animals showed an increase in the rate of mitosis (Figure R1.15C) and a decrease in rate of apoptosis (Figure R1.15D). Furthermore, during this 3-week period RNAi animals grew faster and reached a larger size (Figure R1.15E) than controls (Figure R1.15F). Quantification of dissociated cells revealed an increase in total cell number in *bls2/3/5* (RNAi) animals after 3 weeks of RNAi (Figure R1.15G).
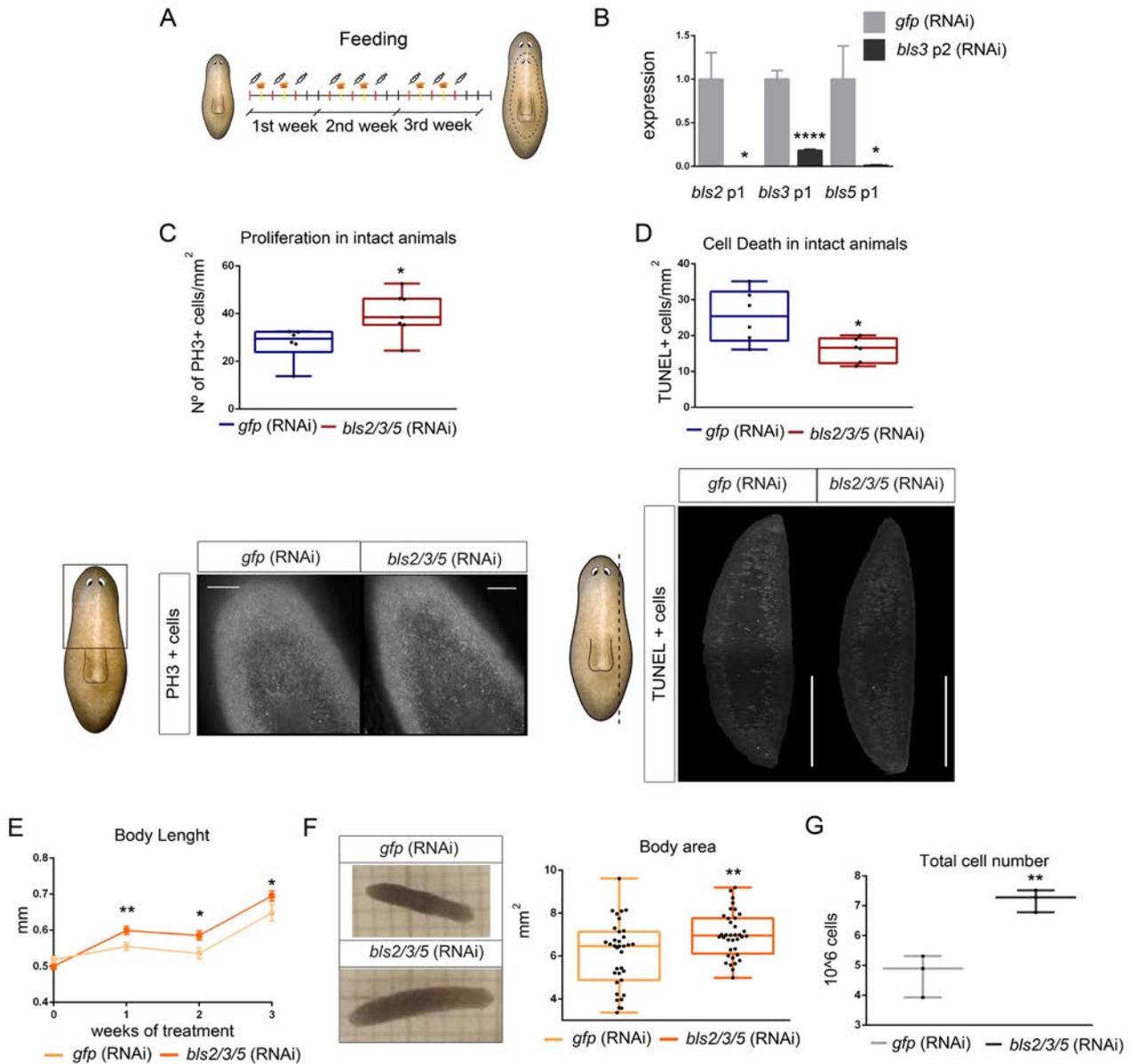
Figure R1.15: In fed conditions, *bls2/3/5* (RNAi) animals show an increase in proliferation and a decrease in cell death, which leads to cell number increment and bigger animals. (A) Scheme of the RNAi procedure to inhibit *bls2/3/5* in fed planarians. Each week, animals were injected 3 non consecutive days, being feed the other two days. At the end of the third week, animals were fixed (three days after the last injection) and analyzed. **(B)** qRT-PCR measuring *bls2*, *bls3* and *bls5* expression after 3 rounds of *bls3* inhibition, demonstrating that all three genes are downregulated. Relative expression is plotted as $2^{-\Delta\Delta CT}$ values. Data are plotted as mean and error bars represent s.e.m. (*P<0.05; ***P<0.001). **(C)** Quantification of PH3+ cells after 3 weeks of RNAi treatment (n of controls=6, n of RNAi=7, *P<0.05). Anti-PH3 immunostaining images of *bls2/3/5* (RNAi) animals and controls showing the increment of the mitotic cells after 3 weeks of the treatment. **(D)** Quantification of TUNEL+ cells after 3 weeks of RNAi treatment (n of controls=6, n of RNAi=6, *P<0.05). TUNEL assay images of *bls2/3/5* (RNAi) animals and controls, corresponding to the posterior region of the animals. **(E)** Length measurement of the RNAi and control animals (n of controls>35, n of RNAi>35, *P<0.05, **P<0.01). **(F)** Area quantification of *in vivo* animals (n of controls=35, n of RNAi=36, **P<0.01). **(G)** Cell number quantification (n of controls=3, n of RNAi=3, **P<0.01). Each biological replicate represents 5 animals. Images from (C) and (D) correspond to Z projections. Scale bars: 200 μm in the entire panel. One side of a square equate to 1 mm in (F).
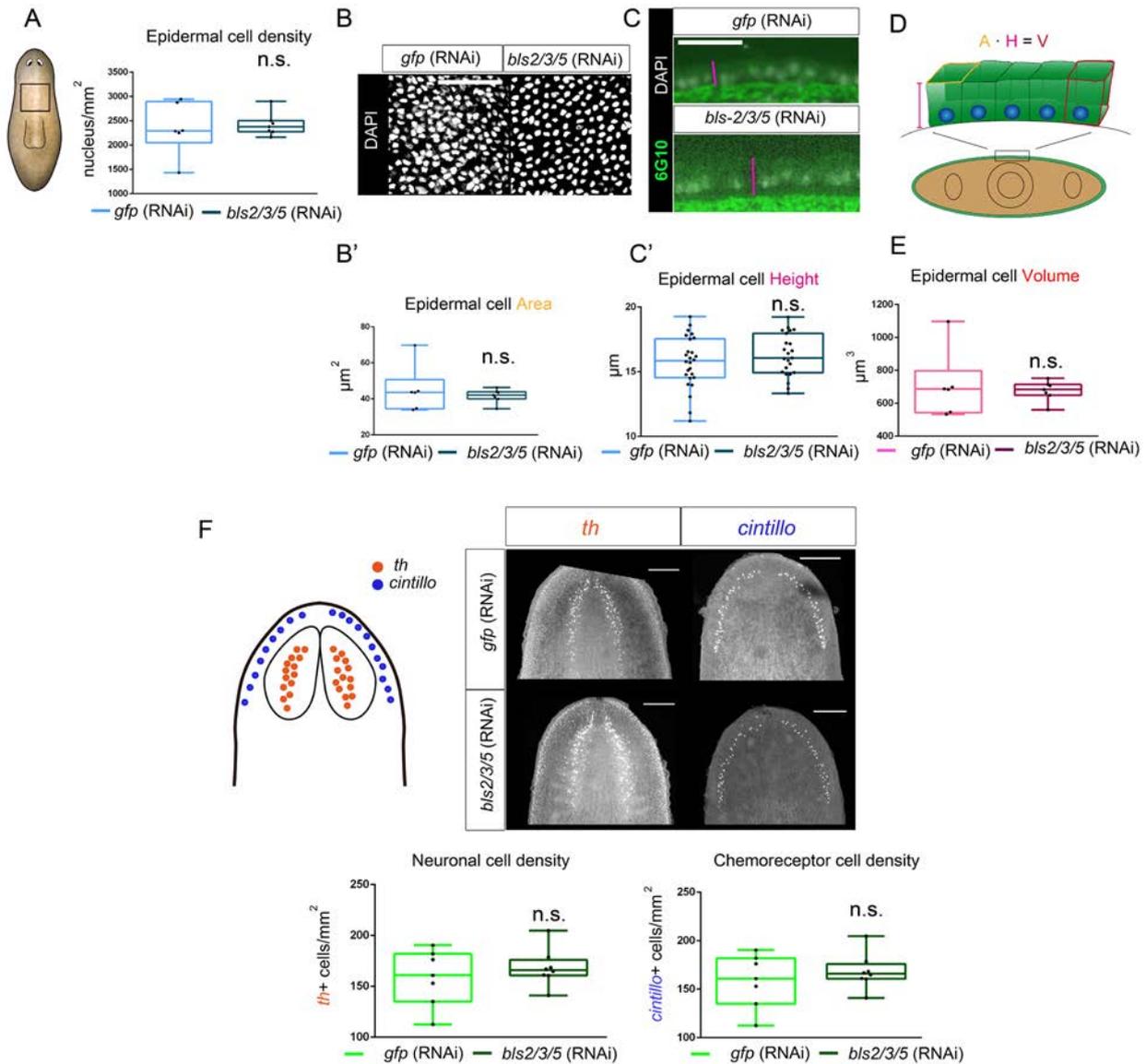
Figure R1.16: In fed conditions, *bls2/3/5* (RNAi) animals show cell number increment, without changing cell size.
(A) Quantification of nuclei of epidermal cells stained with DAPI per area (n of controls=15, n of RNAi=15, ***P<0.001). The illustration indicates the area quantified with a black square. **(B)** DAPI staining of epithelial cells of the prepharyngeal region. **(B')** Epidermal cell average Area (A) quantification (n of controls=6, n of RNAi=7, n.s.). **(C)** Transversal sections of planarian epidermis immunostained with anti-6G10. The distance from the basal to the apical part of the cells measured (epidermal cell high, H) is indicated with a pink line. **(C')** Quantification of the H (n of controls=26, n of RNAi=23, n.s.). **(D)** Illustration of the measurements performed to quantify the V. **(E)** Quantification of the Epidermal Cell Volume (V) (n of controls=6, n of RNAi=7, n.s.). (F) Quantification of *th*+ cells (n of controls=9, n of RNAi=6, n.s.) and *cintillo*+ cells (n of controls=7, n of RNAi=8, n.s.). Confocal images of the expression of neural (*th*) and chemoreceptor (*cintillo*) markers in *bls2/3/5* (RNAi) animals and controls. The illustration shows the expression of *th* and *cintillo*. Images from (B') and (F) correspond to Z projections. Scale bars: 20 µm in (B) and (C); 200 µm in F.

In contrast to the results obtained for starved planarians, no differences in epidermal cell area or volume were observed in fed animals with respect to controls (Figure R1.16-E). Furthermore, quantification of neural and chemoreceptor cells revealed no differences in cell density between RNAi and control planarians (Figure R1.16F).

These data indicate that *Smed-bls2/3/5* also promotes cell death and attenuates the rate of mitosis during growth periods, resulting in an increase in cell number. Remarkably, in fed animals this increase in cell number translates to an increase of body size, since cell size is maintained in this nutrient rich context.

4.8. *bls* transcription is regulated accordingly nutrient intake

Our results demonstrate that *bls2/3/5* subfamilies control the balance of cell proliferation and cell death in planarians not only after injury but also during normal homeostasis. We hypothesize that *bls2/3/5*-mediated signalling may constitute a general mechanism required to attenuate cell proliferation and promote cell death. According to our hypothesis, *bls2/3/5* activity is required in nutrient-poor environments but not when food is readily available. As previously mentioned, planarian growth is sustained by increasing mitosis and decreasing cell death. After feeding, apoptosis remains very low and changes little (Figure R1.17A), but proliferation increases and mitosis peaks at 3 hours post-feeding (hpf) (79,134) (Figure R1.17B).
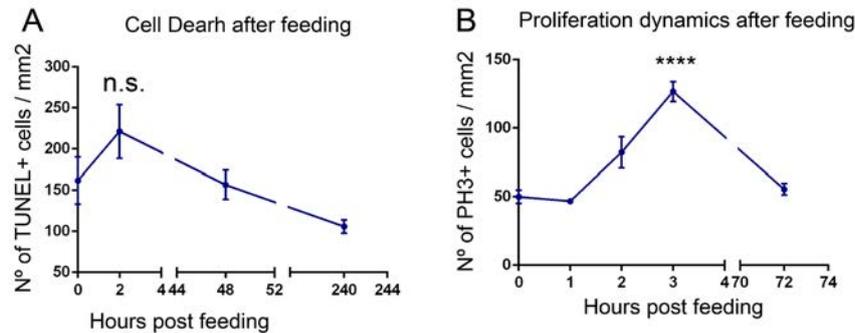
Figure R1.17: Cell death and cell proliferation dynamics after feeding. (A) TUNEL assay at different points after feeding reveals no changes in wild type planarians (n of starved=7, n of 2hpf=6, n.s.). (B) PH3+ cells quantification at different time points after feeding reveals a proliferative peak at 3hpf (n of starved=6, n of 3hpf=6, ****$P<0.0001$).

According to our hypothesis, *bls* expression is actively regulated a few hours after food ingestion to enable subsequent growth. Quantification of mRNA levels of *bls2*, *bls3*, and *bls5* by qPCR at 3 hpf and 24 hpf revealed down-regulation of all 3 *bls* mRNAs (Figure R1.18A). This down-regulation was also confirmed by FISH expression analysis: after feeding (24 hpf) expression of all 3 genes had decreased and/or the expression pattern had expanded with respect to starved conditions (Figure R1.18B, 18C).
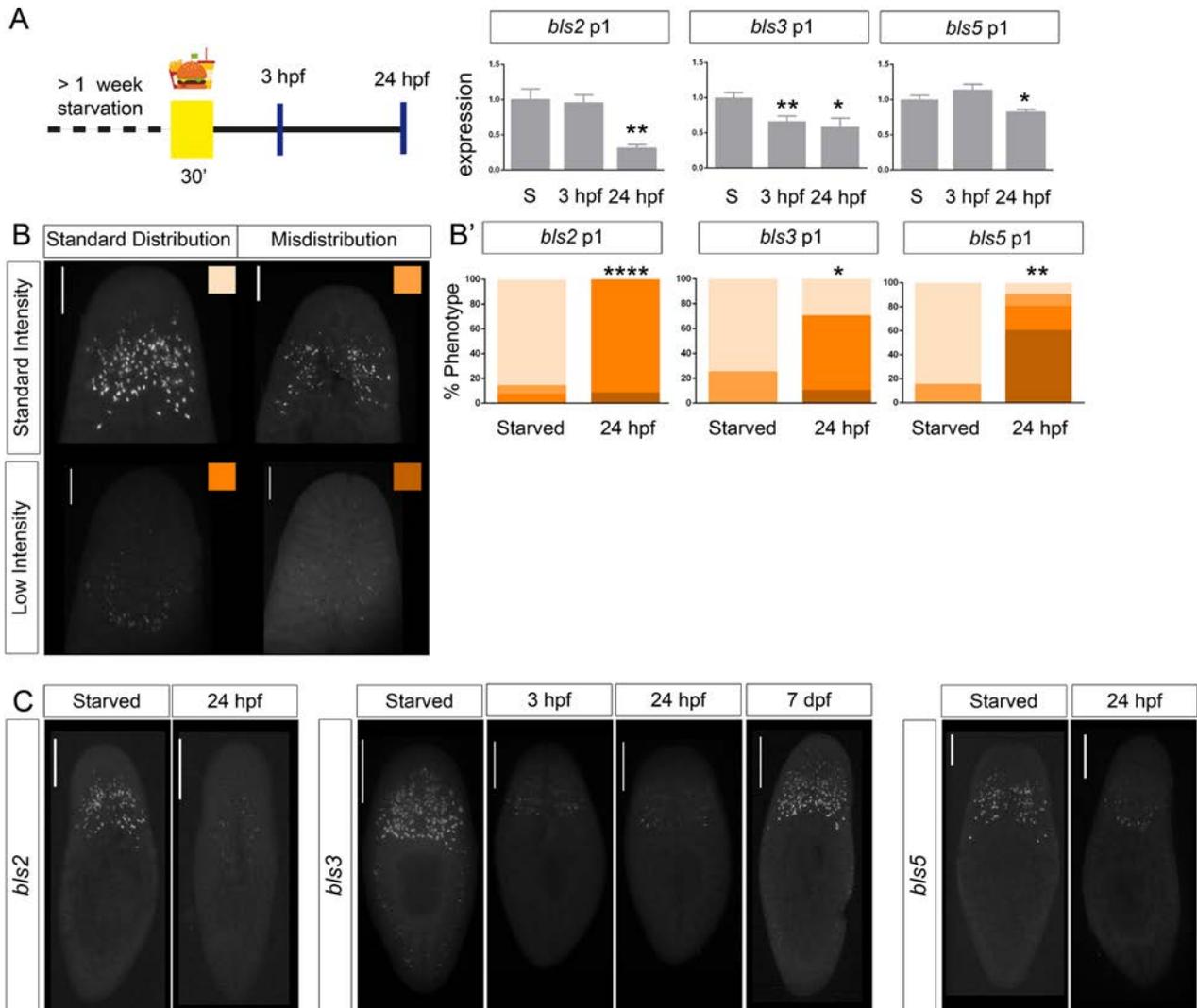
Figure R1.18: *bls2, bls3* and *bls5* are down-regulated by food ingestion. (A) Scheme of the experimental procedure. Animals starved for more than one week, were fed during 30 minutes and fixed at different time points. qRT-PCR quantification of *bls2*, *bls3* and *bls5* expression of starved animals (S) and after 3 hours (3 hpf) and 24 hours (24hpf) post feeding. Relative expression is plotted as $2^{-\Delta\Delta CT}$ values. Data are plotted as mean and error bars represent s.e.m. (*$P<0.05$; **$P<0.01$). **(B)** Representative FISH images of the four expression pattern categories found of *bls2*, *3* and *5* after feeding: high intensity and localized, high intensity and delocalized, low intensity and localized, and low intensity and delocalized. The percentage of each category is shown **(B')** (n of starved>7, n of 24hpf>10, *$P<0.05$, **$P<0.01$****, $P<0.0001$). **(C)** Representative FISH images of *bls2*, *bls3* and *bls5* before feeding and after feeding, demonstrating its expression at 3 hpf and 24 hpf was reduced. But, it was recovered at 7 dpf.

Furthermore, forced reduction of *bls2/3/5* expression by RNAi at 3 hpf enhanced the prolife-rative response, demonstrating that *bls2/3/5* is essential to attenuate proliferation in homeos-tatic conditions (Figure R1.19).
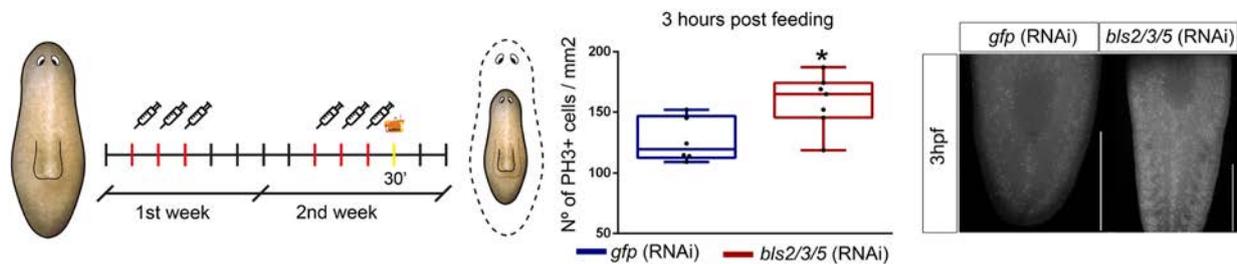
Figure R1.19: *bls2/3/5* (RNAi) animals show an increase in proliferation after feeding. Immunostaining with anti-PH3 in *bls2/3/5* (RNAi) animals 3 hpf (n of controls=6, n of RNAi=7, *$P<0.05$). A scheme with the experimental design is shown. anti-PH3 immunostaining images of control and *bls2/3/5* (RNAi) animals. Scale bars: 500 µm in the entire panel.

Overall, these results show that subfamilies *bls2/3/5* attenuate the mitotic response triggered by feeding. During planarian homeostasis *Smed-bls* transcription may be constantly required to regulate the cell proliferation/cell death ratio, and actively down-regulated after nutrient intake to allow planarian growth. *bls2/3/5* may represent a novel molecular mechanism to regulate cell number in response to nutrient intake.

5. Chapter II. Genomic and transcriptomic analysis reveals new cWnt-pathway related elements required for posterior identity specification

During the last years, our group has been interested in the activity of the planarian's organizing tips and its relation with the cWNT signalling pathway. As explained in the introduction, the role of *wnt1* and *β-catenin1* in the establishment of the posterior organizer has been deeply studied (59,60,150). However, while working in this thesis, new molecular tools came to illuminate new aspects of the field. This is why we sought to start new projects to understand the cWNT pathway, using genome wide approaches. The new results and their impact on our knowledge will be presented in the next chapter.

5.1. *wnt1* (RNAi) RNA-seq reveals transcriptomic changes during posterior planarian regeneration

5.1.1. Strategy of *wnt1* (RNAi) RNA-seq in regenerating planarians to study the establishment of the posterior organizer

To identify cWNT-related genes and their putative functions in establishing the posterior organizer, we performed RNA-seq of *wnt1* (RNAi) regenerating planarians, since it had been demonstrated that *β-catenin1* is having a regulatory function in posterior wounds (59). Previous experiments demonstrate that after strong inhibition of *wnt1* (two rounds of inhibition and amputation), 20% of regenerating planarians regenerate a head instead of a tail (referred to as two-headed planarians), while the rest (80%) are able to close the wound without establishing any posterior or anterior identity (referred to as tailless phenotype) (60,150,151,154) (Figure R2.1A). In order to study the elements involved in the establishment of the posterior organizer, in this thesis, we decided to modify the previously described protocol and perform a milder inhibition of *wnt1* to study the tailless and not the two-headed phenotype, which could interfere with the interpretation of the results (Figure R2.1B). To avoid the formation of a posterior head we only performed one round of inhibition; planarians were injected on 3 consecutive days (1200ng/µl). Amputation of planarians was carried out on the fourth day, at postpharyngeal level and amputated pieces were soaked for 3 hours in dsRNA. Performing this protocol, 7 days after amputation, 40% of the animals were obtained with a tailless phenotype, while the rest of the animals appeared normal. However, *wnt1* expression analysis at 12 hR and 3 dR demonstrates a reduction of *wnt1* at both time points in all the animals (Figure R2.1C). Analysis of the posterior marker (*fz4*) (227), also showed a decrease in 85% of planarians at 3 dR (Figure R2.1C). Thus, with the new RNAi protocol, 40% of the animals show a tailless phenotype, but none of them regenerates a functional posterior or anterior organizer.
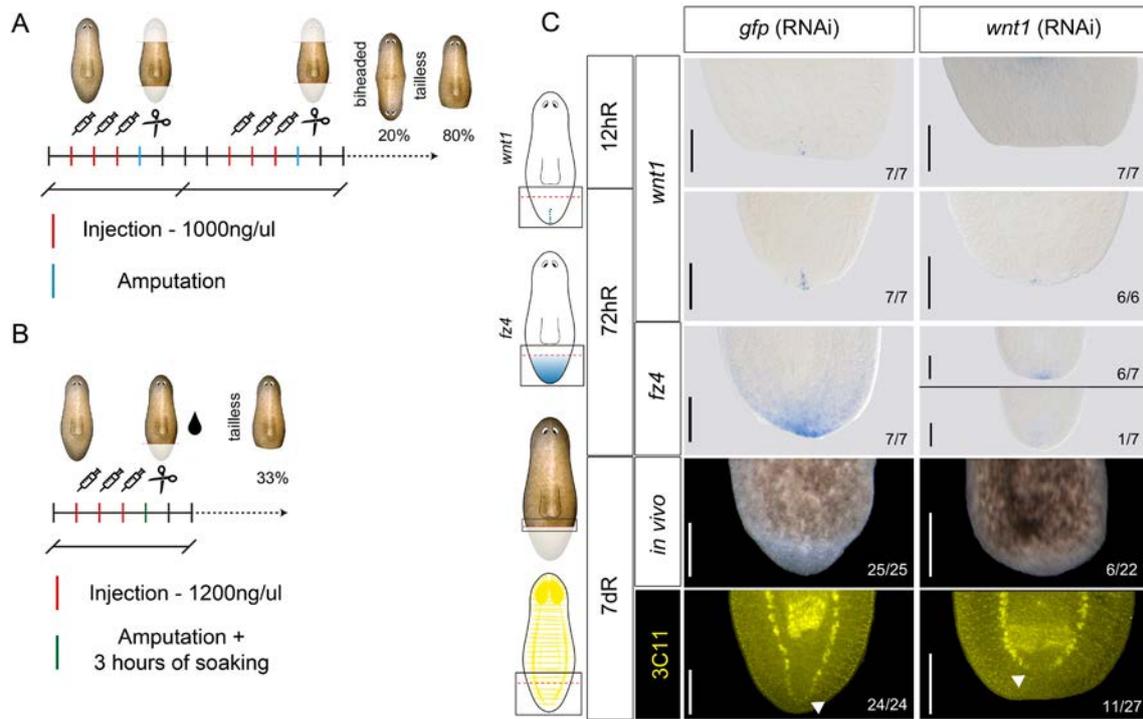
Figure R2.1: RNAi soaking protocol in *wnt1* (RNAi) animals. (A) Cartoon illustrating the injection protocol of *wnt1* (RNAi). One week starved planarians were injected 3 on consecutive days and amputated the following day. The following 3 days, planarians were allowed to regenerate (this is what is called one round of inhibition). Planarians were amputated anterior and posteriorly, and a second round of injections was performed only with trunk fragments. Those trunks were then amputated anterior and posteriorly. A small percentage of inhibited animals presented a biheaded phenotype and the rest of them were tailless. **(B)** Cartoon illustrating the soaking protocol of RNAi. Starved animals were injected on 3 consecutive days and amputated the following day. After that, pieces were soaked for 3 hours in dsRNA diluted in PAM water. As a result, a third of the animals presented a tailless phenotype and no two-headed animal was obtained. **(C)** WISH of *wnt1* at 12 hR and 3 dR indicates indicates the decrease of *wnt1* RNA in *wnt1* (RNAi) animals. The posterior marker *fz4* is also reduced. *in vivo* images of planarians at 7 dR shows a tailless phenotype, which is corroborated by synapsin (3C11) immunohistochemistry , showing the ventral nerve cords fused in the midline in U shape (white arrows). Scale bar are 200 μm in all panel.

Since we were interested in the formation of the organizer and how tissue around it responds, we analyzed blastema and postblastema regions (carefully detailed in Material and Methods). The timepoints studied included early and late *wnt1* expression phases. As explained in the introduction, after amputation, *wnt1* shows two phases: the early (0-36hR) one shows a cell-dispersed pattern and is considered a wound response since its activation is SC independent (Figure R2.2A). It has been suggested that it could trigger the activation of the regeneration itself, as well as that it could be involved in the decision to make a head or a tail. The second *wnt1* expression phase (from 36hR), is SC dependent, and corresponds to the formation of the organizer itself (Figure R2.2B). It has been suggested that this second expression participates in the organization of posterior identity (155). To study both *wnt1* expression waves and the process of how the organizer is formed we analyzed 5 regenerating time points: 0hR, 12hR, 24hR, 36hR, 48hR and 72hR. We assumed that by comparing control with *wnt1* (RNAi) animals, we should be able to determine how each *wnt1* phase controls different sets of genes. This would allow to determine which genes could be putative cWNT targets in planarian, and how transcriptional dynamics are affected when lacking the posterior organizer. The 0hR time point allowed us to have a reference of expression and to compare the relative expression of each gene.
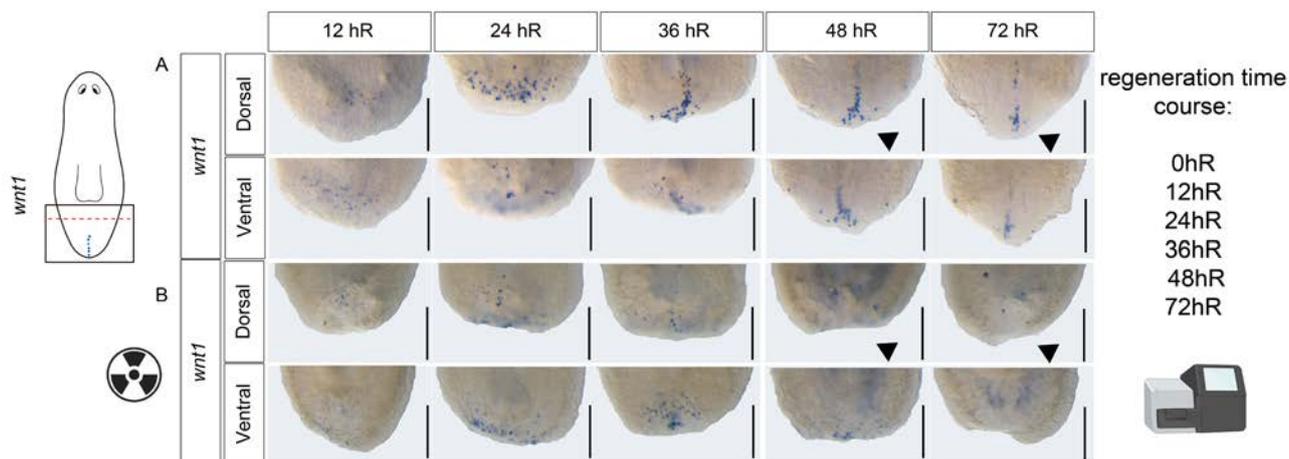
Figure R2.2: *wnt1* **expression during regeneration in wild type and irradiated animals.** The illustration shows the expression of *wnt1*, the amputation plane and the area analyzed (dashed red line and black square, respectively). **(A)** After amputation, *wnt1* is expressed both dorsally and ventrally. After 36 h it is restricted to the dorsal part. **(B)** In irradiated animals, *wnt1* expression is observed during the first hours, both in dorsal and ventral parts, but from 48 hR on it is not present anymore (dark arrows). The timepoints analyzed are the same as used for the RNA-seq experiment. Scale bar are 250 µm in all panel.

Three replicates per time and condition were analyzed, ending up with 36 libraries with more than 45M PE reads. Instead of using another reference transcriptome (90) or ensemble the reads to produce a new transcriptome, we sought to use the last assembly version of the planarian genome (89); since the assembling and annotation of the last published versions were substantially improved (89). Most of the reads mapped in a unique region in the genome, and 60% of the total mapped in a gene.

5.1.2. *wnt1* inhibition produces the same transcriptomic profile that can be seen in just-amputated control animals

To decipher the genetic profiles among samples and between the two experimental conditions, we performed hierarchical clustering of expression profiles using correlation distance of the most expressed genes in each sample. Firstly, we analyzed control libraries to observe the genetic profile compared to other (Figure R2.3A). We could clearly identify two groups: libraries of 0 hR, 12 hR and 24 hR grouped together, and libraries of 36 hR, 48 hR and 72 hR also grouped together (Figure R2.3A). These results agree with the two phases of *wnt1* expression already described.

Secondly, we added RNA-seq libraries to identify whether the lack of organizer differs among samples and compared results to controls. In the second analysis, we could observe that just after the amputation (0 hR), control and RNAi tissues share the same profile since they replicate clusters together. Interestingly, RNAi regenerating samples at 72 hR seem to cluster with the previous ones, suggesting their similarity at the gene expression profile. These results indicate similarity between a just amputated tissue and a completly regenerated *wnt1* (RNAi) planarians. A feature of both scenarios is the lack of organizer (Figure R2.3B). In contrast, during regeneration (from 12 hR to 48 hR) the lack of organizer produces huge differences among samples and replicates, which could be explained by the fact that multiple signalling pathways are involved in the regenerating steps, amongst these the cWNT pathway.

Overall, these results indicate that *wnt1* (RNAi) animals show a transcriptomic profile similar to just-amputated animals. Thus, the lack of organizer leads animals to close the wound but not regenerate any patterned tissue
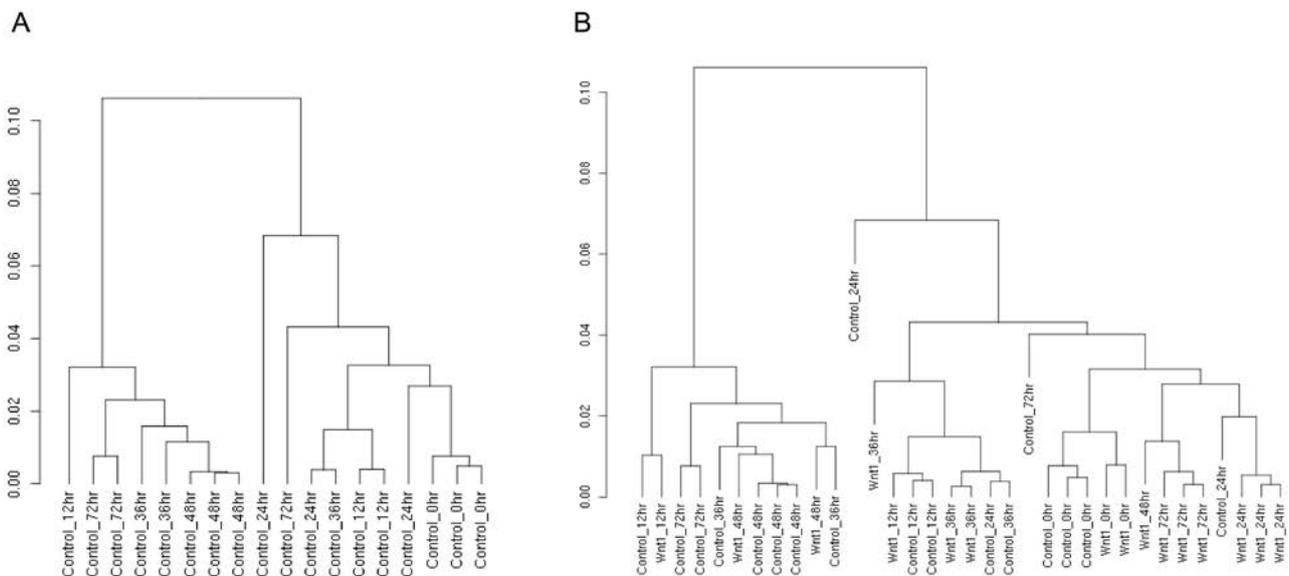
Figure R2.3: Hierarchical clustering of the RNA-seq libraries. (A) Hierarchical clustering of regenerating control libraries using the correlation distance with the most expressed genes. Two main groups appear: 0 hR, 12 hR and 24 hR; and 36 hR, 48 hR and 72 hR. **(B)** Hierarchical clustering using the correlation distance with the most expressed genes of regenerating control and *wnt1* (RNAi) libraries. The result shows clustering of regenerating and RNAi conditions. Yellow line indicates one of those correlations.

5.1.3. RNA-seq of *wnt1* (RNAi) reveals WNT targets in planarian

Plotting *wnt1* expression in all regenerating time points, we could characterize that control organisms show an increment of *wnt1* expression that reaches a peak at 24hR and subsequently decreases stopping at 48hR in a plethora (Figure R2.5A). In *wnt1* (RNAi), the same dynamics appear, however at each time point *wnt1* expression is below the control levels (Figure R2.8A). Although statistical analyses reveal that *wnt1* was only significantly down-regulated at 48hR and 72hR (Figure R2.4).

By comparing the transcriptomes of *wnt1* and control point by point (Ctrl 0hR vs *wnt1* (RNAi) 0hR; Ctrl 12hR vs *wnt1* (RNAi) 12hR; …), we identified a different set of genes that were differentially expressed in *wnt1* (RNAi) animals compared to the control (padj<0.05, fc>0.5). Overall, 2708 genes were down- and 1422 were up-regulated. Among the down-regulated genes, we found some genes previously described as *wnt1* target genes, such as: *wnt11-1*, *wnt11-2* and *fz4* (60), which were down-regulated at 48hR, and in the top 6% of down-regulated genes at 72hR (Figure R2.4). In addition to them, we identified evolutionary conserved genes related with the cWNT signal as: *stat*, *TNF factor*, *fz4*, *sp5*, *axinB*, *Sox*, *Rnf43*, *nkx*, *ets*, *dvl* and *MMPs*. We considered that the new soaking inhibition protocol was a good strategy to inhibit *wnt1*, their targets and investigate putative transcriptomic changes

In order to discriminate whether the genes affected by the inhibition of *wnt1* could be related with the first or the second regeneration wave, we separately analyzed genes down-regulated at 0hR, 12hr and 24hR, and the ones at 36 hR, 48 hR and 72 hR. Even thought that not many genes were down-regulated at first stages, we could identify some interesting ones; such as: zinc finger NFX1 genes (*SMESG000034856.1*, *SMESG000034891.1*, *SMESG000081218.1* and *SMESG000047510.1*) that encoded for proteins related that contain DNA and RNA binding domains; two cWNT target genes were *stat* (228) and (*SMESG000014122.1*), TNF (229). Interestingly, retinoic acid inducible protein (*SMESG000043533.1*) and IGFALS (*SMESG000057458.1*) were also down-regulated (Figure R2.4A).

Some genes related to the second wave were also identified. Besides the described above (*fz4*, *wnt11-1* and *wnt11-2*), we also identified different transcription factors, all described to be expressed in posterior, as Hox genes *lox5a* (149), *post2c* (149), *hox4b* (61), or *axinB* (61), *sp5* (149) and *tsh* (171,172) (Figure R2.4). *tsh* has been described to exert a role regulating posterior identity, being controlled by *βcat1* (175,176). Thus, after *wnt1* inhibition, described posterior genes were down-regulated, confirming that the new protocol appears to efficiently inhibit the WNT pathway. Performing GO term analysis unveils that most down-regulated genes were related to regulation of transcription, signal transduction, immune response, protein ubiquitnation and chromatin remodelling (Figure R2.4B).
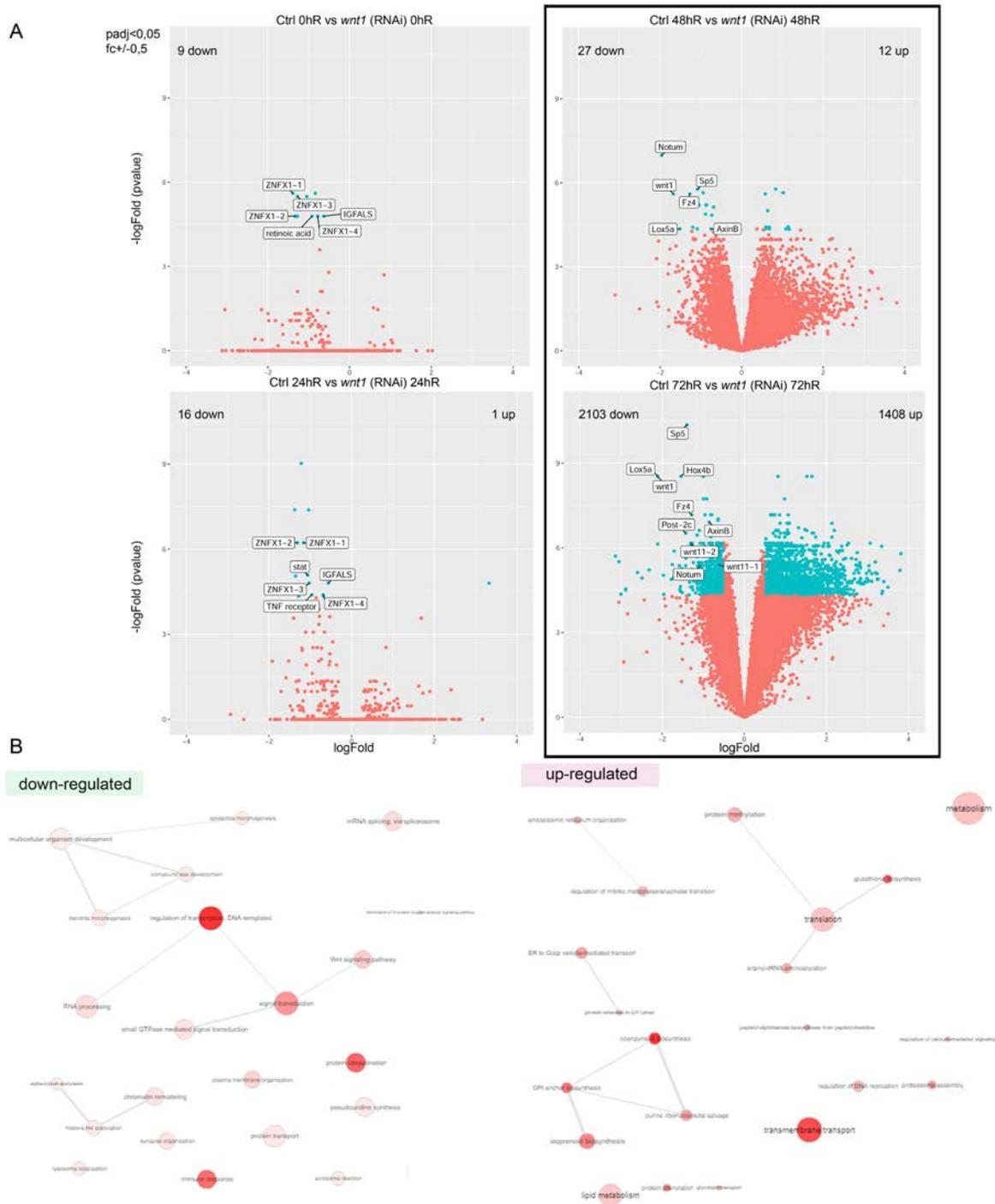
Figure R2.4: RNA-seq of *wnt1* (RNAi) and control animals reveals differentially expressed genes at different regenerating time points. (A) Volcano plots illustrating log2 of fold change expression versus −log10 of adjusted p value (padj). Each volcano plot represents a regenerating time point. Up-regulation (up) after *wnt1* (RNAi) compared to control RNAi were transcripts with adjusted pvalue < 0.05 and log2 fold change > 0.5. Down-regulation (down) after *wnt1* (RNAi) compared to control RNAi were transcripts pvalue < 0.05 and log2 fold change < −0.5. Each dot represents a transcrit, differencially expressed (blue) or not (pink). **(B)** Gene Ontology term enrichment of differentially expressed genes at 48 hR and 72 hR (dark box), visualized by ReviGO. The size of the circles denotes the number of genes; circle colour indicates the p-value of each term. Highly similar GO terms are linked in the graph.

When analyzing up-regulated genes at later stages of regeneration, we did not identify any anterior markers neither brain nor eye markers, confirming that a lack of posterior organizer was not replaced by anterior ones. GO term analysis revealed that up-regulated genes were related to metabolic pathways and transmembrane transport (Figure R2.4B).

Overall, we designed a new inhibition protocol which we used to produce *wnt1* (RNAi) planarians and study the transcriptomic profile of the posterior organizer. The functional study of newly identified candidate targets of the WNT pathway is currently carried out and new interesting insights are expected to be gained. In additon, our transcriptomic results were compared with published transcriptomic data related with the specification of the AP axis, as presented in the next sections.
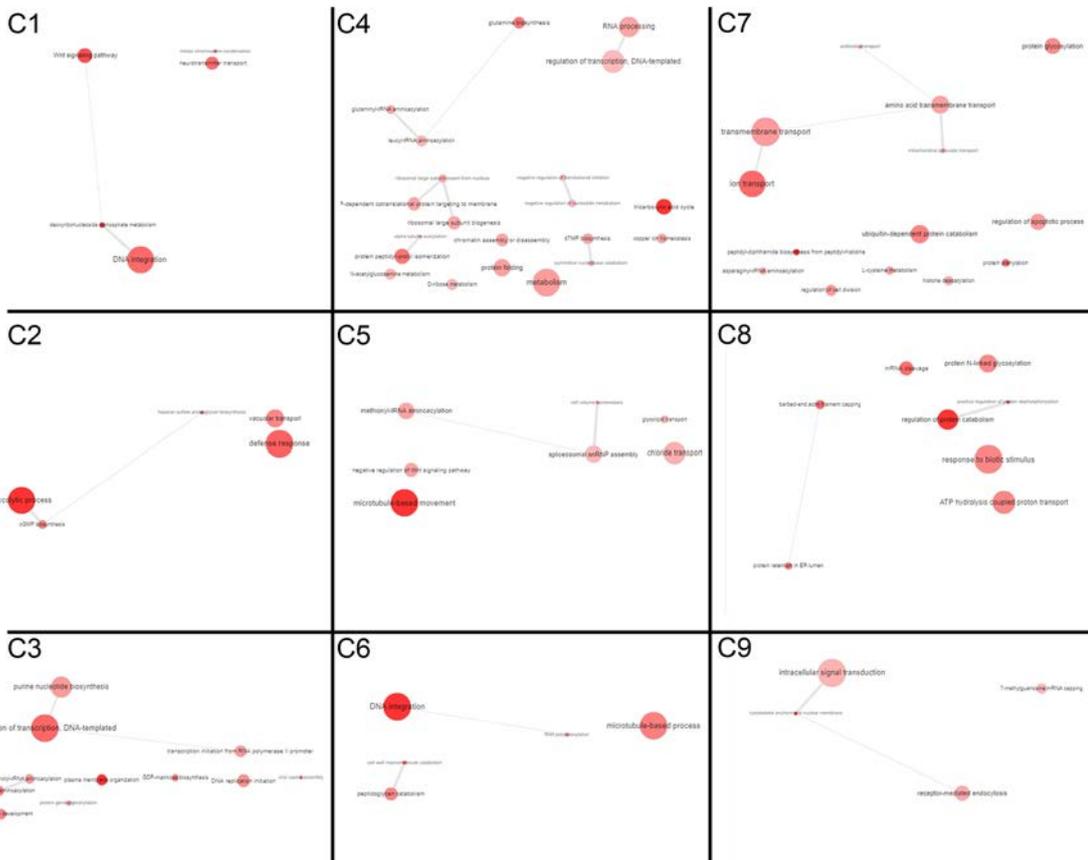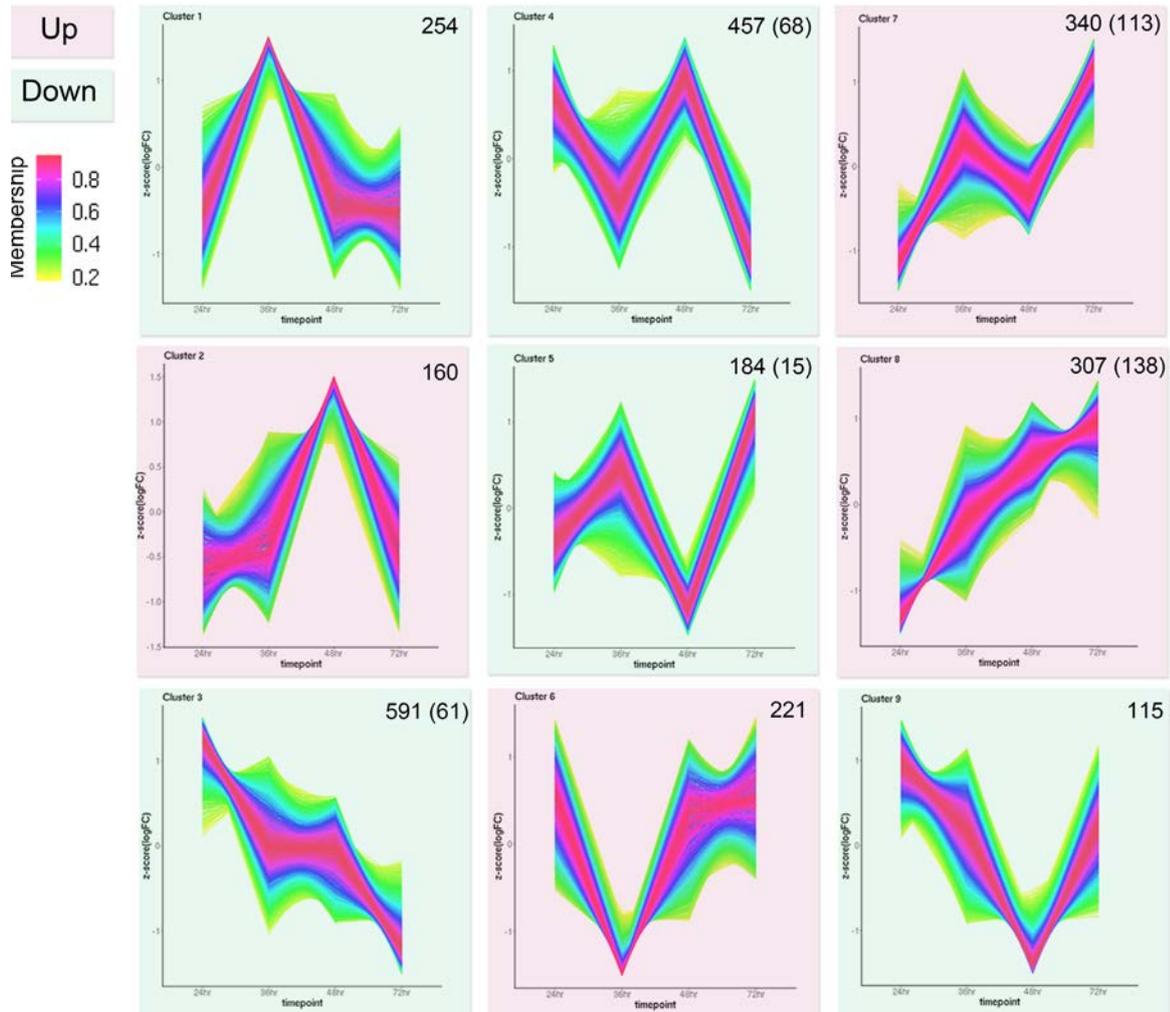
5.1.4. *wnt1* controls gene expression during all regenerative stages

We analyzed the expression of down-regulated genes in control and *wnt1* (RNAi) conditions and during regeneration. To achieve that, we plotted the expression of each gene at each time point. Checking the patterns of specific genes, we found that some of the down-regulated genes such as *axinB*, started to be expressed between 24hR and 36hR in controls, but in knockdown animals the expression was reduced at every single time point (Figure R2.8A). Other genes such as *SMESG34891*, *SMESG81218*, *SMESG47510* or *SMESG34905* started to be down-regulated at first stages of regeneration between 0hR and 24hR and remained so until 48hR. To better understand the gene expression dynamics, we used TCseq to check temporal patterns in the differentially expressed genes. To perform the analysis, we decided to pick the following regeneration time points: 24hR, 36hR, 48hR and 72hR, and omited the first two (0hR and 12hR) due to the huge variance of gene expression in both conditions. By clustering the different dynamics in 9 clusters (named c1 to c9) (Figure R2.5), we were able to identify different groups of genes that were down-regulated in every single point as c3, or being up-regulated as c8. There were others that were down-regulated in two time points as c4 (Figure R2.5). Previously described down-regulated genes were indentified in c3 and c4 since these were the groups with the biggest negative tendency. In order to decipher the most affected putative functions affected, we performed a GO analysis of each cluster. Clusters with up-regulated genes (C2/C6/C7/C8) were linked with different processes of catabolism. Whereas clusters with down-regulated genes (C1/C3/C4/C5/C9) were connected with GO terms such as the Wnt signalling pathway, transcriptional regulation and intracellular signal transduction (Figure R2.6).

Figure R2.5: *wnt1* (RNAi) during posterior regeneration generates different expression pattern profiles. In all the clusters, each line represents a single gene expression; colour intensity indicates membership degree. log-Fold change between samples per time point was plotted in all the clusters. Positive slopes indicate genes increasing expression in knockdown animals and negative slopes indicated genes reducing expression in knockdown animals. The number of genes belonging to each cluster is indicated. When possible the number of genes differentially present in each cluster, it is also indicated between brackets

Figure R2.6: *wnt1* (RNAi) produces genetic changes during regeneration Gene Ontology term enrichment of differentially expressed genes in different clusters visualized by ReviGO. The size of the circles denotes the number of genes; circle colour indicates the p-value of each term. Highly similar GO terms are linked in the graph.

The previously present temporal patterns indicate that *wnt1* RNAi affects genes related to the three phases of the regenerative process: wound response (early), regenerating response (late) and differentiating phase (91) (Figure R2.7). Checking the genes representing each phase (95) we observe the deregulation of genes corresponding to the three phases (Figure R2.8).

Altogether, these data suggest that *wnt1* is controlling gene expression during all to the regenerative process.

Genome Id	Transcriptome Id	Gene Name	Class	logFC
SMESG000005130.1	dd_Smed_v6_3257_0_1		Generic	-1.37909510982015
SMESG000023227.1	dd_Smed_v6_7824_0_1	sp5	Generic	-1.04902654795941
SMESG000017137.1	dd_Smed_v6_4012_0_1	pim-1	Generic	-0.686830982639051
SMESG000045089.1	dd_Smed_v6_14289_0_1		Regeneration	-1.28129505194586
SMESG000015861.1	dd_Smed_v6_10783_0_1		Regeneration	-0.96495048871418
SMESG000004222.1	dd_Smed_v6_24180_0_1	notum	Regeneration	-0.919474403727862
SMESG000065993.1	dd_Smed_v6_6259_0_1		Regeneration	-0.813850913618512
SMESG000031290.1	dd_Smed_v6_4279_0_1	X1.A2.3	Regeneration	-0.749809767554055
SMESG000004519.1	dd_Smed_v6_11650_0_1	fz4	Regeneration	-0.633456051584457
SMESG000002557.1	dd_Smed_v6_7811_0_1	col4-5	Differentation	-0.838309593250152
SMESG000056472.1	dd_Smed_v6_5999_0_1		Differentation	-0.78552075869285
SMESG000071674.1	dd_Smed_v6_6053_0_1		Generic	1.08676605814864
SMESG000071674.1	dd_Smed_v6_9642_0_1		Generic	1.08676605814864
SMESG000038731.1	dd_Smed_v6_14219_0_1		Generic	1.64678393667912
SMESG000027529.1	dd_Smed_v6_9204_0_1		Regeneration	0.595050697770548
SMESG000054253.1	dd_Smed_v6_2324_0_1		Regeneration	0.816757778646011
SMESG000021620.1	dd_Smed_v6_5781_0_1	5F04 PDZ ring finger protein 4	Regeneration	1.02236327979854
SMESG000027528.1	dd_Smed_v6_36283_0_1		Differentation	0.678692885441719
SMESG000065075.1	dd_Smed_v6_102_0_1		Differentation	2.1346191796702

Figure R2.7: *wnt1* (RNAi) affects genes involved in different stages of the regenerative response. (A) Schematic representation of three genetic responses that occurs after any planarian amputation. Adapted from (91). In the upper part of the graph, the percentage of genes affected was included. An example of each condition was plotted, indicating whether the gene was up or down-regulated. Plots presented in the panel represent the expression of a gene in control conditions and *wnt1* (RNAi). Mean expression per each regenerating time point is represented by a dot. Log-fold change was represented on the X axis. **(B)** Table was added with all the genes affected indicating their genome and transcriptome reference number, planarian gene name, time point where they were affected and the log fold change. Orange rows show genes previously described in planarian

5.1.5. *wnt1* regulates the expression of posterior expressed genes

In order to check weather after *wnt1* inhibition, not just the main posterior markers were reduced, but other posterior genes were affected, we compared our transcriptomic data with the one generated by Stückemann and collaborators, which performed a regional RNA-seq in planarians. In his study 14 sets of genes were differentially expressed along the AP axis of planarians (149). Among these, one group of genes (12) was differently in the tail, and other one (14) was was expressed at body edges (Figure R2.8A). Analyzing the behaviour of those gene sets in *wnt1* (RNAi) animals, we observed that 13.9% of genes from group 12 and 13% from group 14 were affected (Figure R2.8A). This data confirmed that not just the previously known posterior markers changed after *wnt1* inhibition, but also the whole posterior specific landscape.

Li et al. studied posterior pole enriched genes that were expressed in the posterior organizer (*wnt1+* and *collagen+* cells) (230) (Figure R2.8B). When checking those genes in our data we found 15 genes (7.6%) affected in knockdown animals, 6 down- and 9up-regulated (Figure R2.8B). Among those genes, we detected: *Smed-myoD* (87,231) and *Smed-NetR* (232), which were down-regulated at 72 hR; and *Smed-slc46a-a* (120,121), which was up-regulated at 72 hR. Overall, this data confirms that in our *wnt1* (RNAi) animals, posterior identity is suppressed.
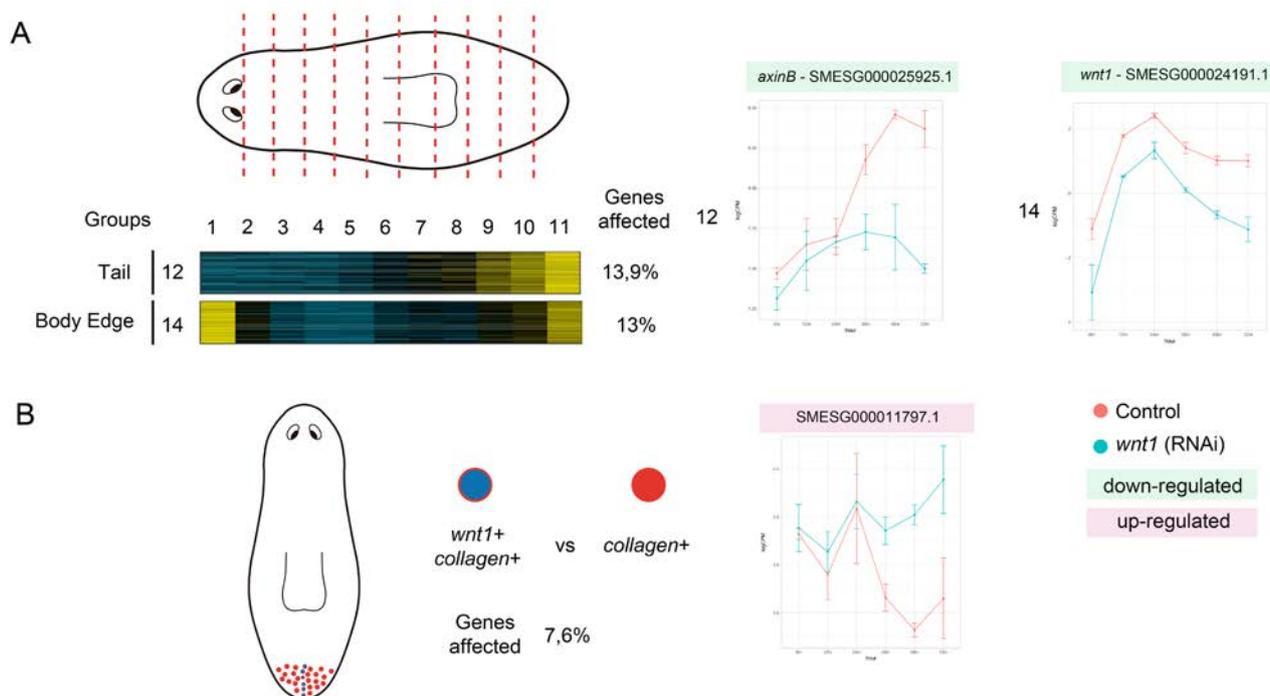
Figure R2.8: RNAi of *wnt1* affects posterior gene expression patterns. (A) Regionalized RNA-seq was performed according to the illustration (149). Genes were grouped in different clusters from head to tail: 12 group genes were expressed exclusively in the tail; and 14 group genes in both tips. From each cluster, the percentage of genes affected is indicated next to each group row. One example per group was plotted **(B)** Schematic experimental design from RNA-seq comparing *wnt1+/collagen+* cells versus *collagen+* cells (230). Percentage of genes affected was indicated and one example plot was added. Plots present in the panel represent the expression of a gene in control conditions and *wnt1* (RNAi). Mean expression per each regenerating time point is represented by a dot. Log-fold change was represented on the X axis. The name of *Smed* was added when it exists.

5.1.6. *wnt1* and *βcat1* (RNA-seq) comparison reveals putative canonical WNT targets

In 2014, Reuter et al. published a RNA-seq of *βcat1* (RNAi) regenerating animals (171), and they were able to describe down- (440 genes) and up- (438 genes) regulated targets of *βcat1* (Figure R2.9A, 9B). We sought to combine this data with ours in order identify putative targets of the cWnt pathway in planarians. With this strategy, we identified 13 down- and 8 up-regulated genes in both datasets. Interestingly, seven out of 13 down-regulated genes were genes previously described to be related with posterior identity in planarians or in other species: *lox5a*, *post2c*, *hox4b*, *axinB*, *tsh*, *sp5* and *fz4* (Figure R2.9A). Among the others (8/13), just one gene encoded for a known protein (*Troponin I - SMESG000002071.1*), and the rest encoded for genes showing no homolog known gene. Thus, *Troponin I* would be a good candidate to be functionally studied.

Among the eight common up-regulated genes, *scl22a-5* was previously described in planarians and encodes a protein located at the proximal tube of the protonephridia system (121). The rest of the genes showed no homology with other known genes (Figure R2.9B).

Interestingly, when we plotted down-regulated genes from *βcat1* (RNAi) and *wnt1* (RNAi) data, it came to our attention that genes were affected at late stages of regeneration. Down-regulated genes started to be affected between 24 hR and 36 hR, as in the case of *sp5*, *tsh*, *lox5a*, *axinB* and *SMESG000015861.1* (Figure R2.10). This suggests that the cWNT pathway could exert a role during the second WNT phase; and could first act through other TF or with the use of other cofactors. With this strategy, we were able to propose a list (Figure R2.9A, 9B) of putative target genes of the cWNT pathway in planarians.

Overall, we used a novel strategy that allowed the identification of cWNT related genes related with the formation and function of the posterior organizer.

A

440 down-regulated genes in *βcatenin 1* (RNAi)

Down

βcat1 (RNAi)   wnt1 (RNAi)

221 genes in SMEDG

13 genes down-regulated in *wnt1* (RNAi)

Genome Id	Transcriptome Id	Gene Name	Time Point	logFC	
SMESG000022667.1		lox5a	72	-2,10875642892328	●
SMESG000015861.1	dd_Smed_v6_10783_0_1		72	-2,1060674119448	
SMESG000057720.1	dd_Smed_v6_34946_0_1		72	-1,75208665645272	
SMESG000044469.1		hox4b	72	-1,54041306152128	
SMESG000022667.1		lox5a	48	-1,5138243541489	●
SMESG000022668.1		post2c	72	-1,42607287807026	
SMESG000023227.1		sp5	72	-1,39668037453989	●
SMESG000065348.1	dd_Smed_v6_1320_0_1		72	-1,34139367510459	
SMESG000037574.1	dd_Smed_v6_288_0_1		72	-1,30504361976233	
SMESG000004519.1		fz4	72	-1,28129505194586	●
SMESG000004519.1		fz4	48	-1,27129922350708	●
SMESG000023227.1		pp5	48	-1,08055028522091	●
SMESG000002071.1	dd_Smed_v6_323_0_1	troponin I	72	-1,05438025033858	
SMESG000008973.1		tsh	48	-0,885490093238136	
SMESG000024991.1	dd_Smed_v6_69_0_2		72	-0,865763125542772	
SMESG000025925.1		axinB	72	-0,852526765254967	●
SMESG000025925.1		axinB	48	-0,749809767554055	●

B

348 up-regulated genes in *βcatenin 1* (RNAi)

Up

βcat1 (RNAi)   wnt1 (RNAi)

216 genes in SMEDG

8 genes down-regulated in *wnt1* (RNAi)

Genome Id	Transcriptome Id	Gene Name	Time Point	logFC
SMESG000075450.1	dd_Smed_v6_1379_0_1		72	2,32657161252991
SMESG000035252.1	dd_Smed_v6_4468_0_1		72	2,08182881300272
SMESG000041818.1	dd_Smed_v6_50434_0_1		72	1,36483427555031
SMESG000045320.1	dd_Smed_v6_15338_0_1		72	0,956110022176061
SMESG000072594.1	dd_Smed_v6_10682_0_1		72	0,865241156440074
SMESG000017151.1	dd_Smed_v6_16489_0_1		72	0,83989965371955
SMESG000077674.1	dd_Smed_v6_25487_0_1		72	0,688828642579803
SMESG000064528.1	dd_Smed_v6_10903_0_1	scl22a-5	72	0,572415827532933

Figure R2.9: *wnt1* (RNAi) and *βcat1* (RNAi) RNA-seq present similar up and down-regulated genes. (A) Screening procedure to identify down-regulated canonical Wnt target genes: an RNA-seq dataset of 440 down-regulated genes in *β-catenin* (RNAi) planarians (171) was mapped in the planarian genome obtaining 221 genes. These candidates were compared with the *wnt1* (RNAi) RNA-seq data obtained under regenerating conditions. **(B)** Screening procedure to identify up-regulated canonical Wnt target genes: an RNA-seq dataset of 348 up-regulated genes in *β-catenin* (RNAi) planarians (171) was mapped in the planarian genome obtaining 216 genes. These candidates were compared with *wnt1* (RNAi) RNA-seq data obtained in regenerating conditions. Tables were added showing up and down-regulated genes in *wnt1* and *βcat1* (RNAi) animals indicating their genome and transcriptome reference number, planarian name, time point at which they were affected and the log fold change. Colour dots indicate shared genes between regenerating time points.

Figure R2.10: *Putative cWNT target genes display different regenerative expression patterns.* Plots represent the expression of a gene in control and *wnt1* (RNAi) conditions. Mean expression per regenerating time point is represented in a dot. log-fold change was represented on the X axis. Name of *Smed* was mentioned if it appeared in the genome. Colored dots follow the same as the previous figure.

5.2. ATAC-seq analysis reveals Fox Family as a key elements for posterior organizer function in Planarians

When using the RNA-seq approach we saw that the establishment of the P organizer requires a specific transcriptional activation. Since transcriptional activity depends on the regulation of the chromatin state, or the epigenome, we next studied the chromatin changes that occur during the establishment of the posterior organizer in regenerating planarians. To that purpose we performed ATAC-seq (transposase-accessible chromatin using sequencing), of regenerating planarians with the aim to compare the chromatin state of anterior versus posterior blastemas (Figure R3.1A). We performed ATAC-seq analysis. ATAC-seq is based on the Tn5 transposase enzyme, which fragments accessible chromatin regions and simultaneously tags the fragments to be sequenced (Figure R3.1B) (89) (descried in Material and Methods). The final outputs of the technique are peaks distributed along the genome. Each peak represents a nucleosome free region, where the chromatin is accessible to transcriptional complexes, that is: promoters and enhancers. Each peak has a particular height and width, and using different statistical methods we could compare and classify them as accessible chromatin regions (open) or non-accessible (closed). ATAC-seq has been used in many organisms to describe developmental processes (223) and tissue differentiation (234), or in humans to characterize diseases (235,236). However, there is no published work of ATAC-seq analysis in regenerating planarians.

5.2.1. Specific chromatin changes occurs at anterior and posterior wounds

We performed ATAC-seq analysis of posterior and anterior wild type wounds at 12hR after post pharyngeal amputation (Figure R3.1B). Two replicates per point were included, to allow statistical analysis. This regenerating point is crucial to determine anterior or posterior identity in *Schmidtea mediterranea* (237). In order to identify chromatin regions specifically open during anterior or posterior regeneration, we compared accessible chromatin regions (ACR) from anterior and posterior blastemas. Sample analysis by PCA reveals that sample replicates cluster together and are differentially distributed, indicating differences among them (Figure R3.1C). We selected differentially opened peaks (FDR<0.01, fc > 1.5) in one of each of the tissues. With this strategy, we identified 2484 specific posterior ACR and 611 specific anterior ACR (Figure R3.1D). Using MACS2, we were able to discriminate peaks that were emergent or increased in each specific population of ACR. Emerging peaks were the ones just present in one of the two tissues; increasing peaks were the ones present in both tissues, but showing a significant increment in one with respect to the other (Figure R3.1E). 91.6% of the anterior peaks were emerging (eAnt), and the rest were increasing (iAnt). During posterior regeneration, 60% of the peaks were emerging (ePost) and the remainder ones were increased (iPost) (Figure R3.1E). These results suggest that some accessible chromatins regions are shared during A and P regeneration (Figure R3.1F), but most of them are specifically open in one scenario.
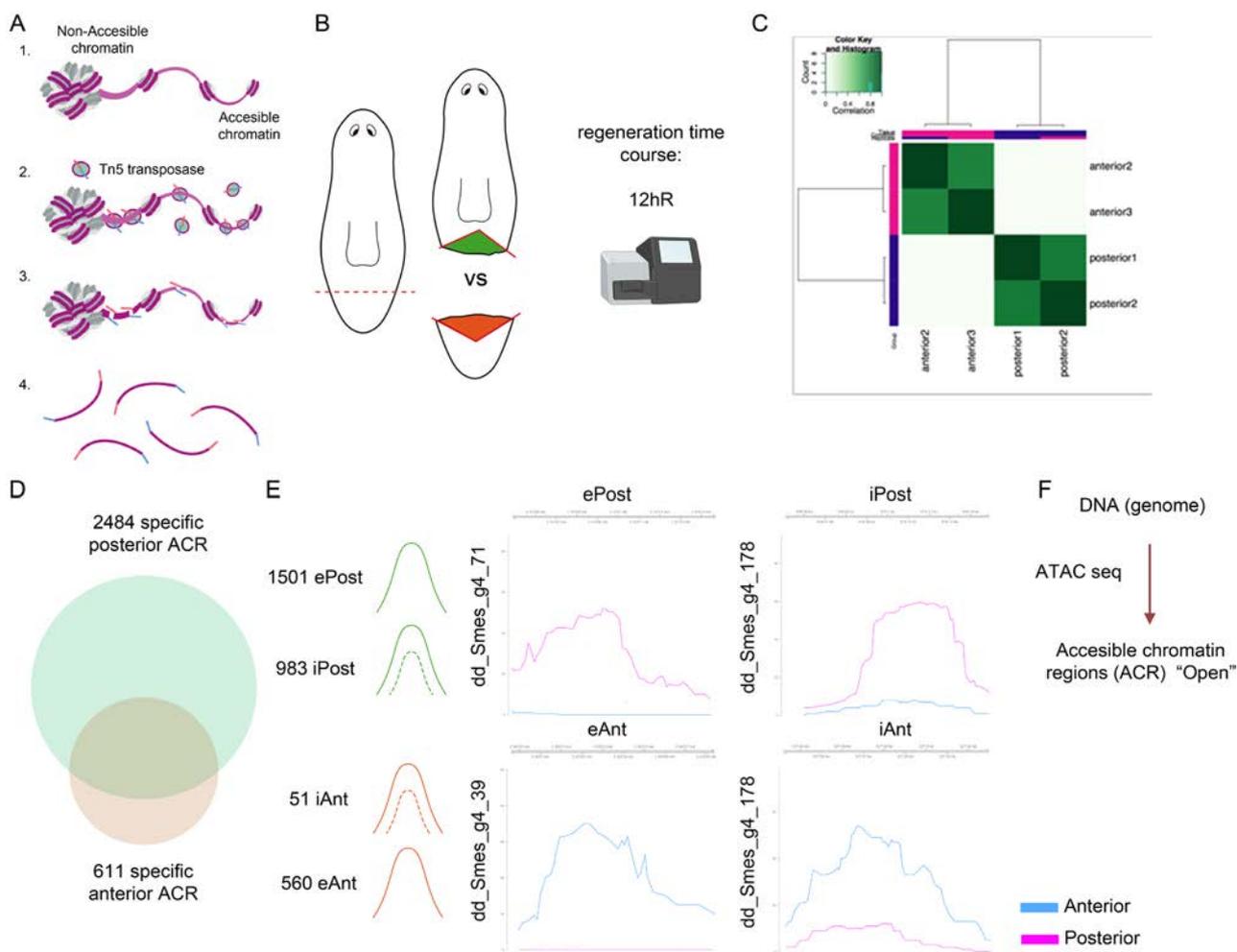
Figure R3.1: ATAC-seq of anterior and posterior blastemas, show specific accessible chromatin regions. (A) Schematic cartoon of ATAC-seq procedure: 1) chromatin could be non-accessible or accessible. 2) Tn5 is able to bind the accessible chromatin regions (ACR). 3) Tn5 tagmentation fragments the genome that are accessible and tags the resulting DNA fragments with sequencing adapters. 4) Fragments and tagged DNA are purified in order to prepare the libraries. Adapted from (432). **(B)** Planarian cartoon indicating were amputation was performed (red dashed line). In the regenerating planarians the area taken for the analysis is indicated (red lines). **(C)** Hierarchical clustering using correlation distance of 12 hR libraries. Anterior replicates cluster together and posterior replicates cluster together, demonstrating their differences between anterior and posterior but similarity between replicates. **(D)** Venn diagram showing anterior and posterior ACR. Regions not present in the intersection were considered specific per each region. Genome distribution of an example of each specific region was added. **(E)** Representation of emerging peaks (e) and increasing (i) peaks for each conditions: posterior (Post) and anterior (Ant). ATAC-seq profiles of each example was included. Genome Browser screenshot showing ATAC-seq profiles of anterior (blue) and posterior (pink) regeneration. **(F)** Summary of procedure to identify accessible chromatin regions specific for each regenerating pole.
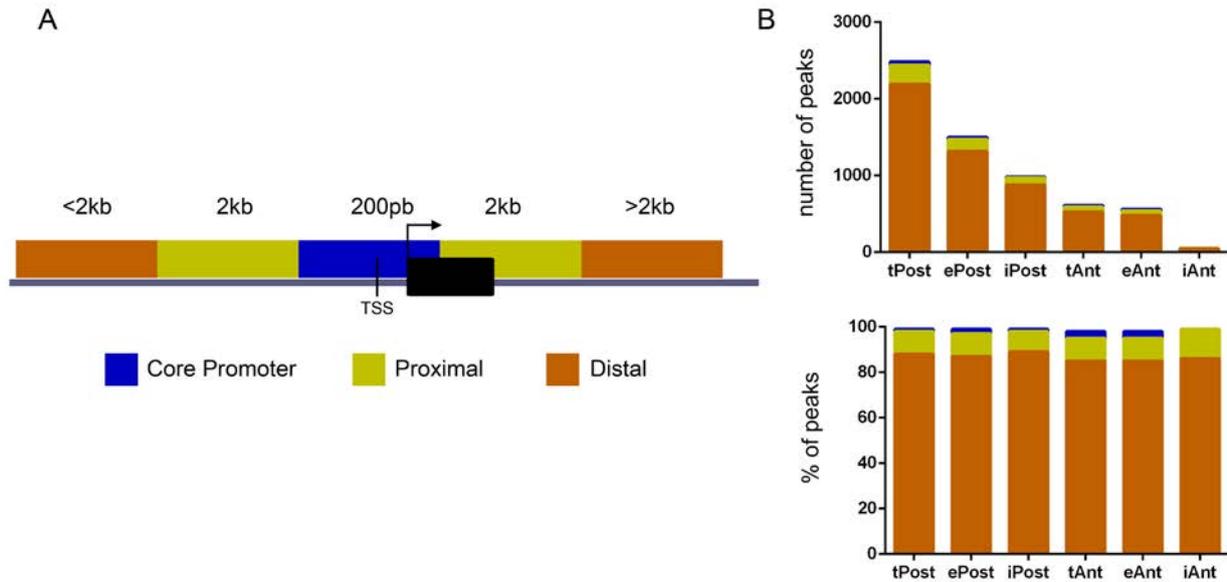
Figure R3.2: Accessible chromatin landscape after amputation at 12 hR. (A) Schematic representation of peak distribution: core promoter, proximal or distal. **(B)** Bar plots representing the total peak distribution in each experimental condition. Down, same values are plotted but using percentages.

Specific ACR were not equally distributed across genome, and we classified them depending on their relative position from the transcription starting site (TSS). Thus, peaks located ±200pb from de TSS were considered core promoter (CP); the ones presented ±2kb from the TSS were considered proximal (Pro) and the ones further ±2kb from the TSS were classified as distal (Dis) (Figure R3.2A). Posterior ACR were enriched in Dis promoters with respect to the TSS (88.16%), 10.06% were located in proximal regions, and the rest of promoters were considered proximal (1.77%) (Figure R3.2B). In the case of the anterior ACR, we found similar percentages, finding that 85.59% of promoters were located at distal parts, and 10,53% were distributed in proximal regions. A small percentage (3.6%) were found at the core promoter. Those results indicate that during regeneration most posterior and anterior ACR are located at distal positions.

5.2.2. Identification of enhancers specifically involved in anterior or posterior planarian regeneration

Chromatin and nucleosomes can be modified allowing transcription or repression of different genes. This ability realize in the capacity to be less or more accessible. Histone residues modifications from nucleosomes are one of the key aspects for chromatin regulation. Some of those modifications are related with active enhancers, such as H3K27ac or H3K27me3. ChIP-seq is a technique that combines chromatin immunoprecipitation and massive sequencing that allow us to detect chromatin modifications and transcription binding sites (Figure R3.3A) (carefully described in Material and Methods). After sequencing, we mapped the reads in the genome obtaining peaks, being each peak related with protein occupancy.
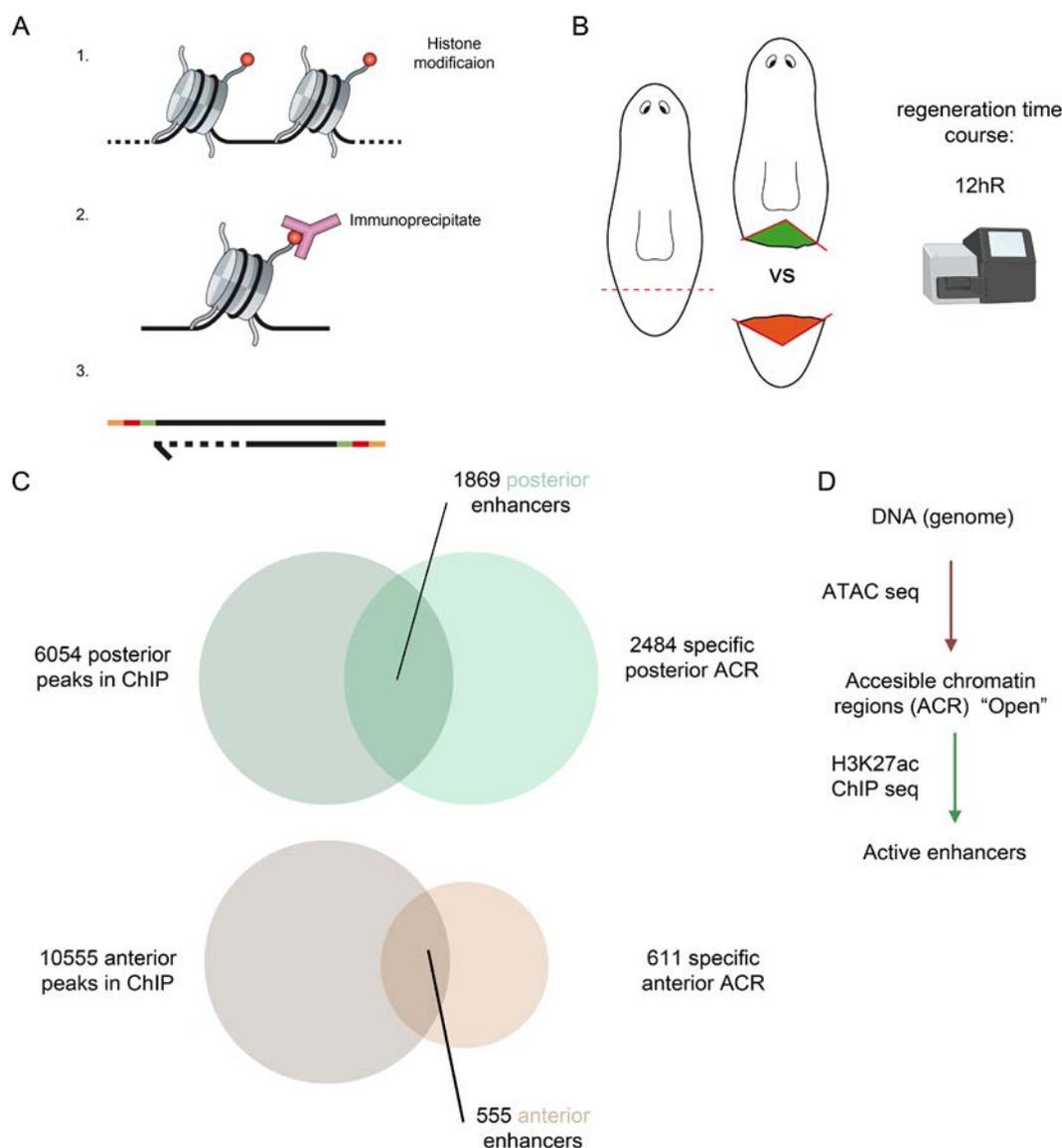
Figure R3.3: ChIP-seq of anterior and posterior blastemas reveal specific enhancers. (A) Schematic cartoon of ChIP-seq procedure: 1) chromatin folds around nucleosomes in which histones could be modified. 2) After nucleosome fragmentation, a specific antibody is used to immunoprecipitate the fragments. 3) DNA is purified and amplified to prepare the libraries. Adapted from (443). **(B)** In an intact planarian cartoon is indicated where amputated was done (red dashed line). In the regenerating planarians is indicated were amputations were done to study blastemas (red lines). **(C)** Venn diagram comparing peaks between ChIP-seq and ATAC-seq. Intersection peaks were considered active enhancers. **(D)** Summary of the procedure to identify active enhancers specifics for each regenerating pole.

To better understand whether ACR found in ATAC-seq analysis could behave as active enhancers, we performed ChIP-seq using H3K27ac antibody at 12hR of posterior and anterior wounds, to characterize putative active enhancers (Figure R3.3B). Samples were collected at the same postpharyngeal level performed in the ATAC-seq analysis. After sequencing and mapping the reads in the genome, we identified 6054 peaks specific of the posterior wound and 10555 peaks of the anterior (Figure R3.3C). Then, we intersected the peaks found with ChIP-seq with the ones found in ATAC-seq analysis, and selected those that were present in both datasets. These should be posterior or anterior ACR that function as active enhancers at 12hR. With this strategy, we identified 1869 posterior putative active enhancers and 555 anterior putative enhancers (Figure R3.3C, 3D).

To determine if the putative enhancers that we detected with the presented strategy were accessible only during early regeneration, at 12hR, or were also open at earlier or later stages, we performed ATAC-seq at 0hR and 48hR (Figure R3.4A). The first time point (0 hR) was selected to analyze a control situation, to see if the enhancers are constitutively open; we collected the samples just after amputation at the same postpharyngeal zone as collected previously. The second time point (48 hR) was selected because at this point planarian anterior or posterior identity has already been determined. The data shows that most of the putative active enhancers were constitutively open during anterior (75.86%) and posterior (52.97%) regeneration, since they were already accessible at 0hR. However, some enhancers appeared regeneration-specific, since they were closed at 0 hR and open at 12 hR (Figure R3.4B). During posterior regeneration, 451 (24.14%) enhancers were required just during regeneration (rPost), and most of them were specifically needed at 12 hR, suggesting that gene regulation at this timepoint is crucial to allow posterior regeneration. In the anterior case, we identified 261 (47.03%) enhancers accessible during regeneration (rAnt), half of them were required at 12 hR and the rest were accessible during all the anterior regenerating process (Figure R3.4B).

With this analysis we have been able to identify enhancers that are specifically active in anterior or in posterior blastemas. Furthermore, some enhancers were regeneration-specific, since they were found open after amputation only (Figure R3.4C).
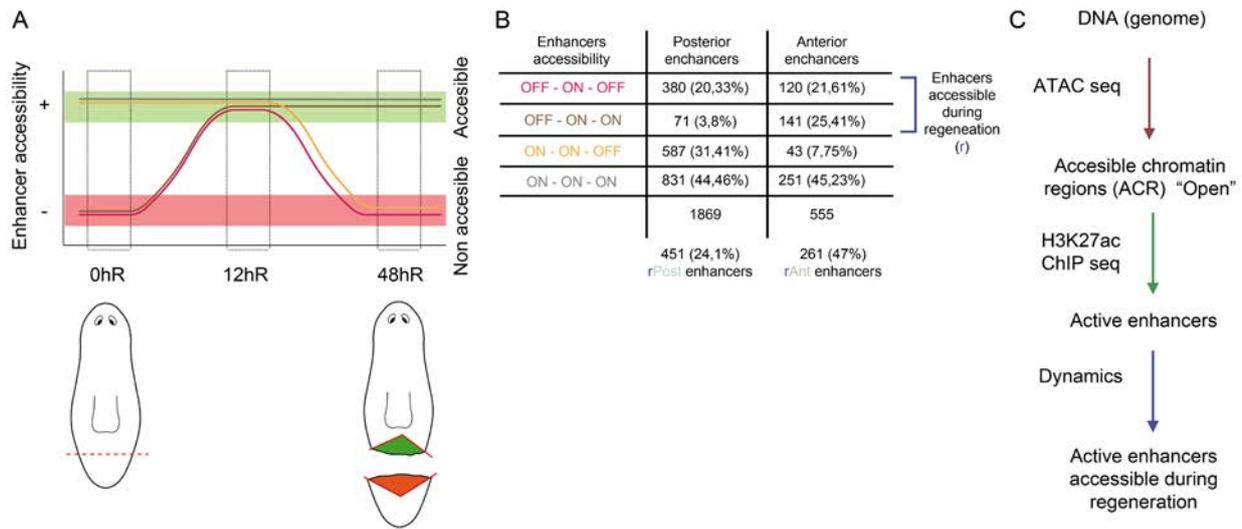
Figure R3.4: Specific chromatin regions regions are open during anterior and posterior regeneration. (A) Schematic fluxplot showing how 12 hR active enhancers change their accessibility at 0 hR and 48 hR. Schematics cartoon visualize where amputations were performed and samples studied. **(B)** Genome browser screenshot showing ATAC-seq of the three regenerating time points. Summary table indicating total (and percentage) peaks of each condition. Peaks, just accessible during regeneration, are indicated in blue (r). **(C)** Summary of procedure to identify active enhancers during regeneration specific for each pole**.**

5.2.3. After *notum* and *wnt1* inhibition chromatin dynamics change

The Wnt pathway is crucial to determine posterior and anterior identity in planarian. *notum* exerts a crucial role to determine anterior identity, and *wnt1* plays a determining role giving posterior identity (Figure R3.5A). To control cell or tissue identity, the chromatin should change its accessible state. To study the chromatin changes triggered by *notum* and *wnt1*, we performed ATAC-seq in anterior *notum* (RNAi) blastemas at 12hR, in planarian which have lost anterior identity; and in posterior *wnt1* (RNAi) blastemas at 12hR, which have lost posterior identity (Figure R3.5A). We analyzed the state of the enhancers found to be open in posterior or anterior regeneration in this loss of function context.

As expected, at 12hR, during anterior regeneration, anterior enhancers appeared open and the posterior ones were closed. However, analyzing the same time point after *notum* inhibition, we observed that only 12.2 % of anterior enhancers were open, and the rest of them were closed or had reduced their accessibility. Moreover, 87.5 % of posterior enhancers, which were not open in a WT situation, were now accessible (Figure R3.5B). These results indicate that in *notum* (RNAi) planarian chromatin state changes as soon as 12hR, being more accessible for specific posterior enhancers, and less accessible for the anterior ones.
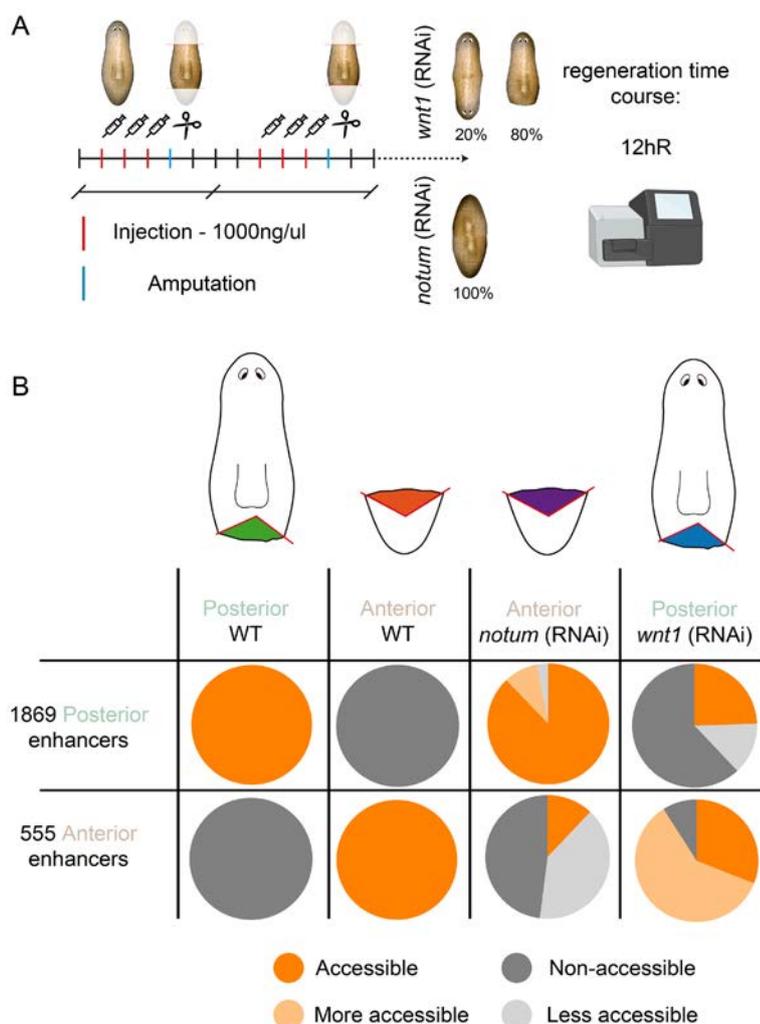
Figure R3.5: ***wnt1*** **and** ***notum*** **(RNAi) change their genomic landscape. (A)** Schematic representation of the experimental design. Two rounds of injection and amputation has performed. After the second amputation, the wound region was isolated and studied at 12 hR. **(B)** Classification of posterior and anterior enhancers in each experimental condition: accessible, more accessible, less accessible or non-accessible.

In the same way, during posterior regeneration, specific posterior enhancers were accessible and anterior ones were closed. Nevertheless, after *wnt1* inhibition only 25.5 % of the specific posterior enhancers were open and the rest were closed or had decreased its accessibility. Specific anterior enhancers, which were closed during posterior regeneration in control animals, now appeared more accessible: 30.8 % of them changed from closed to accessible, and 59.6 % became more accessible, (Figure R3.5B). With these findings, we can suggest that after *wnt1* inhibition, specific posterior enhancers are less accessible, and anterior ones are more accessible.

Overall, we used the previously identified specific CREs to determine chromatin dynamics after axis identity shifting. At 12 hR, in *notum* and *wnt1* (RNAi), chromatin changes had already occurred, overall it is in the *notum* gene expression that shows (RNAi) huge differences between times.

5.2.4. cWNT pathway specifically regualtes posterior CREs

We described that 12hR after *wnt1* inhibition 62% of posterior enhancers were closed (Figure R3.5B), suggesting that they could be related with the WNT pathway. We sought to identify how many of the 451 posterior enhancers that were open during regeneration, were affected by *wnt1* inhibition (Figure R3.6A). We identified 335 (74%) that were closed or less accessible in *wnt1* (RNAi) animals, meaning that in wild type conditions, when the WNT the pathway is not disturbed, those enhancers are open and have a function in posterior identity specification.
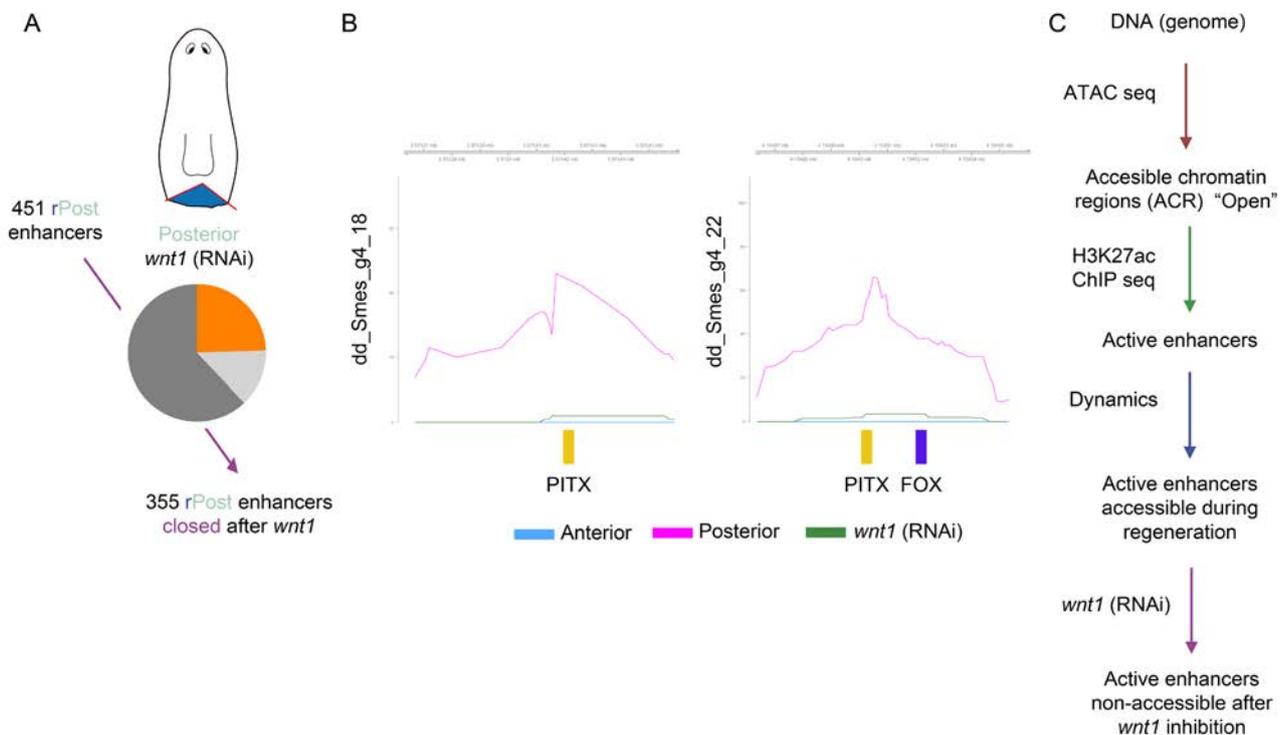
Figure R3.6: *wnt1* inhibition change transcription factor motif accessibility (A) Screening procedure to identify regenerating posterior (rPost) enhancers closed after *wnt1* inhibition. **(B)** Genome screenshot showing closed ATAC-seq peaks in *wnt1* (RNAi) (green) and anterior (blue), but open in posterior (pink). DNA binding sites of PITX and FOX are indicated. **(C)** Summary of the procedure to identify active enhancers during regeneration regulated by *wnt1*.

To further understand the function of those enhancers and their functional relationship with the WNT pathway, we analyzed the presence of specific transcription factor (TF) motifs binding sites using Homer. The most represented motifs were the ones recognized by the HMG domain, associated with TCF, SOX and LEF TFs; Otx2 and pitx domains (bicoid homeobox class) where OTX and PTX can bind; and the forkhead domain, associated with the family of TFs Fox.

TCF has a well described role as a βCAT1 nuclear co-factor, and its motif is normally located at the enhancers of cWNT target genes (238). In planarians, 5 TCFs have been described (239,240). Particularly, TCF-2 participates in planarian eye modulation, trough βCAT1 (239); and TCF-3 is crucial for specifies GABA neural cells (240). Thus, planaria TCF seem to not participate in posterior identity. *pitx* has already described regulating posterior identity, but nothing is known about Fox TF and posterior identity. In the next chapters, their function will be discussed.

Combining ATAC-seq with the loss-of-function of *wnt1* (RNAi), we were able to identify putative TFs related with posterior regeneration, posterior identity specification and the WNT pathway (Figure R3.6C).

5.2.5. *pitx* is required for *wnt1* expression and posterior identity specification

From the ATACseq analysis, binding sites for OTX-PITX TFs appeared to be highly represented in posterior regeneration. This result suggests that through the ATAC-seq analysis we were able to identify a previously reported *pitx* gene in planarians, which is required for serotonergic neuron differentiation, and also for expression of *wnt1* during posterior regeneration. *pitx* inhibition impedes *wnt1* expression and a consequence produces a tailless phenotype (105,106). Moreover, it has been also shown that *pitx* and *wnt1* coexpress in the same cells at 2 and 3 dR, suggesting that *pitx* plays a role regulating the second *wnt1* expression stage. Since in previous studies, it could not be identified whether posterior identity was affected in *pitx* (RNAi) animals, we studied this further. Our results show that after one round of inhibition (Figure R3.7A), at 7 dR, 33.3 %of the animals presented an *in vivo* tailless phenotype as it was previously described (Figure R3.7B). Additionally, in head fragments the percentage increased to 75% (data not shown). We also checked that *wnt1* expression was not present at 3dR; and we could demonstrate that the first *wnt1* wave, at 12 hR, was not affected in posterior nor anterior (Figure R3.7C). We also show that the tailless phenotype previously reported *in vivo*, is supported by synapsin immunostaining, since 80% of the animals showed U shaped nerve cords (Figure R3.7B), and by the disappearance or down-regulation of posterior markers such as *fz4*, *hox4b* and *post2d* at 3 dR, or *wnt11-1* and *wnt11-2* at 6 dR (Figure R3.7C).
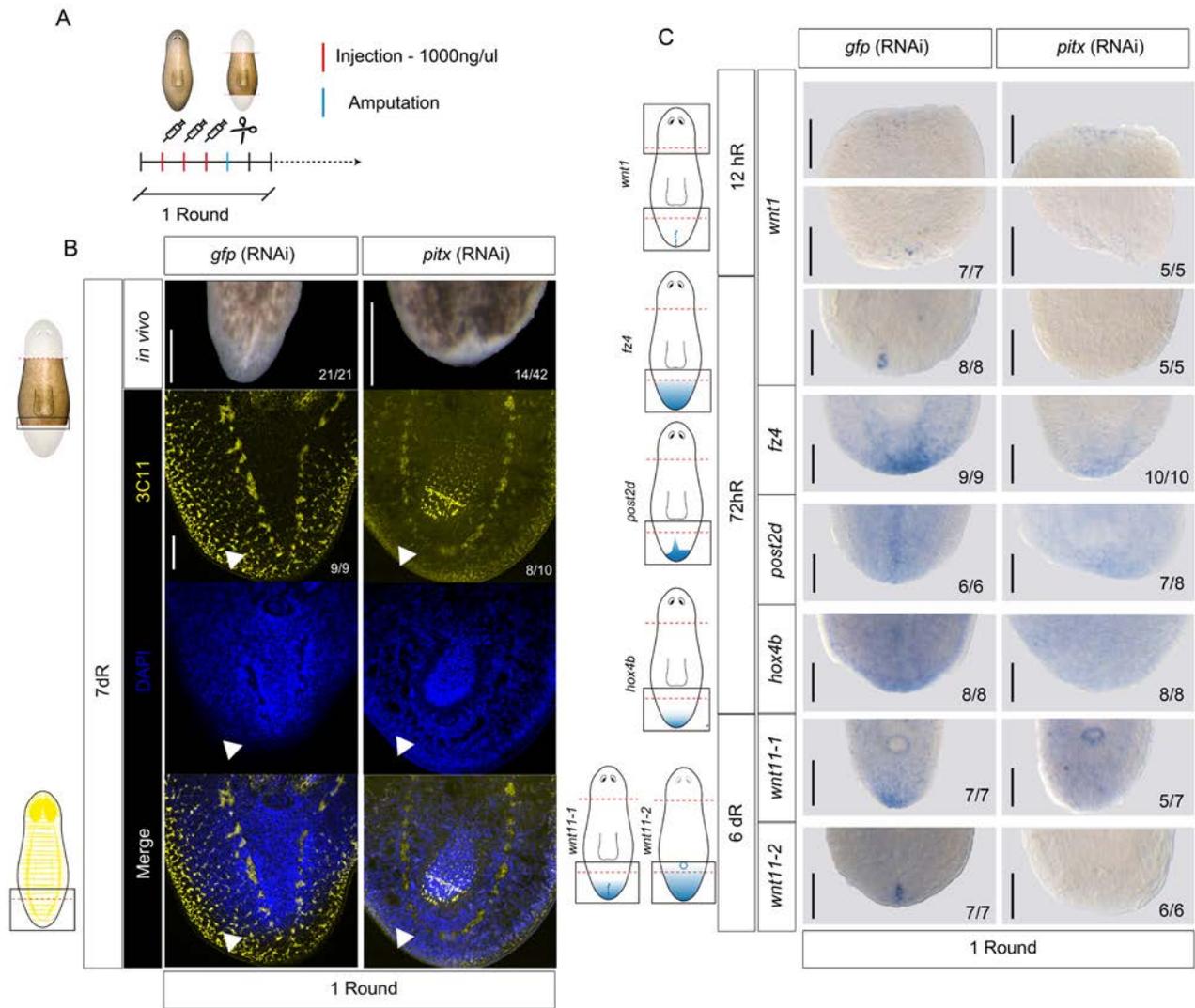
Figure R3.7: *pitx* (RNAi) animals lack *wnt1* and present a tailless phenotype. (A) Experimental design: one round of inhibition and amputation. After the amputation, animals regenerate and the phenotype is studied at different timepoints. **(B)** *in vivo* images of *pitx* (RNAi) animals showing the half of them a tailless at 7 dR. anti-3C11 immunostaining images confirms the tailless phenotype (white arrows). Nuclei are stained with DAPI. **(C)** Illustrations, indicating where *wnt1*, *fz4*, *post2d*, *hox4b*, *wnt11-1* and *wnt11-2* are expressed in intact animals. WISH of *wnt1*, *fz4*, *post2d*, *hox4b, wnt11-1* and *wnt11-2* in regenerating animals demonstrates their lack of expression in *pitx* (RNAi) animals, with the exception of *wnt1* at 12 hR. Scale bars: 100 µm in (B) and 200 µm in (C).

These findings indicate that combining genome approaches and RNAi functional strategies we could identify TFs required for posterior identity specification. Specifically, *pitx* is controlling *wnt1* expression, and its inhibition does not just perturb *wnt1* expression but also that of downstream genes.

5.2.6. *foxG* is required for early and late *wnt1* expression, and for posterior identity specification

One of the most representative TF binding sites found with our strategy was the Fox family. In the planarian transcriptome, there are several transcripts that contain a forkhead domain. BLAST searches with the specific domain allowed us the identification of a large number of Fox TFs in *Smed* (see next chapter). After a first screening of some of them, we identified *Smed-foxG* (*foxG*), a gene not described in planarian yet, which when inhibited lead to the production of the tailless phenotype.

foxG is expressed in a subset of cells all along the D/V margin, in the dorsal midline and in some scattered cells in the dorsal and ventral planarian part (Figure R3.8A). SCS data reveals that *foxG*+ cells were muscle and neurons (Figure R3.8B, Annex III). Moreover, it has been recently published that *foxG* is coexpressed with posterior organizer cells (*wnt1*+ and *collagen*+) in intact animals and posterior regenerating blastemas at 72 hR (Figure R3.8C) (230). These data demonstrate that *foxG* is expressed in two different subtypes of muscle and neuronal cells
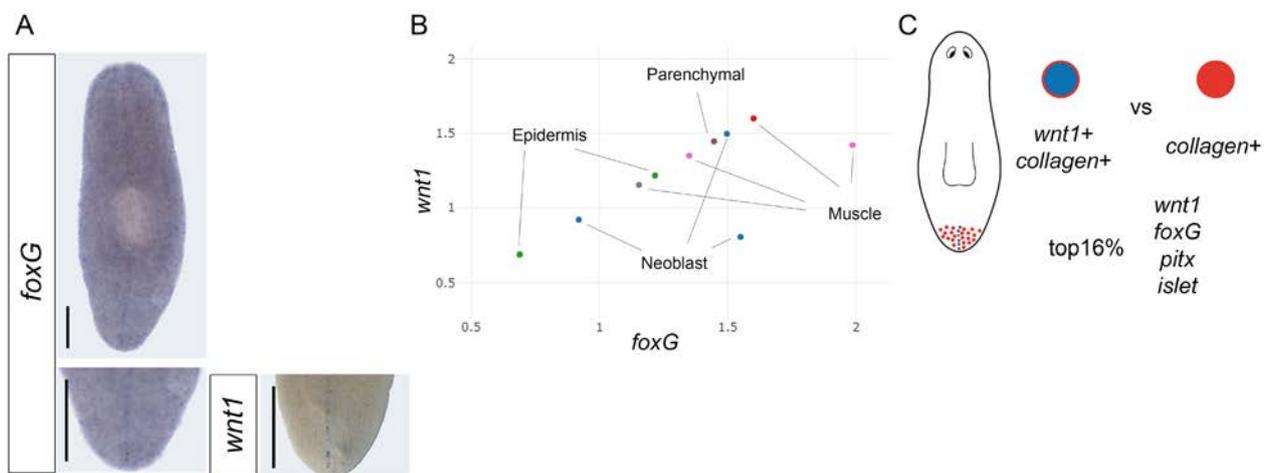
Figure R3.8: *foxG* is expressed in a subset of muscle cells and coespress with *wnt1*. (A) *foxG* and *wnt1* WISH expression in intact animals. **(B)** *foxG* coexpresses with *wnt1* in muscle cells and in neoblasts. **(C)** Schematic experimental design from (230) shows that *foxG* is present in the top 16% genes in *wnt1*+ and collagen+ population (the organizing region). Scale bar in (A) is 200 µm

To study its function, we produced two rounds of RNAi inhibition and amputation (Figure R3.9A); the 75% of knockdown regenerating animals presented a tailless phenotype (Figure R3.9B). Analyzing the phenotype through labelling the nerve cords by immunohistochemistry, we observed that 80% of the animals showed posteriors with fussioned nerve cords in U shape (Figure R3.9B) as it has been described after inhibition of other key posterior genes such as *wnt1* (60), *wnt11-2* (60), *islet* (163) or *pitx* (105,106). Furthermore, ISH with *fz4*, *post2d* and *hox4b* riboprobes at 3dR showed that all markers were under-expressed at posterior blastemas (Figure R3.10).
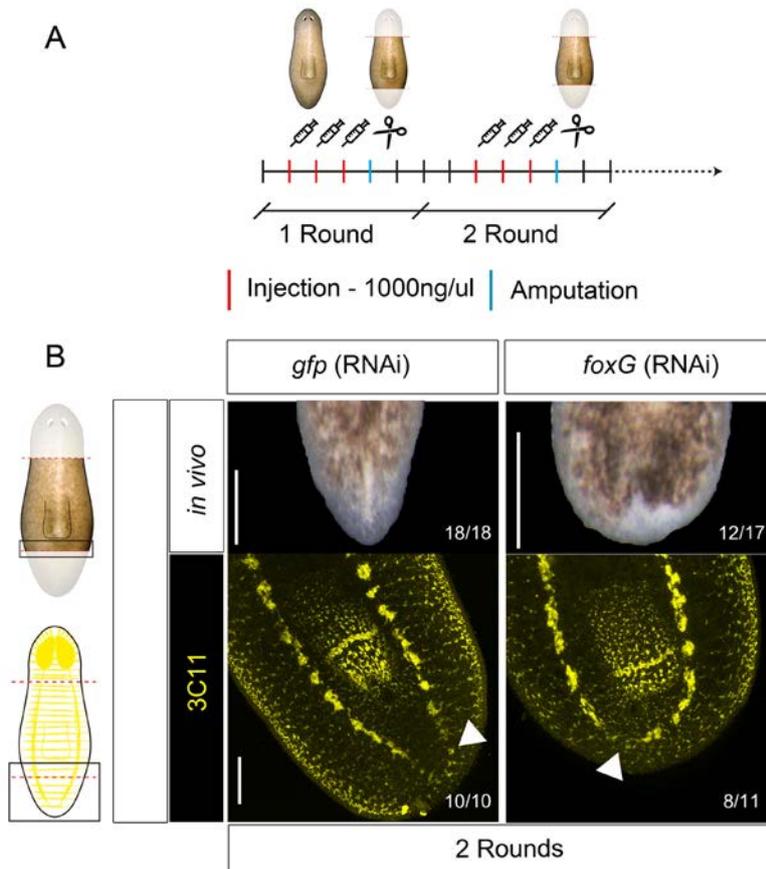
Figure R3.9: *foxG* **(RNAi) animals show a tailless phenotype. (A)** Experimental design with two rounds of inhibition and amputation. After the second amputation, animals regenerate and phenotype was studied. **(B)** *in vivo* images of planarian showing tailless phenotype in *foxG* (RNAi) animals. anti-3C11 immunostaining images of control and *foxG* (RNAi) animals corroborates the tailless phenotype (white arrows). Scale bars:100 µm in (B).

Since *foxG* coexpresses with *wnt1*, we sought to test whether *foxG* could also regulate its expression. ISH of *wnt1* demonstrated that it was absent in *foxG* knockdowns at 3 dR, suggesting that this gene could participate in the second stage of *wnt1* expression, which is stem cell dependent. Interestingly, the first *wnt1* expression was also affected after *foxG* inhibition, in both blastemas. This result is important, since it is the first gene reported to date that regulates the early *wnt1* SC independent expression (Figure R3.10). Overall, *foxG* is regulating *wnt1* expression at any region and stage. However, we cannot discriminate whether the observed tailless phenotype is a consequence of affecting the early or late *wnt1* expression; neither if the second *wnt1* wave of expression is affected because the first one does not take place.
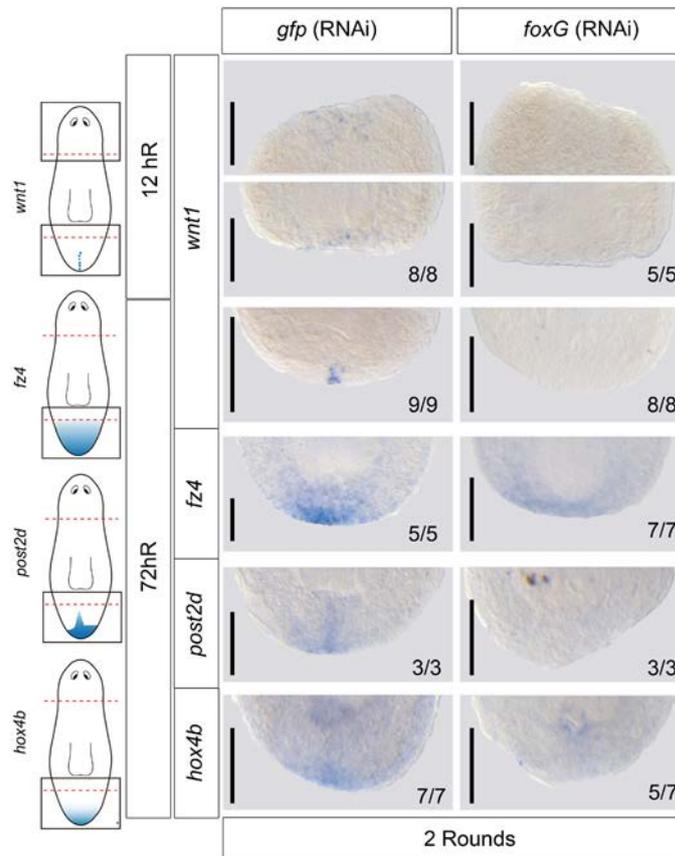
Figure R3.10: *foxG* (RNAi) animals lack *wnt1* and posterior markers expression. Illustrations, indicating where *wnt1*, *fz4*, *post2d* and *hox4b* are expressed in intact animals. WISH of *wnt1*, *fz4*, *post2d* and *hox4b* in regenerating animals demonstrates a lack of expression in *foxG* (RNAi) animals. Scale bars: 200 µm in all the panels.

All together, these results demonstrate that *foxG* is a new element of the posterior organizer. *foxG* is required for *wnt1* expression in posterior organizer cells and its inhibition produces a tailless phenotype with suppressed posterior identity.

5.2.7. FoxK family plays a role regulating the posterior organizer

During the screening to find Fox genes with a putative role regulating the posterior organizer formation, we found another interesting family of Fox genes, the FoxK family, which after its inhibition produce a posterior phenotype. It could hance play a role in regulating the posterior identity. In planarians, there are 3 *foxK* genes; and thanks to the phylogenetic analysis (see chapter III), we were able to name them as *Smed-foxK1-2.1* (*foxK1-2.1*), *Smed-foxK1-2.2* (*foxK1-2.2*) and *Smed-foxK1-1* (*foxK1-1*). Analysing the expression pattern of the three *foxK* genes shows their expression in the CNS and ubiquitously all over the animal (Figure R4.1 A; Annex III).
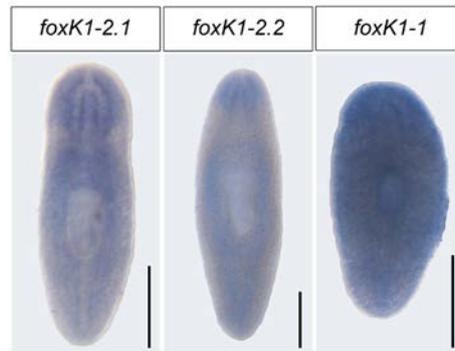
Figure R4.1: *foxK* genes are expressed in the nervous system. *foxK1-2.1*, *foxK1-2.2* and *foxK1-1* WISH in whole mount inatct animals. Scale bars: 250 µm in all the panels.

We inhibited separately all three *foxK* genes for two rounds (Figure R4.2A). The three RNAis presented some anterior defects in the head and in the eyes (Figure R4.2B). All *foxK1-2.1* knockdown animals showed a delay of anterior blastema formation, and improper eye formation. *foxK1-2.2* (RNAi) animals presented a less sever phenotype, with more developed blastemas but without properly regenerated eyes. And after the inhibition of *foxK1-1*, animals showed a delay in regeneration without major defects. Thus, the three *foxK* genes play a role in regulating anterior regeneration.

foxK knockdown animals also presented a posterior phenotype. After *foxK1-2.1* and *foxK1-2.2* (RNAi), trunks showed an *in vivo* tailless phenotype at 7dR. However after *foxK1-1* ablation, animals did not show the tailless phenotype, but they regenerate smaller tails than control animals. To better characterize whether the FoxK family could play a role in regulating the posterior organizer, we analyzed the *wnt1* expression at 7dR. *foxK1-2.2* and *foxK1-1* (RNAi) animals did not show differences in *wnt1* expression (Figure R4.2B). However, *foxK1-2.1* (RNAi) showed an increment of *wnt1* expression. This increment was just detected in the second *wnt1* expression wave, since the first *wnt1* expression (at 12hR) was not affected in posterior neither anterior (Figure R4.2B). Observing that *foxK1-2.1* was controlling *wnt1* expression, and the knockdown animals presented a tailless phenotype, we decided to further characterize *foxK1-2.1*.
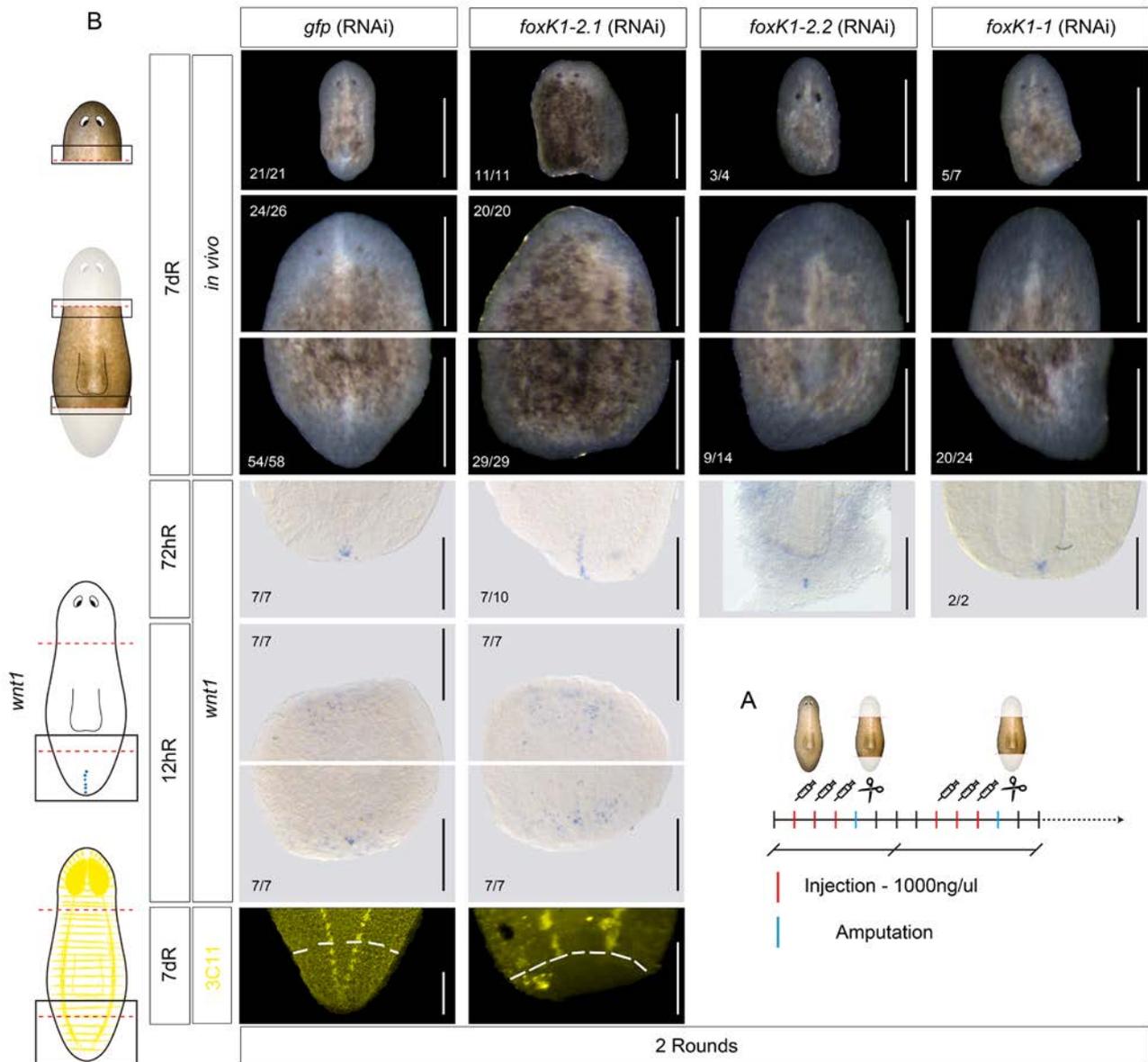
Figure R4.2: RNAi of *foxk* genes generates anterior and posterior defects. *in vivo* images of planarian showing anterior (trunks) and posterior (trunks and heads) phenotype. *wnt1* WISH at 3dr showing differences in *foxk1-2.1* (RNAi) animals. *wnt1* WISH at 12 hR did not present differences in the same group of animals. anti-3C11 immunostaining images of control and *foxK1-2.1* (RNAi) animals show a lack of posterior nervous system regeneration. Illustration, indicating where *wnt1* is expressed in intact animals, is shown. Scale bars: 100 µm in all the panels.

5.2.7.1. *foxK1-2.1* regulates *wnt1* expression and impars, posterior specicification

Considering that after two rounds of inhibition, most of the *foxK1-2.1* RNAi animals died and their nervous system was highly affected, we sought to inhibit one round only to avoid the huge mortality and to properly study the possible role of the genes influence within the posterior identity. After one round of inhibition (Figure R4.3A), mortality was reduced, but head and eye regeneration were still affected (Figure R4.3B). The proportion of tailless phenotype was maintained, suggesting that one round could be enough to study the *foxK1-2.1* role in the posterior identity specification (Figure R4.3A, B, C). By ISH, we could describe that at 3 dR, *fz4* and *hox4b* were reduced in knockdown animals, but *post2d* was increased. On the other hand, after 6 days of amputation, *wnt11-2* was unaffected and *wnt11-1* expression was decreased in *foxK1-2.1* (RNAi) animals (Figure R4.4). These results suggest that *foxK1-2.1* is controlling some posterior cWNT target genes.
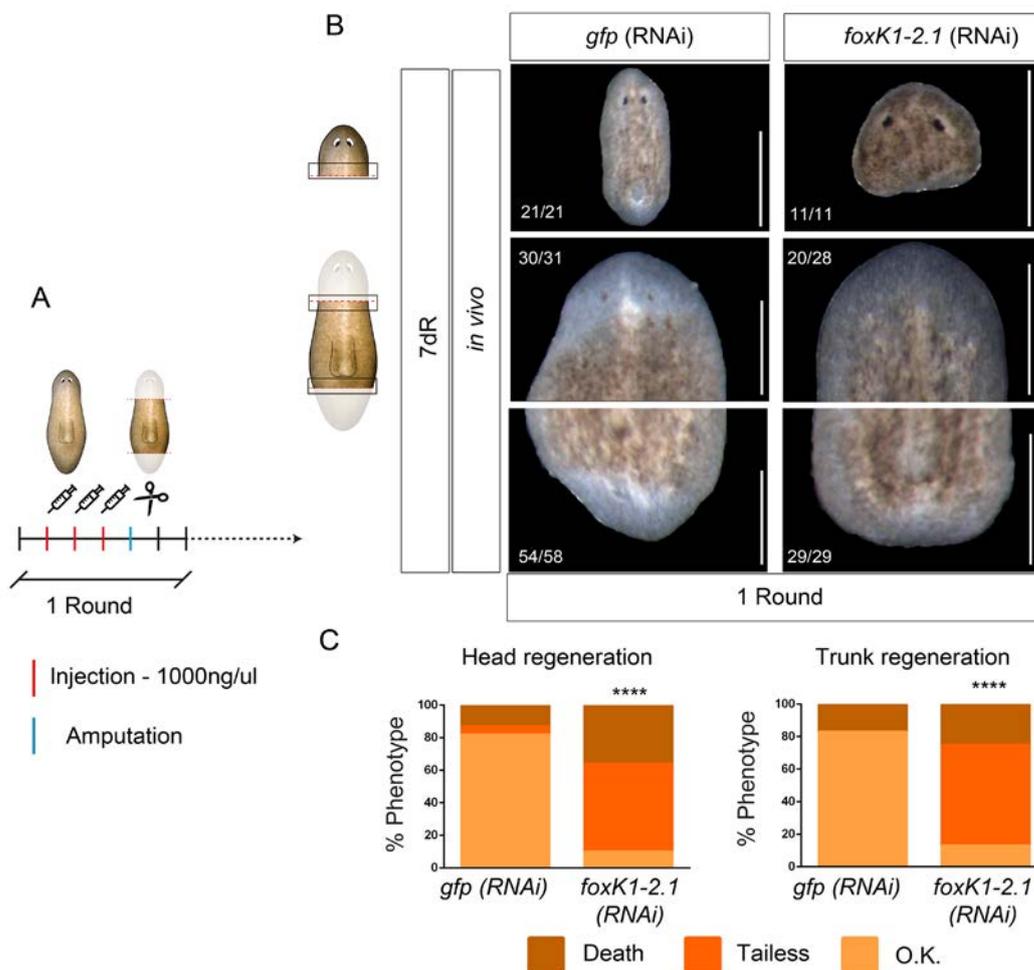
Figure R4.3: *foxK1-2.1* (RNAi) animals show a tailless phenotype. (A) Experimental design with one round of inhibition and amputation. After the amputation, animals regenerate and the phenotype is studied. **(B)** *in vivo* images of planarian showing anterior (trunks) and posterior (trunks and heads) phenotype of control and *foxK1-2.1* (RNAi). Posterior blastemas show a tailless phenotype and the anterior show a lack of eye regeneration. **(C)** Bar plots showing phenotype penetrance in posterior heads and trunk regeneration.
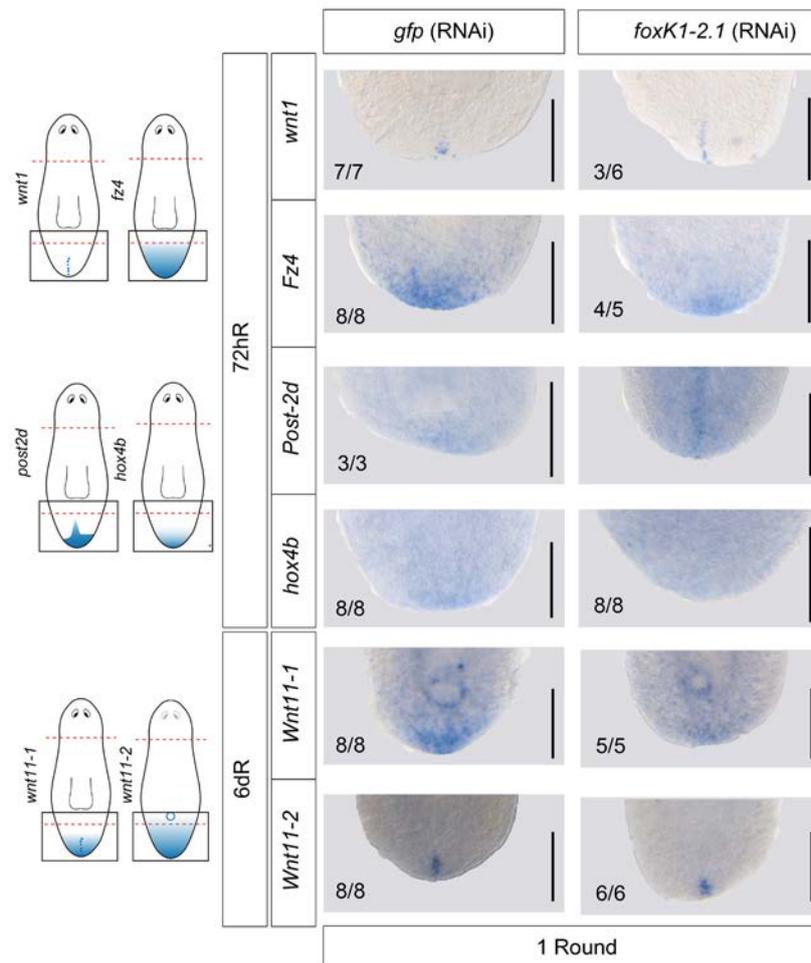
Figure R.4.4: *foxK1-2.1* (RNAi) animals show differences in posterior markers expression. Illustrations, indicating where *wnt1*, *fz4, post2d*, *hox4b*, *wnt11-1* and *wnt11-2* are expressed in intact animals, are shown. WISH of *wnt1*, *fz4*, *post2d*, *hox4b wnt11-1* and *wnt11-2* in regenerating demonstrates an increment of *wnt1* and *post2d* expression, and a reduction of *fz4*, *hox4* and *wnt11-1*. Scale bars: 100 μm in all the panels.

By immunostaining, we observed that *foxK1-2.1* (RNAi) animals showed a non-well regenerated ventral nerve cords (VNC); and even thought they presented a U shape, they did not fussion at the midline (Figure R4.5). For those reasons we named the emerging verion as a "tailless-like" phenotype. Another interesting aspect was that the pre-existing nervous system was disrupted and the synapsin (3C11) was under-expressed compared to the controls, suggesting that there was no proper regeneration either maintenance of the VNCs.
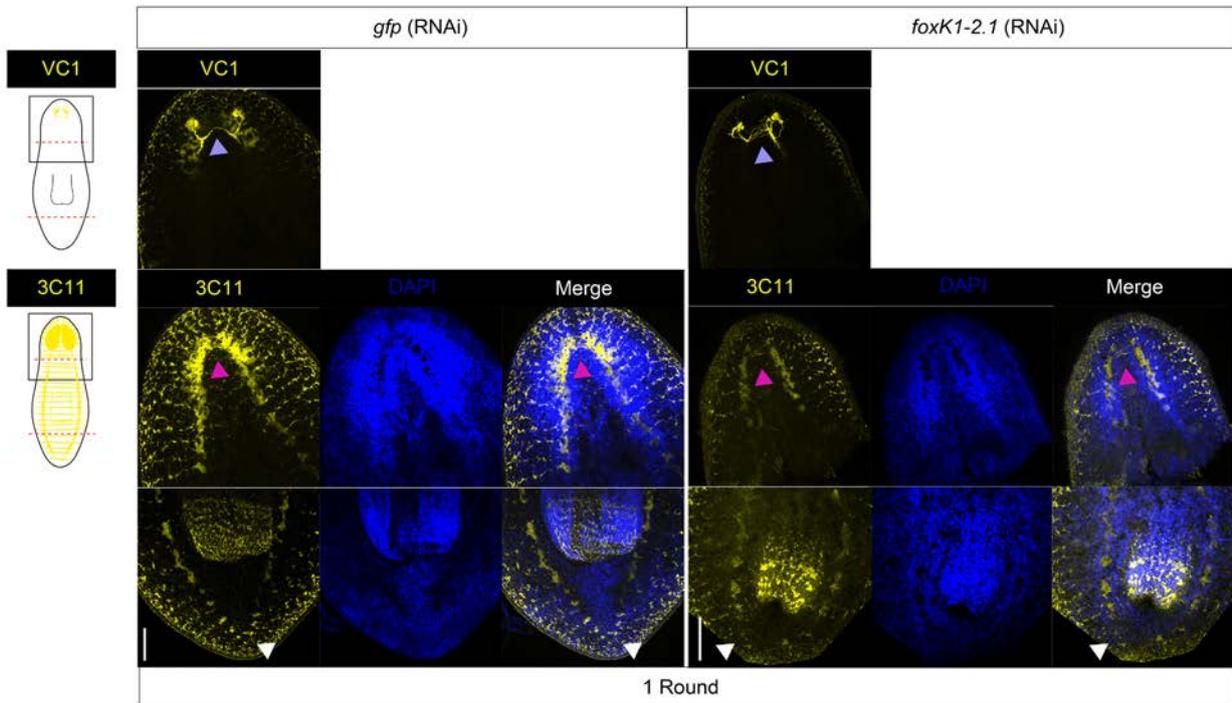
Figure R.4.5: *foxK1-2.1* (RNAi) animals show defects in the nervous system. At 7 dR, anti-VC1 immunostaining images of control and *foxK1-2.1* (RNAi) animals show a bad formation of the optic chiasm (blue arrows). In anterior, anti-3C11 immunostaining images of control and *foxK1-2.1* (RNAi) animals show a bad formation of the brain (purple arrows) and in the posterior a lack of fussion of the ventral nerve cords in the tip of the animal (white arrows). Nuclei are stained with DAPI. Scale bars: 100 µm in all the panels

5.2.7.2. *foxK1-2.1* and *foxk1.2.2* act synergistically affecting the posterior identity

The analysis of the aminoacid sequences of each *foxK* gene (see next Chapter; Annex III), shows that all three genes presented the two typical domains of the FoxK family: forkhead (FKD), forkhead associated domain (FHA) and different nuclear localization signalling (NLS) (Figure R4.6A). Analyzing sequence identity, we identified that FOXK1-2.1 and FOXK1-2.2 sequences share more identity in both domains than the other *foxK* gene (Figure R4.6B). Moreover, after each gene inhibition, we observed a similar posterior phenotype with different penetrance. This is why we sought to performed a double RNAi experiment in order to observe a putative strongest posterior phenotype.

We performed a FISH using *Smed-pc2* to visualize the nervous system, since with the usual marker 3C11; the nerve cords were not properly visualized. Both single RNAi animals showed the rounded U shape of the nerve cords without reaching the posterior midline (Figure R4.7). Double RNAi organisms also showed that same phenotype. DAPI staining confirms the U shape of the digestive system in the all knockdown animals. We also checked *wnt1* expression, since single inhibition of *foxK1-2.1* showed an increment of its expression. At 3dR, double RNAi planarians also showed an increment of *wnt1* expression (Figure R4.7).
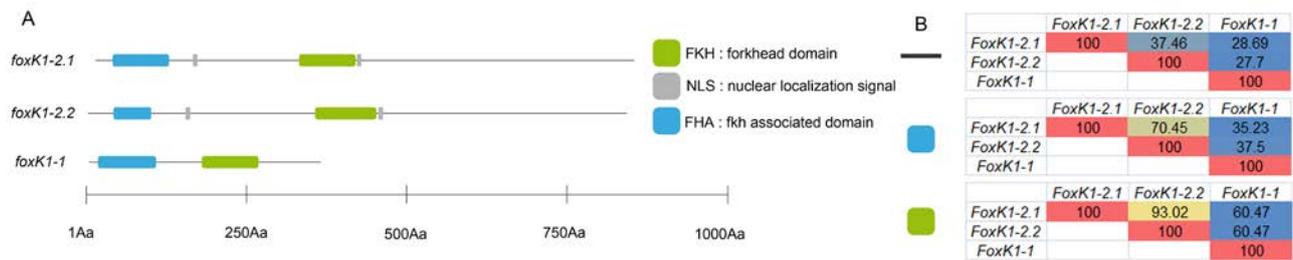
Figure R4.6: FOXK1-2.1 and FOXK1-2.2 are similar at aminoacidic level . (A) Schematic cartoon of aminoacidic sequence and domains in *Schmidtea mediterranea* proteins: FOXK1-2.1, FOXK1-2.2 and FOXK1-1. **(B)** Table comparing the identity among three *foxK* gene members: at the whole sequence level, at the forkhead domain (FKH) level and the forkhead associated domain (FHA) level..
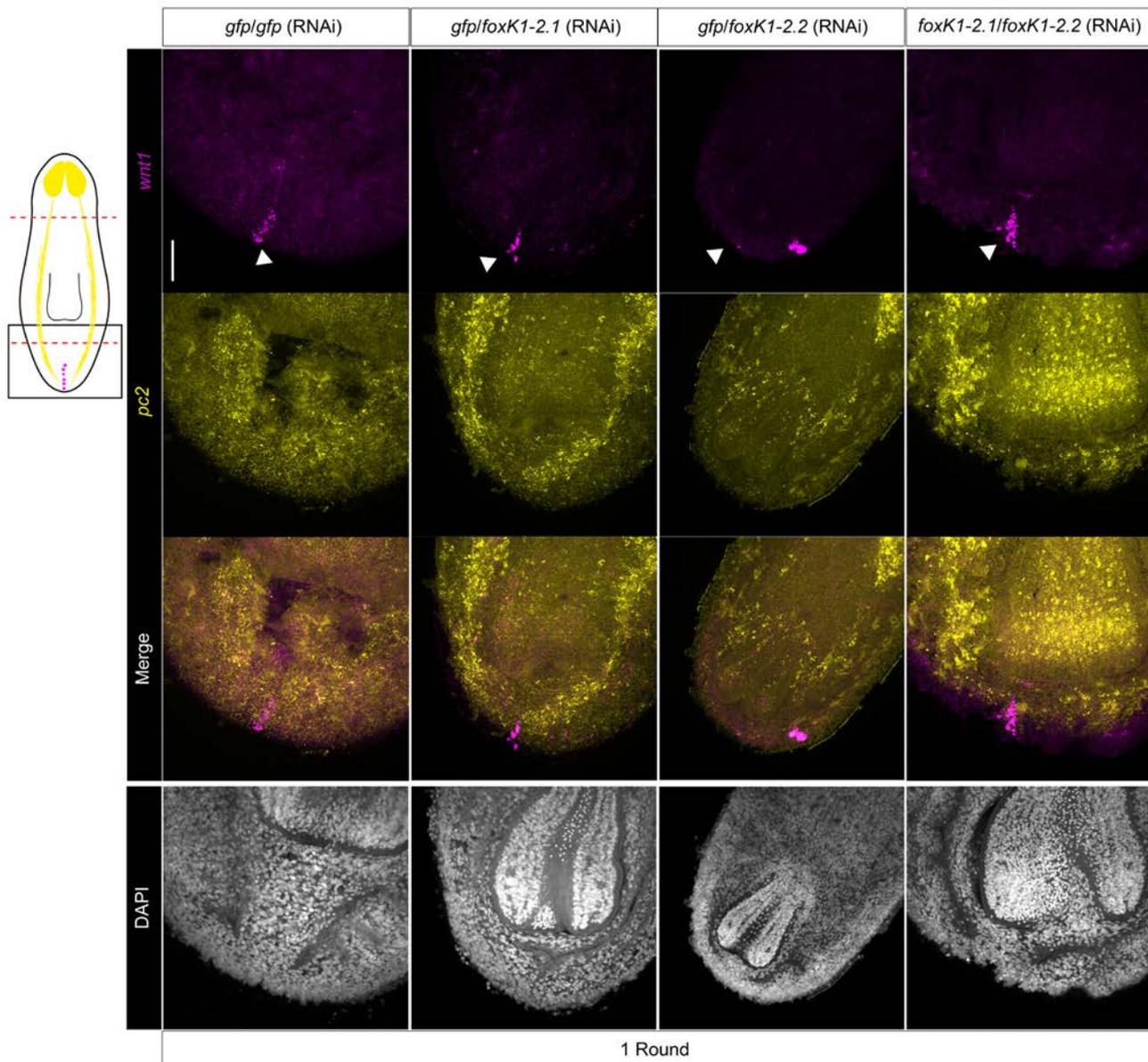
Figure R4.7: *foxK1-2.1* and *foxk1-2.2* double (RNAi) animals show a tailless-like phenotype and an increment of *wnt1* expression. At 7 dR, dFISH of *wnt1* and *pc2* demonstrates the lack of fusion of the ventral nerve chords in the midline. DAPI staining (nuclei) levels shows intestine shape, being not elongated and fussioned in the midline. *wnt1* is increased in double RNAi compared to the control animals (white arrow).

Interestingly, single RNAi animals showed similar phenotype proportions as in the previous experiments, being more tailless-like and showing more anterior defects in *foxK1-2.1* (RNAi) animals than *foxK1-2.2*, respectively (Figure R4.8). Any double RNAi trunks presented eyes, and most of them also showed a tailless-like phenotype, increasing the proportion of the taillless-like phenotype. Death animal percentage was no modified among the inhibitions (Figure R4.8).
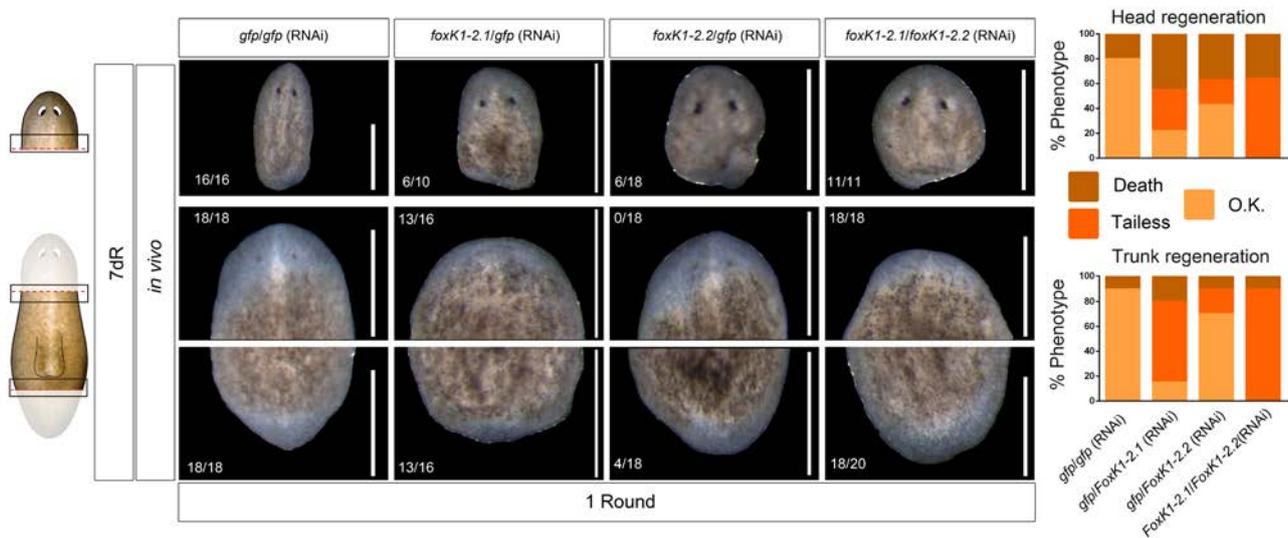
Figure R4.8: *foxK1-2.1* and *foxk1-2.2* (RNAi) animals show a stronger phenotype. *in vivo* images of planarian showing anterior (trunks) and posterior (trunks and heads) phenotype of control and *foxK1-2.1/gfp* (RNAi), *foxK1-2.2/gfp* (RNAi) and *foxK1-2.1/foxK1-2.2* (RNAi). Double RNAi animals presented more percentage of phenotype. Bar plots corroborates the increment phenotype penetrance in posterior of double RNAi heads and trunk fragments. Scale bars: 100 μm in all the panels.

Altogether, this data suggest that both *foxK* genes could act synergistically in the posterior organization, since their double inhibition increased the penetrance of the phenotype.

5.2.7.3. FOXK1-2.1 could interact with DVL regulating WNT target genes

We have shown that *foxK* genes regulate *wnt1* expression and their RNAi impairs tail rege-neration. Wang et al. (241) described that *foxK* genes and the WNT pathway were interacting in vertebrates. They described that FOXK interacts with DVL in the cytoplasm allowing its nuclear translocation and transcriptional activation together with βCAT (cWNT signalling). To allow FOXK-DVL interaction, FOXK proteins require three conserved hydrophobic ami-noacids located adjacent to FHA (241); in *Schmidtea mediterranea* two amino acids are conserved and the third conserves its polarity (Figure R4.9A; Annex II). DVL protein contains at least 4 domains: DAX; Dishevelled (DVL) domain which is a cyctoplasmic phosoprotein that acts down frizzled; PDZ domain, playing a key role in anchoring receptor proteins in the membrane to cytoskeletal; and DEP, a domain being a G-protein regulator. Moreover, DSH_C is a domain related to DSH usually found in the C terminal position and only found in vertebrates. In order to permit DVL interaction, it needs PDZ domain and some particular residues which are able to be phosphorylated (242). Two DVL proteins in *Schmidtea medi-terranea* contain the four main domains, including the PDZ; also preservinf two complete phosphorylation sites (2 and 4), and the others conserve their polarity (Figure R4.9B; Annex II). This suggests that in *Schmidtea mediterranea* FOXK and DVL show the domains allowing their functional interaction.

To test whether FOXK and DVL proteins were expressed in the same cells we analyzed the SCS data. Besides the three *foxK* and the two *dvl*, we also included *βcat1* in the study as the TF that FOXK should be the cofactor; and the two cWNT receptor: *fz1* and *LRP*. Analyzing the presence of these genes in the most differentially expressed genes per cell reveals as third common combination *foxK1-2.1* and *dvl1*, indicating that both genes are in the same cell (Figure R4.9C). *foxK1-2.1* and *βcat1*, *foxK1-2.1*, *fz1* and *LRP* combinations were also present. Altogether, these data also agrees with the hypothesis that *foxK1-2.1* could also be related with cWNT through DVL regulation, as described in vertebrates.

To prove this hypothesis we could have inhibited *dvl* with *foxK* to see whether the phenotype is stronger. However *dvl-1* and *dvl-2* (RNAi) produce a strong anteriorized phenotype that would interfere in the interpretation of the results (174).
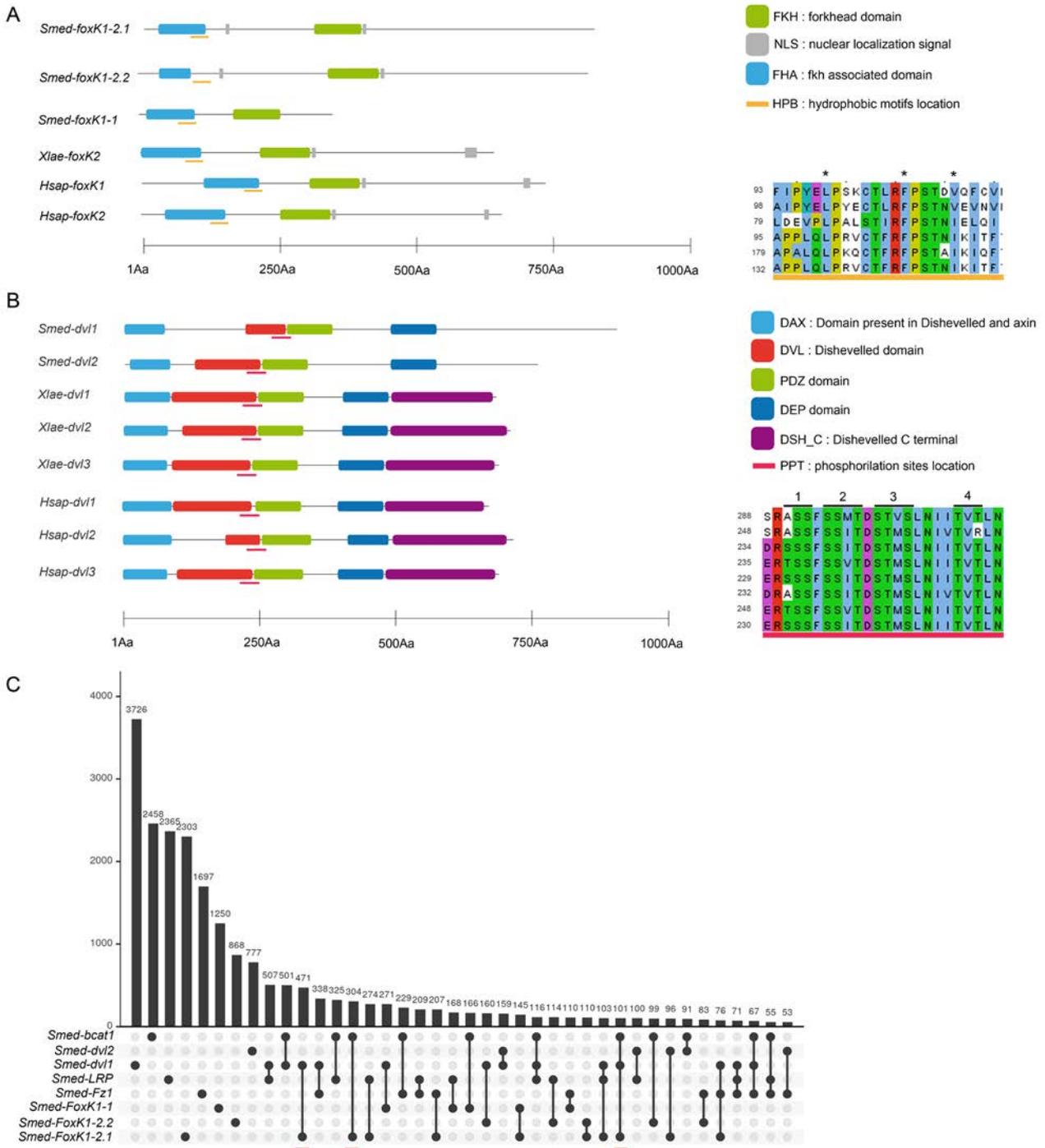
Figure R4.9: *Smed-foxK* genes could interact with *Smed-dvl* genes. (A) FOXK aminoacidic analysis of *Schmidtea mediterranea* (*Smed*), *Xenopus laevis* (*Xlae*) and *Homo sapiens* (*Hsap*). Domains are marked in different colours: forkhead domain (green), nuclear localization signal (gray) and forkhead associated domain (blue). Yellow line marks the localization of hydrophobic motifs. Sequence magnification of this region (yellow line) demonstrated its high level of conservancy. **(B)** DVL aminoacidic analysis of *Schmidtea mediterranea* (*Smed*), *Xenopus laevis* (*Xlae*) and *Homo sapiens* (*Hsap*). Domains are marked in different colours: DAX domain (light blue), DVL domain (red), PDZ domain (green), DEP domain (dark blue) and DSH_C domain (dark purple). Magenta line marks the localization of phosphorilation sites. Sequence magnification of this region (magenta line) demonstrated its high level of conservancy. **(C)** Upset plot intersection showing the presence of a single gene and the combination of them in planarian cells from (98). On top of each column, number of cells was indicated. Dark red lines indicated the colocalization at single cell level of different gene combination.

Since one round of *wnt1* (RNAi) produces a mild phenotype, we performed a double *foxK1-2.1/ wnt1* (RNAi) to test whether *foxK-2.1* increases the posteriorization of *wnt1* (RNAi) animals. The result shows that each single RNAi showed a tailless phenotype in low percentage, but the simultaneous silencing of *wnt1* and *foxK1-2.1* produces an increment of animals with tailless phenotype (Figure R4.10A). To validate whether *foxK1-2.1* could be acting as a cofactor of the βCAT1, we characterize some posterior target genes after double inhibition. At 3dR, *post2d* was absent in *wnt1* (RNAi), increased in *foxK1-2.1* (RNAi), but their double inhibition shows its absence. *hox4b* was absent after the ablation of *wnt1*, reduced after the inhibition of *foxK1-2.1*, and missing again after the double inhibition. At 6dR, meanwhile half of the single RNAi animals presented a reduction of *fz4*, double RNAi animals showed severe reduction of *fz4* or its absence (Figure R4.10B). This result together with *in vivo* tailless phenotype suggests that *foxK1-2.1* could be acting as a cofactor of βCAT1, and as a consequence it is acting through the cWNT pathway. This action could be via DVL interaction as it has been proposed in vertebrates.
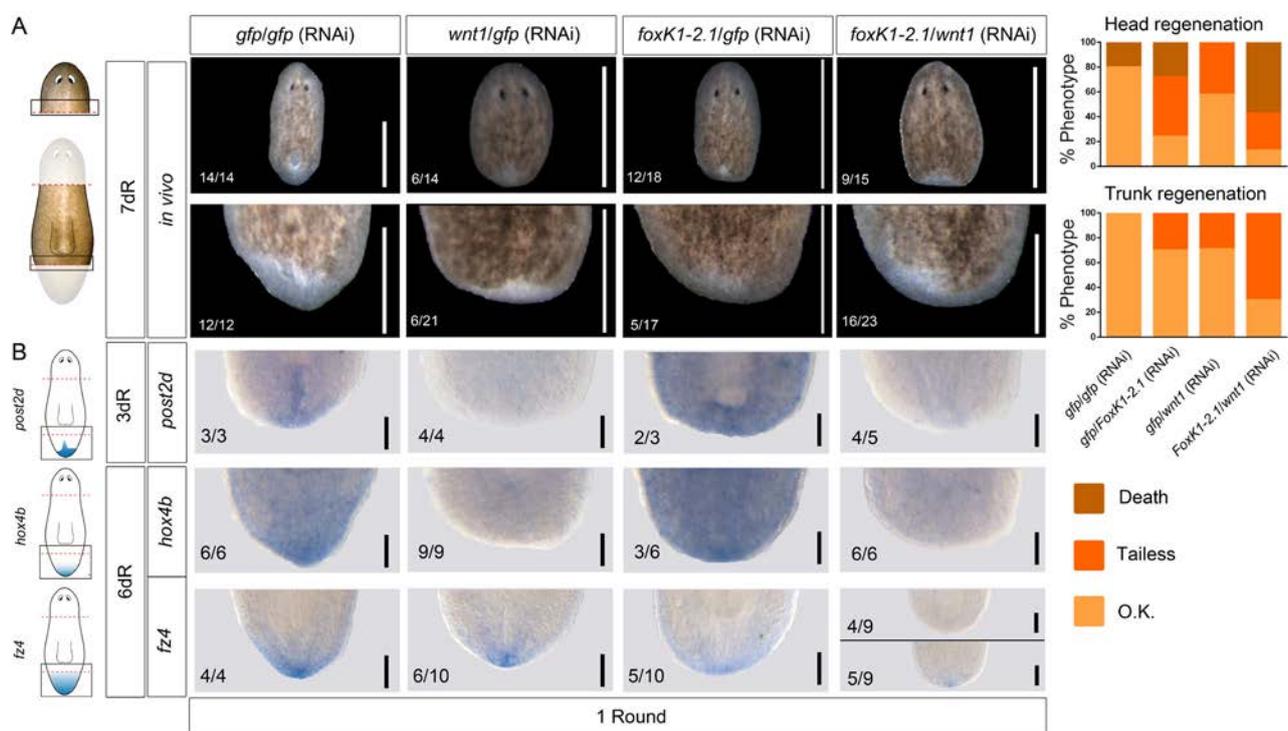
Figure R4.10: *foxK1-2.1* and *wnt1* (RNAi) animals show an increment of tailless phenotype. (A) *in vivo* images of planarian showing posterior (trunks and heads) phenotype of control and *wnt1/gfp* (RNAi), *foxK1-2.1/gfp* (RNAi) and *foxK1-2.1/wnt1* (RNAi). Double RNAi animals presented more percentage of tailless phenotype. Bar plots corroborates the increment phenotype penetrance in posterior of double RNAi heads and trunk fragments. **(B)** Illustrations, indicating where *fz4*, *post2d* and *hox4b* are expressed in intact animals, are shown. WISH of *fz4*, *post2d* and *hox4b* in regenerating demonstrates a reduction of all posterior markers in double RNAi animals. Scale bars: 100 μm in all the panels.

Overall, we demonstrated that *foxK1-2.1* could have two putative roles. First, it could be regulating the regeneration and homeostasis of the nervous system. The second could be linked with posterior identity, helping with the nuclear translocation of DVL and act as a cofactor of cWNT signalling and regulating some of their targets.

6. Chapter III: Characterization of the Fox family of transcription factors in *Schmidtea mediterranea*

Transcription factors (TFs) are key elements to regulate DNA transcriptional activity. TFs interact with DNA at the Cis Regulatory Elements (CREs) to regulate transcription (199). TFs respond to cellular signals activating or repressing specific target genes. Specific TFs are grouped according to the structure and degree of homology of their DNA binding domain (DBD). Forkhead genes are a group of specific TFs of the 'winged helix' superfamily (DBDs). The Forkhead family comprise over 2000 proteins identified in 108 animal and fungi species (243). Kaestner phylogenetically classified all Forkhead genes as Fox (Forkhead box), and grouped them in letters (A to S) indicating their membership to a family (244).

The role of *fox* genes during embryonic development of many organisms has been described. They are playing multiple roles during different developmental steps, influencing gastrulation or regulating of the differentiation and maintenance of several tissues. Specific functions have been described for FoxO, which is related with metabolism and cell growth, or FoxM which is associated to cell cycle and proliferation (245). Despite the vast number of *fox* genes identified in many species, very little is known about their presence in *Smed* and any other Lophocotrozoan clade. For that reason, and because Fox families appeared to be important for specifying planarian poles, according to our data in the previous chapter, we sought to identify the TF Fox family in *Smed* and analyze their phylogeny as well as their function during regeneration.

6.1. Identification and phylogenetic analysis of *Smed-fox* genes

In order to identify *fox* genes in *Smed*, we used TBLASTN searches in planarian genome (89) and transcriptome (90) for a Forkhead domain (FKH); we were able to identify 27 predicted *fox* genes, containing a FKH domain (Annexe III). In order to classify *Smed fox* in each family, we performed a phylogenetic analysis of the 27 FHK domains found in *Smed* (Figure R5.1). In the analysis, sequences available from different species were considered, including the Deuterostomia, Lophotrocozoa and Ecdysozoa clade. Since planarians are Platyhelminthes we specifically included Platyhelminthes species such as *Macrostomum lignano* (*Mli*) and *Schistosoma mansoni* (*Sman*) in the analysis. We also included three basal species contai-ning *fox* genes. These were *Nematostella vectensis* (*Nvec*) (246), *Amphimedon queenslan-dica* (*Amq*) (247) and *Suberites domuncula* (*Sdo*) (248). The result of the analysis shows that the 27 *fox* genes found in *Smed* can be grouped in 11 families: A, C, D, F, G, Q/D, J1, K, N, O and P. Two *fox* genes were not able to classify in any Fox family (Figure R5.1, 5.2, 5.3).

According to their evolution, Fox families are subdivided in two clades: Clade II was proposed to be the ancestor Clade, from which all Fox TFs evolved, since families belonging to it were present in stem opisthokont and fungi. In this second Clade, eight families have identified: J1, J2/3, K, M, N1/4, N2/3, O and P. Five out of these eight families of Clade II were found to be present in *Smed*: J1, K, N, O and P (Figure R5.2, 5.3). Clade I includes Fox families that appeared more recently during evolution. In this family we find families: A, AB, B, C, D, E, F, G, L1, L2, Q1, Q2 (247,249). In *Smed*, six out of these twelve families were present: A, C, D, F, G and Q/D (Figure R5.2, 5.3). Additionally, the two unclassified *fox* genes were related to FoxAB and FoxI families, respectively (Figure R5.1). Interestingly, eight *fox* genes were previously described in planarian (Figure R5.8). All of them were also presented in our data, being: FoxD (166,167,250), FoxJ1 family with four genes (251), FoxA (252,253), FoxP (254) and FoxF (123).

Figure R5.1: The ML phylogenetic tree reveals gene and family losses and some gene duplications in *Schmidtea mediterranea.* *Smed* genes are in purple. Dark asterisks indicate classified genes, and unclassified are indicated by an orange asterisk. Scale bar indicates amino acid substitution. At nodes are showed values for the approximate likelihood ratio test. Scale indicates expected aminoacidic substitution per site = 0.7. Species used are the following ones. Bilateria species: *Homo sapiens* (*Hsa*), *Xenopus tropicals* (*Xtr*), *Branchiostoma lanceolatum* (*Bla*), *Strongylocentrotus purpuratus* (*Spu*), *Saccoglossus kowalevskii* (*Sko*) and *Ptychodera flava* (*Pfl*). Protosomia: Ecdysozoa clade: *Drosophila melanogas-ter* (*Dme*) and *Tribolium castaneum* (*Tca*). Lophochotrozan clade: *Crassostrea gigas* (*Cgi*), *Lottia gigantean* (*Lgi*), *Octopus bimaculoides* (*Obi*), *Lingula anatine* (*Lan*), *Intoshia linei* (*Ili*), *Capitella teleta* (*Cap*), *Helobdella robusta* (*Hbo*). Ecdysozoa clade: *Drosophila melanogaster* (*Dme*) and *Tribolium castaneum* (*Tca*). Platyhelminthes: *Macrostomum lignano* (*Mli*) and *Schistosoma mansoni* (*Sman*). Porifera: *Nematostella vectensis* (*Nvec*), *Amphimedon queenslandica* (*Amq*) and *Suberites*
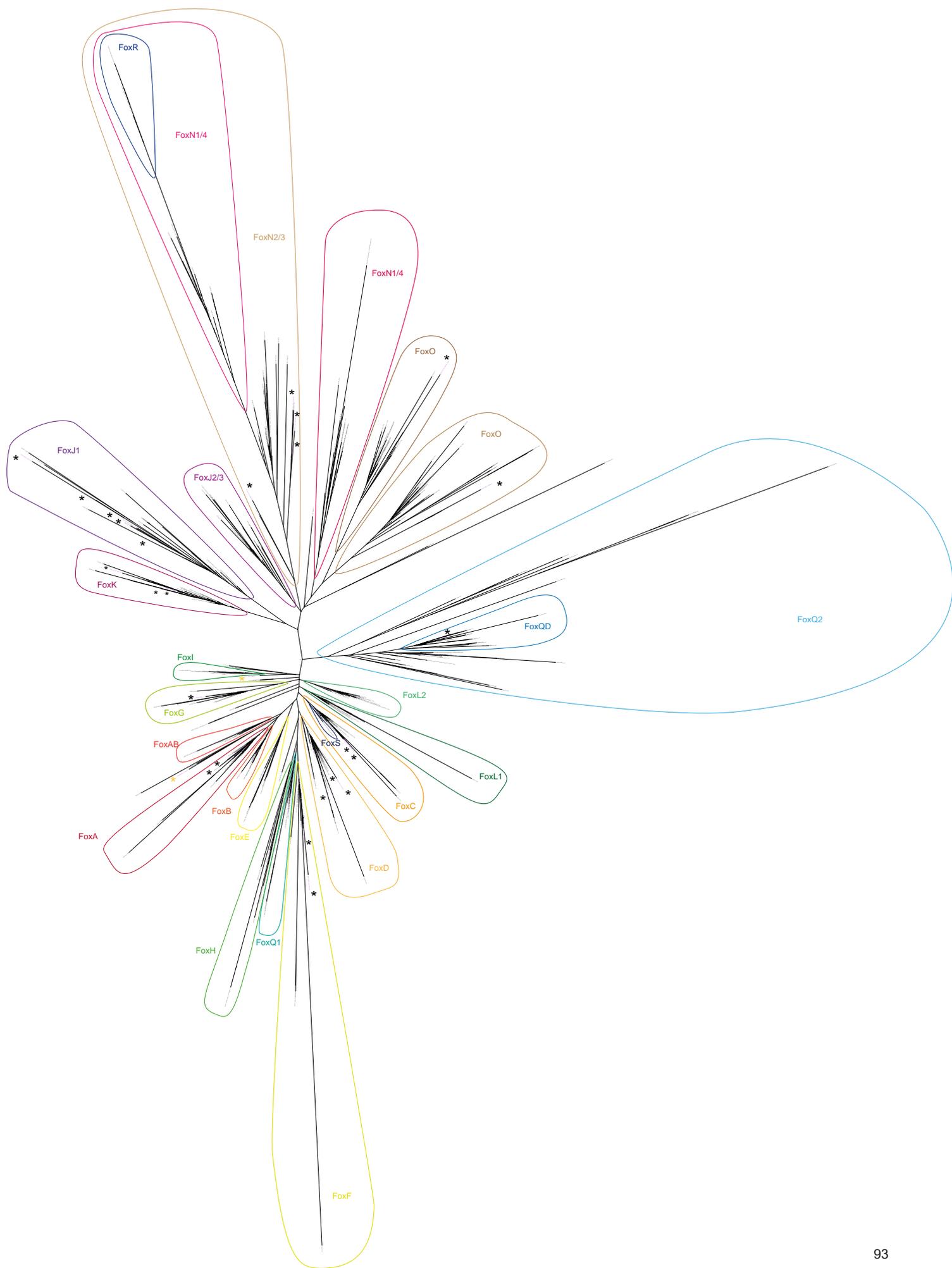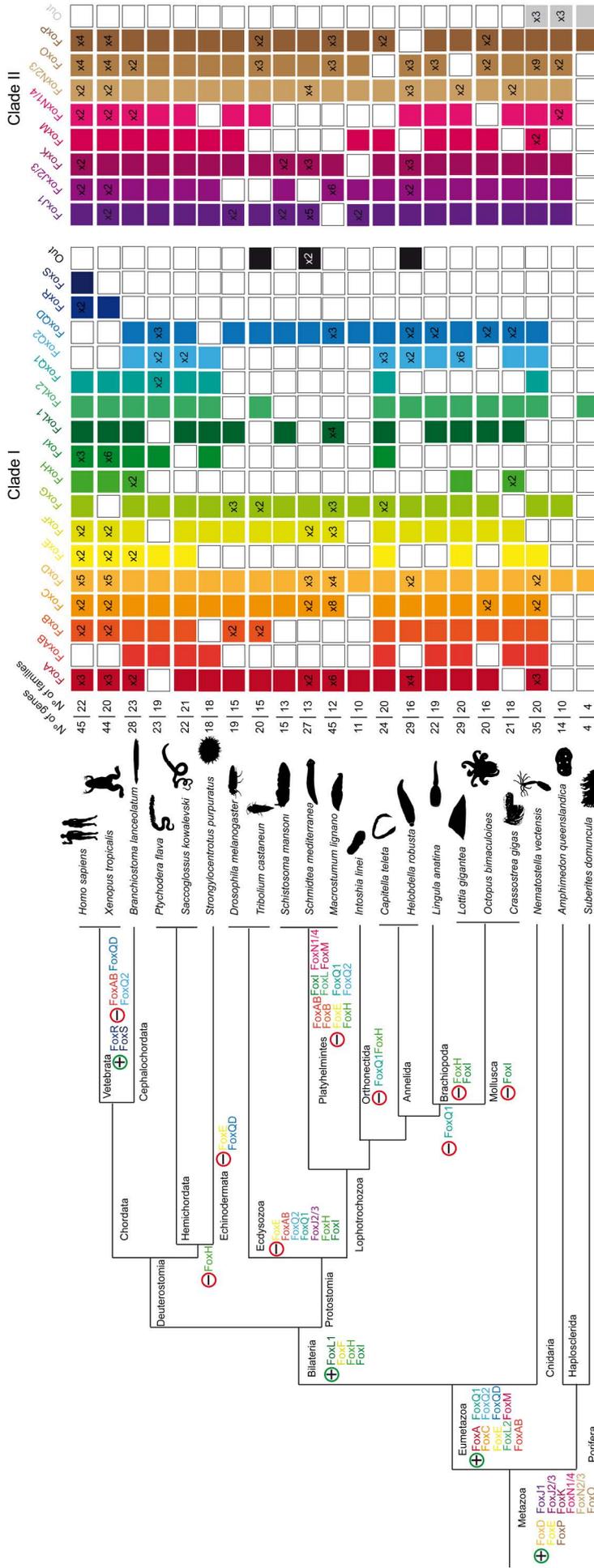
Figure R5.2: Distribution of Fox homologs in Metazoan clade. Coloured boxes indicate the presence of an ortholog based on the phylogenetic analysis. When there were no evidences of ortholog, box remains white. A number (x N°) inside a box indicates paralogs per family and specie. Families are divided in two main clades. Number of genes and number of families per specie are indicated. Metazoan (428) and Lophotrochozoa (436) phylogenies were used. Gains (+) and a losses (-) of genes are indicated. Main Clade I Fox acquisition was at base of Eumetazoa and different events of gains and losses happened thought evolution. Animal silhouettes were obtained from PhyloPic (http://phylopic.org/)

Figure R5.3: Distribution of Fox homologs in Metazoan species. Coloured boxes indicate the presence of an ortholog based on the phylogenetic analysis. When there were no evidences of ortholog, box remains white. A number (x Nº) inside a box indicates paralogs per family and specie. Families are divided in two main clades. Number of genes and number of families per specie are indicated.

Thanks to our analysis, we found that the Fox family in *Schmidtea mediterranea* is highly diverse, having 27 *fox* genes divided in 13 families. Comparing those numbers with other Lophocotrozoan species, we observe no differences in gene number but family number of Fox is reduced.

6.2. Phylogenetic analysis of Platyhelminthes *fox* genes

Thanks to our phylogenetic study we have been able to classify *Smed fox* genes. However, there were two genes in *Smed* not properly classified with the completed phylogenetic analysis. We sought to build a new phylogenetic tree using just Platyhelminthes species data (Figure R5.4, 5.5). To that aim, we included a minimum of one species per Platyhelminth order. Additionally, we also included additional Fox sequences from Tricladida species, which is the order to which *Smed* belong.
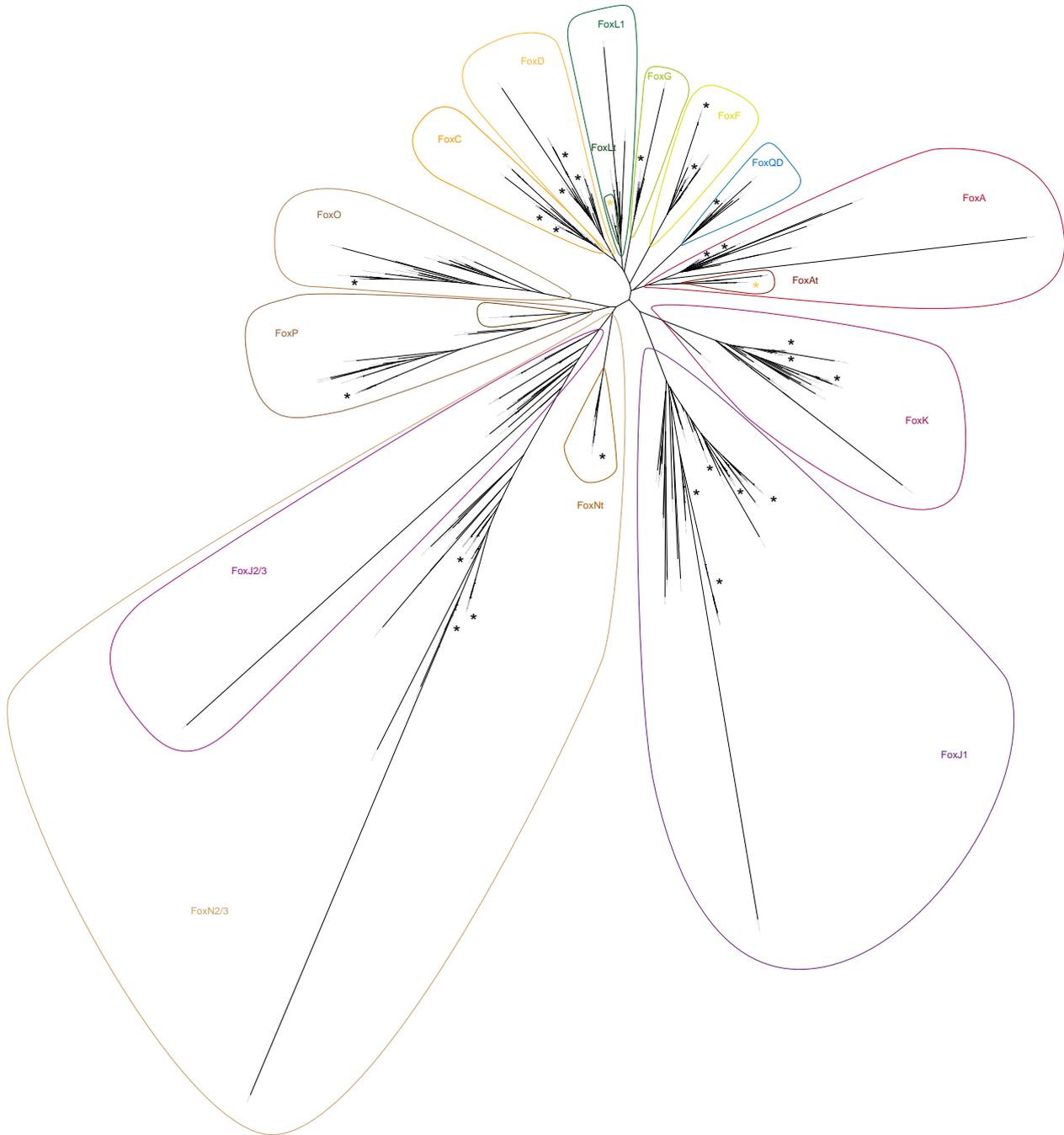
Figure R5.4: Fox family evolution in Lophotrochozoan clade. The ML phylogenetic tree reveals a specific gain Fox gene formation in *Schmidtea mediterranea*. Previous unclassified genes (orange) are classified in specific families. Scale bar indicates amino acid substitution. At nodes are showed values for the approximate likelihood ratio test. Species used are the following ones. Scale indicates expected aminoacidic substitution per site = 0.7 Species used are the following ones. Plathylemintes: *Taenia solium* (*Tso*), *Echinoccocus multiocularis* (*Emu*), *Gyrodactylus salaris* (*Gsa*), *Bothrioplana semperi* (*Bose*), *Monocelis sp.*(*Mosp*), *Mesostoma lingua* (*Mosp*), *Leptoplana lingua* (*Leli*), *Geocentrophora applanta* (*Geap*) and *Catenulia* (*Cate*). Tricladida: *Planaria torva* (*Pto*), *Polycelis nigra* (*Pni*), *Polycelis tenius* (*Pte*), *Dendrocoelum lacteum* (*Dla*), *Dugesia japonica (Dja)*, the sexual strain of *Schmidtea mediterranea* (*Smes*) and *Schmidtea polychroa* (*Spol*).

Figure R5. 5: Distribution of Fox homologs in Platyhelminthes clade. Coloured boxes indicate the presence of an ortholog based on the phylogenetic analysis. When there were no evidences of ortholog, box remains white. A question mark is disposal for putative ortholog. A number (x Nº) inside a box indicates paralogs per family and specie. Families are divided in two main clades. Number of genes and number of families per specie are indicated. Platyhelminthes phylogeny was used (80) to classify all the Orders. Grey box indicates Tricladida Order, species were classified accordingly to (12). Species belonging to Tricladida Order show different losses, gains and specialization events.

Comparing *Smed* families within Platyhelminthes, we could evaluate their presence in each Clade. FoxM and FoxN1/4 families were missing in all analysd Platyhelminthes species (Figure R5.5, 5.6). However they were present in other Lophocotrozon species (Figure R5.2). FoxJ1, FoxK and FoxN families presented a higher number of paralogs in Platyhelminthes (from 2 to 8) compared with other Lophocotrozon species (Figure R5.2, 5.5). These results suggest that during evolution Platyhelminthes lost FoxM and FoxN1/4 families, and gained gene copies of the FoxJ1, FoxK and FoxN2/3 families.

In most animal species the FoxN family is divided in FoxN1/4 and FoxN2/3 families. In Platyhelminthes, only FoxN2/3 is present (Figure R5.4). Furthermore, an interesting aspect was the presence of a homolog which was only present in the Tricladida order (Figure R5.5, 5.6). Thus, we named it as *foxNt*, and T is for Tricladida.

In this phylogenetic study, the presence of the FoxJ2/3 family inside the FoxN2/3 family was observed (Figure R5.4). Interestingly using general species analysis (Figure R5.1), FoxJ2/3 was a differentiated group (Family). Carefu analysis (Figure R5.4), allowed to observe that all FoxJ2/3 genes rise separately missing a unifying branch for all of them. This lead us to suggest that even some Platyhelminthes present FoxJ2/3 genes, those are genetically diverse among them.

Regarding Clade I, it was interesting to note the absence of several families in Platyhelminthes: FoxAB, FoxB, FoxE, FoxL2, FoxQ1 and FoxQ2 (Figure R5.5, 5.6). In addition, four families (FoxA, FoxC, FoxD, FoxF) presented two or more paralogs in different Platyhelminth species (Figure R5.5, 5.6). The new phylogenetic analysis allowed us the classification of the two previously unclassified Fox genes. One of them is related with the A family and is only present in the Dugesiidae Family. The second one falls into the FoxL family and was present in most of the Tricladida species, as exception of *Ptor* and *Pten*. Thus, we decided to name them *foxAt* and *foxL1t,* where T is for Tricladida (Figure R5.5, 5.6).

	Nº genes	Nº families	Clade I																			Clade II									
			FoxA	FoxAt	FoxAB	FoxB	FoxC	FoxD	FoxE	FoxF	FoxG	FoxH	FoxI	FoxL	FoxL1t	FoxL2	FoxQ1	FoxQ2	FoxQD	FoxR	FoxS	FoxJ1	FoxJ2/3	FoxK	FoxM	FoxN1/4	FoxN2/3	FoxNt	FoxO	FoxP	FoxPt
Schistosoma mansoni	15	13	●				●	●		●	●			●					●			x2		x2			●		●	●	
Taenia solium	17	11	x3				?	x2						●					●			x2		x2			x2		●	●	
Echinococcus multilocularis	17	13	x2					x2						●					●			x2		x2			●		●	●	
Gyrodactylus salaris	17	11	x2					x2						x2					●			x2		x2			x2		●	●	
Bothrioplana semperi	10	7	●				x2	?											●			x3					x2				
Planaria torva	31	11	x2				x2	x2		x3	x3			●					●			x6		x4			x5		x2	x2	
Polycelis nigra	34	11	x3				x3	x3		x2	x2			●					●			x8		x4			x3		x2	x2	
Polycelis tenius	27	10	x3					●						●					●			x7		x5			x3		x2	x2	
Dendrocoelum lacteum	30	12	x2				x4	x4						●					●			x6		x3			x3		x2		
Dugesia japonica	23	12	x2	●			?	x3						●					●			x5		x3			x3				
Schmidtea mediterranea	27	13	x2	●			x2	x3		x2				●					●			x5		x3			x3				
Schmidtea mediterranea (sex)	27	13	x2	●			x2	x3		x2				●					●			x5		x3			x3				
Schmidtea polychroa	27	13	x2	●			x2	x3		x2				●					●			x5		x3			x3				
Monocelis sp.	21	12					●	x2						●					●					x2			x3			x2	
Mesostoma lingua	18	10	●					●		x2				●								x3		x4			x3				
Leptoplana tremellaris	10	8	x3					●						●													?				
Geocentrophora applanta	19	8						●						●								x6	x3							x2	
Macrostumum lignano	45	12	x6				x8	x4		x3	x3			x4					x3				x6						x3	x3	
Catenulida	24	11	●				x2							x2					x4			x2	x2				x3		x4	x2	

Figure R5. 6: Distribution of Fox homologs in Platyhelminthes species. Coloured boxes indicate the presence of an ortholog based on the phylogenetic analysis. When there were no evidences of ortholog, box remains white. A question mark is disposal for putative ortholog. A number (xN) inside a box indicates paralogs per family and specie. Families are divided in two main clades. Number of genes and number of families per specie are indicated.

Comparing the number of Fox families and *fox* genes among Platyhelminthes and Lophoco-trozoan species, we could observe differences in the number of families, since most of the families were lost in Platyhelminthes. These results indicate evolutionary loss events of Fox families in Platyhelminthes, although the gene number is maintained.

The presented analysis allowed us to identify, classify and name all *Smed fox* genes. In previous reports some *fox* genes had already been identified, but with our phylogenetic study we renamed some of them. In (Figure R5.7) we summarize the previous and current nomenclature.

Genome id	transciptome id	Family	New Gene Name	Previous Name	Other homolgs	Published in	Clade
SMESG000065670.1/SMESG000065671.1/SMESG000065672.1	dd_Smed_v6_10718_0_1	A	*Smed-foxA1-1*	*Smed-FoxA*	*Djap-FoxA / Spol-FoxA*	Adler, et al., 2014; Koinuma, et al., 2000; Martín-Durán, et al., 2010	
SMESG000067799.1	dd_Smed_v6_39758_0_1		*Smed-foxA1-2*				
SMESG000028135.1	dd_Smed_v6_30453_0_1		*Smed-foxAt*				
SMESG000015673.1	dd_Smed_v6_16297_0_1	C	*Smed-foxC2-1*				
SMESG000004239.1	dd_Smed_v6_18389_0_1		*Smed-foxC2-2*				
SMESG000077075.1	dd_Smed_v6_23249_0_1	D	*Smed-foxD3-1*	*Smed-FoxD*	*Djap-FoxD*	Vogg, et al., 2014; Koinuma, et al., 2002; Vásquez-Doorman, et al., 2014	I
SMESG000065690.1	dd_Smed_v6_17749_0_1		*Smed-foxD2*				
SMESG000021761.1	dd_Smed_v6_30720_0_1		*Smed-foxD3-2*				
SMESG000075929.1	dd_Smed_v6_15035_0_1	F	*Smed-foxF1-1*	*Smed-FoxF*		Scimone et al., 2014	
SMESG000066497.1	dd_Smed_v6_6910_0_1		*Smed-foxF1-2*	*Smed-FoxF-1*		He, et al., 2017; Scimone, et al., 2018	
SMESG000010270.1	dd_Smed_v6_16466_0_1	G	*Smed-foxG*		*Djap-FoxG*	Koinuma, et al., 2002	
SMESG000021434.1	dd_Smed_v6_19255_0_1	L1	*Smed-foxL1t*				
SMESG000062929.1	dd_Smed_v6_50245_0_1	Q2/D	*Smed-foxQ/D*	*Smed-FoxQ2*		Lapan, et al. 2002; Scimone et al., 2014	
SMESG000072183.1	dd_Smed_v6_14635_0_1	J1	*Smed-foxJ1-1*	*Smed-FoxJ1-1*		Vij et al. 2012	
	no transcrit		*Smed-foxJ1-2*	*Smed-FoxJ1-2*		Vij et al. 2012	
	dd_Smed_v6_103874_0_1		*Smed-foxJ1-3*	*Smed-FoxJ1-3*		Vij et al. 2012	
SMESG000010030.1	dd_Smed_v6_10152_0_1		*Smed-foxJ1-4*	*Smed-FoxJ1-4*		Vij et al. 2012	
SMESG000017088.1	dd_Smed_v6_13009_0_1		*Smed-foxJ1-5*				
SMESG000061695.1	dd_Smed_v6_4500_0_1	K	*Smed-foxK1-2.1*			van Wolfswinkel et al., 2014	
SMESG000040645.1	dd_Smed_v6_5767_0_1		*Smed-foxK1-2.2*			van Wolfswinkel et al., 2014	
SMESG000064173.1	dd_Smed_v6_7583_0_1		*Smed-foxK1-1*				II
SMESG000048193.1	dd_Smed_v6_11337_0_1	N2/3	*Smed-foxN2-1*			van Wolfswinkel et al., 2014	
SMESG000015734.1	dd_Smed_v6_13005_0_1		*Smed-foxN2-2*				
	dd_Smed_v6_4078_0_1		*Smed-foxN2-3*			van Wolfswinkel et al., 2014	
SMESG000044628.1	dd_Smed_v6_12170_0_1		*Smed-foxNt*				
SMESG000037781.1	dd_Smed_v6_3040_0_1	O	*Smed-foxO*			van Wolfswinkel et al., 2014	
SMESG000068148.1	dd_Smed_v6_6316_0_1	P	*Smed-foxP*	*Smed-Albino*		van Wolfswinkel et al., 2014; He, et al., 2017	

Figure R5.7: Summary of *fox* genes in *Schmidtea mediterranea*. In the table is indicated the genome and transcriptome id of each gene. For each gene, it is also added the family, the new name, the previous name and orthologs of close planarian species. Literature of previous *fox* mentions in planarian was included.

6.3. The new FoxQD family

FoxQ2 family was widely described in many species (255,256). Furthermore, a couple of Fox families could derive from it, such as: FoxQD (257) or FoxQM (258), which were described in *Sko* and *Ech*, respectively. However, while we were carrying out phylogenetic analysis at different levels, we realized that inside the branch of the FoxQ2 family, a secondary branch including *foxQD* genes existed. Some of the *foxQD* genes were previously described as *foxQ2* genes, and others were newly annotated (Figure R5.1). Thus, the presence of the Fox-QD family is suggested in different species, which were never described. To further validate the existence of this family, we sought to build a phylogenetic tree just using *foxQ2* and putative *foxQD* sequences. Two clear groups appeared, dividing *foxQ2* and *foxQD* genes (Figure R5.7). Thus, we decided to rename all old previously refeered to as *foxQ2* genes and unnamed genes, and gave them the name as *foxQD* genes. For instance, genes that previously belonged to another Fox family, such as: FoxQ2, FoxI or FoxD, turned into being FoxQD. As happened in *Smed*, *Bla* and *Cgi*. Thanks to the new search and classification of *fox* genes, we could annotate new *foxQD* genes never reported from different species, such as: *Pfl*, *Sko*, *Dme*, *Tca*, *Sman*, *Mli*, *Ili*, *Cte*, *Hro*, *Lan*, *Lgi*, *Obi* and *Nve*.

From an evolutionary point of view, FoxQD wass lost in the Chordata phylum, and it seems to be lost completely in Echinodermata. Particularly, some Platyhelminthes species do not present it. As for other families, the FoxQD family also presents loss and duplication events.
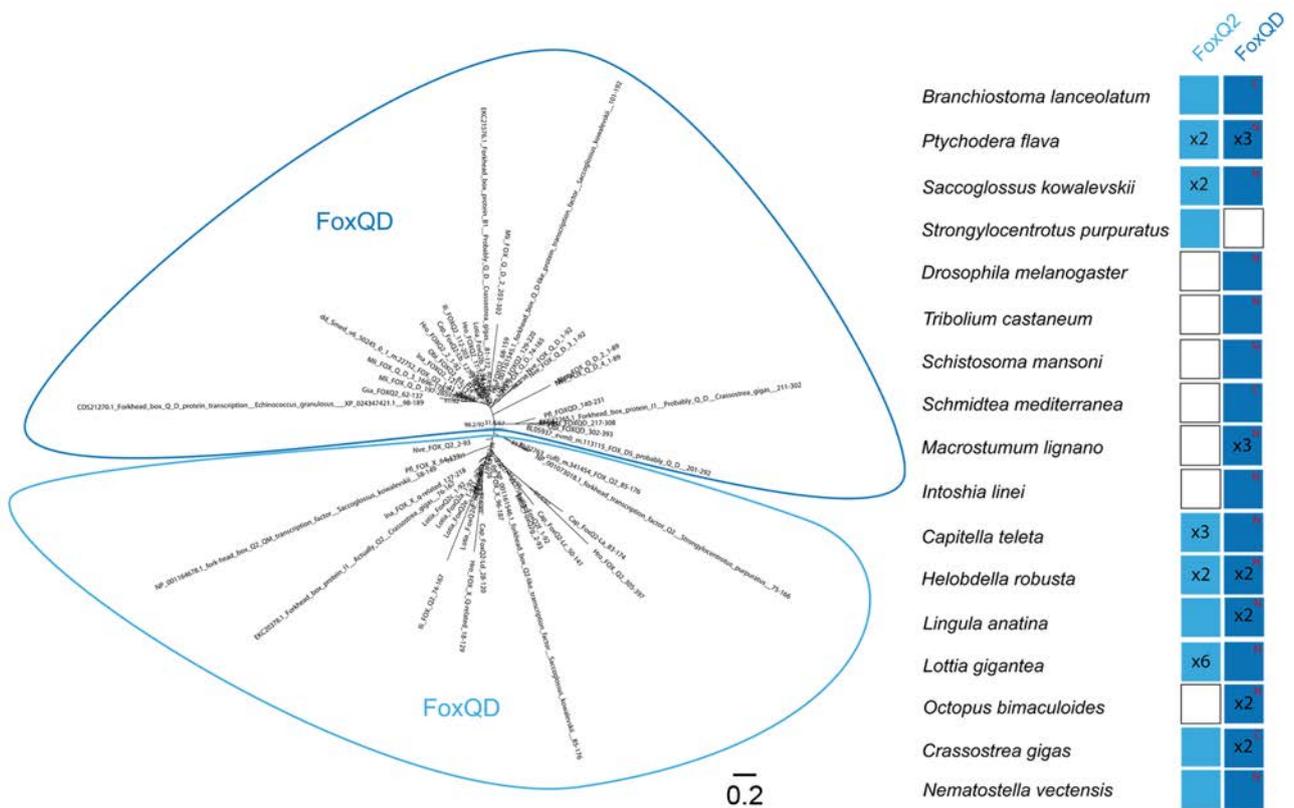
Figure R5.8: Evolution of FoxQ2 and FoxQD families. (A) The ML phylogenetic tree reveals two clear families distributed in different species trough evolution. Scale bar indicates amino acid substitution. At nodes are showed values for the approximate likelihood ratio test. Species used are the following ones. Species used are listed in the figure. **(B)** Coloured boxes indicate the presence of an ortholog based on the phylogenetic analysis. When there were no evidences of ortholog, box remains white. A number (xN) inside a box indicates paralogs per family and specie. New (red N) FoxQD annotations were indicated. And previous annotated FoxQ2 converted (red C) to FoxQD were also indicated.

6.4. Genomic distribution of *fox* genes in *Smed*

The analysis of the genomic distribution of *fox* genes from different species, allowed the identification of two Fox family clusters: FoxD-FoxE, and FoxC-FoxF-FoxL1-FoxQ1 (259). We sought to investigate whether this or other cluster could be found in *Smed*, although this is currently a difficult task. The sequencing and ensemble of the planarian genome has recently been improved, but not at the level of chromosome organization. Analyzing the available genomic database, we identified that just two *fox* genes (*foxA1-1* and *foxD2*) were found in the same genomic scaffold, and were separated by 200kB (Figure R5.9). Although, they are far from each other, the distance between them is 0.1 to 1% of the planarian genome size (259), and as a consequence they could be considered a cluster. This gene cluster had never been reported before. The rest of *fox* genes were found to be spread in different scaffolds, being impossible to relate to each other.

Family	New Gene Name	Scaffold
A	Smed-foxA1-1	61 *
	Smed-foxA1-2	67
	Smed-foxAt	207
C	Smed-foxC2-1	153
	Smed-foxC2-2	113
D	Smed-foxD3-1	88
	Smed-foxD2	61 *
	Smed-foxD3-2	181
F	Smed-foxF1-1	85
	Smed-foxF1-2	63
G	Smed-foxG	133
L1	Smed-foxL1t	17
Q2/D	Smed-foxQ/D	57
J1	Smed-foxJ1-1	76
	Smed-foxJ1-2	78
	Smed-foxJ1-3	65
	Smed-foxJ1-4	132
	Smed-foxJ1-5	15
K	Smed-foxK1-2.1	55
	Smed-foxK1-2.2	298
	Smed-foxK1-1	
N2/3	Smed-foxN2-1	36
	Smed-foxN2-2	154
	Smed-foxN2-3	238
	Smed-foxNt	31
O	Smed-foxO	26
P	Smed-foxP	68

Figure R5.9: *fox* are not clustered in planarian genome. Table showing scaffold disposition of each *fox* gene. Asterisks mark the two genes in the same scaffold.

6.5. Protein domains of *Smed* Fox family proteins

All *Smed fox* genes analyzed contained a forkhead domain (FKD). This was the property used to select the initial transcripts. Moreover, *Smed* FOXK proteins also showed a forkhead associated domain (FHA) at the N terminal part (Annexe III), mostly associated with the FoxK family but also present in other proteins (260). The *Smed* FOXP protein also shows an evolutionary conserved FOXP coiled-coil domain at the N terminal part (Annexe III), which allows dimerization and stabilization (261). Additionally, nuclear localization signals (NLS) were found in 13 *Smed* Fox, according to their nuclear function (Figure R5.9; Annexe III). The non-conserved regions appeared highly divergent, indicating their different protein interactions and functions. This variation could be explained by their different cellular functions.
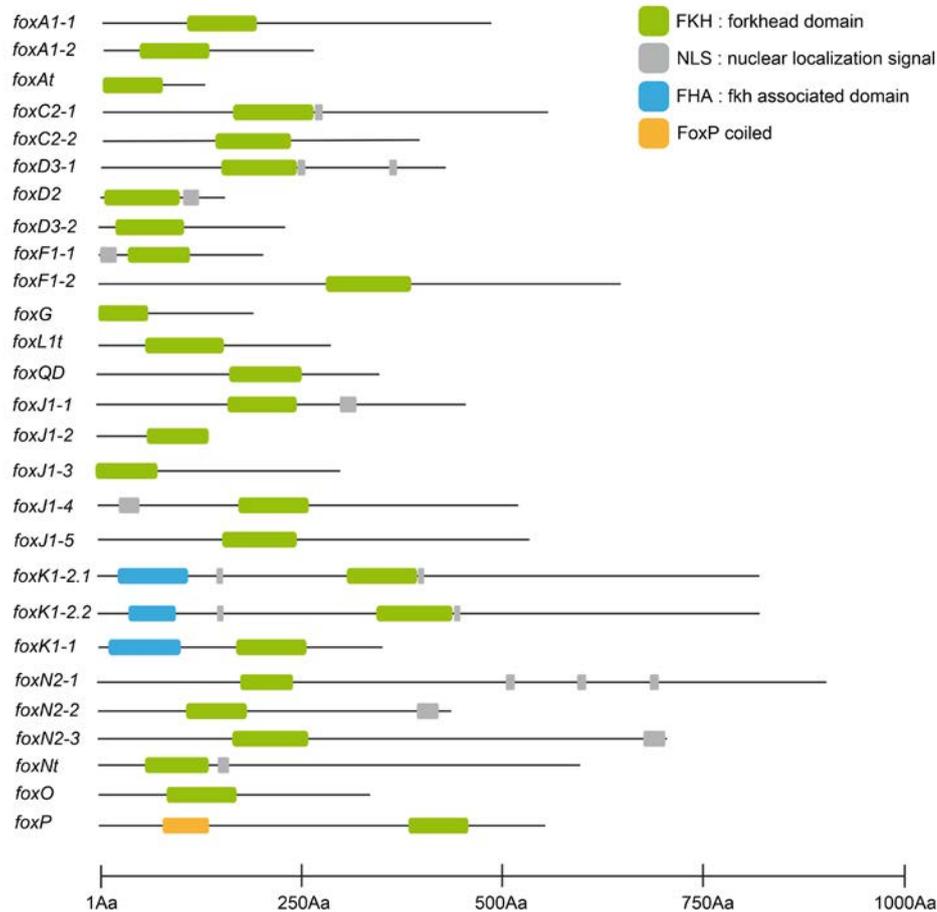
Figure R5.10: Domains of FOX proteins in *Schmidtea mediterranea.* Conserved domains in planarians: forkhead (green), forkhead associated (blue), FoxP coiled (yellow) and nuclear localization signal.

6.6. *fox* genes are tissue specific in *Smed*

Previous studies of *fox* genes expression in *Schmidtea mediterranea* showed that they were tissue and cell type specific. To support this information and to better understand *fox* gene expression, we performed ISH with riboprobes corresponding to the new *fox* genes identified (Figure R5.11).

Within the FoxA family, *foxA1-1* was expressed in the pharynx and its progenitors (253). *foxA1-2* was marginally expressed in a dotted pattern all along the animal body. We were not able to detect the new *foxAt* using ISH, even tought SCS data suggests that it could be expressed in early epidermal progenitors and/or non-ciliated neurons (Annex III).

foxC genes were expressed around the pharynx. Additionally, *foxC2-1* was also expressed at the pharynx itself.

foxD3-1 was expressed in muscle cells at the anterior tip (165,166). The other two *foxD* genes were not detected by FISH. SCS data reveals that *foxD3-2* could be present in some neural progenitors and non-ciliated neuronal cell type. *foxD2* could be found at different muscle cell types, early epidermal progenitors and neurons otf+2 cells (Annex III).

One member of the FoxF family, *foxF1-2*, was previously described to be expressed in muscle (non-body wall) and pigment cells (123). *foxF1-1* was expressed in cells in the margin of the head and in the lateral dorsal part of the animal, between the pharynx and the margin of the organism.

foxG was expressed in a subset of muscles all along the DV margin, in the dorsal midline and some scattered cells in the dorsal and ventral part.

foxL1t was not detected by ISH. Thanks to SCS data, it seems that it could be found in a muscular pharynx cell type (Annex III).

foxQ/D was expressed in differentiated eye cells (rhabdomeric photoreceptor neurons), some brain progenitors and in ventral nerve cords (262).

All previous *foxJ1* described genes were expressed in ciliated cells, located in different patterns and being more dorsally or more ventrally located depending on the gene (251). The non-previously described *foxJ1-5* was also expressed in the epidermis more concentrated in the head area and in the pharynx.

The three FoxK family genes were expressed ubiquitously and specifically in the CNS.

The FoxN family contains four genes. *foxN2-2* and *foxN2-1* were expressed ubiquitously, with the latter one was also being expressed in the SNC. *foxN2-3* was not detected by ISH. *foxNt* was expressed in the brain branches.

foxO was expressed ubiquitously, but not in the SNC neither the pharynx.

foxP is ubiquitously expressed throughout the body of a whole worm (254).
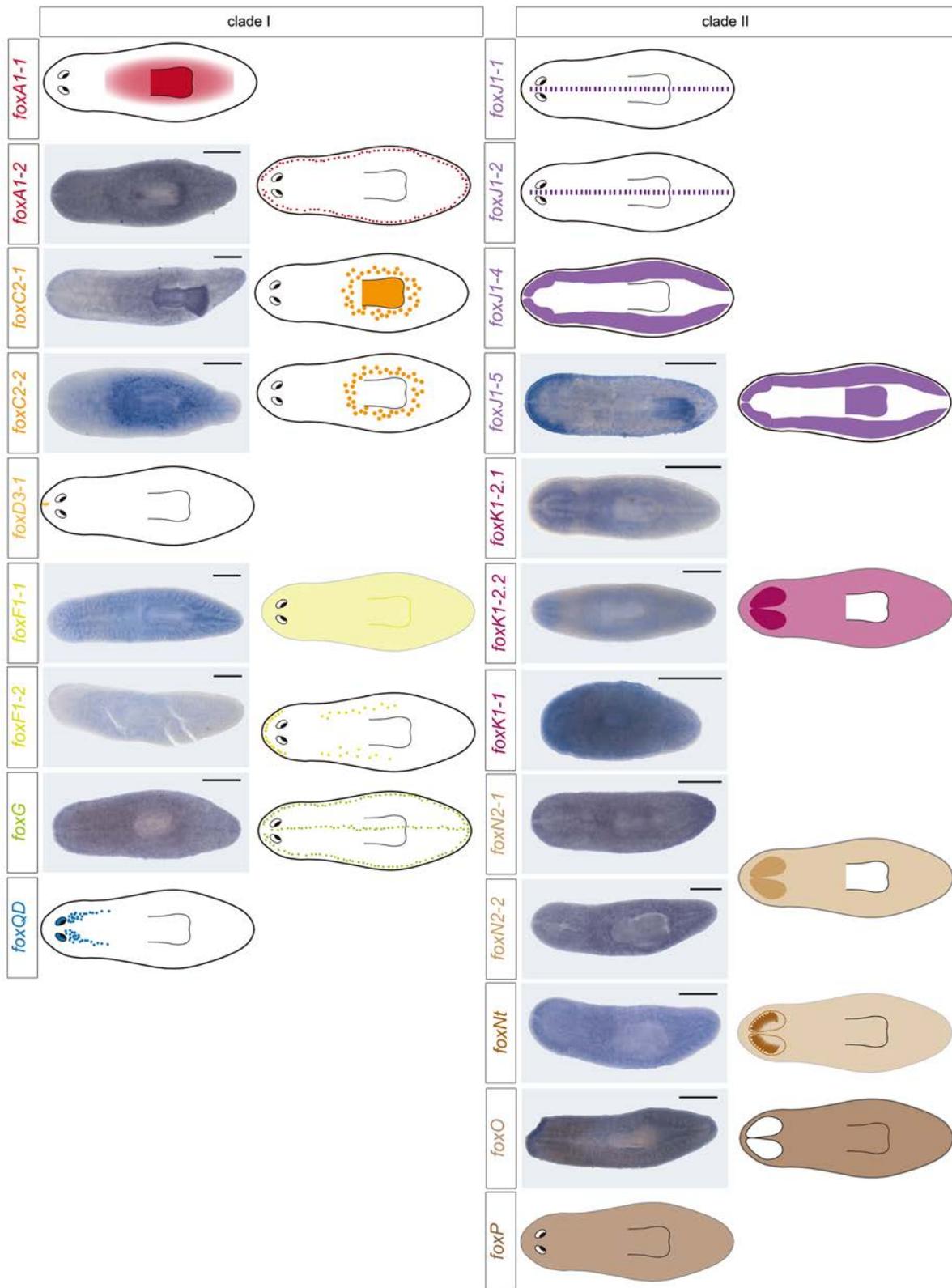
Figure R5.11: *fox* genes present different expression patterns. WISH of new *fox* genes in intact animals. For the new and the previously described genes a schematic cartoon showing where the genes is expressed, is added. Gene names are located laterally of each image. Scale bar: 250 µm.

The expression of some *fox* genes was not possible to be detected, although we designed at least two riboprobes

6.7. Role of *Smed fox* genes during regeneration

Fox families have been related with regeneration and in fact they control some specific aspects of tissue regeneration and/or turn over (263). We decided to use planarian to study whether some *fox* genes could be involved in the regeneration process. In the previous chapter we described for the first time the *foxG* and *foxK* function during regeneration processes. We choose another family which was never reported in planarian: FoxN. The reason to do so is that this family has been described to act downstream of Wnt5 in other species (264) and *wnt5* function has a well described function in regulating the mediolateral (ML) axis in planarian (150,155,160).

In the case that FoxN regulates the ML axis in planarian downstream of Wnt5, this should coexpress with *ror* (*wnt5* receptor) or *roboC* (*slit* receptor). We sought to investigate if *foxN* genes were expressed in the same neuron cells. We used SCS data to validate this hypothesis. 25% of the neurons expressed a *foxN* gene and a receptor (Figure R5.12), suggesting that they could play a role in the ML axis specification. We performed functional experiments with the three *foxN* genes, and double knockdown with *foxN2-1* and *foxNt*, since they were the two paralogs showing higher colocalization with the receptors in neurons (Figure R5.12).
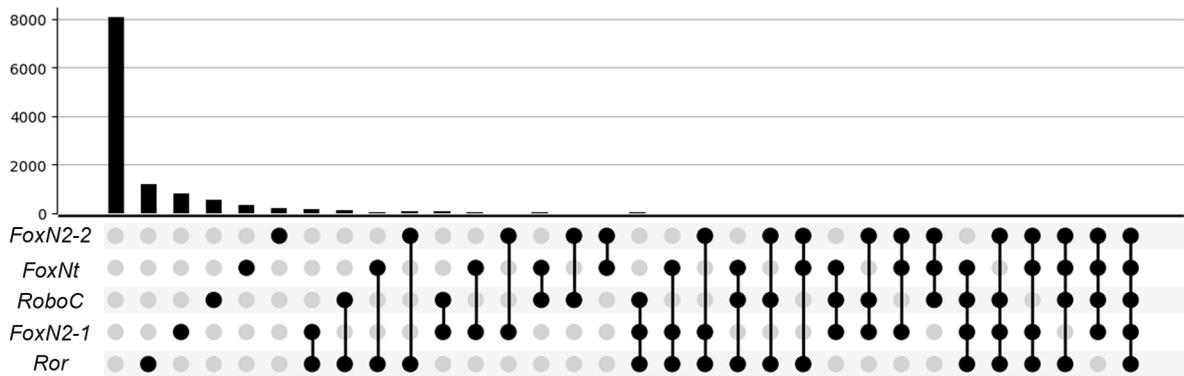
Figure R5.12: *Semd-foxN* **genes are coexpressed with** *Smed-ror* **and** *Smed-roboC.* Upset plot intersection showing the presence of each gene and the combination of them, in neuronal cells.

After two rounds of inhibition and amputation, most of the animals showed eye defects and anterior delay regeneration compared to controls (Figure R5.13). Since the inhibition of any element of the ML axis generates mistargeting of axons, we sought to perform an immunostaining, using synapsin (3C11) and arrestin (VC1) to label brain and eyes, respectively. RNAi animals showed reduced size brains and optic chiasm misconnections (Figure R5.14). These results suggest that FoxN members could play a role during nervous system regeneration and ML axis identity development.
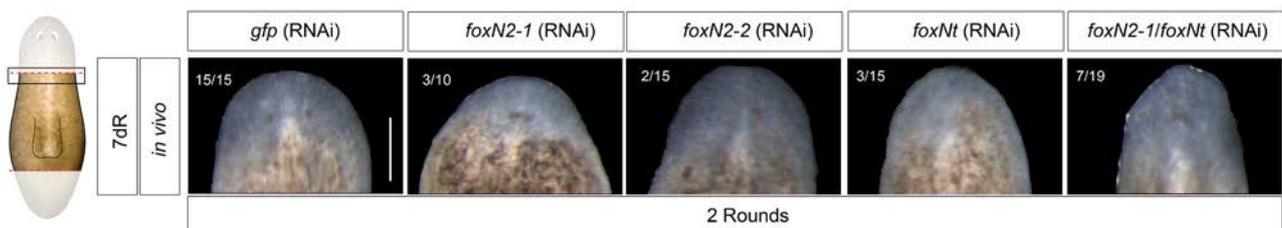
Figure R5.13: Inhibition of *foxN* **genes generate animal with bad anterior regeneration.** *in vivo* images of planarian showing anterior (trunks) phenotype of control and *foxN* (RNAi) genes. Anterior blastemas seem smaller and some of them present bad eye formation or not present eyes. Scale bar: 250 μm
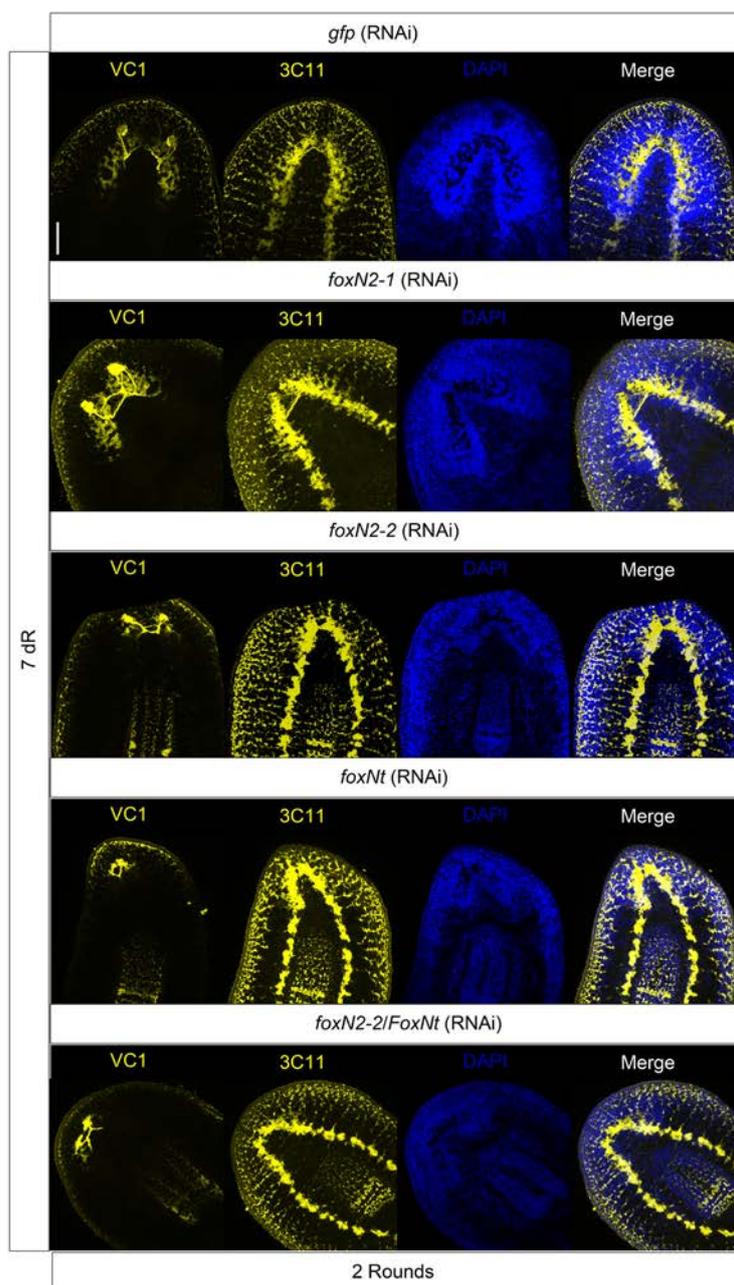
Figure R5.14: *foxN* (RNAi) animals show small brains and mistargeting of the optic chiasm. anti-VC1 immunostaining images of control and *foxN* (RNAi) animals show bad eyes formation and mistargeting of optic chiasm connections. anti-3C11 immunostaining images of control and *foxN* (RNAi) animals show smaller brains than controls. Nuclei are stained with DAPI. Illustration, indicating where arrestin (VC1) and synapsin (3C11) are detected, is shown. The amputation level and the area analyzed. Scale bar for the entire panel is 100 μm.

In this results chapter, we have identified 27 *fox* genes in *Smed*. The phylogenetic analysis including Lopochotrozoan and Platyhelminthes species, allowed the classification of 13 families and their designation. We have analyzed them at genomic and sequence level confirming their diversity. This diversity was also confirmed by ISH, since *fox* genes display tissue specificity. Finally, we functionally analyzed *foxN* genes, which were found to be related with brain regeneration.

DISCUSSION

DISCUSSION

This thesis comprises three subprojects: 1) identification of *bls* as a novel gene family that controls cell number by balancing cell proliferation and cell death; 2) study planarian epigenome during regeneration and identify new elements that regulates posterior organizer formation and function; and 3) characterization of the transcription factor family Fox in planarian. Here, the three subprojects will be discussed independently, and at the end a final and integrative discussion will be presented.

7. Chapter I - Planarian size depends on *Blitzschnell*, a novel gene family that controls cell number by balancing cell proliferation and cell death.

Cell number and cell size regulation has been studied during embryonic development and in tissue renewal in adulthood. Particularly, cell number regulation is based the cell proliferation and cell death modulation. Planarians became an extraordinary tool to understand how those mechanisms regulate their body size, because 1) cell number mainly drive body size and 2) cell proliferation and cell death are well described and are technically easy to study.

7.1. *bls* is a *de novo* gene family taxonomically restricted to the order Tricladida (planarians)

In this chapter, we have identified a new gene family, *blitzschnell* (*bls*), which appears to be an evolutionary novelty of Triclads (planarians), and is essential for the control of cell number in response to nutrient intake. In *S. mediterranea*, *bls* family is composed by 15 members, grouped in five subfamilies (*bls1- 5*). Members of *bls1* and *bls4* subfamilies are pseudogenes, while members of *bls2*, *bls3* and *bls5* subfamilies encode for short peptides that contain a signal peptide (SP) and a coiled coil domain (CC). FISH analysis with specific riboprobes, demonstrates that *bls2*, *3* and *5* are all expressed in a subset of secretory cells, seeming tissue specific. Furthermore, we have only been able to find homologs of *bls* in species of the Tricladida order. Although the genomic databases of Platyhelminthes are incomplete, *bls* family appears to be Taxonomically restricted (265). All described features: gene duplication and presence of pseudogenes (266), short open reading frame with a signal peptide and a ISD (267–270), being expressed in specific cell types (275–277), and being Taxonomically restricted (274), are shared by genes that originated *de novo* during evolution. *de novo* genes, previously known as orphan genes (275), could originate from an existing gene in the genome (276), from non codifying genomic regions (275), or from transposon domestication (277). Although further phylogenetic studies are required to understand the origin of *bls* family, our data favours the last two possibilities, since we could not find any homolog in species outside Tricladida, and we found transposable elements in the same genomic region where *bls* family is found.

7.2. *bls* is required to restrict cell number during planarian starvation, and is downregulated in response to nutrient intake to enable increases in cell number and body size

Our results demonstrate that *bls2/3/5* attenuates proliferation and triggers apoptosis in all scenarios analysed. In homeostatic animals, the imbalance in the mitotis:apoptosis ratio pro-

duced by *bls2/3/5* inhibition led to an increase in cell number. Strikingly, this increase in cell number resulted in normal body size but smaller cell size in starved planarians, and in larger body size and normal cell size in fed animals (Figure D1.1). We observed down-regulation of *bls2*, *bls3*, and *bls5* a few hours after food ingestion. In other organisms *de novo* genes have been shown to play an important role in the response to biotic and/or abiotic stresses (273,278,279). Our results suggest that the appearance of *bls* in Tricladida during evolution may be linked to the requirement for continuous modulation of cell number in response to nutrient availability in these organisms. According to this hypothesis, *bls* expression is required to restrict cell number (and maintain cell size) in starvation conditions, but is down-regulated after nutrient intake to allow for increases in cell number and body size (Figure D1.1). The increase in the number of copies of *bls* family members and their tandem disposition suggest that they may be regulated by the same promoter, facilitating rapid regulation of their protein levels according to cell energy status.
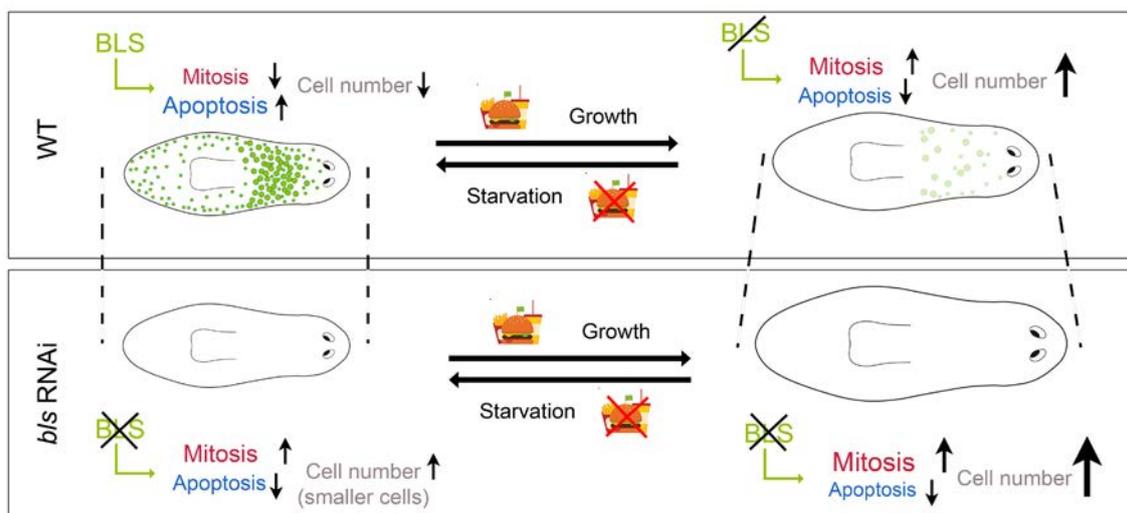
Figure D1.1: Model of BLS function in controlling cell number. In starving conditions, *bls* is expressed and limits cell number and body size by attenuating the mitotic rate and promoting apoptosis. After feeding, *bls* is down-regulated allowing the increase in cell number. In starved *bls* (RNAi) animals, the mitotic/apoptotic rate increases as does total cell number. However, cells cannot maintain their size and body size does not increase. In fed *bls* (RNAi) animals, the mitotic/apoptotic rate is even higher than in controls and cell number increases, as does body size, since cells maintain their normal size.

Because *bls* genes share 70–100% of identity at the nucleotide level, we were unable to inhibit specific copies using RNAi, and were therefore unable to determine which gene copies perform the described function. For this reason, in this study we have described this function to "*bls2/3/5*". However, because all *bls* genes appear to follow the same expression dynamics and share almost identical amino acid sequences, we hypothesize that the gene copies encoding the SP and CC domains may perform the same function. This is in agreement with the aforementioned hypothesis of simultaneous regulation enabling rapid changes in expression. Nonetheless, we cannot rule out the possibility that copies that do not encode the CC domain may act as inhibitors.

Because *bls2/3/5* appears to function as a sensor of cell energy status, it may interact with members of the insulin/Akt/mTOR pathways, which regulate growth in all organisms in response to nutrient intake. It has been described that *de novo* and TRG lack catalytic domains and normally interact with proteins in conserved networks (280). The presence of a SP suggests that *bls2*, *bls3*, and *bls5* may be secreted and interact with components of those conserved pathways. TORC-1 is down-regulated during starvation and its inhibition in planarians

decreases proliferation without affecting cell death. mTOR is up-regulated in response to food intake in planarians, and its inhibition decreases proliferation and increases cell death, impeding growth (68,281). mTOR hyper-activation, through *PTEN* or *smg-1* (RNAi), does not give rise to larger animals but promotes over-proliferation and outgrowths (225,226). Thus, *bls2*, *bls3*, and *bls5* may inhibit the mTOR signalling pathway.
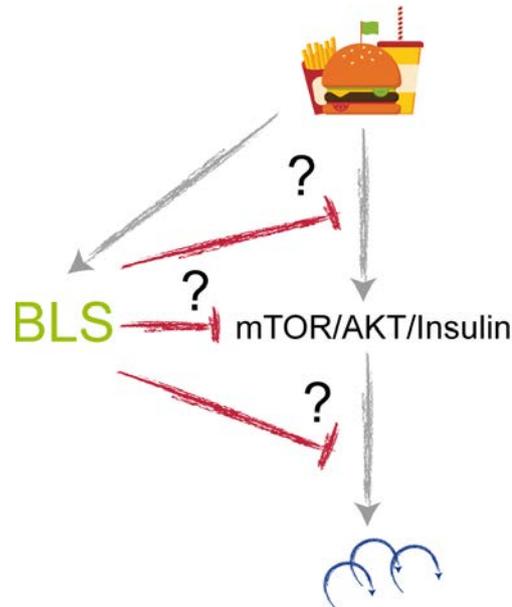
Figure D1.2: Model of mTOR/Akt/Insulin regulation by BLS. BLS could be sensing the nutrient intake and as a consequence inhibiting mTOR/Akt/Insulin pathway at different levels. The lack of this regulation would decrease the number of cycling cells.

One possible explanation for the inability of starved animals to maintain cell size is that *bls2/3/5* silencing in these conditions may promote entry into M phase before cells reach their proper size. It is possible that in wild-type planarians cell cycle length varies according to nutritional status. Recent data suggest the existence of crosstalk between cell division and mitochondrial dynamics and metabolic pathways (282). For example, yeast grown in nutrient-poor conditions adjust their cell-cycle duration to accommodate slower growth, so that the size at which cells divide is similar to that observed in nutrient-rich environments (283). It is possible that the duration of the cell cycle is longer in starved than fed animals, thereby ensuring that daughter cells reach the appropriate size. Promoting entry into M phase after *bls2/3/5* silencing could give rise to smaller cells in starved but not in fed animals. Since a mechanism through which mTOR signals regulate cell size is by controlling cell cycle (284), and as described above planarian mTOR activity is regulated by food intake, interaction of *bls2/3/5* with this pathway could account for the smaller size of cells in starved animals (Figure D1.2)

7.3. *bls* is a tumour suppressor, inhibition of which favours regeneration

As observed in tumoral processes, the hyperplasia promoted after sustained inhibition of *bls2/3/5* in starved animals leads to the formation of overgrowths (Figure R1.14). *bls* thus acts as a tumour suppressor during planarian degrowth, acting as a break for proliferation. This observation presents us with a paradox: although caloric restriction extends lifespan (285), in *bls2/3/5* (RNAi) planarians food deprivation promotes hyperplasia and the formation of overgrowths, while fed *bls2/3/5* (RNAi) animals only increase body size with no apparent

changes in patterning. A second key observation is that while *bls2/3/5* inhibition in starved animals promotes overgrowths, it favours regeneration after any kind of injury. This is consistent with the view that tumour suppressors evolved not to suppress tumour growth but to control cellular processes such as proliferation, cell death, and cell differentiation, which are essential during embryogenesis and are activated during regeneration of complex tissues (286). Perturbation of tumour suppressor function can enhance the regeneration of somatic stem cells in the hematopoietic system or endocrine cells (287). Furthermore, inhibition of the Hippo pathway and consequent YAP/TAZ activation results in increases in organ size and promotes tumour formation in adult mice, but also promotes regeneration of the liver, gut, muscle, and heart in mouse models (288).

Silencing of several known vertebrate tumour suppressors including mTOR, p53, and Hippo, also induces the formation of overgrowths in planarians, but despite increasing proliferation does not promote proper regeneration. While TOR hyper-activation results in larger blastemas, these remain undifferentiated (226). Hippo hyper-activation also enhances the wound response and promotes expansion of the epidermal and muscle cell populations and regeneration of larger structures such as the eyes (74). However, this new tissue is not properly patterned (75). *bls* is the first gene described whose inhibition promotes faster but apparently normal regeneration. One possible explanation is that *bls* specifically controls cell number (through regulation of the cell proliferation:cell death ratio) but not cell differentiation, as described for other signalling pathways such as Hippo (75). In this scenario, an increase in the number of cells during early stages of regeneration could accelerate the expression of wound-induced genes (138), as we observed for *pitx*, and thereby promote more rapid appearance of regenerated structures.

7.4. *bls* family represents an evolutionary strategy to increase planarian fitness in changing environments

In most animal species the adult stage is distinguished from the embryonic stage by the maintenance of body size, cell number, and proportions. However, long-lived animals such as planarians continuously regulate body size in adulthood by controlling cell number according to nutrient availability. Thus, the mechanisms described for other organisms, such as *Drosophila*, in which tissues "know" their final size, may not apply to planarians (289). Given that nutrient availability always fluctuates in nature, the *bls* family may represent an example of *de novo* genes that evolved in planarians to fulfil the requirement for continuous regulation of cell number according to nutrient availability. Other *de novo* genes have been implicated in increasing the fitness of the organism (275,290). Examples are described in cnidarians, in which *Hym301* regulates tentacle number (291), and in molluscs, in which each species expresses a unique set of secreted proteins that drives shell diversity (292). *de novo* genes are usually integrated into existing pathways, adding additional levels of regulation. *bls* genes may interact with members of the insulin/Akt/mTOR signalling pathways, which regulate growth in response to nutrient intake in planarians and in vertebrates (Figure D1.2). RNAi of components of these pathways does not fully phenocopy *bls2/3/5* (RNAi). However, this signalling pathway should be thought of as a network in which each of these signals functions in a complex and dynamic manner, as opposed to a linear pathway. Future studies will need to determine whether the primary function of *bls* is to control proliferation, apoptosis, or both, and to elucidate the molecular integration of *bls* within the insulin/Akt/mTOR network. Given that the molecular signals controlling body and organ growth are also key players in most human cancers, understanding the mechanism by which *bls* genes act as tumour suppressors would help identify novel targets for the design of therapeutic strategies to modulate tissue growth.

8. Chapter II - A posterior wound induces a posterior organizer formation, evocating tissue surround

Organizers were described for the first time by Spemann and Mangold (29). During the last century many studies were published describing organizer features in embryos of other model organisms as: birds, chicken, zebrafish or mouse (293,293). Anderson defines organizers in the context of an embryo as "group of cells that harbour the ability to instruct fates and morphogenesis in surrounding cells, steering their development into specific organs and tissues" (295). This definition is mainly useful for all vertebrates' embryos. However, it can also be applied to the regenerating planarian tips. In this thesis, we focused on the study of the posterior organizer, covering aspects ranging fom its formation until interpreting genetic changes induced in the surrounding tissue. The reason to study the posterior and not the anterior organizer is because posterior regeneration shows less complexity at tissue level. During anterior regeneration, the formation of new eyes, chemoreceptors, brain and other organs involves the appearance of additional signalling centres (163, 189,296) that would compliacte the interpretation of the results

8.1. *wnt1* inhibition changes the genetic profile during regeneration

Regeneration implies huge changes in gene expression. Nowadays, transcriptomic approaches such as microarrays (297) or RNA sequencing (RNA-seq) (298) allow us to studying those changes. The RNAseq technique has been applied to study different regenerating time points, allowing to identify genes involved in different steps the of regeneration process, e.g. in *Drosophila* imaginal disc (208), *Nematostella* (299,300), *Hydra* (301), in limb bud of axolotls (302,303) zebrafish (207) and also planarians (95). Our RNA-seq data reveals that two transcriptomic profiles are activated during posterior regeneration. An early one from 0 to 24 hR and a late one from 36 to 72 hR. These data agrees with two previous studies in planarian, from Kao et al. (304) and Wurtzel et al. (91), which also observed these two phases of transcriptomic changes during P regeneration.

A goal of this thesis was to indentify genes related with posterior identity. This is why we also studied how the inhibition of *wnt1* affects genetic profiles. Our results demonstrate that the genetic profile of *wnt1* (RNAi) animals at 0 hR and 72 hR is similar to the one of controls at 0 hR (just after amputation). This result indicates that in *wnt1* (RNAi) animals the transcriptomic changes required for P regeneration do not occur. This suggests that during regeneration *wnt1* is required to achive P transcriptomic profile.

As observed in other organisms, the canonical WNT pathway regulates directly (target genes) or indirectly (many genes) as compiled in (306,307). In planarian, we found that after *wnt1* inhibition, thousands of genes were up- and down-regulated. In the down-regulated genes group, we found genes known to be expressed in the posterior part or related with posterior identity, such as: *wnt11-1*, *wnt11-2*, *tsh*, *fz4*, *hox4b*, *post2d*, *lox5a* or *post2c*. Some of them were previously described by other groups (150,154,155), confirming that our inhibiting strategy was powerful enough to inhibit *wnt1* and detect target genes. Interestingly, studying GO terms of specificly up and down-regulated genes reveal some specific terms related to metabolism, such as: glutamine biosynthesis, glycolysis or the Krebs cycle. cWNT pathway has related with metabolic regulation (308). In hepatocytes, the cWNT regulates glutamine metabolism genes (309). TCF/LEF binding sites (βcatenin cofactor) have been found in promoters of metabolic genes related with the carbohydrate and glutamine metabolism (310).

Lipid mobilization in adipose tissue have been also related with cWNT (311). These features could also be useful in planarians not just to obtain energy to regenerate as it has been proposed by Plass et al. (99), but also to grow and store energy (186). Additionally, our results also indicate that *wnt1* is controlling posterior pole genes (organizer).

8.2. Cis-Regulatory Elements (CRE) dynamics during planarian posterior regeneration

During the period that this thesis was completed, the planarian genome quality increased its reliability since it was better assembled (reducing the number of scaffolds) and better annotated (89). This aspect leads us to board new questions about the chromatin landscape and the epigenome during planarian regeneration. Part of this thesis was focused on the study of organizers and the genetic changes that they produce in the surrounding tissue. Thus, the question where: Do A and P planarian wounds show specific epigenomes? Are there specific A or P active enhancers? Are those related with regeneration? In order to solve these questions, we performed ATAC-seq (312) and ChIP-seq (313) at A and P blastemas at different regenerating time points. This strategy allowed us to: 1) describe regions differentially open in A and P blastemas and classify them according to their relative increment, being emerging or increasing regions; 2) identify which of those regions were active enhancers specifically open during A or P regeneration; and 3) determine which transcription factors could bind to those enhancers.

Once we analyzed the chromatin landscape during A and P regeneration, we could compare it with the chromatin landscape after *wnt1* and *notum* inhibition at 12 hR, when an identity decision was taken. Importantly, notum knockdown animals showed a drastic change, since they presented a P chromatin profile during A regeneration. This result confirms its master role as an A to identity specifying factor in planarians (156). Although, *wnt1* (RNAi) animals also presented changes during P regeneration, they were not as severe as in *notum* (RNAi). The reason could be that the lack of organizers in a tailless phenotype does not imply a complete change in the chromatin landscape

In this thesis, we have used for the first time ATAC-seq and ChIP-seq to describe the genome landscape during planarian regeneration. Analyzing this landscape in *notum* and *wnt1* (RNAi) reveals that chromatin changes are happening globally in a tissue context. Particularly, *notum* seems to exert a crucial role in the chromatin defining A identity. Overall, these results suggest that organizers are very powerful and can define planarian identities.

8.3. Organizers and tissue competence

The organizer acts as a force defining identities. During recent years many studies started to describe the genetic profile of organizers. Anderson identified genes expressed in different chicken organizers as: Hense's node, notochord and ventral tube, and wing bud (295). The laboratory of Brigitte Galliot also tried to decipher the molecular profile of the *Hydra* organizer (301). In *Nematostella* some studies tried to identify genes related to organizer activity (299,300). And recently, Atekin and collaborators describe the regenerating-organizing cells (ROC) in *Xenopus laevis* (314). In the last months, planarian community also tried to identify the genetic profile of P and A organizer (234). They described genes specifically expressed in the posterior pole (*wnt1* + and *collagen*+ cells). In the top expressed genes. there were previously described *wnt1* regulators, such as *pitx* (105,106) and *islet* (163). Nevertheless, other secreted molecules important for patterning regulation were also present. Some exam-

ples are: BMP (62,63), SLIT (155,160) or FGF(126). Interestingly, the genetic profile of the planarian P organizer does not differ from other organizers previously described, suggesting that organizers are not different among them as it has been proposed by Aztekin et al. and Martinez Arias et al. (314,315).

Then, even tought the organizers are not different among each other, what is it that makes the difference to instruct the tissue? Martinez Arias (315) reviews that one important aspect of the function of organizers relies on the competence of the surrounding tissue to be instructed. Whether the tissue is no already prepared to receive the information, then organizer function would not be accomplished. Following those terms, Waddington suggests that organizers rather "evocate" than "induce" (293) (Figure D2.1). This idea could be confirmed with our data, since one of the most regulated genes was *fz4*, which is the putative receptor of the cWNT in planarian. *axinB* was also down-regulated and could be another example, since it belongs to cWNT and its inhibition produced an over activation of cWNT and as a consequence two tail planarian regeneration (316). To properly investigate the organizer induction idea, we could try to identify TCF binding sites in promoter and enhancer regions of those genes.

In case a feature of an organizer would evocate the surrounding tissue, the first step of the organizer function would be to prepare the tissue surrounding it to make it competent to itself (Figure D2.1). This idea is along the lines with the observation that organizers are transient structures that need to be functional in a certain time of the developmental or regenerative process. This idea is also linked with the fact that early developmental stages should have a broad competence allowing to respond to the signals of an organizer. However, as development progresses, tissue starts to be more competence restricted. In planarian, after postpharyngeally amputation both injured facing tissues will be similar, and during the first stage of regeneration they will present the same genetic profile. In that moment, tissues are plastic and could regenerate a head or a tail, as it has been demonstrated inhibiting genes of the cWNT pathway (154–156,316). As regeneration takes place, A and P organizer are well defined and tissue is less competent to change its fate. This is why we also propose that in planarians, the decision to regenerate a head or a tail (gain the identity) is taken during the first stages of regeneration, when organizers are active and tissue is more plastic.
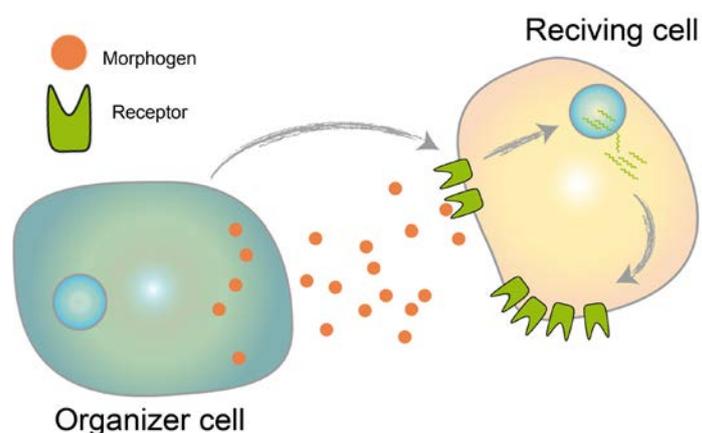
Figure D2.1: Organizer evocates surrounding tissue. Ligands secreted from the organizers could induce the expression of its own receptors in receiving cells.making them more sensitive to the signal itself.

8.4. *pitx* and *foxG* regulate *wnt1* expression

One of the objectives of this thesis was to identify new transcription factors able to identity posterior organizer activity. We could determine that some enhancers were non-accessible after *wnt1* inhibition, indicating that in normal wild type conditions they could be open. After performing motif discovery in those enhancers, we noted that one of the most present were forkhead and homeobox domains, suggesting that Fox and Homeobox transcription factors can bind in those regions. In other organisms, it has been reported that cWNT ligands presented homeobox domains in their proximal enhancers, as pax in fruit flies (317,318), zebrafish (319) and mice (320), and otx in fruit flies (321). This result suggests that at least one homeobox gene could be directly regulating *wnt1* expression in planarians.

After screening of genes, we could determine that at least one member of each family has a relation with *wnt1* expression, and as a consequence with organizer formation and activity. *pitx* (Homeobox family) and *foxG* (Fox family) regulates the second *wnt1* wave expression SC dependent). Additionally, *foxG* also regulates *wnt1* early expression at the wounds (SC independent), which leads to new questions, such as: when early *wnt1* expression is affected, is the second one also affected? Is *foxG* participating in both *wnt1* regulations? Interestingly, after the inhibition of *foxG* and *pitx*, we could show that all cWNT target genes were reduced or absent indicating the strength of the phenotype.

PITX is a transcription factor which belongs to the subfamily of bicoid class (Homeodomain) (322). As described in *C. elegans* PITX and OTX can bind redundantly to the *bicoid* domain (323), which would explain the huge presence of the otx domain in our data. Yasuoka et al. demonstrated that OTX could regulate a huge battery of genes in the head organizer in *Xenopus laevis* (324), as it might have happened with the posterior organizer (*wnt1*+) in planarians (Figure D2.2A). Moreover, in this regulation Lim1 was also present (*islet* in planarian); both genes were specifically up-regulated in the anterior planarian organizer. *islet* also regulates second *wnt1* expression after amputation (SC dependent) (163,230). Interestingly *islet* is also expressed in a subset of neuronal cells (163).
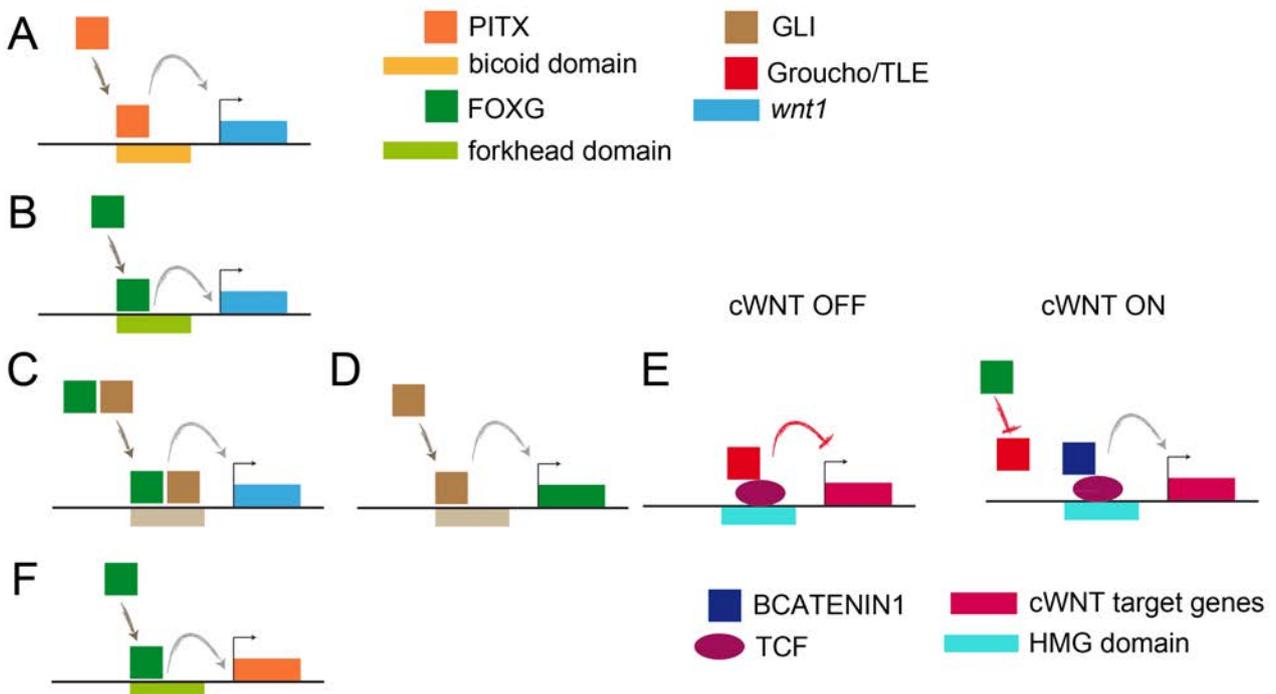
We reported that the transcription factor (FOXG) regulates the expression of *wnt1*. This could be mediated in different manners: 1) *wnt1* presents its promoter and enhancers binding sites for FOXG. 2) FOXG interacts with another factor that regulates *wnt1* expression; and 3) FOXG regulates other *wnt1* TF regulators. To further investigate first the hypothesis, we could identify whether *wnt1* promoters or enhancers present forkhead binding site (Figure D2.2B). Related to the second explanation, we could suggest different scenarios. *foxG* could be related with the Hh pathway, since in planarians this signalling pathway regulates the first *wnt1* expression (178,179). Then, FOXG could act as a cofactor of GLI (TF of the Hh pathway) (Figure D2.2C), or alternativeli it could be a target of Hh (Figure D2.2D). Another possible scenario related to point two is the putative groucho/TLE (*gro*) inhibition by *foxG*, as it has been demonstrated in zebrafish, *amphioxous* (325), and *Xenopus* (326,327). GROUCHO/TLE is a family of TFa factor that co-represses different pathways: cWNT, Notch, TFGβ and EGF (328–333). When cWNT pathway is OFF, GROUCHO/TLE binds to TCF (co-transcription factor of βCAT) acting as a repressor (Figure D2.2E). When the cWNT pathway is ON, βCAT displaces the union between GROUCHO/TLE and TCF, and binds to TCF (341). In planarian, FOXG would be acting as a suppressor of GROUCHO/TLE, and as a cWNT activator (Figure D2.2E). This regulation could take place in *wnt1* expressing cells where FOXG could be regulating its expression, or in cWNT receiving cells acting as a βCAT1 cofactor. Finally, linked to the third explanation, FOXG could be regulating other factors that regulate *wnt1* expression (Figure D2.2F). In order to advance in this idea we could look for homeodomains in promoters or enchancers of previously *wnt1* regulators as such *pitx* or *islet*.

OTX (PITX) could also bind GROUCHO/TLE to avoid some specific gene expression (324), suggesting a main GROUCHO/TLE role in posterior identity regulation. As demostrated in the thesis of Elliot, planarian present two *groucho/TLE* homologs (*gro1* and *gro2*), both of which conserve two interaction domains (334). Both homologs were broadly expressed. *gro1* was also specifically expressed in the brain (Figure D2.3A). *gro2* was also expressed in the margin of the cells and in the midline, resembling *foxG* expression (Figure D2.3A). SCS data confirms that *gro2* is expressed in different cell types, including progenitors and differentiated cells of muscle cells (Figure D2.3B). Moreover, we could confirm that *gro2* and *foxG* highly coexpress in muscle cells *in silico* (Figure D2.3C). Unfortunately, functional analysis has not yet been carried out in planarian. A future line of research could focus on the analysis of P regeneration and *wnt1* expression after *gro2* inhibition.

Altogether, we confirmed the role of *pitx* regulating *wnt1* and also added *foxG* as a new TF that participates in this complex GRN regulating *wnt1* expression.

Figure D2.2: Working model for the PITX, FOXG and Groucho/TLE regulating *wnt1* expression. (A) PITX binds to bicoid domain, in *wnt1* promoter or enhancer, actively regulating its expression. (B) FOXG binds to forkhead domain located in a promoter or enhancer of *wnt1*, regulating its expression. (C) FOXG acts as a cofactor of GLI binding in a *wnt1* promoter or enhancer, regulating its expression. (D) GLI binds to a promoter or enhancer of *foxG*, and modulates its expression. (E) When cWNT is OFF. Groucho/TLE is binding TCF repressing cWNT target gene expression. When is ON, βCATENIN shifts Groucho/TLE and binds with TCF allowing cWNT target genes expression. FOXG would be also inhibition Groucho/TLE in order to favour cWNT target gene expression. (F) FOXG might bind to a forkhead domain located in a promoter or enhancer of *pitx*, modulating its expression.
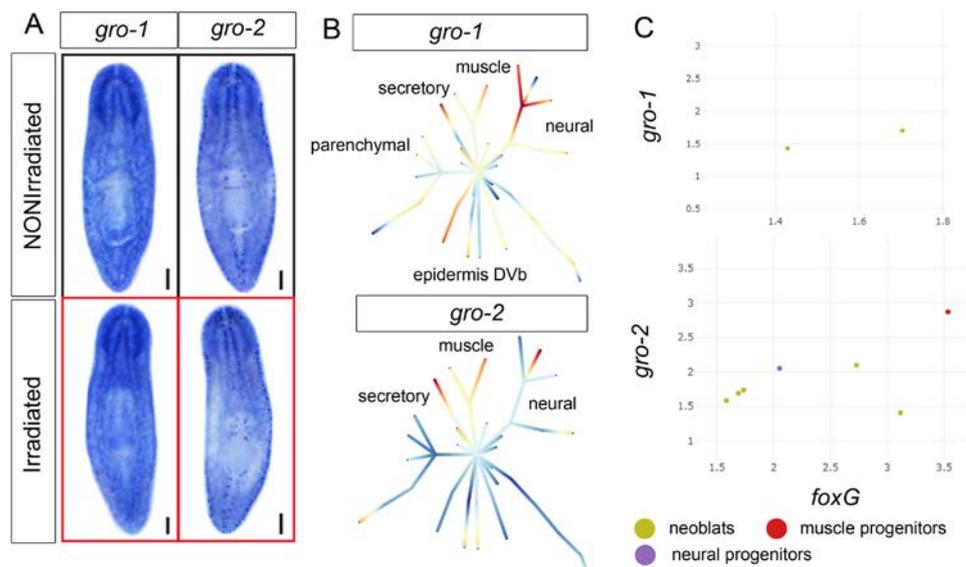
Figure D2.3: *gro-1* and *gro-2* coexpress with *foxG* in muscle cells. (A) WISH of *gro-1* and *gro-2* in intact animals from (334). **(B)** SCS data reveals that both genes are broadly expressed in all the planarian cell types, particularly in the muscle and some neural populations. **(C)** *gro-1* coexpress with *foxG* in neoblast; and *gro-2* coexpress with *foxG* in neoblast and neural and muscle progenitors.

8.5. *foxk* regulates neural differentiation and would act as a cofactor of *βcat1* in planarians

Most of times, the cWNT pathway is presented as a continuous linear of events that ends when βCATENIN is able to interact with TCF and mediate the expression of target genes. However, many molecules are able to interact and regulate different steps of it. For instance, it has been demonstrated that βCAT could also interact with other TFs such as FOXO, SOX or OCT4 (335). In the same way, other elements of the pathway are regulated by other molecules. In vertebrates, DVL can interact with FOXK, enabling its nuclearization and stabilizing βCAT, allowing for some cWNT target genes expression (336). We could demonstrate that motifs to allow the interaction between FOXK and DVL are conserved in planarians. Indeed, both genes coexpress in same cells. This is why we could explain that after the inhibition of *foxk1-2.1* some cWNT target genes were down-regulated, these are *fz4*, *hox4b* and *wnt11-1*. Thus, *foxk1-2.1* could interact with the cWNT pathway regulating the effect of the posterior organizer. However, the possible interaction between all three *foxk* genes and the fact that two *foxK* genes also presented a tailless-like phenotype; make it impossible to determine *foxk1-2.1*'s unique role with posterior identity.

Other *foxk* gene functions have reported. In mammals, *foxK* have been related with antiviral gene regulation (337). In nutrient rich environment, it is regulated by the mTOR pathway entering to the nucleus and transcriptionally repressing autophagy genes (338). Moreover, *foxK* has been related with glioma (brain cancer), which throught the cWNT pathway could regulate cell proliferation, cell cycle and apoptosis (339). Regarding *foxK* expression in brains of mammals, it has also been described that *foxK* plays a role regulating ectoderm and mesoderm tissues in *Xenopus* (340,341). This last ectoderm relation could link with the fact that our results show f*oxk1-2.1* as a key element of the nervous system regeneration (Figure D2.4). This hypothesis is also confirmed by unpublished data no published from Dr. Francesc Cebrià's group, demonstrating that strong inhibtion of *foxk1-2.1* produces a reduction in progenitors and differentiated neural cells.

foxk1-2.1 seems to control *wnt1* expression. One possible explanation is that *foxK* only acts as a cofactor for certain cWNT target genes. A second interpretation would be that the lack

of cWNT target expression acts as a call for the source (*wnt1*), and as a consequence it would increase its expression. As a third explanation, it should be considered that neuronal and muscle cells derived from the same progenitor as it is demonstrated in planarians (99) (Figure D2.4), and *Xenopus* (340,341); and the muscular cells expressing *wnt1* (Figure D2.4). Hence, the lack of neural cells produced by the inhibition of *foxK* would change the balance between the two differential paths increasing muscular cells. Excess of muscular cells would contribute to increment *wnt1* expression, which could be linked to the increment of *post2d*.
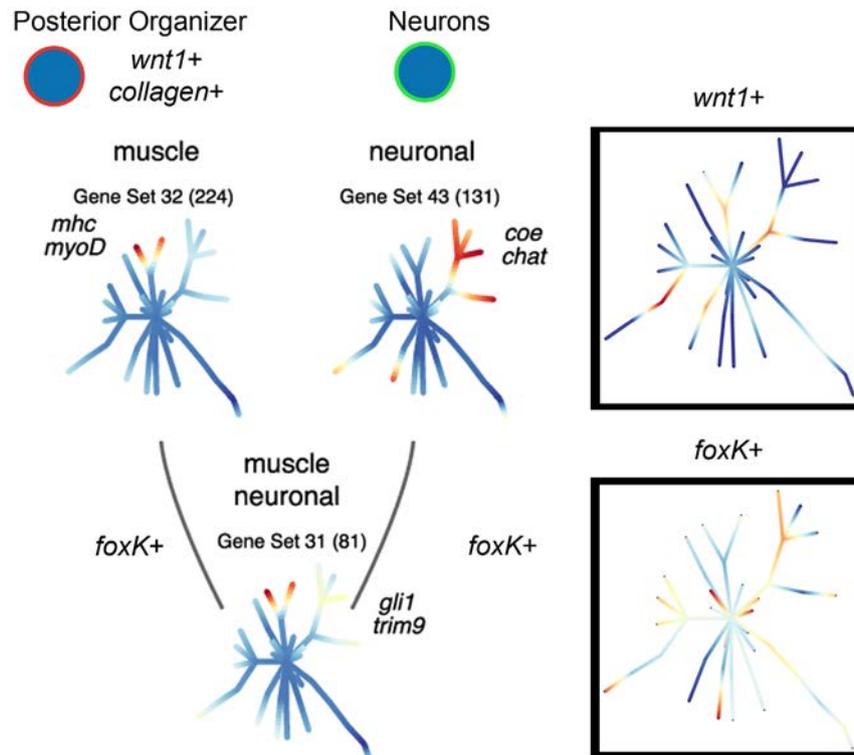
Figure D2.4: Posterior muscle organizer might have a muscle/neural progenitor. Graphical representation of set of genes expression that participates in muscle and neural differentiation. Graphical representation of *wnt1* and *foxk1-2.1* of gene expression; presenting their expression in muscle and neural progenitors.

A final explanation, and not excluding any of the previous ones, is the possibility that FOXK could be related with the Notch pathway. This particular pathway regulates cell-cell interaction (342). Notch is a receptor located in the membrane, when it interacts with its ligand (Delta); cellular cleavage results in intercellular Notch (NICD) formation. Then, NICD can move to the nucleus, bind with Su(H) and promote target gene expression (343). The relation between WNT and Notch pathways in organisms such as: *Drosophila*, sea urchin and vertebrate (343) has been previously demostrated. Notch could negatively regulate βCATENIN, and DVL; or GSK3 (elements of Wnt pathway) which could negatively regulate NICD (344). Moreover, during developmental stages, the WNT pathway is related with stemness behaviour (345), as it could also be related in planarian since its inhibition was misregulated by SC markers. The Notch pathway is related to transient amplifying (TA) cells. Specifically, during mES differentiation Wnt and Notch pathways act to determine neuroectodermal (by Notch) (346,347) or endomesodermal (by Wnt) fate (348–350). FOXK trough DVL could be involved in Notch regulation and as a consequence modulate differentiation states between muscular and neural fate, and modulate *wnt1* expression (as discussed above). Additionally, the inhibition of elements of the Notch pathway (*notch-2*) generates changes in genes expressed in the planarian muscle middle (125), such as *slit*, which is absent, or *wnt1*, which increases its expression.

In planarians, *foxK* seems to have a dual role regulating neural differentiation and cWNT target genes, and these functions could be medited by cWNT and Notch pathways.

8.6. Posterior planarian organizer and mesoectodermal origin

The first organizer described by Spemann and Mangold was only related with new axis formation, but nowadays it is known that it underlies gastrulation, being is the embryo reorganization that allows the appearance of the three germ layers (for triplobastic organisms) (315). Thus, the organizer does not only define an axis but induces neural and mesodermal derivates (351). In planarians, this idea could help to understand why genes specifically expressed in the nervous system such as *hh* (178), *pitx* (105), *islet* (105,163), *foxG* or *foxK* could have a role in the regulating posterior organizer. This idea was previously reported by (163,178) suggesting that the Hh pathway could induce *wnt1* expression. In fact Yazawa et al. (178) demonstrates that at 18 hR *wnt1* is expressed in the surrounding ventral nerve cords. In the same way, during planarian regeneration it has been demonstrated that *wnt1* colocalizes with *pitx*, *islet* and *foxG* (neural markers). However, this colocalization is not reported in homeostatic conditions. In addition, *wnt1* is expressed in different neoblast progenitor, such as: muscle, neural or parenchymal (Figure D2.4). These results suggest that the origin of the posterior organizer is a progenitor that could turn into muscle or neural cell type.

The progenitor is defined by the capacity to proliferate. It was reported that a certain percentage of *islet* cells coexpress *piwi* (neoblast marker), suggesting that a certain *wnt1* precursor could proliferate (163). Nevertheless, this does not imply that *wnt1*+ and *islet*+ cells also express *piwi*. We could demonstrate that at 3 dR some *wnt1*+ cells coexpress *wnt1* with *h2B* (neoblast marker) (Figure D2.5). This fact leads us to think that the organizer is acting as a growth zone (GZ). This behaviour has described in *Tribolium castaneum* to allow posterior elongation (352), and in spider development regulating posterior specification (353). Furthermore, in somitogenesi, GZ induces production of new paraxial mesoderm cells (354). In those last two examples, WNT signalling has been reported for the specification and maintenance of growth-zone cells.
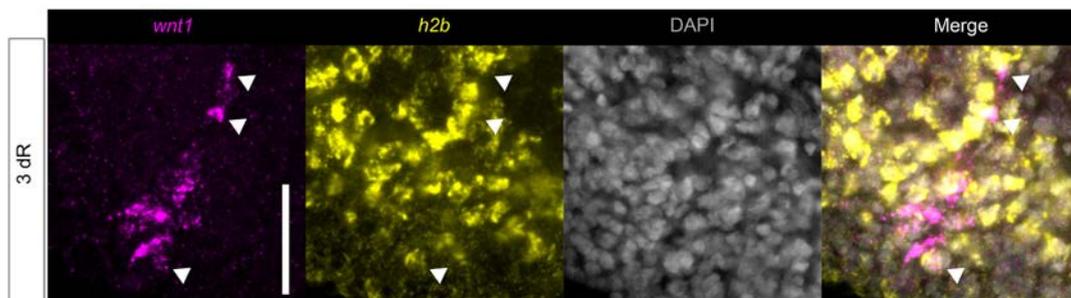
Figure D2.5: *wnt1*+ cells proliferate during regeneration. Double FISH combining *wnt1* and *h2b* riboprobes. Riboprobes colocalize in few cells (white arrows) at 3 dR. Scale bar is 50 µm.

In this thesis, it is suggested that the posterior organizer origin could have neural origin acting as a *infintinum* GZ. And since it is formed and established, it remains in muscle cells which express positional control genes (PCG) in planarians.

8.7. *notum* wound determines the anterior epigenome

The cWNT pathway could be regulated by different inhibitors, the most important one in planarian is *notum* (156), which by WISH is only detected in the anterior pole of intact planarian. After amputation *notum*, as well as *wnt1* are expressed in A- and P- facing wounds. It was proposed that *wnt1* (trough βCAT1) allows wound *notum* expression.

These findings suggest that induction of *wnt1* and *notum* result in a high Wnt-signaling environment at posterior-facing wounds, which will lead a tail formation. In anterior facing wounds, a Wnt-inhibitory (low Wnt) environment will be created promoting head formation (156). Our RNA-seq results reveal that *notum* is also expressed during posterior regeneration (Figure D2.6A). First during the wound response; but as regeneration goes, it is continuously expressed. This findings has so far never been reported by WISH. A personal communication of Dr. Sureda may clarify this situation. He states that he could detect *notum* expression in posterior blastemas at 3 dR in the sister planarian species *Schmidtea polychroa* (Figure D2.6B). This leads to suggest that in one hand WISH in *Smed* is not a sensitive technique to detect those transcriptomic changes.
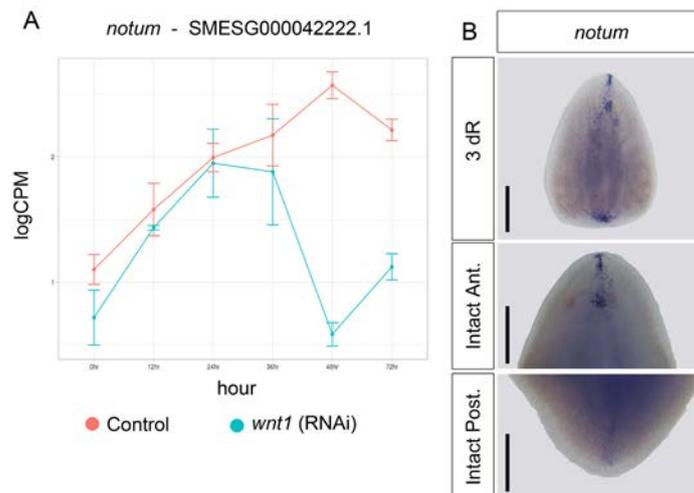
Figure D2.6: *wnt1* **regulates posterior** *notum* **expression during posterior regeneration. (A)** *notum* expression plotted during posterior regeneration in control animals and *wnt1* (RNAi) animals showing a reduction of its expression at late stages of regeneration. Log fold change was represented in X axis. **(B)** WISH of *notum* in *Schmidtea polychroa* in intact animals and regenerating heads at 3 dR. Showing anterior and posterior expression. Scale bar are 100 μm.

And on the other hand, that NOTUM could present post-translational modifications (not detectable by WISH either RNA-seq) that inhibits its function in the posterior (Figure D2.6C). An antibody against NOTUM might help us to determine if our detection of *notum* by RNA-seq correlated with its translation and its extracellular localization in posterior.
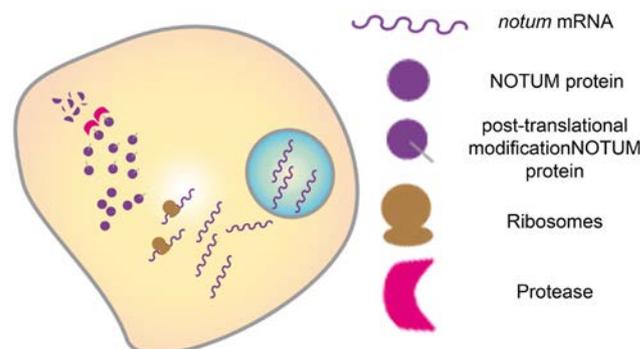
Figure D2.7: Working model for NOTUM post-translational modification. After translation, NOTUM is modified and degradated before its secration.

Interestingly, late *notum* expression in posterior blastemas could be regulated to by *wnt1* (Figure D2.6A), as reported in anterior. However, the function of this second *notum* expression was never reported in planarian. One transcriptomic analysis was performed by Reuter et al. (171) of *notum* (RNAi) anterior-facing wound, but they did not further analyse differentially expressed genes. Understanding the *notum* and *wnt1* relation is crucial to understand when the decision to regenerate a head or a tail is taken. Our study of the epigenome reveals that just after 12 hours post amputation the regenerating blastemas has already take this decision (Figure R3.5B). Indeed, it was after *notum* inhibition that we observe major changes in anterior regenerating blastemas. In the following analysis, we investigate which genes are close to active enhancers that are less accessible in *notum* (RNAi) and we identify which TF could bind there. Additionally to that we also need to elucidate which TF could regulate *notum* expression. Is this behaviour related with wound signalling? Is *notum* truly regulated by cWNT? Or might it be regulated indirectly?

8.8. Cell death and organizer formation

The cell death process related with development and homeostasis is apoptoic, which implies a genetic program that triggers cell death. Cells condensate chromatin and generate apoptotic bodies, which will attract phagocytes. Apoptosis is a tightly regulated event, that can also affect surrounding tissue (355).

Apoptosis has been described as a crucial step at different developmental stages in tissue remodelling and regeneration. After amputation, apoptosis wound response has been reported necessary in different organisms for cell proliferation induction and whole regeneration process triggering (356). *Xenopus* tadpoles needs apoptosis to trigger cell proliferation (357) and in mice liver it plays a crucial role regulating wound healing and regeneration. Recently, apoptosis was reported to be crucial for *Nematostella* regeneration, where the whole cell death program was involved in triggering proliferation and regeneration (300). During imaginal disc regeneration, Serras groups demonstrated that dying cells trigger regeneration in living cells (358,359). In *Hydra*, a similar connection was made where apoptotic cells produce *wnt3* (cWNT gene), which was responsible to trigger cell proliferation in the neighbouring cells (360).

In planarians only few publication tried to relate cell death and *wnt1* expression, or vice versa. In my opinion there are two ways when cell death and *wnt1* could be related: 1) cell death at wound response induces *wnt1* expression (Figure D2.78A). Or 2) *wnt1* regulated cell death triggers regeneration (Figure D2.8A). Even though that in planariansit is impossible to specifically induce cell death, we have different evidence that allow to discuss them. Related with the first hypothesis, ERK signalling is related with differentiation (361), and activated ERK (pERK) rapidly increases after amputation (141). Pharmacologically inhibition of the ERK pathway demonstrates that pERK is controlling apoptosis responses. Interestingly, wound induced and pole expression of *wnt1* and *notum* were also reduced. Indeed, *βcat1* seemed to be activated in an ERK-dependent manner. This might suggest that wound expression of *wnt1* and *notum* (tip organizers) might be regulated by pERK. Related with the second hypothesis, previous results of our group might help to elucidate such a behaviour. We inhibited *wnt1* and investigated cell death affections. After *wnt1* inhibition, it was determined that first and second apoptotic peaks were increased (Figure D2.8B). This suggested that *wnt1* was attenuating cell death response during regeneration. Interestingly, those animals presented less proliferation at 6 hR and an increment at 48 hR (Figure D2.8C).

These could suggest that *wnt1* inhibition produces a slower regeneration kinetic, generating posterior smaller blastemas.

These findings leads to propose that both process seem interconnected during the first stage of regeneration. And Perrez-Carrijo (355) proposed that apoptosis could be a key source of signals required for wound signalling to organize cell formation and tissue regeneration. It could be considered as a first step to rewire genetic connections and reestablish the new cell pattern and tissue identity.
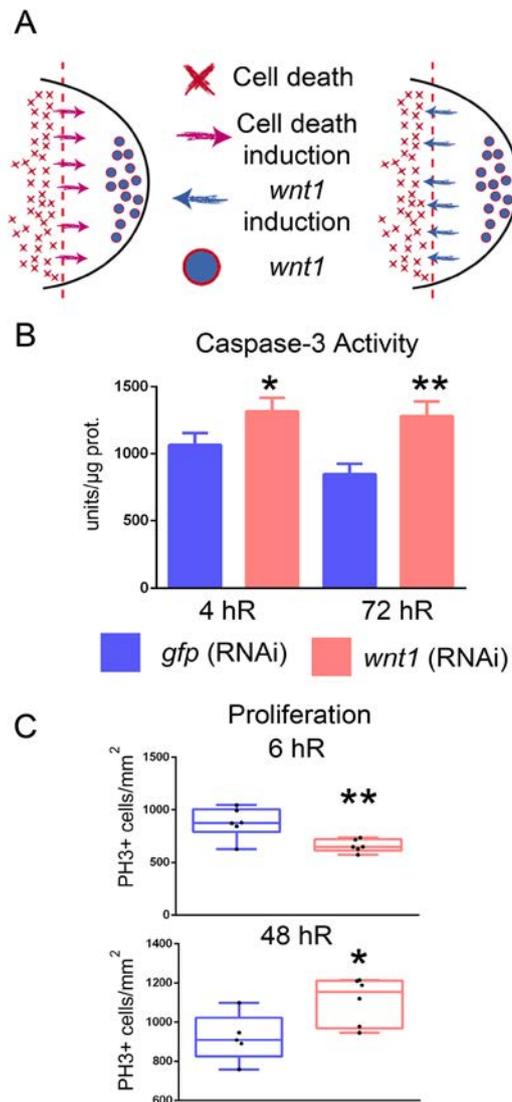
Figure D2.8: **Cell death could control organizer formation. (A)** Two putative models of how cell death induces *wnt1* expression and *wnt1* regulates cell death. **(B)** Quantification of caspase-3 activity in *wnt1* (RNAi) animals and controls (n of controls=3, n of RNAi=3, *$P<0.05$, **$P<0.01$) at 4 hR and 72 hR. **(C)** Quantification of PH3+ cells at after amputation shows misregulation of mitotic cells/mm^2 in *wnt1* (RNAi) animals at 6 hR (n of controls=6, n of RNAi=6, **$P<0.01$) and 48 hR (n of controls=5, n of RNAi=6, *$P<0.05$).

9. Chapter III - Fox family evolution

Transcription factors (TF) regulate patterning events and developmental specification. During metazoan evolution TFs participate in the new evolutionary events as multicellularity and embryogenesis (247). Comparative analysis of genomes from the species at the base of Metazoan clade demonstrate the presence of most of the TF families and the signalling pathways with influence on developmental processes (247,362). Characterizing how those elements evolved, and are present in the lives species of nowadays helps us to understand the mechanisms that underlie developmental processes.

9.1. The loss of Fox in evolution

Previous phylogenetic analyses were performed to understand Fox family relationship in different organisms (259) and within the Lophotrocochoan clade (249). Those studies collected published data and focused their attention on understanding how evolutionary events succeeded. Additionally, they performed some phylogenetic analysis using particular close species. In this thesis, we used genome and transcriptome databases to annotate planarian *fox* genes and phylogenetically analyzed the relation of these between species.

fox genes previously described at the base of the Metazoan clade, these were also present in our study as *Amq* (247) and *Sdo* (248) from Porifera, and *Nvec* (246) from Cnidaria. Additionally, Fox in the Chordata species were identified as *Blan* (363), *Xlae* (340) and *Hsa* (364) were identified during this thesis.

In opitstokon species, the FoxJ1 family has been identified as the last common ancestor of fungi and metazoan. It was proposed that after the Fox family appearance, it grew by gene expansion and duplication (247). Our results also confirm this hypothesis since most of the species presented FoxJ1 family, and Eumetazoan presented a huge variety of gene number, gene families and paralogs within the families (Figure R5.2). Interestingly, some gene loss events were are also present during Fox evolution. The lost of FoxQ2 in *Drosophila* (365) and FoxE family lost in *Spur* (366) has been reported.

Even tought FoxH and FoxI families appear to be restricted to Deutorostomia (365), previous data demonstrate the presence of both families in some Lophocotrozoan species. *Cgi* and *Lotia* present FoxH, and *Cte* presents a FoxI (367). Our data confirms these presences suggesting different independent loses during Bilateria evolution, as it was proposed for the FoxQ1 family in Protostomia clade (249) (Figure R5.2). FoxH has been described to determine mesoderm in *Xenopus* (368) through TFGβ/SMAD regulation; the same function could be occurring in *Cgi* and *Lotia*. FoxI has been reported to regulate endoderm specification (366) in Sea urchin, which could also be in *Cte*.

FoxN is usually subdivided in two main families: FoxN1/4 and FoxN2/3. It has been proposed that FoxN1/4 was first generated and the FoxN2/3 family appeared based on it (247). However, our phylogenetic analysis suggested that the process could have happened the other way around, since in the first analysis, the principal FoxN branch is related with the N2/3 family, and later genes related with FoxN1/4 appear (Figure R5.1). We could suggest that the division between both families is not that clear and it could be just a huge family with a very big interspecific variation inside.

To summarize, we could confirm some predicted previous hypotheses of gain and loss of genes, which give the confirmation that our phylogenetic analysis was well performed. Moreover, it is clear that the Fox family rapidly changes within species shuffling its family composition in the GRn and their development.

9.2. Fox diversification in Platyhelminthes

Our results show that despite the number of *fox* genes in *Smed* is maintained compared to other species, the number of families is reduced. The average family number in Platyhelminthes is around 12 (Figure R5.4), but close species of the Lophochotrozoan clade present higher family number: *Cte*, 24; *Hro*, 29; *Lingulia*, 22; *Lottia*, 29; *Octopus*, 20 or *Crassostea*, 21. Even Ecdyzozoan species present even more families than Platyhelminthes (Figure R5.2). What it is interesting, is that the amount of *fox* genes does not change within Platyhelminthes, neither Lophocotrozoan species (Figure R5.2, 5.4). In flatworms different gene family loss events have been reported, such as: Wnt (369,370) and Hox genes (371-373). This suggests that the reduction of genes is compensated with an increment of paralogs of each present family which would have lead to and explain the increased tissue and function diversification as it has been proposed by different authors (276,374).

In this thesis, we could also confirm that some *fox* genes seem to be hardly classified and some of them were presented in out groups (Figure R5.2). After the second phylogenetic analysis, we could determine their relation with some families and confirm their presence just in the Tricladida order, suggesting that those were taxonomically restricted to the Tricladida order (Figure R5.4). It has been proposed that gene duplication could be the origin of orphan genes, (266) or as we known taxonomically restricted genes (TRG). This gene variability in side families could be explained by the fact that those families were subdued to selective pressure and erased specific functions. Functional analysis should be done to confirm this idea.

Further analysing families in Tricladida, we could clearly determine that some of the families were lost: B, E, L2, Q2, M and N1/4. The loss of FoxB could be explained since it was related with FoxQ2 regulating neural development in humans (245), *Drosophila* (375) and *Clytia hemisphaerica* (256). The lack of FoxQ2 in *Smed* together with the missing of FoxQD with axis formation (112) would explain FoxB absence. FoxE is related with thyroid (376–378) and lents development (379–382) in humans. Planarians lack homolog organs for thyroid and their eye develop without lents (112). In chordates, FoxM is related with cell cycle progression through cyclinA regulation and in adults, it was reported as an oncogene (245). The lack of FoxM in planarian could be justified by the fact that cyclins have been hardly found in planarian, suggesting an independent cell cycle regulation. In Platyhelminthes, none of the checked species presented FoxN1/4 genes. However, they presented at least 4 paralogs of the FoxN2/3 family, which in our analysis seems to be diversified including the presence of *foxNt*. This family expansion could explain the absence of FoxJ2/3. FoxJ2/3 and FoxN families are similar in our Platyhelminthes and Metazoan phylogenetic analysis. This resemblance might suggest the co-option of FoxJ2/3 by FoxN as it has been described to have happened in other Fox families (257), or other TFs (383,384).

In the genome, genes can arrange in a cluster, a group of two or more genes that encode for similar proteins. Analysing different genomes from insects to chordates, two Fox clusters have been identified: FoxD-FoxE and FoxC-FoxF-FoxL1-FoxQ1 (259). Even thought the first cluster is not further investigated; the second demonstrated to be involved in the develop-

ment of the endo-mesodermal tissue. Members of this cluster are expressed and in those tissues and functional analysis demonstrated their relation with endo-mesodermal structures (385). Our analyses do not reveal similar scaffold position of those genes, but not having a linear genome does not allow us to confirm the negative result.

We further characterize the FoxC-FoxF-FoxL1-FoxQ1 cluster since most genes are present in *Smed*. We have already discussed the loss of the FoxQ1 in the Protostomia Clade, and its putative independent loss. FoxL1 has previously been detected in mesoderm tissues in *Drosophila* (386), mouse (387,388), and humans (389). SCS data reveals that *foxL1t* is expressed muscle pharynx cells in *Smed* (Annex III). FoxC is also expressed in mesodermal tissues in different organisms such as *Drosophila* (386,390), zebrafish (391,392), *Xenopus laevis* (393,394), chicken (395) and mouse (396,397). Its function in humans is linked with cardiac muscle, skeletal iris and lymphatic system development (245). *Smed* shows two paralogs expressed in muscle cells (Annex III) around the pharynx and the pharynx itself. Finally, FoxF is also involved in mesodermal tissue in *C. elegans* (398), *Drosophila* (399) or mouse and humans (245). Both FoxF paralogs are expressed in different types of muscle cells. *foxF1-2* regulates non-body wall muscle in *Smed* (123). Overall, we could demonstrate that even thought no gene cluster arrangement of FoxC-FoxF-FoxL1 was demostrated in *Smed*; those genes might conserve their function determining endo-mesodermal tissues in *Schmidtea mediterranea*.

Conclusively, *Smed* has suffered family Fox event, that might be compensated by family diversification. Which could increase the plasticity of the GRN, modifying developmental pathways.

10. General discussion

In this thesis, we have studied different autonomous and non-autonomous mechanisms that regulate different aspects of planarian growth and regeneration. First, we studied a secreted molecule that controls cell number through balancing cell proliferation and cell death ratio. Next we analyzed how *wnt1*, a secreted ligand of the cWNT pathway, could be regulated; and how its diffusion could genetically change cells that are far from the source. We deeply discussed how those secreted and non-secreted changes could be mediated by interactions of cellular components of cWNT and TFs, or how other TFs could cooperate with βCAT1 to modulate cWNT target genes. Finally, we described and classified all *fox* genes in *Schmidtea mediterranea*. Moreover, we discuss why some of them were gained or lost during evolution.

Studying the Fox family and the new *bls* family that we identified in *Schmidtea mediterranea* indicate us that *de novo* gene formation has occurred in the Platyhelminthes clade. Specifically, those genes regulate developmental process that increases their diversity to better adapt to biotic and abiotic changes. Discovering new TRG and their function help us to understand evolution as a not straight forward process. TFs, signalling pathways and GRN are conserved among the species, but every single evolutionary path would have their specific genes that can interact with them. Species will use TRG and GRN to increase their fitness in the environment where they live. In particular, this thesis results could be helpful to decipher new elements of GRN in planarians. We have done an important step to relate TFs and GRN, and the expression of certain genes related to the regeneration and organizer activity.

Nowadays, organizers are well studied and are considered fundamental for developmental processes. However, their essential role in regeneration has not been studied in depth. The study of organizers in regenerative medicine is becoming indispensable since many laboratories all over the world for instance are establishing organoids as a model system. Organoids are useful to model diseases, drug testing, cell therapies and to study organ development (400). SCs have a remarkable ability to self-organize and reproduce in culture forming some homologous structures and broad array of functionality, including muscle contractility, epithelial barrier function, neuronal activity, hepatocyte detoxification, gastric acid secretion and insulin secretion. However, there is still a lack of control to finally produce an organ or proper size organoids with 3D cellular structure (400). In order to increase organoids complexity and produce 3D structures, researches mainly focus in the properties of stem cells and signalling pathways, but we think that the knowledge concerning the organizer must be applied. If we are able to instruct a tissue to grow, differentiated and pattern correctly, then we will be able to transplant it, replacing missing structures such as the skin. Additionally, little is known about organizer activity in regenerative model systems such as spinal cord in zebra fish and mice. In my opinion, organizer activity function should be further investigated in these regenerative scenarios. Finally, this thesis can also be integrated in cancer knowledge. Tumour suppressors and oncogenes are TFs, such as the Fox family (243,245,364). Studying how these TFs evolve helps to understand how cancer evolves and why some genes were selected through evolution.

CONCLUSIONS

CONCLUSIONS

1. We found a new family of genes in *Schmidtea mediterranea* (*Smed-bls*) made up of 15 members, grouped in five subfamilies. Members of *bls2*, *bls3* and *bls5* subfamilies are transcriptionally active, while *bls1* and *bls3* are pseudogenes.

2. *Bls* family genes are de novo genes taxonomically restricted to the Tricladida Order.

3. *Bls* genes are expressed in secretory cells and their function is to control cell number through regulating cell proliferation/cell death ratio. Nutrient intake down-regulates *bls*, enabling the increase in cell number and body size. During starvation periods, *bls* attenuates proliferation, acting as a tumour suppressor. *Bls* could appear in Tricladida during evolution and could respond to their requirement of continuous regulation of cell number in a nutrient-fluctuating environment.

4. RNA-seq transcriptomic analysis of *wnt1* RNAi animals allowed the identification of new genes required for P regeneration at different stages.

5. ATAC-seq analysis reveals that anterior and posterior planarian early regeneration presents two different genomic landscapes, regulated by *notum* and *wnt1*, respectively. Analysis of ATAC-seq data allowed us to the identification of specific transcription factors required for posterior specification: *pitx* and *foxG*.

6. *pitx* and *foxG* regulate *wnt1* expression and are essential for posterior identity especification. Particularly, *pitx* only regulates the stem cell dependent response, and *foxG* is the first gene reported to regulate the early stem cell independent response

7. 27 *fox* genes were found in *Schmidtea mediterranea*, which could be phylogenetically classified in 13 families: A, At, C, D, F, G, L1t, QD, J1, N2/3, Nt, O and P.

8. The Fox TF family has suffered different lost events during evolution. Although, Platyhelminthes clade has lost 10 families, the number of genes has been maintained through diversification of genes of each family, such as the apparence of three *fox* genes just present in Tricladida.

MATERIAL AND METHODS

MATERIAL AND METHODS

While working on this thesis different methodologies have been used. Since some methodologies were used in several projects these are grouped and referred to as general methodologies. Apart from that techniques specific to individual chapters are introduced after the general part relative to the chapter they appear in.

Planarian culture

The planarians used in this study are the asexual clonal strain of *S. mediterranea* BCN-10 biotype and were maintained as previously described (401) in PAM water (232). Animals were fed twice per week with liver, and those used in starvation experiments were starved for 1 week.

Whole-mount *in situ* hybridization (WISH)

Probes were synthesised *in vitro* using SP6, T7 or T3 polymerase and DIG- or FITC- modified (Roche). RNA probes were purified by ethanol precipitation and the addition of 7.5 M ammonium acetate. For colorimetric whole-mount *in situ* hybridization (WISH) animals were sacrificed with 5% N-acetyl-L-cysteine (NAC), fixed with 4% formaldehyde (FA), and permeabilized with Reduction Solution. The fixative and WISH protocol used has been previously described (402). For whole-mount fluorescent *in situ* hybridization (FISH) animals were sacrificed with 7.5% NAC and fixed with 4% FA. FISH was carried out as described previously (96). For double FISH (dFISH) an azide step (150 mM sodium azide for 45 min at room temperature [RT]) was added. Nuclei were stained with DAPI (1:5000; Sigma). For FISH of paraffin sections animals were sacrificed with 2% HCl and fixed with 4% PFA. Paraffin embedding and sectioning were carried out as previously described (43) and slides were de-waxed, re-hydrated; and an antigen retrival step was performed as previously described (59). Sections were hybridized with the corresponding probes for 16 hours and incubated with antibody diluted 1%BSA, for 16 hours. Both steps were carried out in a humidified chamber (43).

Immunohistochemistry

Whole-mount immunohistochemistry was performed as previously described (403). Animals were killed with 2% HCl and fixed with 4% FA. The following antibodies were used in these experiments: anti-synapsin (anti-SYNORF1, 1:50; Developmental Studies Hybridoma Bank, Iowa City, IA, USA) and anti-phospho-histone H3 (Ser10) (D2C8) (pH3) (1:500; Cell Signaling Technology). The secondary antibodies used were Alexa 488-conjugated goat anti-mouse (1:400; Molecular Probes, Waltham, MA, USA) and Alexa 568-conjugated goat anti-rabbit (1:1000; Molecular Probes). Nuclei were stained with DAPI (1:5000). For immunohistochemistry of paraffin sections animals were killed and treated as described previously. Sections were blocked in 1% bovine serum albumin (BSA) in 1X PBS for 1 h at RT and then incubated with primary antibodies diluted in blocking solution (mouse anti-muscle fibre antibody, 6G10, 1:400; Developmental Studies Hybridoma Bank) for 16 h at 4ºC in a humidified chamber. Subsequently, the sections were washed in 1X PBS and incubated with secondary antibodies (anti-mouse Alexa 488-conjugated antibody, 1:400; Molecular Probes) in blocking solution for 3 h at RT in a humidified chamber. Nuclei were stained with DAPI (1:5000; Sigma).

dsRNA synthesis

Double strand RNA (dsRNA) was synthesised by *in vitro* transcription (Roche) as previously described (92). dsRNA (3 × 32.2 nl) was injected into the digestive system of each animal on 3 consecutive days (1 round).

TUNEL assay

For the whole-mount TUNEL assay animals were sacrificed with 10% NAC, fixed with 4% FA, and permeabilized with 1% sodium dodecyl sulfate (SDS) solution. TUNEL assay was carried out as described previously (147) using the ApopTag Red In situ Apoptosis Detection Kit (CHEMICON, S7165). Nuclei were stained with DAPI (1:5000; Sigma). For TUNEL assay on paraffin sections animals were killed and treated as described above. Sections were treated as described previously (147) and after the dewaxing step a proteinase K step was added for permeabilization. Next, we used the ApopTag Red In situ Apoptosis Detection Kit (CHEMICON, S7165). Positive cells were counted in at least 6 representative sagittal sections per animal and the overall mean value was determined.

Caspase-3 activity assay

For each condition protein extraction was performed in 5 planarians. The protein concentration of the cell lysates was measured using BioRad protein reagent. Fluorometric analysis of caspase-3 activity was performed as described previously (405) using 20 mg of protein extract, which was incubated for 2 hours at 37ºC with 20 µM caspase-3 substrate Ac-DEVD-AMC or 2 ml from a stock of 1 mg/ml for a final volume of 150 µl. Using a Fluostar Optima microplate fluorescence reader (BMG Labtech) fluorescence was measured in a luminescence spectrophotometer (Perkin- Elmer LS-50), applying the following settings: excitation, 380 nm; emission, 440 nm. Three technical replicates were analysed per condition.

Imaging

Whole-mount WISH, FISH, and immunohistochemistry images were captured with a ProgRes C3 camera from Jenoptik (Jena, TH, Germany). A Leica MZ16F microscope (Leica Microsystems, Mannhiem, BW, Germany) was used to observe the samples and obtain FISH, immunostaining, and TUNEL images. A Leica TCS SPE confocal microscope (Leica Microsystems, Mannhiem, BW, Germany) was used to obtain confocal images of whole-mount FISH, immunostaining, and TUNEL assays. Representative confocal stacks for each experimental condition are shown.

Statistical analyses

Statistical analyses were performed using GraphPad Prism 6. Two-sided Student's t-tests (α = 0.05) were performed to compare the means of 2 populations. Two-sided Fisher's exact tests were used to compare 2 phenotypic variants between 2 populations. fisher.test (R function) was used to compare more than 2 phenotypic variants between 2 populations.

Statistical data presentation

Results were plotted using GraphPad Prism 6. To compare 2 populations, we used box plots depicting the median, the 25th and 75th percentiles (box), and all included data points (black

dots). Whiskers extend to the largest data point within the 1.5 interquartile range of the upper quartile and to the smallest data point within the 1.5 interquartile lower range of the quartile. To plot data points over time we used XY plots, in which each dot represents the mean and bars represent the standard error. Each dot is connected with the next in an arbitrary manner. To visualize the percentage phenotype in each population we used the Stacked Bars plot in R. Each phenotype is assigned a distinct colour.

Chapter I

Sequence and phylogenetic analyses

A fragment of *Smed-bls3* was identified from Li et al (406). Other members of the families were identified from the genome (89) and amplified using specific primers (Annex IV). The signal peptide was identified with SigalP v5.0 (407) and the coiled-coil domain was characterized using the PRABI tool (Pole Rhone-Alpes de Bioinformatique) avaliable online; https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_lupas.html (408). Sequence identity comparison was carried out using the pairwise alignment tool in Jalview suite v2.11 (409).

To determine which members of each *bls* subfamily were expressed, we mapped the RNA-seq paired reads from adult wild-type animals (75) against assembly 2 of the *S. mediterranea* genome (89) using Bowtie2 (410) v2.3.4, selecting the -end-to-end option. After alignment, we extracted the reads mapping the scaffolds of interest using samtools view (411) v1.9. The final assessment was performed manually using the Integrative Genomics Viewer (412) (IGV v2.4.4) to verify the families with mapped reads.

Sequence comparison against the GenBank database was performed using the NCBI BLAST network server (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Potential orthologs were searched for using tBLASTx where possible to allow a certain level of tolerance in case of a high degree of divergence. The search for orthologs was performed against transcriptomes and genomes from several Platyhelminthes species (Figure R1.4B) (80).

The IQ-tree web server (413) was used to reconstruct the phylogenetic relationships between *Smed-bls* families. The nucleotide or protein sequences were first aligned using the alignment servers in JalView suite (MUSCLE for nucleotides and MAFFT for amino acids). Substitution model selection was performed automatically by the software, the number of bootstrap iterations was set to 1500 and default options were selected for the remaining parameters. The trees were visualized using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) with the default parameters.

RNAi experimental design

The experiments in which regeneration was studied consisted of 2 consecutive rounds of injections and an amputation at the end of each round (Figure R1.8A). In experiments in which planarians were starved animals underwent three or four consecutive rounds of injection, without amputation (Figure R1.12A). In experiments involving fed animals, planarians received dsRNA injections on three non-consecutive days per week and were fed on the 2 intervening days. This process was repeated for three weeks in total. All control animals were injected with dsRNA of green fluorescent protein (GFP). RNAi of subfamilies *bls2*, *bls3*, and *bls5* was carried out using 2 different RNA sequences (Annexe III), both of which produced the same phenotype when injected in regenerating planarians. The sequences correspon-

ded to the full bls3 sequence and to a smaller region showing the greatest similarity between the *bls2*, *bls3*, and *bls5* subfamilies. In the case of the second RNA sequence, inhibition of all members of the transcribed families was demonstrated by qPCR analysis (Figure R1.8C, 12B, 15B)

Quantitative real-time PCR

Total RNA was extracted from a pool of 5 planarians per condition using TRIzol reagent (Invitrogen). cDNA was synthesized as previously described in (67). Expression levels were normalized to that of the housekeeping gene ura4. All experiments were performed using 3 biological and 3 technical replicates for each condition. The design of specific primers corresponding to the 5' region for subfamilies *bls2*, *bls3*, and *bls5* allowed verification of the inhibition of the 3 gene families after RNAi (Annexe I). All primers used in this study are shown in Annex IV.

Feeding experiments

In long term growth experiments involving RNAi, animals were fed twice per week: food was provided in the morning and removed at the end of the day (Figure R1.15A). PAM water (planarian artificial medium) was replenished three times per week. In RNAi experiments, after 2 weeks of injections in starvation conditions animals were fed for 30 minutes (Figure R1.18A**).** Next, food was removed and PAM water replenished. To study gene expression after feeding we analysed planarians that had been starved for 1 week and then fed for 30 minutes. Next, we removed the food and replenished the PAM water (Figure R1.19). Hours post feeding (hpf) were counted from the moment of removal of the last piece of food.

Cell number and cell volume analyses

To quantify total cell number planarian cells were dissociated with trypsin and the nuclei stained with DAPI (97). The cell suspension was transferred to a Neubauer chamber, cells were manually counted on three occasions, and the mean value calculated. Five planarians were analysed per biological replicate, and three replicates were analysed per condition. Mean cell volume (V) was calculated by multiplying mean epidermal cell area (A) by epidermal cell height (H). To quantify the mean epidermal cell area, the prepharyngeal epidermal area was imaged and the number of nuclei per area was quantified. To determine the mean epidermal cell height, the distance between the apical to the basal part of the cell was measured. Measurements were taken in three different regions of the same section and the mean value obtained (Figure R1.13 A-E, 16A-E).

Chapter II

RNAi experimental design

The experiments in which regeneration was studied were divided in two strategies. Injecting, consisted of one or two consecutive rounds of injections (1000 ng/µl) and an amputation at the end of each round (Figure R2.1A, 5A). Socking, consisted of one round of injection at higher concentration (1200 ng/µl), amputation, and right after sock the pieces in dsRNA diluted in PAM water (1000 ng/µl) per three hours in the dark (Figure R2.1B). The soaking step was done in double folded parafilm wax forming a cross. In the centre, a drop of dsRNA is placed and animals are deposited thanks to a brush. Next, animals were removed and place in a

new container where PAM water is replaced two times to remove the remaining dsRNA. The socking step was also performed for control animals and dsRNA green fluorescent protein (GFP) was diluted in PAM water. All control animals were injected with dsRNA of GFP. In double gene-silencing experiments, the total amount of dsRNA injected for each gene and also the total amount of dsRNA injected in each animal was maintained constant by injecting the amount of GFP required

RNA-sequencing, sample preparation

RNA-sequencing samples were obtained after the soaking step, except for 0 hR samples. First animals were displaced in a petri dish with cold 1%HCl diluted in water for 2'. Next, animals were transferred in a new petri dish with cold PBS 1X. Two washedes were performed. Then, animals were transferred in cold RNAlater for 20'. Afterwards, planarians were transfer and amputated in a Peltier Cell with a clean blade, to obtain the blastemas and post-blastemas. Fragments were washed with RNAlater and 50% RNAlater /Trizol. Finally, liquids were removed and 100 µl of Trizol was added. Total mRNA extraction was performed and the final diluted in 20 µl of water. Three biological replicates were used per time point. Each biological replicate was composed by eight pieces. Libraries preparation and sequencing was carried out by Centre Nacional d'Anàlisi Genòmic (CNAG).

RNA-sequencing, analysis

RNA reads were mapped against the planarian genome version S2F2 (98) using the STAR software tool (414). Lowly expressed genes were filtered by removing genes with less than 1 count-per-million (CPM). Two biological replicates were removed due to ineffective Wnt1 inhibition. Differentially expressed genes were detected using the lima-voom pipeline (415), using an FDR cut-off of 0.05 and a log fold change cut-off of ± 0.5. Gene Ontology enrichment analyses were performed with the package TopGO (416), using the "weight01" algorithm and a Fisher statistic cut-off of 0.05. The GO annotations for the planarian genes were obtained from PlanMine (90).

TCseq (417) was used to determine clusters of gene behaviour based on their expression patterns over the analyzed time points. The z-score transformed log fold changes between Wnt1-RNAi samples and control samples were used as input, and the algorithm "cmeans" was selected for computing the cluster scores. Clusters were assigned to genes by introducing a cut-off of 0.7 to the product of the maximum cluster score (as reported by the cmeans algorithm) and the standard deviation of all scores.

Assay for transposase-accessible chromatin sequencing (ATAC-seq)

ATAC-seq samples were obtained after RNAi treatment and amputation or from wild type animals after amputation (Figure R2.5A). Planarian mucous was removed by washing in 2%L-Cystein (pH7) for 2'. Afterwards, animals were transferred in a petri dish with CMFH (2.56mM $NaH_2PO_4x2H_2O$, 14.28mM NaCl, 10.21mM KCl, 9.42mM $NaHCO_3$, 1%BSA, 0.5%Glucose, 15mM HEPES pH 7.3). Post-blastemas and blastemas were obtained to be analyzed. Next, they were transferred in an eppendorf tube to be dissociated using a solution of liberase/ CMFH (1:10) at room temperature for ten minutes. Twenty animals were used per biological replicate. ATAC-sequencing was carried out as first described in (418) and then adapted by (419)

ChIP-mentation

ChIPmentation combines ChIP with library preparation using Tn5 transposase, similar to ATAC-sequencing. ChIP-mentation samples were obtained from wild type animals after amputation. Planarians were placed in Peltier cells to amputate post-blastemas. Obtained pieces were transferred to 1M MgCl2 solution, for 15-30'' rocking, RT. PBS 1X was added to remove salts. Next, blastemas were fixed with formaldhyde 1,85% for 15' rocking, RT. Glycine was added to obtain a final concentration of 0.125M to quench formaldehyde, for 5' at RT, rocking. Then, blastemas were washed 3X with cold PBS1X. Finally PBS excess was removed, and samples were stored at -80°C. 2000 anterior and posterior blastemas were used. Groups of 100 blastemas were done at a time. ChIP-mentation was carried out as described by (420).

ATAC-seq and ChIP-seq analysis

Reads were aligned using bowtie1 using -m 3 -k 1 arguments. Bam reads were filtered using a <=100bp insert size threshold to identify nucleosome free regions (NFR) (421). Bam files were converted to bed and then the coordinates were shifted +4 and -5 positions to overcome the Tn5 cut position. MACS2 were used for peak calling and HOMER for motif discovery. Differential binding analysis was carried out using DiffBind (R function) (422).

Chapter III

Sequence and phylogenetic analyses

For generating the phylogenetic trees, we had to obtain the FOX protein sequences from several different sources. In some of the cases we were able to collect them from the public databases, like in the case of *Hsa* or *Xtr*. For the rest of the organisms a manual annotation was required. If the only resource available was a trancriptome, like in the case of *Tsol*, we used Transdecoder (v5.5.0) to obtain the translated proteins. Using Hammer (v3.1b2) and the Pfam (423) motive for the Forkhead domain, we extracted the Forkhead-containing proteins. For annotation purposes, we grouped the organisms phylogenetically and aligned the proteins obtained using MAFFT with the L-INS-I strategy with the amphioxus FOX set. The *amphioxus* set was selected due to the existence of at least one member of every FOX family described in this cephalochordate. The alignment was cropped to select the Forkhead domain and used as input in the webserver of IQ-TREE with all the parameters left by default. We then named the proteins according to their relationship with *amphioxus* FOX proteins. For *Smed* FOX domains disposition was used (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) to identify FKD, FHA and FOXP coiled coild; and (424) to identify NLS.

The whole set of named proteins acquired by the described methods, was aligned again using MAFFT (425) with the L-INS-i strategy and the aligning Forkhead domain was selected. This alignment was the input used for IQ-TREE (413) to generate the definitive phylogenetic tree. The options use to run the webserver of IQ-TREE were the ones set by default, including the automatic substitution model selector and the ultrafast bootstrap analysis, except for the number of bootstrap alignments (set at 2000) and the single branch test number of replicates (set at 1500). The trees were visualized using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) with the default parameters.

Organism Abbreviation	Source of FOX Proteins
Hsa	Public Domain
Xtr	Public Domain
Bla	Manual Annotation
Spu	Public Domain
Sko	Public Domain
Cca	Manual Annotation
Dme	Manual Annotation
Tca	Manual Annotation
Hro	Manual Annotation
Ili	Manual Annotation
Lgi	Manual Annotation
Mli	Manual Annotation
Nve	Manual Annotation
Obi	Manual Annotation
Pfl	Manual Annotation
Cgi	Public Domain
Sman	Manual Annotation
Smed	Manual Annotation
Tsol	Manual Annotation
Ina	Manual Annotation
Sdo	Public Domain
Amq	Public Domain

RNAi experimental design

The experiments in which regeneration was studied were performing injections in two consecutive rounds (1000 ng/µl) and the amputation was caried out at the end of each round.

REFERENCES

Annexes

1.   Gilbert SF. Developmental Biology 12th edition. 2019.

2.   Müller WA, Müller WA. Model Organisms in Developmental Biology. In: Developmental Biology [Internet]. New York, NY: Springer New York; 1997 [cited 2019 Sep 13]. p. 21–121. Available from: http://link.springer.com/10.1007/978-1-4612-2248-4_3

3.   Jameson JL, DeGroot LJ, De Kretser DM (David M., Giudice L, Grossman A, Melmed S, et al. Endocrinology : adult and pediatric [Internet]. [cited 2019 Sep 10]. Available from: https://www.sciencedirect.com/book/9780323189071/endocrinology-adult-and-pediatric

4.   Galliot B, Ghila L. Cell plasticity in homeostasis and regeneration. Mol Reprod Dev. 2010;77(10):837–55.

5.   Pellettieri J, Alvarado AS. Cell Turnover and Adult Tissue Homeostasis: From Humans to Planarians. Annu Rev Genet [Internet]. 2007;41(1):83–105. Available from: http://www.annualreviews.org/doi/abs/10.1146/annurev.genet.41.110306.130244

6.   Rosenfeld CS. The Epigenome and Developmental Origins of Health and Disease. The Epigenome and Developmental Origins of Health and Disease. 2015. 1–542 p.

7.   Cook SF, Bies RR. Disease Progression Modeling: Key Concepts and Recent Developments [Internet]. Vol. 2, Current Pharmacology Reports. Springer International Publishing; 2016 [cited 2019 Sep 9]. p. 221–30. Available from: http://link.springer.com/10.1007/s40495-016-0066-x

8.   Sims-Lucas S, Good M, Vainio SJ. Editorial: Organogenesis: From development to disease. Frontiers in Cell and Developmental Biology [Internet]. 2017 Sep 20 [cited 2019 Sep 9];5(SEP):85. Available from: http://journal.frontiersin.org/article/10.3389/fcell.2017.00085/full

9.   Miller CJ, Davidson LA. The interplay between cell signalling and mechanics in developmental processes [Internet]. Vol. 14, Nature Reviews Genetics. 2013 [cited 2019 Sep 10]. p. 733–44. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24045690

10.   Wilson MR, Close TW, Trosko JE. Cell population dynamics (apoptosis, mitosis, and cell-cell communication) during disruption of homeostasis. Exp Cell Res [Internet]. 2000 Feb 1 [cited 2019 Sep 8];254(2):257–68. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10640424

11.   Charras G, Sahai E. Physical influences of the extracellular environment on cell migration. Nat Rev Mol Cell Biol [Internet]. 2014 Oct 30 [cited 2014 Oct 30];(Box 1). Available from: http://www.nature.com/doifinder/10.1038/nrm3897

12.   Sluys R, Riutort M. Planarian Diversity and Phylogeny. In 2018. p. 1–56. Available from: http://link.springer.com/10.1007/978-1-4939-7802-1_1

13.   Miyaoka Y, Ebato K, Kato H, Arakawa S, Shimizu S, Miyajima A. Hypertrophy and Unconventional Cell Division of Hepatocytes Underlie Liver Regeneration. Curr Biol [Internet].

2012;22(13):1166–75. Available from: http://dx.doi.org/10.1016/j.cub.2012.05.016

14.   Hariharan IK. Organ Size Control: Lessons from Drosophila. Dev Cell [Internet]. 2015;34(3):255–65. Available from: http://dx.doi.org/10.1016/j.devcel.2015.07.012

15.   Guertin D a, Sabatini DM. Cell Size Control. Encycl Life Sci [Internet]. 2006;1–10. Available from: http://doi.wiley.com/10.1038/npg.els.0003359

16.   Dhanasekaran DN, Reddy EP. JNK-signaling: A multiplexing hub in programmed cell death. Genes Cancer [Internet]. 2017 [cited 2019 Aug 2];8(9). Available from: www.impactjournals.com/Genes&Cancer

17.   Udan RS, Kango-Singh M, Nolo R, Tao C, Halder G. Hippo promotes proliferation arrest and apoptosis in the Salvador/Warts pathway. Nat Cell Biol [Internet]. 2003 Oct 21 [cited 2019 Aug 1];5(10):914–20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14502294

18.   Zeng Q, Hong W. The Emerging Role of the Hippo Pathway in Cell Contact Inhibition, Organ Size Control, and Cancer Development in Mammals [Internet]. Vol. 13, Cancer Cell. 2008 [cited 2019 Aug 1]. p. 188–92. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18328423

19.   Tumaneng K, Schlegelmilch K, Russell RC, Yimlamai D, Basnet H, Mahadevan N, et al. YAP mediates crosstalk between the Hippo and PI(3)K-TOR pathways by suppressing PTEN via miR-29. Nat Cell Biol [Internet]. 2012 Dec 11 [cited 2019 Aug 2];14(12):1322–9. Available from: http://www.nature.com/articles/ncb2615

20.   Willsey HR, Zheng X, Pastor-Pareja J, Willsey AJ, Beachy PA, Xu T. Localized JNK signaling regulates organ size during development. Elife. 2016;5(MARCH2016):1–18.

21.   Saxton RA, Sabatini DM. mTOR Signaling in Growth, Metabolism, and Disease [Internet]. Vol. 168, Cell. Cell Press; 2017 [cited 2019 Aug 1]. p. 960–76. Available from: https://www.sciencedirect.com/science/article/pii/S0092867417301824#fig5

22.   Manning BD, Toker A. AKT/PKB Signaling: Navigating the Network. Cell [Internet]. 2017;169(3):381–405. Available from: http://dx.doi.org/10.1016/j.cell.2017.04.001

23.   Yoon MS. The role of mammalian target of rapamycin (mTOR) in insulin signaling. Nutrients. 2017;9(11).

24.   Turing AM. The chemical basis of morphogenesis. Bull Math Biol. 1990;52(1–2):153–97.

25.   Tabata T, Takei Y. Morphogens, their identification and regulation. Development [Internet]. 2004 Feb 15 [cited 2019 Sep 9];131(4):703–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14757636

26.   Smith JC. Mesoderm-inducing factors in early vertebrate development. EMBO J [Internet]. 1993 Dec [cited 2019 Sep 9];12(12):4463–70. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8223456

27.   Godsave SF, Isaacs H V., Slack JMW. Mesoderm-inducing factors: A small class of molecu-

les. Development [Internet]. 1988 [cited 2019 Sep 9];102(3):555–66. Available from: https://dev.biologists.org/content/102/3/555

28.   Smith JC, Howard JE. Mesoderm-inducing factors and the control of gastrulation. Dev Suppl [Internet]. 1992 [cited 2019 Sep 9];127–36. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1299357

29.   Spemann H, Mangold H. über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. Arch für Mikroskopische Anat und Entwicklungsmechanik [Internet]. 1924 Sep [cited 2019 Aug 28];100(3–4):599–638. Available from: http://link.springer.com/10.1007/BF02108133

30.   Waddington CH. Experiments on the Development of Chick and Duck Embryos, Cultivated in vitro [Internet]. Vol. 221, Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character. Royal Society; [cited 2019 Sep 9]. p. 179–230. Available from: https://www.jstor.org/stable/92208

31.   WADDINGTON CH. Induction by the Primitive Streak and its Derivatives in the Chick. J Exp Biol. 1933;10(1).

32.   Stemple DL. Structure and function of the notochord: An essential organ for chordate development [Internet]. Vol. 132, Development. The Company of Biologists Ltd; 2005 [cited 2019 Sep 9]. p. 2503–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8162859

33.   Rubin L, Saunders JW. Ectodermal-mesodermal interactions in the growth of limb buds in the chick embryo: Constancy and temporal limits of the ectodermal induction. Dev Biol [Internet]. 1972 May 1 [cited 2019 Sep 9];28(1):94–112. Available from: https://www.sciencedirect.com/science/article/pii/0012160672901297

34.   Anderson C, Stern CD. Organizers in Development. In: Current Topics in Developmental Biology [Internet]. 2016 [cited 2019 Sep 8]. p. 435–54. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26969994

35.   Browne EN. The production of new hydranths in Hydra by the insertion of small grafts. J Exp Zool [Internet]. 1909 Aug 1 [cited 2019 Aug 28];7(1):1–23. Available from: http://doi.wiley.com/10.1002/jez.1400070102

36.   Heldin CH, Moustakas A. Signaling receptors for TGF-β family members. Cold Spring Harb Perspect Biol [Internet]. 2016 Aug 1 [cited 2019 Sep 8];8(8):a022053. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27481709

37.   Hogan BLM. Bone morphogenetic proteins: Multifunctional regulators of vertebrate development [Internet]. Vol. 10, Genes and Development. 1996 [cited 2019 Sep 8]. p. 1580–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8682290

38.   Bandyopadhyay A, Tsuji K, Cox K, Harfe BD, Rosen V, Tabin CJ. Genetic analysis of the roles of BMP2, BMP4, and BMP7 in limb patterning and skeletogenesis. PLoS Genet [Internet]. 2006 Dec [cited 2019 Sep 8];2(12):2116–30. Available from: http://www.ncbi.nlm.nih.

gov/pubmed/17194222

39. Thisse B, Thisse C. Formation of the vertebrate embryo: Moving beyond the Spemann organizer [Internet]. Vol. 42, Seminars in Cell and Developmental Biology. 2015 [cited 2019 Sep 3]. p. 94–102. Available from: http://dx.doi.org/10.1016/j.semcdb.2015.05.007

40. De Robertis EM, Wessely O, Oelgeschläger M, Brizuela B, Pera E, Larraín J, et al. Molecular mechanisms of cell-cell signaling by the Spemann-Mangold organizer. Int J Dev Biol. 2001;45(1):189–97.

41. Tanaka EM, Reddien PW. The Cellular Basis for Animal Regeneration [Internet]. Vol. 21, Developmental Cell. 2011 [cited 2019 Sep 3]. p. 172–85. Available from: https://www.cell.com/action/showPdf?pii=S1534-5807%2811%2900250-4

42. Gemberling M, Bailey TJ, Hyde DR, Poss KD. The zebrafish as a model for complex tissue regeneration [Internet]. Vol. 29, Trends in Genetics. NIH Public Access; 2013 [cited 2019 Sep 10]. p. 611–20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23927865

43. Morgan T. Regeneration. 1901.

44. Imokawa Y, Yoshizato K. Expression of Sonic hedgehog gene in regenerating newt limb blastemas recapitulates that in developing limb buds. Proc Natl Acad Sci U S A [Internet]. 1997 Aug 19 [cited 2019 Sep 8];94(17):9159–64. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9256452

45. Bryant S V, Endo T, Gardiner DM. Vertebrate limb regeneration and the origin of limb stem cells. Int J Dev Biol [Internet]. 2002 [cited 2019 Sep 8];46(7):887–96. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12455626

46. Pearl EJ, Barker D, Day RC, Beck CW. Identification of genes associated with regenerative success of Xenopus laevis hindlimbs. BMC Dev Biol [Internet]. 2008 Jun 23 [cited 2019 Sep 8];8(1):66. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18570684

47. Wang Y-H, Beck CW. Distal expression of sprouty (spry) genes during Xenopus laevis limb development and regeneration. Gene Expr Patterns [Internet]. 2014 May [cited 2019 Sep 8];15(1):61–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24823862

48. Bryant DM, Johnson K, DiTommaso T, Tickle T, Couger MB, Payzin-Dogru D, et al. A Tissue-Mapped Axolotl De Novo Transcriptome Enables Identification of Limb Regeneration Factors. Cell Rep [Internet]. 2017 Jan 17 [cited 2019 Sep 8];18(3):762–76. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28099853

49. Hutchins ED, Markov GJ, Eckalbar WL, George RM, King JM, Tokuyama MA, et al. Transcriptomic analysis of tail regeneration in the lizard Anolis carolinensis reveals activation of conserved vertebrate developmental and repair mechanisms. McGregor AP, editor. PLoS One [Internet]. 2014 Aug 20 [cited 2019 Sep 8];9(8):e105004. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25140675

50. Mathew LK, Sengupta S, Franzosa JA, Perry J, La Du J, Andreasen EA, et al. Comparative

expression profiling reveals an essential role for Raldh2 in epimorphic regeneration. J Biol Chem [Internet]. 2009 Nov 27 [cited 2019 Sep 8];284(48):33642–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19801676

51.   Binari LA, Lewis GM, Kucenas S. Perineurial glia require notch signaling during motor nerve development but not regeneration. J Neurosci [Internet]. 2013 Mar 6 [cited 2019 Sep 8];33(10):4241–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23467342

52.   Carlson MRJ, Komine Y, Bryant S V., Gardiner DM. Expression of Hoxb13 and Hoxc10 in developing and regenerating axolotl limbs and tails. Dev Biol [Internet]. 2001 Jan 15 [cited 2019 Sep 8];229(2):396–406. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11150241

53.   Millimaki BB, Sweet EM, Riley BB. Sox2 is required for maintenance and regeneration, but not initial development, of hair cells in the zebrafish inner ear. Dev Biol [Internet]. 2010 Feb 15 [cited 2019 Sep 8];338(2):262–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20025865

54.   Snippert HJ, Clevers H. Tracking adult stem cells. EMBO Rep [Internet]. 2011;12(2):113–22. Available from: http://dx.doi.org/10.1038/embor.2010.216

55.   Gattazzo F, Urciuolo A, Bonaldo P. Extracellular matrix: A dynamic microenvironment for stem cell niche [Internet]. Vol. 1840, Biochimica et Biophysica Acta - General Subjects. 2014 [cited 2019 Sep 8]. p. 2506–19. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24418517

56.   Xin T, Greco V, Myung P. Hardwiring Stem Cell Communication through Tissue Structure [Internet]. Vol. 164, Cell. 2016 [cited 2019 Sep 8]. p. 1212–25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26967287

57.   Saló E. The power of regeneration and the stem-cell kingdom: freshwater planarians (Platyhelminthes). Bioessays [Internet]. 2006 May [cited 2014 Jan 23];28(5):546–59. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16615086

58.   Reddien PW, Sánchez Alvarado A. Fundamentals of planarian regeneration. Annu Rev Cell Dev Biol. 2004;20:725–57.

59.   Sureda-Gómez M, Martín-Durán JM, Adell T. Localization of planarian βCATENIN-1 reveals multiple roles during anterior-posterior regeneration and organogenesis. Development [Internet]. 2016;(October):dev.135152. Available from: http://dev.biologists.org/lookup/doi/10.1242/dev.135152

60.   Sureda-Gómez M, Pascual-Carreras E, Adell T. Posterior Wnts Have Distinct Roles in Specification and Patterning of the Planarian Posterior Region. Int J Mol Sci [Internet]. 2015;16(11):26543–54. Available from: http://www.mdpi.com/1422-0067/16/11/25970/

61.   Iglesias M, Gomez-Skarmeta JL, Saló E, Adell T. Silencing of Smed-betacatenin1 generates radial-like hypercephalized planarians. Development [Internet]. 2008 Apr [cited 2014 Mar 27];135(7):1215–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18287199

References

62.   Molina MD, Neto A, Maeso I, Gómez-Skarmeta JL, Saló E, Cebrià F. Noggin and noggin-like genes control dorsoventral axis regeneration in planarians. Curr Biol [Internet]. 2011 Feb 22 [cited 2014 Jan 23];21(4):300–5. Available from: http://www.sciencedirect.com/science/article/pii/S096098221100039X

63.   Gaviño MA, Reddien PW. A Bmp/Admp regulatory circuit controls maintenance and regeneration of dorsal-ventral polarity in planarians. Curr Biol [Internet]. 2011 Feb 22 [cited 2014 Jan 23];21(4):294–9. Available from: http://www.sciencedirect.com/science/article/pii/S0960982211000406

64.   Emili E, Pallarès ME, Romero R, Cebrià F. Smed-egfr-4 is required for planarian eye regeneration. 2019;15:9–15.

65.   Fraguas S, Barberán S, Cebrià F. EGFR signaling regulates cell proliferation, differentiation and morphogenesis during planarian regeneration and homeostasis. Dev Biol [Internet]. 2011 Jun 1 [cited 2014 Aug 21];354(1):87–101. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21458439

66.   Barberán S, Fraguas S, Cebrià F. The EGFR signaling pathway controls gut progenitor differentiation during planarian regeneration and homeostasis. Dev. 2016;143(12):2089–102.

67.   Almuedo-Castillo M, Crespo X, Seebeck F, Bartscherer K, Salò E, Adell T. JNK controls the onset of mitosis in planarian stem cells and triggers apoptotic cell death required for regeneration and remodeling. PLoS Genet [Internet]. 2014 Jun [cited 2014 Aug 6];10(6):e1004400. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4055413&tool=pmcentrez&rendertype=abstract

68.   Tu KC, Pearson BJ, Sánchez Alvarado A. TORC1 is required to balance cell proliferation and cell death in planarians. Dev Biol [Internet]. 2012;365(2):458–69. Available from: http://dx.doi.org/10.1016/j.ydbio.2012.03.010

69.   González-Estévez C, Felix DA, Smith MD, Paps J, Morley SJ, James V, et al. SMG-1 and mTORC1 act antagonistically to regulate response to injury and growth in planarians. PLoS Genet. 2012;8(3).

70.   Iglesias M, Felix DA, Gutiérrez-Gutiérrez Ó, De Miguel-Bonet M del M, Sahu S, Fernández-Varas B, et al. Downregulation of mTOR Signaling Increases Stem Cell Population Telomere Length during Starvation of Immortal Planarians. Stem Cell Reports [Internet]. 2019;13:1–14. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2213671119302310

71.   Miller CM, Newmark P a. An insulin-like peptide regulates size and adult stem cells in planarians. Int J Dev Biol. 2012;56(1–3):75–82.

72.   Peiris TH, Ramirez D, Barghouth PG, Oviedo NJ. The Akt signaling pathway is required for tissue maintenance and regeneration in planarians. BMC Dev Biol [Internet]. 2016;16(1):7. Available from: http://bmcdevbiol.biomedcentral.com/articles/10.1186/s12861-016-0107-z

73.   Oviedo NJ, Pearson BJ, Levin M, Sánchez Alvarado A. Planarian PTEN homologs regulate

stem cells and regeneration through TOR signaling. Dis Model Mech. 2008;1(2–3):131–43.

74. Lin AYT, Pearson BJ. Yorkie is required to restrict the injury responses in planarians. PLoS Genet. 2017;13(7):1–30.

75. de Sousa N, Rodríguez-Esteban G, Rojo-Laguna JI, Saló E, Adell T. Hippo signaling controls cell cycle and restricts cell plasticity in planarians. PLoS Biol [Internet]. 2018 [cited 2019 Aug 2];16(1). Available from: https://doi.org/10.1371/journal.pbio.2002399

76. Baguna J, Salo E, Auladell C. Regeneration and pattern formation in planarians. III. Evidence that neoblasts are totipotent stem cells and the source of blastema cells. Development. 1989;107(1):77–86.

77. Wagner DE, Wang IE, Reddien PW. Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. Science [Internet]. 2011 May 13 [cited 2014 Jan 23];332(6031):811–6. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3338249&tool=pmcentrez&rendertype=abstract

78. Baguñà J. The planarian neoblast: the rambling history of its origin and some current black boxes. Int J Dev Biol [Internet]. 2012 Jan [cited 2014 Sep 4];56(1–3):19–37. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22252540

79. Newmark P a, Sánchez Alvarado a. Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. Dev Biol [Internet]. 2000 Apr 15 [cited 2014 Aug 19];220(2):142–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10753506

80. Egger B, Lapraz F, Tomiczek B, Müller S, Dessimoz C, Girstmair J, et al. A transcriptomic-phylogenomic analysis of the evolutionary relationships of flatworms. Curr Biol. 2015;25(10):1347–53.

81. Alvarez-Presas M, Baguñà J, Riutort M. Molecular phylogeny of land and freshwater planarians (Tricladida, Platyhelminthes): from freshwater to land and back. Mol Phylogenet Evol [Internet]. 2008 May [cited 2014 Dec 16];47(2):555–68. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18359250

82. Chandebois R. Histogenesis and morphogenesis in planarian regeneration. Monogr Dev Biol [Internet]. 1976 [cited 2019 Sep 9];11:1–182. Available from: http://www.ncbi.nlm.nih.gov/pubmed/775318

83. Zack RS. Microscopic Anatomy of Invertebrates. Volume 7—Annelida. Ann Entomol Soc Am [Internet]. 1994 [cited 2019 Sep 9];87(4):489–489. Available from: https://books.google.es/books/about/Microscopic_Anatomy_of_Invertebrates_Pla.html?id=HZEWAQAAIAAJ&redir_esc=y

84. Conn DB. The biology of flatworms (Platyhelminthes): Parenchyma cells and extracellular matrices. TRANSAMMICROSCSOC [Internet]. 1993 Oct [cited 2019 Sep 9];112(4):241–61. Available from: https://www.jstor.org/stable/3226561?origin=crossref

85. Stunkard HW. The Invertebrates: Platyhelminthes and Rhynchocoela. The Acoelomate Bi-

lateria. Vol. II . Libbie Henrietta Hyman. Q Rev Biol [Internet]. 1951 Dec 22 [cited 2019 Sep 9];26(4):403–403. Available from: https://www.journals.uchicago.edu/doi/10.1086/398483

86.   Cebrià F, Nakazawa M, Mineta K, Ikeo K, Gojobori T, Agata K. Dissecting planarian central nervous system regeneration by the expression of neural-specific genes. Dev Growth Differ [Internet]. 2002 Apr [cited 2018 Jun 18];44(2):135–46. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11940100

87.   Cebrià F. Planarian Body-Wall Muscle: Regeneration and Function beyond a Simple Ske-letal Support. Front Cell Dev Biol [Internet]. 2016;4(February). Available from: http://journal.frontiersin.org/article/10.3389/fcell.2016.00008

88.   Vreys C, Michiels NK. Sperm trading by volume in a hermaphroditic flatworm with mutual penis intromission. Anim Behav [Internet]. 1998 Sep [cited 2019 Sep 9];56(3):777–85. Avai-lable from: http://www.ncbi.nlm.nih.gov/pubmed/9784230

89.   Grohme MA, Schloissnig S, Rozanski A, Pippel M, Young G, Winkler S, et al. The genome of Schmidtea mediterranea highlights the plasticity of cellular core mechanisms. Nat Publ Gr [Internet]. 2018;1–24. Available from: http://dx.doi.org/10.1038/nature25473

90.   Brandl H, Moon HK, Vila-Farré M, Liu SY, Henry I, Rink JC. PlanMine - A mineable resource of planarian biology and biodiversity. Nucleic Acids Res. 2016;44(D1):D764–73.

91.   Wurtzel O, Cote LE, Poirier A, Satija R, Regev A, Reddien PW. A Generic and Cell-Type-Specific Wound Response Precedes Regeneration in Planarians. Dev Cell [In-ternet]. 2015;35(5):632–45. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1534580715007182

92.   Sanchez Alvarado A, Newmark PA. Double-stranded RNA specifically disrupts gene expres-sion during planarian regeneration. Proc Natl Acad Sci [Internet]. 1999 Apr 27 [cited 2014 Aug 27];96(9):5049–54. Available from: http://www.pnas.org/content/96/9/5049.long

93.   Forsthoefel DJ, Ross KG, Newmark PA, Zayas RM. Fixation, Processing, and Immunofluo-rescent Labeling of Whole Mount Planarians. In 2018. p. 353–66. Available from: http://link.springer.com/10.1007/978-1-4939-7802-1_10

94.   Ross KG, Omuro KC, Taylor MR, Munday RK, Hubert A, King RS, et al. Novel monoclonal antibodies to study tissue regeneration in planarians. BMC Dev Biol [Internet]. 2015 Jan [ci-ted 2015 Sep 11];15:2. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4307677&tool=pmcentrez&rendertype=abstract

95.   Pearson BJ, Eisenhoffer GT, Gurley KA, Rink JC, Miller DE, Sánchez Alvarado A. Formal-dehyde-based whole-mount in situ hybridization method for planarians. Dev Dyn [Internet]. 2009 Feb [cited 2014 Jul 23];238(2):443–50. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2640425&tool=pmcentrez&rendertype=abstract

96.   King RS, Newmark PA. In situ hybridization protocol for enhanced detection of gene expres-sion in the planarian Schmidtea mediterranea. BMC Dev Biol [Internet]. 2013;13(1):8. Availa-

ble from: http://bmcdevbiol.biomedcentral.com/articles/10.1186/1471-213X-13-8

97. Moritz S, Stöckle F, Ortmeier C, Schmitz H, Rodríguez-Esteban G, Key G, et al. Heteroge-
neity of planarian stem cells in the S/G2/M phase. Int J Dev Biol [Internet]. 2012 Jan [cited
2014 Feb 4];56(1–3):117–25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22450999

98. Fincher CT, Wurtzel O, de Hoog T, Kravarik KM, Reddien PW. Cell type transcriptome atlas
for the planarian *Schmidtea mediterranea*. Science (80-) [Internet]. 2018;5:eaaq1736. Avai-
lable from: http://www.sciencemag.org/lookup/doi/10.1126/science.aaq1736

99. Plass M, Solana J, Wolf FA, Ayoub S, Misios A, Glažar P, et al. Cell type atlas and line-
age tree of a whole complex animal by single-cell transcriptomics. Science (80-) [Inter-
net]. 2018;1723(April):eaaq1723. Available from: http://science.sciencemag.org/content/
early/2018/04/18/science.aaq1723.abstract%0Ahttp://www.sciencemag.org/lookup/
doi/10.1126/science.aaq1723

100. Fraguas S, Barberán S, Ibarra B, Stöger L, Cebrià F. Regeneration of neuronal cell types
in Schmidtea mediterranea: an immunohistochemical and expression study. Int J Dev Biol
[Internet]. 2012 Jan [cited 2014 Dec 2];56(1–3):143–53. Available from: http://www.ncbi.nlm.
nih.gov/pubmed/22451002

101. Cebrià F. Regenerating the central nervous system: How easy for planarians! [Internet]. Vol.
217, Development Genes and Evolution. 2007 [cited 2019 Sep 8]. p. 733–48. Available from:
http://www.ncbi.nlm.nih.gov/pubmed/17999079

102. Roberts-Galbraith RH, Brubacher JL, Newmark PA. A functional genomics screen in planari-
ans reveals regulators of whole-brain regeneration. Elife. 2016;5(September2016):1–31.

103. Wang IE, Lapan SW, Scimone ML, Clandinin TR, Reddien PW. Hedgehog signaling regula-
tes gene expression in planarian glia. Elife [Internet]. 2016 Sep 9 [cited 2019 Sep 8];5(Sep-
tember2016). Available from: https://elifesciences.org/articles/16996

104. Cebrià F. Organization of the nervous system in the model planarian Schmidtea mediter-
ranea: an immunocytochemical study. Neurosci Res [Internet]. 2008 Aug [cited 2014 Dec
2];61(4):375–84. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18499291

105. März M, Seebeck F, Bartscherer K. A Pitx transcription factor controls the establishment and
maintenance of the serotonergic lineage in planarians. Development [Internet]. 2013 Nov
[cited 2014 Feb 28];140(22):4499–509. Available from: http://www.ncbi.nlm.nih.gov/pub-
med/24131630

106. Currie KW, Pearson BJ. Transcription factors lhx1/5-1 and pitx are required for the mainte-
nance and regeneration of serotonergic neurons in planarians. Development [Internet]. 2013
Sep [cited 2014 Aug 4];140(17):3577–88. Available from: http://www.ncbi.nlm.nih.gov/pub-
med/23903188

107. Nishimura K, Kitamura Y, Inoue T, Umesono Y, Yoshimoto K, Takeuchi K, et al. Identification
and distribution of tryptophan hydroxylase (TPH)-positive neurons in the planarian Dugesia

japonica. Neurosci Res [Internet]. 2007 Sep [cited 2019 Sep 8];59(1):101–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17624455

108. Nishimura K, Kitamura Y, Inoue T, Umesono Y, Sano S, Yoshimoto K, et al. Reconstruction of dopaminergic neural network and locomotion function in planarian regenerates. Dev Neurobiol [Internet]. 2007 Jul [cited 2019 Sep 8];67(8):1059–78. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17565705

109. Cebrià F, Kudome T, Nakazawa M, Mineta K, Ikeo K, Gojobori T, et al. The expression of neural-specific genes reveals the structural and molecular complexity of the planarian central nervous system. Mech Dev [Internet]. 2002 Aug [cited 2019 Sep 8];116(1–2):199–204. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12128224

110. Nishimura K, Kitamura Y, Inoue T, Umesono Y, Yoshimoto K, Taniguchi T, et al. Characterization of tyramine β-hydroxylase in planarian Dugesia japonica: Cloning and expression. Neurochem Int [Internet]. 2008 Dec [cited 2019 Sep 8];53(6–8):184–92. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18926867

111. Nishimura K, Kitamura Y, Umesono Y, Takeuchi K, Takata K, Taniguchi T, et al. Identification of glutamic acid decarboxylase gene and distribution of GABAergic nervous system in the planarian Dugesia japonica. Neuroscience [Internet]. 2008 Jun 2 [cited 2019 Sep 8];153(4):1103–14. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18440152

112. Lapan SW, Reddien PW. Transcriptome analysis of the planarian eye identifies ovo as a specific regulator of eye regeneration. Cell Rep [Internet]. 2012 Aug 30 [cited 2014 Aug 21];2(2):294–307. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3785364&tool=pmcentrez&rendertype=abstract

113. Pineda D, Rossi L, Batistoni R, Salvetti A, Marsal M, Gremigni V, et al. The genetic network of prototypic planarian eye regeneration is Pax6 independent. Development [Internet]. 2002 Mar [cited 2019 Sep 8];129(6):1423–34. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11880351

114. Rompolas P, Patel-King RS, King SM. An outer arm dynein conformational switch is required for metachronal synchrony of motile cilia in planaria. Mol Biol Cell [Internet]. 2010 Nov 1 [cited 2019 Sep 8];21(21):3669–79. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20844081

115. Wurtzel O, Oderberg IM, Reddien PW. Planarian Epidermal Stem Cells Respond to Positional Cues to Promote Cell-Type Diversity. Dev Cell [Internet]. 2017;40(5):491-504.e5. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1534580717300771

116. Kobayashi C, Kobayashi S, Orii H, Watanabe K, Agata K. Identification of Two Distinct Muscles in the Planarian Dugesia japonica by their Expression of Myosin Heavy Chain Genes. Zoolog Sci [Internet]. 1998 Dec 1 [cited 2019 Sep 8];15(6):861–9. Available from: http://www.bioone.org/doi/abs/10.2108/zsj.15.861

117. Bueno D, Baguñà J, Romero R. Cell-, tissue-, and position-specific monoclonal antibodi-

es against the planarian Dugesia (Girardia) tigrina. Histochem Cell Biol [Internet]. 1997 Feb 12 [cited 2019 Sep 8];107(2):139–49. Available from: http://link.springer.com/10.1007/s004180050098

118. Bowen ID, Ryder TA, Thompson JA. The fine structure of the planarianPolycelis tenuis Iijima. Protoplasma [Internet]. 1974 Mar [cited 2019 Sep 8];79(1–2):1–17. Available from: http://link.springer.com/10.1007/BF02055779

119. Forsthoefel DJ, Cejda NI, Khan UW, Newmark PA, Disease H. Cell-type diversity and regionalized gene expression in the planarian intestine revealed by laser-capture microdissection transcriptome profiling. bioRxiv. 2019;

120. Scimone ML, Srivastava M, Bell GW, Reddien PW. A regulatory program for excretory system regeneration in planarians. Development [Internet]. 2011 Oct [cited 2014 Aug 8];138(20):4387–98. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3177309&tool=pmcentrez&rendertype=abstract

121. Thi-Kim Vu H, Rink JC, McKinney S a, McClain M, Lakshmanaperumal N, Alexander R, et al. Stem cells and fluid flow drive cyst formation in an invertebrate excretory organ. Elife [Internet]. 2015;4:1–4. Available from: http://elifesciences.org/lookup/doi/10.7554/eLife.07405

122. Rink JC, Vu HT-K, Sánchez Alvarado A. The maintenance and regeneration of the planarian excretory system are regulated by EGFR signaling. Development [Internet]. 2011 Sep [cited 2014 Aug 8];138(17):3769–80. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3152929&tool=pmcentrez&rendertype=abstract

123. Scimone ML, Wurtzel O, Malecek K, Fincher CT, Oderberg IM, Kravarik KM, et al. foxF-1 Controls Specification of Non-body Wall Muscle and Phagocytic Cells in Planarians. Curr Biol [Internet]. 2018 [cited 2018 Nov 26];28(23):3787-3801.e6. Available from: https://doi.org/10.1016/j.cub.2018.10.030

124. Cebrià F, Bueno D, Reigada S, Romero R. Intercalary muscle cell renewal in planarian pharynx. Dev Genes Evol [Internet]. 1999 Apr [cited 2019 Sep 8];209(4):249–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10079368

125. Witchley JN, Mayer M, Wagner DE, Owen JH, Reddien PW. Muscle cells provide instructions for planarian regeneration. Cell Rep [Internet]. 2013 Aug 29 [cited 2014 Feb 26];4(4):633–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23954785

126. Scimone ML, Cote LE, Rogers T, Reddien PW. Two FGFRL-Wnt circuits organize the planarian anteroposterior axis. Elife [Internet]. 2016 Apr 11 [cited 2016 Apr 13];5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27063937

127. Cote LE, Simental E, Reddien PW. Muscle functions as a connective tissue and source of extracellular matrix in planarians. Nat Commun [Internet]. 2019 [cited 2019 Apr 8];10(1). Available from: https://doi.org/10.1038/s41467-019-09539-6

128. Morgan TH. Experimental studies of the regeneration of Planaria maculata. Arch für Entwic-

kelungsmechanik der Org [Internet]. 1898 Oct [cited 2019 Sep 9];7(2–3):364–97. Available from: http://link.springer.com/10.1007/BF02161491

129. Coward SJ. Chromatoid bodies in somatic cells of the planarian: Observations on their behavior during mitosis. Anat Rec [Internet]. 1974 Nov [cited 2019 Sep 9];180(3):533–45. Available from: http://www.ncbi.nlm.nih.gov/pubmed/4371241

130. Reddien PW, Oviedo NJ, Jennings JR, Jenkin JC, Sánchez Alvarado A. SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. Science [Internet]. 2005 Nov 25 [cited 2014 Jan 23];310(5752):1327–30. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16311336

131. Solana J, Kao D, Mihaylova Y, Jaber-Hijazi F, Malla S, Wilson R, et al. Defining the molecular profile of planarian pluripotent stem cells using a combinatorial RNAseq, RNA interference and irradiation approach. Genome Biol [Internet]. 2012 [cited 2019 Sep 4];13(3):R19. Available from: http://genomebiology.com/content/pdf/gb-2012-13-3-r19.pdf

132. Wenemoser D, Reddien PW. Planarian regeneration involves distinct stem cell responses to wounds and tissue absence. Dev Biol [Internet]. 2010 Aug 15 [cited 2014 Jul 19];344(2):979–91. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2950745&tool=pmcentrez&rendertype=abstract

133. Zeng A, Li H, Guo L, Gao X, McKinney S, Wang Y, et al. Prospectively Isolated Tetraspanin + Neoblasts Are Adult Pluripotent Stem Cells Underlying Planaria Regeneration. Cell [Internet]. 2018 [cited 2019 Apr 26];173(7):1593-1608.e20. Available from: https://doi.org/10.1016/j.cell.2018.05.006

134. Baguñà J. Dramatic mitotic response in planarians after feeding, and a hypothesis for the control mechanism. J Exp Zool. 1974;190(1):117–22.

135. Child CM. Studies on the dynamics of morphogenesis and inheritance in experimental reproduction - VI. The nature of the axial gradients in Planaria and their relation to antero-posterior dominance, polarity and symmetry. Arch für Entwicklungsmechanik der Org [Internet]. 1913 Jul [cited 2019 Sep 9];37(1):108–58. Available from: http://link.springer.com/10.1007/BF02275089

136. Dalyell JG. Observations on some interesting phenomena in animal physiology, exhibited by several species of Planariae. 1894.

137. CHANDEBOIS R. THE DYNAMICS OF WOUND CLOSURE AND ITS ROLE IN THE PROGRAMMING OF PLANARIAN REGENERATION. II — DISTALIZATION. Dev Growth Differ [Internet]. 1980 Aug 1 [cited 2019 Sep 12];22(4):693–704. Available from: http://doi.wiley.com/10.1111/j.1440-169X.1980.00693.x

138. Wenemoser D, Lapan SW, Wilkinson AW, Bell GW, Reddien PW. A molecular wound response program associated with regeneration initiation in planarians. Genes Dev [Internet]. 2012 May 1 [cited 2014 Jul 14];26(9):988–1002. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3347795&tool=pmcentrez&rendertype=abstract

139. Gaviño MA, Wenemoser D, Wang IE, Reddien PW. Tissue absence initiates regeneration through Follistatin-mediated inhibition of Activin signaling. Elife. 2013;2013(2):1–13.

140. Roberts-Galbraith RH, Newmark PA. Follistatin antagonizes activin signaling and acts with notum to direct planarian head regeneration. Proc Natl Acad Sci U S A [Internet]. 2013 Jan 22 [cited 2014 Jan 23];110(4):1363–8. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3557015&tool=pmcentrez&rendertype=abstract

141. Owlarn S, Klenner F, Schmidt D, Rabert F, Tomasso A, Reuter H, et al. Generic wound signals initiate regeneration in missing tissue contexts. Nat Commun [Internet]. 2017;1–13. Available from: http://dx.doi.org/10.1038/s41467-017-02338-x

142. Pirotte N, Stevens A-S, Fraguas S, Plusquin M, Van Roten A, Van Belleghem F, et al. Reactive Oxygen Species in Planarian Regeneration: An Upstream Necessity for Correct Patterning and Brain Formation. Oxid Med Cell Longev [Internet]. 2015;2015:1–19. Available from: http://www.hindawi.com/journals/omcl/2015/392476/

143. Baguñá J, Salo E. Regeneration and pattern formation in planarians I. The pattern of mitosis in anterior and posterior regeneration in Dugesia (G) tigrina, and a new proposal for blastema formation. Development [Internet]. 1984 [cited 2019 Aug 19];83(1):63–80. Available from: https://dev.biologists.org/content/develop/83/1/63.full.pdf

144. Baguñà J. Mitosis in the intact and regenerating planarian Dugesia mediterranea n.sp. II. Mitotic studies during regeneration, and a possible mechanism of blastema formation. J Exp Zool [Internet]. 1976 Jan 1 [cited 2019 Sep 9];195(1):65–79. Available from: http://doi.wiley.com/10.1002/jez.1401950107

145. Guedelhoefer OC, Alvarado a. S. Amputation induces stem cell mobilization to sites of injury during planarian regeneration. Development. 2012;139(19):3510–20.

146. Abnave P, Aboukhatwa E, Kosaka N, Thompson J, Hill MA, Aboobaker AA. Epithelial-mesenchymal transition transcription factors control pluripotent adult stem cell migration in vivo in planarians. Dev [Internet]. 2017 Oct 1 [cited 2019 Sep 12];144(19):3440–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28893948

147. Pellettieri J, Fitzgerald P, Watanabe S, Mancuso J, Green DR, Sánchez Alvarado A. Cell death and tissue remodeling in planarian regeneration. Dev Biol [Internet]. 2010 Feb 1 [cited 2014 Aug 19];338(1):76–85. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2835816&tool=pmcentrez&rendertype=abstract

148. González-Estévez C. Gtdap-1 promotes autophagy and is required for planarian remodeling during regeneration and starvation. Proc Natl Acad Sci [Internet]. 2007 [cited 2014 Aug 7]; Available from: http://www.pnas.org/content/104/33/13373.short

149. Stückemann T, Cleland JP, Werner S, Vu HT-K, Liu SY, Friedrich BM, et al. Antagonistic self-organizing patterning systems control maintenance and regeneration of the anteroposterior axis in planarians. Dev Cell. 2017;in press:248–63.

150. Adell T, Salò E, Boutos M, Bartscherer K. Smed-Evi/Wntless is required for β-catenin-dependent and -independent processes during planarian regeneration. Development. 2009;136(6):905–10.

151. Petersen CP, Reddien PW. A wound-induced Wnt expression program controls planarian regeneration polarity. Proc Natl Acad Sci U S A. 2009;106(40):17061–6.

152. Iglesias M, Gomez-Skarmeta JL, Saló E, Adell T. Silencing of Smed-βcatenin1 generates radial-like hypercephalized planarians . [cited 2014 Mar 27]; Available from: http://dev.biologists.org/content/135/7/1215/F2.expansion.html

153. Petersen CP, Reddien PW. Smed-betacatenin-1 is required for anteroposterior blastema polarity in planarian regeneration. Science. 2008;319(5861):327–30.

154. Sureda-Gómez M, Pascual-Carreras E, Adell T. Posterior wnts have distinct roles in specification and patterning of the planarian posterior region. Int J Mol Sci. 2015;16(11):26543–54.

155. Gurley K a., Elliott S a., Simakov O, Schmidt H a., Holstein TW, Alvarado AS. Expression of secreted Wnt pathway components reveals unexpected complexity of the planarian amputation response. Dev Biol. 2010;347(1):24–39.

156. Petersen CP, Reddien PW. Polarized notum activation at wounds inhibits Wnt function to promote planarian head regeneration. Science. 2011;332(6031):852–5.

157. Molina MD, Saló E, Cebrià F. The BMP pathway is essential for re-specification and maintenance of the dorsoventral axis in regenerating and intact planarians. Dev Biol [Internet]. 2007 Nov 1 [cited 2014 Jan 23];311(1):79–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17905225

158. Reddien PW, Bermange AL, Kicza AM, Sánchez Alvarado A. BMP signaling regulates the dorsal planarian midline and is needed for asymmetric regeneration. Development [Internet]. 2007 Oct 17 [cited 2019 Sep 8];134(22):4043–51. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17942485

159. Orii H, Watanabe K. Bone morphogenetic protein is required for dorso-ventral patterning in the planarian Dugesia japonica. Dev Growth Differ [Internet]. 2007 May 9 [cited 2019 Sep 8];49(4):345–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17501910

160. Cebrià F, Guo T, Jopek J, Newmark PA. Regeneration and maintenance of the planarian midline is regulated by a slit orthologue. Dev Biol [Internet]. 2007 Jul 15 [cited 2015 Aug 29];307(2):394–406. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2148499&tool=pmcentrez&rendertype=abstract

161. Oderberg IM, Li DJ, Scimone ML, Gavino MA, Reddien PW. Landmarks in existing tissue at wounds are utilized to generate pattern in regenerating tissue. 2017;1–10.

162. Atabay KD, LoCascio SA, de Hoog T, Reddien PW. Self-organization and progenitor targeting generate stable patterns in planarian regeneration. Science (80-). 2018;360(6387):404–9.

163. Hayashi T, Motoishi M, Yazawa S, Itomi K, Tanegashima C, Nishimura O, et al. A LIM-home-obox gene is required for differentiation of Wnt-expressing cells at the posterior end of the planarian body. Development [Internet]. 2011;138(17):3679–88. Available from: http://dev.biologists.org/cgi/doi/10.1242/dev.060194

164. Petersen CP, Reddien PW. Smed- catenin-1 Is Required for Anteroposterior Blastema Polarity in Planarian Regeneration. Science (80-) [Internet]. 2008;319(5861):327–30. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.1149943

165. Scimone ML, Lapan SW, Reddien PW. A forkhead transcription factor is wound-induced at the planarian midline and required for anterior pole regeneration. PLoS Genet [Internet]. 2014 Jan [cited 2014 Aug 27];10(1):e1003999. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3886891&tool=pmcentrez&rendertype=abstract

166. Vogg MC, Owlarn S, Pérez Rico Y a, Xie J, Suzuki Y, Gentile L, et al. Stem cell-dependent formation of a functional anterior regeneration pole in planarians requires Zic and Forkhe-ad transcription factors. Dev Biol [Internet]. 2014 Jun 15 [cited 2014 Oct 2];390(2):136–48. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24704339

167. Vásquez-Doorman C, Petersen CP. zic-1 Expression in Planarian Neoblasts after Injury Controls Anterior Pole Regeneration. PLoS Genet. 2014;10(7).

168. Felix DA, Aboobaker AA. The TALE class homeobox gene Smed-prep defines the anterior compartment for head regeneration. PLoS Genet. 2010;6(4).

169. Blassberg RA, Felix DA, Tejada-Romero B, Aboobaker AA. PBX/extradenticle is required to re-establish axial structures and polarity during planarian regeneration. Development [Internet]. 2013 Feb 15 [cited 2014 Jan 23];140(4):730–9. Available from: http://dev.biologists.org/content/140/4/730.long

170. Chen C-CG, Wang IE, Reddien PW. pbx is required for pole and eye regeneration in planarians. Development [Internet]. 2013 Feb 15 [cited 2014 Jan 23];140(4):719–29. Available from: http://dev.biologists.org/content/140/4/719.long

171. Reuter H, März M, Vogg MC, Eccles D, Grífol-Boldú L, Wehner D, et al. β-Catenin-Dependent Control of Positional Information along the AP Body Axis in Planarians Involves a Teashirt Family Member. Cell Rep [Internet]. 2014 Dec 30 [cited 2015 Jan 9];1–13. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25558068

172. Owen JH, Wagner DE, Chen C-C, Petersen CP, Reddien PW. Teashirt Is Required for Head-Versus-Tail Regeneration Polarity in Planarians. Development [Internet]. 2015;(February):1–11. Available from: http://dev.biologists.org/cgi/doi/10.1242/dev.119685

173. Gurley KA, Rink JC, Alvarado AS. β-catenin defines head versus tail identity during planarian regeneration and homeostasis. Science (80-) [Internet]. 2008 Jan 18 [cited 2019 Aug 21];319(5861):323–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18063757

174. Almuedo-Castillo M, Saló E, Adell T. Dishevelled is essential for neural connectivity and pla-

nar cell polarity in planarians. Proc Natl Acad Sci U S A [Internet]. 2011 Feb 15 [cited 2014 Oct 2];108(7):2813–8. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3041082&tool=pmcentrez&rendertype=abstract

175. Lander R, Petersen CP. Wnt, Ptk7, and FGFRL expression gradients control trunk positional identity in planarian regeneration. Elife [Internet]. 2016 Apr 13;5. Available from: https://elifesciences.org/articles/12850

176. Hill EM, Petersen CP. Wnt/Notum spatial feedback inhibition controls neoblast differentiation to regulate reversible growth of the planarian brain. Development [Internet]. 2015;(November). Available from: http://dev.biologists.org/cgi/doi/10.1242/dev.123612

177. Ingham PW, Nakano Y, Seger C. Mechanisms and functions of Hedgehog signalling across the metazoa [Internet]. Vol. 12, Nature Reviews Genetics. Nature Publishing Group; 2011 [cited 2019 Sep 9]. p. 393–406. Available from: http://www.nature.com/articles/nrg2984

178. Yazawa S, Umesono Y, Hayashi T, Tarui H, Agata K. Planarian hedgehog/patched establishes anterior-posterior polarity by regulating Wnt signaling. Proc Natl Acad Sci U S A [Internet]. 2009 [cited 2019 Apr 30];106(52):22329–34. Available from: https://www-pnas-org.sire.ub.edu/content/pnas/106/52/22329.full.pdf

179. Rink JC. Planarian Hh signaling regulates regeneration polarity and links Hh pathway evolution to cilia. Science (80-). 2010;326(5958):1–9.

180. van Wolfswinkel JC, Wagner DE, Reddien PW. Single-Cell Analysis Reveals Functionally Distinct Classes within the Planarian Stem Cell Compartment. Cell Stem Cell [Internet]. 2014 Jul [cited 2014 Jul 12]; Available from: http://www.sciencedirect.com/science/article/pii/S1934590914002550

181. Zhu SJ, Pearson BJ. (Neo)blast from the past: new insights into planarian stem cell lineages. Curr Opin Genet Dev [Internet]. 2016;40:74–80. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0959437X16300818

182. Molinaro AM, Pearson BJ. In silico lineage tracing through single cell transcriptomics identifies a neural stem cell population in planarians. Genome Biol [Internet]. 2016;17(1):87. Available from: http://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-0937-9

183. Zhu SJ, Hallows SE, Currie KW, Xu C, Pearson BJ. A mex3 homolog is required for differentiation during planarian stem cell lineage development. Elife. 2015;4(JUNE 2015):1–23.

184. Scimone ML, Cote LE, Reddien PW. Orthogonal muscle fibres have different instructive roles in planarian regeneration. Nature [Internet]. 2017; Available from: http://www.nature.com/doifinder/10.1038/nature24660

185. Bagunyà J, Romero R. Quantitative analysis of cell types during growth, degrowth and regeneration in the planarians Dugesia mediterranea and Dugesia tigrina. Hydrobiologia [Internet]. 1981 Oct [cited 2014 Aug 21];84(1):181–94. Available from: http://link.springer.com/10.1007/BF00026179

186. Thommen A, Werner S, Frank O, Philipp J, Knittelfelder O, Quek Y, et al. Body size-dependent energy storage causes Kleiber's law scaling of the metabolic rate in planarians. Elife. 2019;8:1–29.

187. Baguñà J. Mitosis in the intact and regenerating planarian Dugesia mediterranea n.sp. I. Mitotic studies during growth, feeding and starvation. J Exp Zool. 1976;195(1):53–64.

188. Kobayashi C, Saito Y, Ogawa K, Agata K. Wnt signaling is required for antero-posterior patterning of the planarian brain. Dev Biol [Internet]. 2007 Jun 15 [cited 2019 Sep 8];306(2):714–24. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17498685

189. Hill EM, Petersen CP. Wnt/Notum spatial feedback inhibition controls neoblast differentiation to regulate reversible growth of the planarian brain. Development [Internet]. 2015;142(24):4217–29. Available from: http://dev.biologists.org/cgi/doi/10.1242/dev.123612

190. Spitz F, Furlong EEM. Transcription factors: from enhancer binding to developmental control. Nat Rev Genet [Internet]. 2012 Sep 7 [cited 2019 Sep 9];13(9):613–26. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22868264

191. Levine M, Tjian R. Transcription regulation and animal diversity [Internet]. Vol. 424, Nature. Nature Publishing Group; 2003 [cited 2019 Sep 9]. p. 147–51. Available from: http://www.nature.com/articles/nature01763

192. Haberle V, Stark A. Eukaryotic core promoters and the functional basis of transcription initiation [Internet]. Vol. 19, Nature Reviews Molecular Cell Biology. 2018 [cited 2018 Nov 2]. p. 621–37. Available from: www.nature.com/nrm

193. Catarino RR, Stark A. Assessing sufficiency and necessity of enhancer activities for gene expression and the mechanisms of transcription activation. Genes Dev [Internet]. 2018;32(3–4):202–23. Available from: http://genesdev.cshlp.org/lookup/doi/10.1101/gad.310367.117

194. Shlyueva D, Stampfel G, Stark A. Transcriptional enhancers: from properties to genome-wide predictions. Nat Rev Genet [Internet]. 2014;15(4):272–86. Available from: http://www.nature.com/doifinder/10.1038/nrg3682

195. Dawson MA, Kouzarides T, Huntly BJP. Targeting Epigenetic Readers in Cancer. N Engl J Med [Internet]. 2012 Aug 16 [cited 2019 Sep 9];367(7):647–57. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22894577

196. Sawan C, Vaissière T, Murr R, Herceg Z. Epigenetic drivers and genetic passengers on the road to cancer. Mutat Res Mol Mech Mutagen [Internet]. 2008 Jul 3 [cited 2019 Sep 9];642(1–2):1–13. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18471836

197. Roeder RG. The role of general initiation factors in transcription by RNA polymerase II. Trends Biochem Sci [Internet]. 1996 Sep [cited 2019 Sep 9];21(9):327–35. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8870495

198. Orphanides G, Lagrange T, Reinberg D. The general transcription factors of RNA polymerase II [Internet]. Vol. 10, Genes and Development. 1996 [cited 2019 Sep 9]. p. 2657–83. Avai-

lable from: http://www.ncbi.nlm.nih.gov/pubmed/8946909

199.   Suryamohan K, Halfon MS. Identifying transcriptional cis-regulatory modules in animal ge-
        nomes. Wiley Interdiscip Rev Dev Biol [Internet]. 2015 Mar [cited 2019 Sep 9];4(2):59–84.
        Available from: http://www.ncbi.nlm.nih.gov/pubmed/25704908

200.   Amano T, Sagai T, Tanabe H, Mizushina Y, Nakazawa H, Shiroishi T. Chromosomal Dyna-
        mics at the Shh Locus: Limb Bud-Specific Differential Regulation of Competence and Active
        Transcription. Dev Cell [Internet]. 2009 Jan [cited 2019 Sep 9];16(1):47–57. Available from:
        http://www.ncbi.nlm.nih.gov/pubmed/19097946

201.   Lawrence M, Daujat S, Schneider R. Lateral Thinking: How Histone Modifications Regulate
        Gene Expression [Internet]. Vol. 32, Trends in Genetics. 2016 [cited 2019 Sep 9]. p. 42–56.
        Available from: http://www.ncbi.nlm.nih.gov/pubmed/26704082

202.   Buecker C, Wysocka J. Enhancers as information integration hubs in development: Lessons
        from genomics [Internet]. Vol. 28, Trends in Genetics. 2012 [cited 2019 Sep 9]. p. 276–84.
        Available from: http://www.ncbi.nlm.nih.gov/pubmed/22487374

203.   Zentner GE, Scacheri PC. The chromatin fingerprint of gene enhancer elements [Internet].
        Vol. 287, Journal of Biological Chemistry. 2012 [cited 2019 Sep 9]. p. 30888–96. Available
        from: http://www.jbc.org/lookup/doi/10.1074/jbc.R111.296491

204.   Calo E, Wysocka J. Modification of Enhancer Chromatin: What, How, and Why? [Internet].
        Vol. 49, Molecular Cell. 2013 [cited 2019 Sep 9]. p. 825–37. Available from: http://www.ncbi.
        nlm.nih.gov/pubmed/23473601

205.   Harris RE, Setiawan L, Saul J, Hariharan IK. Localized epigenetic silencing of a damage-
        activated WNT enhancer limits regeneration in mature Drosophila imaginal discs. Elife.
        2016;5(FEBRUARY2016):1–28.

206.   Kang J, Hu J, Karra R, Dickson AL, Tornini VA, Nachtrab G, et al. Modulation of tissue
        repair by regeneration enhancer elements. Nature [Internet]. 2016 Apr 6 [cited 2019 Sep
        8];532(7598):201–6. Available from: http://www.nature.com/articles/nature17644

207.   Goldman JA, Kuzu G, Lee N, Karasik J, Gemberling M, Foglia MJ, et al. Resolving Heart
        Regeneration by Replacement Histone Profiling. Dev Cell [Internet]. 2017 [cited 2019 Aug
        28];40(4):392-404.e5. Available from: http://dx.doi.org/10.1016/j.devcel.2017.01.013

208.   Guigó R, Corominas M, Mishra RK, Serras F, Vizcaya-Molina E, Klein CC. Damage-respon-
        sive elements in Drosophila regeneration . Genome Res. 2018;28(12):1852–66.

209.   Gehrke AR, Neverett E, Luo YJ, Brandt A, Ricci L, Hulett RE, et al. Acoel genome reveals
        the regulatory landscape of whole-body regeneration. Science (80-) [Internet]. 2019 [cited
        2019 Mar 14];363(6432). Available from: http://science.sciencemag.org/

210.   Jaber-Hijazi F, Lo PJKP, Mihaylova Y, Foster JM, Benner JS, Tejada Romero B, et al. Plana-
        rian MBD2/3 is required for adult stem cell pluripotency independently of DNA methylation.
        Dev Biol [Internet]. 2013 Dec 1 [cited 2019 Sep 9];384(1):141–53. Available from: http://

www.ncbi.nlm.nih.gov/pubmed/24063805

211.   Ohki I, Shimotake N, Fujita N, Jee JG, Ikegami T, Nakao M, et al. Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. Cell [Internet]. 2001 May 18 [cited 2019 Sep 9];105(4):487–97. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11371345

212.   Scimone ML, Meisel J, Reddien PW. The Mi-2-like Smed-CHD4 gene is required for stem cell differentiation in the planarian Schmidtea mediterranea. Development [Internet]. 2010 Apr 15 [cited 2019 Sep 9];137(8):1231–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20223763

213.   Zhu SJ, Pearson BJ. The Retinoblastoma pathway regulates stem cell proliferation in freshwater planarians. Dev Biol. 2013;373(2):442–52.

214.   Eisenhoffer GT, Kang H, Alvarado AS. Molecular Analysis of Stem Cells and Their Descendants during Cell Turnover and Regeneration in the Planarian Schmidtea mediterranea. Cell Stem Cell. 2008;3(3):327–39.

215.   Robb SMC, Alvarado AS. Histone Modifications and Regeneration in the Planarian Schmidtea mediterranea. In: Current Topics in Developmental Biology [Internet]. 2014 [cited 2019 Sep 9]. p. 71–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24512706

216.   Hubert A, Henderson JM, Cowles MW, Ross KG, Hagen M, Anderson C, et al. A functional genomics screen identifies an Importin-α homolog as a regulator of stem cell function and tissue patterning during planarian regeneration. BMC Genomics [Internet]. 2015;16(1):769. Available from: http://www.biomedcentral.com/1471-2164/16/769

217.   Bonuccelli L, Rossi L, Lena A, Scarcelli V, Rainaldi G, Evangelista M, et al. An RbAp48-like gene regulates adult stem cells in planarians. J Cell Sci [Internet]. 2010 Mar 1 [cited 2019 Sep 9];123(5):690–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20124416

218.   Vásquez-Doorman C, Petersen CP. The NuRD complex component *p66* suppresses photoreceptor neuron regeneration in planarians. Regeneration [Internet]. 2016;3(3):168–78. Available from: http://doi.wiley.com/10.1002/reg2.58

219.   Duncan EM, Chitsazan AD, Seidel CW, Sa A. Set1 and MLL1 / 2 Target Distinct Sets of Functionally Different Genomic Loci In Vivo Set1 and MLL1 / 2 Target Distinct Sets of Functionally Different Genomic Loci In Vivo. 2015;1–15.

220.   Hubert A, Henderson JM, Ross KG, Cowles MW, Torres J, Zayas RM. Epigenetic regulation of planarian stem cells by the SET1/MLL family of histone methyltransferases. Epigenetics [Internet]. 2013 Jan [cited 2019 Sep 8];8(1):79–91. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23235145

221.   Mihaylova Y, Abnave P, Kao D, Hughes S, Lai A, Jaber-Hijazi F, et al. Conservation of epigenetic regulation by the MLL3/4 tumour suppressor in planarian pluripotent stem cells. Nat Commun [Internet]. 2018 Dec 7 [cited 2018 Sep 10];9(1):3633. Available from: http://www.

nature.com/articles/s41467-018-06092-6

222. Simon JA, Kingston RE. Occupying Chromatin: Polycomb Mechanisms for Getting to Genomic Targets, Stopping Transcriptional Traffic, and Staying Put [Internet]. Vol. 49, Molecular Cell. 2013 [cited 2019 Sep 9]. p. 808–24. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23473600

223. Wagner DE, Ho JJ, Reddien PW. Genetic regulators of a pluripotent adult stem cell system in planarians identified by RNAi and clonal analysis. Cell Stem Cell [Internet]. 2012 Mar 2 [cited 2014 Jan 23];10(3):299–311. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3338251&tool=pmcentrez&rendertype=abstract

224. Dattani A, Kao D, Mihaylova Y, Abnave P, Hughes S, Lai A, et al. Epigenetic analyses of planarian stem cells demonstrate conservation of bivalent histone modifications in animal stem cells. Genome Res. 2018;

225. Oviedo NJ, Newmark PA, Sánchez Alvarado A. Allometric scaling and proportion regulation in the freshwater planarian Schmidtea mediterranea. Dev Dyn [Internet]. 2003 Feb [cited 2014 Apr 28];226(2):326–33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12557210

226. González-Estévez C, Felix D a., Rodríguez-Esteban G, Aziz Aboobaker a. Decreased neoblast progeny and increased cell death during starvation-induced planarian degrowth. Int J Dev Biol. 2012;56(1–3):83–91.

227. Gurley KA, Rink JC, Sánchez Alvarado A. β-Catenin Defines Head Versus Tail Identity During Planarian Regeneration and Homeostasis. Science (80-). 2008;319(5861):323–7.

228. Patel AK, Park KK, Hackam AS. Wnt signaling promotes axonal regeneration following optic nerve injury in the mouse. Neuroscience [Internet]. 2017 Feb 20 [cited 2019 Sep 9];343:372–83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28011153

229. Tice DA, Szeto W, Soloviev I, Rubinfeld B, Fong SE, Dugger DL, et al. Synergistic induction of tumor antigens by Wnt-1 signaling and retinoic acid revealed by gene expression profiling. J Biol Chem [Internet]. 2002 Apr 19 [cited 2019 Sep 9];277(16):14329–35. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11832495

230. Li DJ, McMann CL, Reddien PW. Nuclear receptor NR4A is required for patterning at the ends of the planarian anterior-posterior axis. Elife. 2019;8.

231. Cebrià F, Vispo M, Newmark P, Bueno D, Romero R. Myocyte differentiation and body wall muscle regeneration in the planarian Girardia tigrina. Dev Genes Evol [Internet]. 1997 Nov 4 [cited 2019 Sep 2];207(5):306–16. Available from: http://link.springer.com/10.1007/s004270050118

232. Cebrià F, Newmark PA. Planarian homologs of netrin and netrin receptor are required for proper regeneration of the central nervous system and the maintenance of nervous system architecture. Development [Internet]. 2005 Aug 15 [cited 2014 Aug 27];132(16):3691–703. Available from: http://dev.biologists.org/content/132/16/3691.long

233. Cusanovich DA, Reddington JP, Garfield DA, Daza RM, Aghamirzaie D, Marco-Ferreres R, et al. The cis-regulatory dynamics of embryonic development at single-cell resolution. Nature [Internet]. 2018 Mar 14 [cited 2019 Sep 9];555(7697):538–42. Available from: http://www.nature.com/articles/nature25981

234. Liu C, Wang M, Wei X, Wu L, Xu J, Dai X, et al. An ATAC-seq atlas of chromatin accessibility in mouse tissues. Sci data [Internet]. 2019 Dec 20 [cited 2019 Sep 9];6(1):65. Available from: http://www.nature.com/articles/s41597-019-0071-0

235. Jia G, Preussner J, Chen X, Guenther S, Yuan X, Yekelchyk M, et al. Single cell RNA-seq and ATAC-seq analysis of cardiac progenitor cell transition states and lineage settlement. Nat Commun [Internet]. 2018 Dec 19 [cited 2019 Sep 9];9(1):4877. Available from: http://www.nature.com/articles/s41467-018-07307-6

236. Bysani M, Agren R, Davegårdh C, Volkov P, Rönn T, Unneberg P, et al. ATAC-seq reveals alterations in open chromatin in pancreatic islets from subjects with type 2 diabetes. Sci Rep [Internet]. 2019 Dec 23 [cited 2019 Sep 9];9(1):7785. Available from: http://www.nature.com/articles/s41598-019-44076-8

237. Tewari AG, Stern SR, Oderberg IM, Reddien PW. Cellular and Molecular Responses Unique to Major Injury Are Dispensable for Planarian Regeneration. Cell Rep [Internet]. 2018;25(9):2577-2590.e3. Available from: https://doi.org/10.1016/j.celrep.2018.11.004

238. Han W, Wang H. Regulation of canonical Wnt/β-catenin pathway in the nucleus. Chinese Sci Bull. 2014;59(28):3530–5.

239. Su H, Sureda-Gomez M, Rabaneda-Lombarte N, Gelabert M, Xie J, Wu W, et al. A C-terminally truncated form of β-catenin acts as a novel regulator of Wnt/β-catenin signaling in planarians. PLoS Genet. 2017;13(10):1–32.

240. Brown DDR, Molinaro AM, Pearson BJ. The planarian TCF / LEF factor Smed-tcf1 is required for the regeneration of dorsal-lateral neuronal subtypes. Dev Biol [Internet]. 2017;433(August):0–1. Available from: http://dx.doi.org/10.1016/j.ydbio.2017.08.024

241. Wang W, Li X, Lee M, Jun S, Aziz KE, Feng L, et al. FOXKs promote Wnt/β-catenin signaling by translocating DVL into the nucleus. 2016;32(6):707–18.

242. Itoh K, Brott BK, Bae GU, Ratcliffe MJ, Sokol SY. Nuclear localization is required for Dishevelled function in Wnt/β-catenin signaling. J Biol [Internet]. 2005 [cited 2019 Sep 9];4(1):3. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15720724

243. Ré B, Benayoun A, Caburet S, Veitia RA. Forkhead transcription factors: key players in health and disease. 2011 [cited 2019 Feb 20]; Available from: http://gero.usc.edu/labs/benayounlab/files/2018/06/Forkhead-transcription-factors.pdf

244. Kaestner KH, Knochel W, Martinez DE. Unified nomenclature for the winged helix/forkhead transcription factors. Genes Dev [Internet]. 2000 Jan 15 [cited 2019 Sep 9];14(2):142–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10702024

245. Golson ML, Kaestner KH. Fox transcription factors: From development to disease. Dev [Internet]. 2016 Dec 15 [cited 2019 Feb 20];143(24):4558–70. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27965437

246. Magie CR, Pang K, Martindale MQ. Genomic inventory and expression of Sox and Fox genes in the cnidarian Nematostella vectensis. Dev Genes Evol [Internet]. 2005 Dec 29 [cited 2018 Aug 24];215(12):618–30. Available from: http://link.springer.com/10.1007/s00427-005-0022-y

247. Larroux C, Luke GN, Koopman P, Rokhsar DS, Shimeld SM, Degnan BM. Genesis and expansion of metazoan transcription factor gene classes. Mol Biol Evol [Internet]. 2008 May 1 [cited 2019 Feb 20];25(5):980–96. Available from: https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/msn047

248. Adell T, Müller WEG. Isolation and characterization of five Fox (Forkhead) genes from the sponge Suberites domuncula. Gene. 2004;334(1–2):35–46.

249. Shimeld SM, Boyle MJ, Brunet T, Luke GN, Seaver EC. Clustered Fox genes in lophotrochozoans and the evolution of the bilaterian Fox gene cluster. Dev Biol [Internet]. 2010;340(2):234–48. Available from: http://dx.doi.org/10.1016/j.ydbio.2010.01.015

250. Koinuma S, Umesono Y, Watanabe K, Agata K. The expression of planarian brain factor homologs, DjFoxG and DjFoxD. [cited 2018 Sep 4]; Available from: www.elsevier.com/locate/modgep

251. Vij S, Rink JC, Ho HK, Babu D, Eitel M, Narasimhan V, et al. Evolutionarily Ancient Association of the FoxJ1 Transcription Factor with the Motile Ciliogenic Program. PLoS Genet. 2012;8(11).

252. Koinuma S, Umesono Y, Watanabe K, Agata K. Planaria FoxA (HNF3) homologue is specifically expressed in the pharynx-forming cells. Gene. 2000;259(1–2):171–6.

253. Adler CE, Seidel CW, McKinney S a, Sánchez Alvarado A. Selective amputation of the pharynx identifies a FoxA-dependent regeneration program in planaria. Elife [Internet]. 2014 Jan [cited 2014 Sep 18];3:e02238. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3985184&tool=pmcentrez&rendertype=abstract

254. Wang C, Han X-S, Li F-F, Huang S, Qin Y-W, Zhao X-X, et al. Forkhead containing transcription factor Albino controls tetrapyrrole-based body pigmentation in planarian. Cell Discov [Internet]. 2016;2:16029. Available from: http://www.nature.com/articles/celldisc201629

255. Yaguchi S, Yaguchi J, Angerer RC, Angerer LM. A Wnt-FoxQ2-Nodal Pathway Links Primary and Secondary Axis Specification in Sea Urchin Embryos. Dev Cell. 2008;14(1):97–107.

256. Chevalier S, Martin A, Leclère L, Amiel A, Houliston E. Polarised expression of FoxB and FoxQ2 genes during development of the hydrozoan Clytia hemisphaerica. Dev Genes Evol [Internet]. 2006 Oct 25 [cited 2019 Aug 28];216(11):709–20. Available from: http://link.springer.com/10.1007/s00427-006-0103-6

257. Fritzenwanker JH, Gerhart J, Freeman RM, Lowe CJ. The Fox/Forkhead transcription factor family of the hemichordate Saccoglossus kowalevskii. Evodevo [Internet]. 2014 May 7 [cited 2018 Aug 24];5(1):17. Available from: http://evodevojournal.biomedcentral.com/articles/10.1186/2041-9139-5-17

258. Koziol U, Jarero F, Olson PD, Brehm K. Comparative analysis of Wnt expression identifies a highly conserved developmental transition in flatworms. BMC Biol [Internet]. 2016;14(1):10. Available from: http://www.biomedcentral.com/1741-7007/14/10

259. Shimeld SM, Degnan B, Luke GN. Evolutionary genomics of the Fox genes: Origin of gene families and the ancestry of gene clusters. Genomics [Internet]. 2010;95(5):256–60. Available from: http://dx.doi.org/10.1016/j.ygeno.2009.08.002

260. Almawi AW, Matthews LA, Guarné A. FHA domains: Phosphopeptide binding and beyond [Internet]. Vol. 127, Progress in Biophysics and Molecular Biology. 2017 [cited 2019 Sep 9]. p. 105–10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27939759

261. Song X, Li B, Xiao Y, Chen C, Wang Q, Liu Y, et al. Structural and Biological Features of FOXP3 Dimerization Relevant to Regulatory T Cell Function. Cell Rep [Internet]. 2012 Jun 28 [cited 2019 Sep 9];1(6):665–75. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22813742

262. Scimone ML, Kravarik KM, Lapan SW, Reddien PW. Neoblast Specialization in Regeneration of the Planarian Schmidtea mediterranea. Stem Cell Reports [Internet]. 2014 Aug [cited 2014 Aug 27];3(2):339–52. Available from: http://linkinghub.elsevier.com/retrieve/pii/S2213671114001866

263. Salih DA, Brunet A. FoxO transcription factors in the maintenance of cellular homeostasis during aging [Internet]. Vol. 20, Current Opinion in Cell Biology. NIH Public Access; 2008 [cited 2019 Sep 9]. p. 126–36. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18394876

264. Hu B, Lefort K, Qiu W, Nguyen BC, Rajaram RD, Castillo E, et al. Control of hair follicle cell fate by underlying mesenchyme through a CSL-Wnt5a-FoxN1 regulatory axis. Genes Dev. 2010;24(14):1519–32.

265. Wilson GA, Bertrand N, Patel Y, Hughes JB, Feil EJ, Field D. Orphans as taxonomically restricted and ecologically important genes [Internet]. Vol. 151, Microbiology. 2005 [cited 2019 Jul 26]. p. 2499–501. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16079329

266. Tautz D, Domazet-Lošo T. The evolutionary origin of orphan genes [Internet]. Vol. 12, Nature Reviews Genetics. Nature Publishing Group; 2011 [cited 2019 Jul 26]. p. 692–702. Available from: http://www.nature.com/articles/nrg3053

267. Neme R, Tautz D. Phylogenetic patterns of emergence of new genes support a model of frequent de novo evolution. BMC Genomics [Internet]. 2013 Feb 21 [cited 2019 Jul 26];14(1):117. Available from: http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-14-117

268. Palmieri N, Kosiol C, Schlötterer C. The life cycle of Drosophila orphan genes. Elife [Internet]. 2014 Feb 19 [cited 2019 Jul 26];3. Available from: https://elifesciences.org/articles/01311

269. Wilson BA, Foy SG, Neme R, Masel J. Young genes are highly disordered as predicted by the preadaptation hypothesis of de novo gene birth. Nat Ecol Evol [Internet]. 2017 [cited 2019 Jul 25];1(6). Available from: http://www.nature.com/authors/editorial_policies/license.html#terms

270. Werner MS, Sieriebriennikov B, Prabh N, Loschko T. Young genes have distinct gene structure , epigenetic profiles , and transcriptional regulation Department of Evolutionary Biology , Max Planck Institute for Developmental Biology , Keywords : Genome Res. 2018;28:1675–87.

271. Toll-Riera M, Bosch N, Bellora N, Castelo R, Armengol L, Estivill X, et al. Origin of primate orphan genes: A comparative genomics approach. Mol Biol Evol [Internet]. 2009 Dec 23 [cited 2019 Jul 26];26(3):603–12. Available from: https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/msn281

272. Carvunis A-R, Rolland T, Wapinski I, Calderwood MA, Yildirim MA, Simonis N, et al. Proto-genes and de novo gene birth. Nature [Internet]. 2012 Jul 19 [cited 2019 Jul 26];487(7407):370–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22722833

273. Zhao L, Saelao P, Jones CD, Begun DJ. Origin and spread of de novo genes in Drosophila melanogaster populations. Science (80-) [Internet]. 2014 [cited 2019 Jul 25];343(6172):769–72. Available from: http://www.sciencemag.org/.The

274. Van Oss SB, Carvunis AR. De novo gene birth. PLoS Genet. 2019;15(5):e1008160.

275. Schlötterer C. Genes from scratch - the evolutionary fate of de novo genes. Trends Genet. 2015;31(4):215–9.

276. Long M, Betrán E, Thornton K, Wang W. The origin of new genes: Glimpses from the young and old [Internet]. Vol. 4, Nature Reviews Genetics. Nature Publishing Group; 2003 [cited 2019 Jul 26]. p. 865–75. Available from: http://www.nature.com/articles/nrg1204

277. McLysaght A, Hurst LD. Open questions in the study of de novo genes: what, how and why. Nat Rev Genet [Internet]. 2016;17(9):567–78. Available from: http://dx.doi.org/10.1038/nrg.2016.78%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/27452112

278. Colbourne JK, Pfrender ME, Gilbert D, Thomas WK, Tucker A, Oakley TH, et al. The Eco-responsive Genome of Daphnia pulex. Science (80-) [Internet]. 2011 Feb 4 [cited 2019 Jul 25];331(6017):555–61. Available from: https://science.sciencemag.org/content/331/6017/555

279. Donoghue MT, Keshavaiah C, Swamidatta SH, Spillane C. Evolutionary origins of Brassicaceae specific genes in Arabidopsis thaliana. BMC Evol Biol [Internet]. 2011 Dec 18 [cited 2019 Jul 25];11(1):47. Available from: http://bmcevolbiol.biomedcentral.com/articles/10.1186/1471-2148-11-47

280. Arendsee ZW, Li L, Wurtele ES. Coming of age: orphan genes in plants. Trends Plant Sci [Internet]. 2014 Nov 1 [cited 2019 Jul 26];19(11):698–708. Available from: https://www.sciencedirect.com/science/article/abs/pii/S1360138514001939

281. Peiris TH, Weckerle F, Ozamoto E, Ramirez D, Davidian D, Garcia-Ojeda ME, et al. TOR signaling regulates planarian stem cells and controls localized and organismal growth. J Cell Sci. 2012;125(7):1657–65.

282. Salazar-Roa M, Malumbres M. Fueling the Cell Division Cycle [Internet]. Vol. 27, Trends in Cell Biology. Elsevier; 2017 [cited 2019 Aug 1]. p. 69–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27746095

283. Lloyd AC. The regulation of cell size [Internet]. Vol. 154, Cell. 2013 [cited 2019 Aug 2]. p. 1194. Available from: http://dx.

284. Fingar DC, Salama S, Tsou C, Harlow E, Blenis J. Mammalian cell size is controlled by mTOR and its downstream targets S6K1 and 4EBP1/eIF4E. GENES Dev. 2002;1472–87.

285. Pifferi F, Aujard F. Caloric restriction, longevity and aging: Recent contributions from human and non-human primate studies. Prog Neuro-Psychopharmacology Biol Psychiatry [Internet]. 2019 Dec 20 [cited 2019 Aug 1];95:109702. Available from: https://www.sciencedirect.com/science/article/pii/S0278584619302702?via%3Dihub

286. Nacu E, Tanaka EM. Limb Regeneration: A New Development? Annu Rev Cell Dev Biol [Internet]. 2011 Nov 10 [cited 2019 Aug 1];27(1):409–40. Available from: http://www.annualreviews.org/doi/10.1146/annurev-cellbio-092910-154115

287. Pomerantz JH, Blau HM. Tumor suppressors: enhancers or suppressors of regeneration? Development. 2013 May 29;140(12):2502–12.

288. Moya IM, Halder G. Hippo–YAP/TAZ signalling in organ regeneration and regenerative medicine [Internet]. Vol. 20, Nature Reviews Molecular Cell Biology. Nature Publishing Group; 2019 [cited 2019 Aug 2]. p. 211–26. Available from: http://www.nature.com/articles/s41580-018-0086-y

289. Nowak K, Seisenbacher G, Hafen E, Stocker H. Nutrient restriction enhances the proliferative potential of cells lacking the tumor suppressor PTEN in mitotic tissues. Elife. 2013;2013(2):1–21.

290. Reinhardt JA, Wanjiru BM, Brant AT, Saelao P, Begun DJ. De Novo ORFs in Drosophila Are Important to Organismal Fitness and Evolved Rapidly from Previously Non-coding Sequences. PLoS Genet [Internet]. 2013 [cited 2019 Jul 26];9(10):1003860. Available from: www.plosgenetics.org

291. Khalturin K, Anton-Erxleben F, Sassmann S, Wittlieb J, Hemmrich G, Bosch TCG. A novel gene family controls species-specific morphological traits in Hydra. Patel N, editor. PLoS Biol [Internet]. 2008 Nov 18 [cited 2019 Aug 1];6(11):2436–49. Available from: http://dx.plos.org/10.1371/journal.pbio.0060278

References

292. Aguilera F, McDougall C, Degnan BM, Irwin D. Co-option and de novo gene evolution underlie molluscan shell diversity. Mol Biol Evol. 2017;34(4):779–92.

293. Waddintong CH. Principles ofEmbryology. New York: The Macmillan Company; 1954.

294. Waddington CH. Experiments on the Development of Chick and Duck Embryos, Cultivated in vitro. Philos Trans R Soc B Biol Sci [Internet]. 1932 Jan 1 [cited 2019 Aug 28];221(474–482):179–230. Available from: http://rstb.royalsocietypublishing.org/cgi/doi/10.1098/rstb.1932.0003

295. Anderson C, Khan MAF, Wong F, Solovieva T, Oliveira NMM, Baldock RA, et al. A strategy to discover new organizers identifies a putative heart organizer. Nat Commun [Internet]. 2016;7(May):12656. Available from: http://www.nature.com/doifinder/10.1038/ncomms12656

296. LoCascio SA, Lapan SW, Reddien PW. Eye Absence Does Not Regulate Planarian Stem Cells during Eye Regeneration. Dev Cell [Internet]. 2017;40(4):381-391.e3. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1534580717300710

297. Taub, Floyd E, Deleo JM, Thompson EB. Sequential Comparative Hybridizations Analyzed by Computerized Image Processing Can Identify and Quantitate Regulated RNAs. DNA [Internet]. 1983 Dec [cited 2019 Aug 28];2(4):309–27. Available from: http://www.ncbi.nlm.nih.gov/pubmed/6198132

298. Wang Z, Gerstein M, Snyder M. RNA-Seq: A revolutionary tool for transcriptomics [Internet]. Vol. 10, Nature Reviews Genetics. Nature Publishing Group; 2009 [cited 2019 Aug 28]. p. 57–63. Available from: http://www.nature.com/articles/nrg2484

299. Warner JF, Guerlais V, Amiel AR, Johnston H, Nedoncelle K, Röttinger E. NvERTx: A gene expression database to compare embryogenesis and regeneration in the sea anemone Nematostella vectensis. Dev [Internet]. 2018 [cited 2019 Aug 28];145(10). Available from: http://nvertx.kahikai.org.

300. Warner JF, Amiel AR, Johnston H, Röttinger E. Regeneration is a partial redeployment of the embryonic gene network. bioRxiv [Internet]. 2019 [cited 2019 Jun 5]; Available from: http://dx.doi.org/10.1101/658930

301. Wenger Y, Buzgariu W, Perruchoud C, Loichot G, Galliot B. Generic and context-dependent gene modulations during Hydra whole body regeneration. bioRxiv [Internet]. 2019;587147. Available from: https://www.biorxiv.org/content/10.1101/587147v1.full

302. King BL, Yin VP. A conserved microRNA regulatory circuit is differentially controlled during limb/appendage regeneration. Gibert Y, editor. PLoS One [Internet]. 2016 Jun 29 [cited 2019 Aug 28];11(6):e0157106. Available from: https://dx.plos.org/10.1371/journal.pone.0157106

303. Dwaraka VB, Smith JJ, Woodcock MR, Voss SR. Comparative transcriptomics of limb regeneration: Identification of conserved expression changes among three species of Ambystoma. Genomics [Internet]. 2018 [cited 2019 Aug 28]; Available from: https://doi.org/10.1016/j.ygeno.2018.07.017

304. Kao D, Felix D, Aboobaker A. The planarian regeneration transcriptome reveals a shared but temporally shifted regulatory program between opposing head and tail scenarios. BMC Genomics. 2013;14(1):1–17.

305. Sureda-Gomez M, Adell T. Planarian organizers [Internet]. Vol. 87, Seminars in Cell and Developmental Biology. 2019 [cited 2018 Nov 2]. p. 95–104. Available from: https://doi.org/10.1016/j.semcdb.2018.05.021

306. Nusse R. Wnt Target genes | The Wnt Homepage [Internet]. 2012 [cited 2019 Aug 28]. Available from: http://web.stanford.edu/group/nusselab/cgi-bin/wnt/target_genes

307. Ramakrishnan A-B, Cadigan KM. Wnt target genes and where to find them. F1000Research [Internet]. 2017;6(May):1–11. Available from: https://f1000research.com/articles/6-746/v1

308. Sethi JK, Vidal-Puig A. Wnt signalling and the control of cellular metabolism [Internet]. Vol. 427, Biochemical Journal. Europe PMC Funders; 2010 [cited 2019 Aug 30]. p. 1–17. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20226003

309. Cadoret A, Ovejero C, Terris B, Souil E, Lévy L, Lamers WH, et al. New targets of β-catenin signaling in the liver are involved in the glutamine metabolism. Oncogene [Internet]. 2002 Nov 25 [cited 2019 Aug 30];21(54):8293–301. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12447692

310. Schwartz DR, Wu R, Kardia SLR, Levin AM, Huang C-C, Shedden KA, et al. Novel candidate targets of beta-catenin/T-cell factor signaling identified by gene expression profiling of ovarian endometrioid adenocarcinomas. Cancer Res [Internet]. 2003 Jun 1 [cited 2019 Aug 30];63(11):2913–22. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12782598

311. Lagathu C, Christodoulides C, Virtue S, Cawthorn WP, Franzin C, Kimber WA, et al. Dact1, a Nutritionally Regulated Preadipocyte Gene, Controls Adipogenesis by Coordinating the Wnt/ -Catenin Signaling Network. Diabetes [Internet]. 2009 Mar 1 [cited 2019 Aug 30];58(3):609–19. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19073771

312. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods [Internet]. 2013 Dec [cited 2019 Aug 28];10(12):1213–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24097267

313. Park PJ. ChIP-seq: Advantages and challenges of a maturing technology [Internet]. Vol. 10, Nature Reviews Genetics. Nature Publishing Group; 2009 [cited 2019 Aug 28]. p. 669–80. Available from: http://www.nature.com/articles/nrg2641

314. Aztekin C, Hiscock TW, Marioni JC, Gurdon JB, Simons BD, Jullien J. Identification of a regeneration-organizing cell in the *Xenopus* tail. Science (80-) [Internet]. 2019;364(6441):653–8. Available from: http://www.sciencemag.org/lookup/doi/10.1126/science.aav9996

315. Martinez Arias A, Steventon B. On the nature and function of organizers. Dev [Internet].

2018 [cited 2019 Mar 28];145(5). Available from: http://dev.biologists.org/content/develop/145/5/dev159525.full.pdf

316. Iglesias M, Almuedo-Castillo M, Aboobaker a A, Saló E. Early planarian brain regeneration is independent of blastema polarity mediated by the Wnt/β-catenin pathway. Dev Biol [Internet]. 2011 Oct 1 [cited 2014 Sep 24];358(1):68–78. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21806978

317. Ingham PW, Hidalgo A. Regulation of wingless transcription in the Drosophila embryo. Development [Internet]. 1993 Jan [cited 2019 Sep 13];117(1):283–91. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8223252

318. Bejsovec A, Martinez Arias A. Roles of wingless in patterning the larval epidermis of Drosophila. Development [Internet]. 1991 Oct [cited 2019 Sep 13];113(2):471–85. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1782860

319. Brand M, Heisenberg CP, Jiang YJ, Beuchle D, Lun K, Furutani-Seiki M, et al. Mutations in zebrafish genes affecting the formation of the boundary between midbrain and hindbrain. Development [Internet]. 1996 Dec [cited 2019 Sep 13];123:179–90. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9007239

320. Rowitch DH, Echelard Y, Danielian PS, Gellner K, Brenner S, McMahon AP. Identification of an evolutionarily conserved 110 base-pair cis-acting regulatory sequence that governs Wnt-1 expression in the murine neural plate. Development [Internet]. 1998 Jul [cited 2019 Sep 13];125(14):2735–46. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9636087

321. Acampora D, Avantaggiato V, Tuorto F, Simeone A. Genetic control of brain morphogenesis through Otx gene dosage requirement. Development [Internet]. 1997 Sep [cited 2019 Sep 13];124(18):3639–50. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9342056

322. Gage PJ, Suh H, Camper SA. The bicoid-related Pitx gene family in development. Mamm Genome. 1999;10(2):197–200.

323. Walton T, Preston E, Nair G, Zacharias AL, Raj A, Murray JI. The Bicoid Class Homeodomain Factors ceh-36/OTX and unc-30/PITX Cooperate in C. elegans Embryonic Progenitor Cells to Regulate Robust Development. PLoS Genet [Internet]. 2015 [cited 2019 Aug 28];11(3). Available from: https://journals.plos.org/plosgenetics/article/file?id=10.1371/journal.pgen.1005003&type=printable

324. Yasuoka Y, Suzuki Y, Takahashi S, Someya H, Sudou N, Haramoto Y, et al. Occupancy of tissue-specific cis-regulatory modules by Otx2 and TLE/Groucho for embryonic head specification. Nat Commun [Internet]. 2014;5:1–14. Available from: http://dx.doi.org/10.1038/ncomms5322

325. Toresson H, Martinez-Barbera JP, Bardsley A, Caubit X, Krauss S. Conservation of BF-1 expression in amphioxus and zebrafish suggests evolutionary ancestry of anterior cell types that contribute to the vertebrate telencephalon. Dev Genes Evol [Internet]. 1998 Oct [cited 2019 Aug 28];208(8):431–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9799423

326. Papalopulu N, Kintner C. A posteriorising factor, retinoic acid, reveals that anteroposterior patterning controls the timing of neuronal differentiation in Xenopus neuroectoderm. Development [Internet]. 1996 Nov [cited 2019 Aug 28];122(11):3409–18. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8951057

327. Roth M, Bonev B, Lindsay J, Lea R, Panagiotaki N, Houart C, et al. FoxG1 and TLE2 act co-operatively to regulate ventral telencephalon formation. Development [Internet]. 2010 [cited 2019 Aug 28];137(9):1553–62. Available from: http://informatics.gurdon.cam.ac.uk/online/xt-fl-db.html

328. Chen G, Courey AJ. Groucho/TLE family proteins and transcriptional repression [Internet]. Vol. 249, Gene. 2000 [cited 2019 Aug 28]. p. 1–16. Available from: www.elsevier.com/locate/gene

329. Zamparini AL, Watts T, Gardner CE, Tomlinson SR, Johnston GI, Brickman JM. Hex acts with β-catenin to regulate anteroposterior patterning via a Groucho-related co-repressor and Nodal. Development [Internet]. 2006 Sep 15 [cited 2019 Aug 28];133(18):3709–22. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16936074

330. Gasperowicz M, Otto F. Mammalian Groucho homologs: Redundancy or specificity? [Internet]. Vol. 95, Journal of Cellular Biochemistry. John Wiley & Sons, Ltd; 2005 [cited 2019 Aug 28]. p. 670–87. Available from: http://doi.wiley.com/10.1002/jcb.20476

331. Hasson P, Paroush Z. Crosstalk between the EGFR and other signalling pathways at the level of the global transcriptional corepressor Groucho/TLE [Internet]. Vol. 94, British Journal of Cancer. Nature Publishing Group; 2006 [cited 2019 Aug 28]. p. 771–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16508633

332. Buscarlet M, Stifani S. The "Marx" of Groucho on development and disease [Internet]. Vol. 17, Trends in Cell Biology. Elsevier; 2007 [cited 2019 Aug 28]. p. 353–61. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17643306

333. Daniels DL, Weis WI. β-catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. Nat Struct Mol Biol [Internet]. 2005 Apr 13 [cited 2019 Aug 28];12(4):364–71. Available from: http://www.nature.com/articles/nsmb912

334. Elliott SA. Studies of conserved cell-cell signaling pathways in the planarian, Schmidtea mediterranea. Thesis. 2015.

335. Valenta T, Hausmann G, Basler K. The many faces and functions of b-catenin. EMBO J. 2012;31(12):2714–36.

336. McCrea PD, Wang W, Jun S, Park J-I, Chen J, Aziz KE, et al. FOXKs Promote Wnt/β-Catenin Signaling by Translocating DVL into the Nucleus. Dev Cell. 2015;32(6):707–18.

337. Panda D, Gold B, Tartell MA, Rausch K, Casas-Tinto S, Cherry S. The transcription factor FoxK participates with Nup98 to regulate antiviral gene expression. Dermody TS, editor. MBio [Internet]. 2015 Apr 7 [cited 2019 Aug 28];6(2). Available from: https://mbio.asm.org/

lookup/doi/10.1128/mBio.02509-14

338. Bowman CJ, Ayer DE, Dynlacht BD. Foxk proteins repress the initiation of starvation-induced atrophy and autophagy programs. Nat Cell Biol [Internet]. 2014 Dec 17 [cited 2019 Aug 28];16(12):1202–14. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25402684

339. Ji ZG, Jiang HT, Zhang PS. FOXK1 promotes cell growth through activating wnt/β-catenin pathway and emerges as a novel target of miR-137 in glioma. Am J Transl Res. 2018;10(6):1784–92.

340. Pohl BS, Knöchel W. Of Fox and Frogs: Fox (fork head/winged helix) transcription factors in Xenopus development [Internet]. Vol. 344, Gene. Elsevier; 2005 [cited 2019 Aug 28]. p. 21–32. Available from: https://www.sciencedirect.com/science/article/abs/pii/S0378111904005967?via%3Dihub

341. Pohl BS, Knöchel W. Isolation and developmental expression of Xenopus FoxJ1 and FoxK1. Dev Genes Evol [Internet]. 2004 Apr 1 [cited 2019 Aug 28];214(4):200–5. Available from: http://link.springer.com/10.1007/s00427-004-0391-7

342. Lai EC. Notch signaling: Control of cell communication and cell fate [Internet]. Vol. 131, Development. The Company of Biologists Ltd; 2004 [cited 2019 Aug 30]. p. 965–73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1769331

343. Hayward P, Kalmar T, Arias AM. Wnt/Notch signalling and information processing during development. Vol. 135, Development. 2008. p. 411–24.

344. Borggrefe T, Lauth M, Zwijsen A, Huylebroeck D, Oswald F, Giaimo BD. The Notch intracellular domain integrates signals from Wnt, Hedgehog, TGFβ/BMP and hypoxia pathways [Internet]. Vol. 1863, Biochimica et Biophysica Acta - Molecular Cell Research. Elsevier; 2016 [cited 2019 Aug 28]. p. 303–13. Available from: https://www.sciencedirect.com/science/article/pii/S0167488915004048

345. Clevers H, Loh KM, Nusse R. An integral program for tissue renewal and regeneration: Wnt signaling and stem cell control [Internet]. Vol. 346, Science. 2014 [cited 2019 Aug 30]. p. 1248012. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25278615

346. Nemir M, Croquelois A, Pedrazzini T, Radtke F. Induction of cardiogenesis in embryonic stem cells via downregulation of Notch1 signaling. Circ Res [Internet]. 2006 Jun 23 [cited 2019 Aug 28];98(12):1471–8. Available from: https://www.ahajournals.org/doi/10.1161/01.RES.0000226497.52052.2a

347. Androutsellis-Theotokis A, Leker RR, Soldner F, Hoeppner DJ, Ravin R, Poser SW, et al. Notch signalling regulates stem cell numbers in vitro and in vivo. Nature [Internet]. 2006 Aug 25 [cited 2019 Aug 28];442(7104):823–6. Available from: http://www.nature.com/articles/nature04940

348. Lindsley RC, Gill JG, Kyba M, Murphy TL, Murphy KM. Canonical Wnt signaling is required for development of embryonic stem cell-derived mesoderm. Development [Internet]. 2006

Oct 1 [cited 2019 Aug 28];133(19):3787–96. Available from: http://www.ncbi.nlm.nih.gov/pub-med/16943279

349. Aubert J, Dunstan H, Chambers I, Smith A. Functional gene screening in embryonic stem cells implicates Wnt antagonism in neural differentiation. Nat Biotechnol [Internet]. 2002 Dec 25 [cited 2019 Aug 28];20(12):1240–5. Available from: http://www.nature.com/articles/nbt763

350. Haegele L, Ingold B, Naumann H, Tabatabai G, Ledermann B, Brandner S. Wnt signa-lling inhibits neural differentiation of embryonic stem cells by controlling bone morpho-genetic protein expression. Mol Cell Neurosci [Internet]. 2003 Nov 1 [cited 2019 Aug 28];24(3):696–708. Available from: https://www.sciencedirect.com/science/article/abs/pii/S104474310300232X?via%3Dihub

351. Kinder SJ, Tsang TE, Wakamiya M, Sasaki H, Behringer RR, Nagy A, et al. The organizer of the mouse gastrula is composed of a dynamic population of progenitor cells for the axial me-soderm. Development [Internet]. 2001 Sep [cited 2019 Aug 28];128(18):3623–34. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11566865

352. Oberhofer G, Grossmann D, Siemanowski JL, Beissbarth T, Bucher G. Wnt/β-catenin sig-naling integrates patterning and metabolism of the insect growth zone. Dev [Internet]. 2014 [cited 2019 Aug 30];141(24):4740–50. Available from: https://dev.biologists.org/content/deve-lop/141/24/4740.full.pdf

353. McGregor AP, Pechmann M, Schwager EE, Feitosa NM, Kruck S, Aranda M, et al. Wnt8 Is Required for Growth-Zone Establishment and Development of Opisthosomal Segments in a Spider. Curr Biol [Internet]. 2008 Oct 28 [cited 2019 Sep 2];18(20):1619–23. Available from: https://www.sciencedirect.com/science/article/pii/S0960982208011238

354. Pourquié O. The chick embryo: A leading model in somitogenesis studies [Internet]. Vol. 121, Mechanisms of Development. Elsevier; 2004 [cited 2019 Sep 2]. p. 1069–79. Available from: https://www.sciencedirect.com/science/article/pii/S0925477304001157

355. Perez-Garijo A, Steller H. Spreading the word: non-autonomous effects of apoptosis during development, regeneration and disease. Development [Internet]. 2015;142(19):3253–62. Available from: http://dev.biologists.org/cgi/doi/10.1242/dev.127878

356. Bergmann A, Steller H. Apoptosis, stem cells, and tissue regeneration. Vol. 3, Science Sig-naling. 2010.

357. Tseng AS, Adams DS, Qiu D, Koustubhan P, Levin M. Apoptosis is required during early stages of tail regeneration in Xenopus laevis. Dev Biol [Internet]. 2007 Jan 1 [cited 2019 Aug 28];301(1):62–9. Available from: https://www.sciencedirect.com/science/article/pii/S001216060601342X?via%3Dihub

358. Bergantinos C, Corominas M, Serras F. Cell death-induced regeneration in wing imaginal discs requires JNK signalling. Development. 2010;137(7):1169–79.

359. Santabárbara-Ruiz P, López-Santillán M, Martínez-Rodríguez I, Binagui-Casas A, Pérez L,

Milán M, et al. ROS-Induced JNK and p38 Signaling Is Required for Unpaired Cytokine Activation during Drosophila Regeneration. Copenhaver GP, editor. PLoS Genet [Internet]. 2015 Oct 23 [cited 2019 Aug 1];11(10):e1005595. Available from: https://dx.plos.org/10.1371/journal.pgen.1005595

360. Chera S, Ghila L, Dobretz K, Wenger Y, Bauer C, Buzgariu W, et al. Apoptotic Cells Provide an Unexpected Source of Wnt3 Signaling to Drive Hydra Head Regeneration. Dev Cell [Internet]. 2009;17(2):279–89. Available from: http://dx.doi.org/10.1016/j.devcel.2009.07.014

361. Tasaki J, Shibata N, Nishimura O, Itomi K, Tabata Y, Son F, et al. ERK signaling controls blastema cell differentiation during planarian regeneration. Development [Internet]. 2011 Jun 15 [cited 2019 Aug 22];138(12):2417–27. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21610023

362. Degnan BM, Vervoort M, Larroux C, Richards GS. Early evolution of metazoan transcription factors. Curr Opin Genet Dev. 2009;19(6):591–9.

363. Paps J, Holland PWH, Shimeld SM. A genome-wide view of transcription factor gene diversity in chordate evolution: Less gene loss in amphioxus? Brief Funct Genomics [Internet]. 2012 [cited 2018 Aug 24];11(2):177–86. Available from: https://academic.oup.com/bfg/article-abstract/11/2/177/213737

364. Hannenhalli S, Kaestner KH. The evolution of Fox genes and their role in development and disease [Internet]. Vol. 10, Nature Reviews Genetics. NIH Public Access; 2009 [cited 2018 Aug 24]. p. 233–40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19274050

365. Mazet F, Yu JK, Liberles DA, Holland LZ, Shimeld SM. Phylogenetic relationships of the Fox (Forkhead) gene family in the Bilateria. Gene [Internet]. 2003 Oct 16 [cited 2018 Aug 24];316(1–2):79–89. Available from: https://www.sciencedirect.com/science/article/pii/S0378111903007418?via%3Dihub

366. Tu Q, Brown CT, Davidson EH, Oliveri P. Sea urchin Forkhead gene family: Phylogeny and embryonic expression. Dev Biol. 2006;300(1):49–62.

367. Yang M, Xu F, Liu J, Que H, Li L, Zhang G. Phylogeny of forkhead genes in three spiralians and their expression in Pacific oyster Crassostrea gigas. Chinese J Oceanol Limnol [Internet]. 2014 [cited 2018 Sep 4];32(6):1207–23. Available from: http://dx.doi.org/10.1007/s00343-015-4009-x

368. Watanabe M, Whitman M. FAST-1 is a key maternal effector of mesoderm inducers in the early Xenopus embryo. Development [Internet]. 1999 [cited 2019 Aug 28]; Available from: https://dev.biologists.org/content/develop/126/24/5621.full.pdf

369. Sullivan JC, Ryan JF, Mullikin JC, Finnerty JR. Conserved and novel Wnt clusters in the basal eumetazoan Nematostella vectensis. Dev Genes Evol [Internet]. 2007 Feb 23 [cited 2019 Aug 28];217(3):235–9. Available from: http://link.springer.com/10.1007/s00427-007-0136-5

370. Cho SJ, Vallès Y, Giani VC, Seaver EC, Weisblat DA. Evolutionary dynamics of the wnt gene family: A lophotrochozoan perspective. Mol Biol Evol [Internet]. 2010 Jul 1 [cited 2019 Aug 28];27(7):1645–58. Available from: https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/msq052

371. Olson PD. Hox genes and the parasitic flatworms: New opportunities, challenges and lessons from the free-living [Internet]. Vol. 57, Parasitology International. 2008 [cited 2019 Aug 28]. p. 8–17. Available from: www.elsevier.com/locate/parint

372. Koziol U, Lalanne AI, Castillo E. Hox genes in the parasitic platyhelminthes Mesocestoides corti, Echinococcus multilocularis, and Schistosoma mansoni: Evidence for a reduced Hox complement. Biochem Genet [Internet]. 2009 [cited 2019 Aug 28];47(1–2):100–16. Available from: https://link-springer-com.sire.ub.edu/content/pdf/10.1007%2Fs10528-008-9210-6.pdf

373. Olson PD, Zarowiecki M, Kiss F, Brehm K. Cestode genomics - progress and prospects for advancing basic and applied aspects of flatworm biology [Internet]. Vol. 34, Parasite Immunology. John Wiley & Sons, Ltd (10.1111); 2012 [cited 2019 Aug 28]. p. 130–50. Available from: http://doi.wiley.com/10.1111/j.1365-3024.2011.01319.x

374. Taylor JS, Raes J. Duplication and Divergence: The Evolution of New Genes and Old Ideas. Annu Rev Genet [Internet]. 2004 Dec 29 [cited 2019 Aug 30];38(1):615–43. Available from: http://www.annualreviews.org/doi/10.1146/annurev.genet.38.072902.092831

375. Lee HH, Frasch M. Survey of Forkhead Domain Encoding Genes in the Drosophila Genome: Classification and Embryonic Expression Patterns. Dev Dyn. 2004;229(2):357–66.

376. Clifton-Bligh RJ, Wentworth JM, Heinz P, Crisp MS, John R, Lazarus JH, et al. Mutation of the gene encoding human TTF-2 associated with thyroid agenesis, cleft palate and choanal atresia. Nat Genet. 1998;19(4):399–401.

377. Macchia PE, Mattei M-G, Lapi P, Fenzi G, Lauro R Di. Cloning, chromosomal localization and identification of polymorphisms in the human thyroid transcription factor 2 gene (TITF2) [Internet]. [cited 2019 Aug 28].

378. Macchia PE, Mattei MG, Lapi P, Fenzi G, Di Lauro R. Cloning, chromosomal localization and identification of polymorphisms in the human thyroid transcription factor 2 gene (TITF2). Biochimie [Internet]. 1999 [cited 2019 Aug 28];81(5):433–40. Available from: http://genetics.nature.com

379. Semina E V. Mutations in the human forkhead transcription factor FOXE3 associated with anterior segment ocular dysgenesis and cataracts. Hum Mol Genet [Internet]. 2001 Feb 1 [cited 2019 Aug 28];10(3):231–6. Available from: https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/10.3.231

380. Eckmiller M, Steinberg R. Investigative Ophthalmology & Visual Science. Invest Ophthalmol [Internet]. 1981 May 1 [cited 2019 Aug 28];21(5):3. Available from: https://iovs.arvojournals.org/article.aspx?articleid=2123145

References

381. Brownell I, Dirksen M, Jamrich M. Forkhead Foxe3 maps to the dysgenetic lens locus and is critical in lens development and differentiation. Genesis [Internet]. 2000 Jun 1 [cited 2019 Aug 28];27(2):81–93.

382. Blixt A, Mahlapuu M, Aitola M, Pelto-Huikko M, Enerbäck S, Carlsson P. A forkhead gene, FoxE3, is essential for lens epithelial proliferation and closure of the lens vesicle. Genes Dev [Internet]. 2000 Jan 15 [cited 2019 Aug 28];14(2):245–54. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10652278

383. Farlie PG, Davidson NM, Baker NL, Raabus M, Roeszler KN, Hirst C, et al. Co-option of the cardiac transcription factor Nkx2.5 during development of the emu wing. Nat Commun [Internet]. 2017 Dec 25 [cited 2019 Sep 14];8(1):132. Available from: http://www.nature.com/articles/s41467-017-00112-7

384. Mohamad Ishak NS, Nong QD, Matsuura T, Kato Y, Watanabe H. Co-option of the bZIP transcription factor Vrille as the activator of Doublesex1 in environmental sex determination of the crustacean Daphnia magna. Desplan C, editor. PLoS Genet [Internet]. 2017 Nov 2 [cited 2019 Sep 14];13(11):e1006953. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29095827

385. Mazet F, Amemiya CT, Shimeld SM. An ancient Fox gene cluster in bilaterian animals. Curr Biol. 2006;16(9).

386. Hacker U, Grossniklaust U, Gehring WJ, Jäckle H. Developmentally regulated Drosophila gene family encoding the fork head domain. Proc Natl Acad Sci U S A [Internet]. 1992 Sep 15 [cited 2019 Aug 28];89(18):8754–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1356269

387. Kaestner KH, Bleckmann SC, Paula Monaghan A, Schlöndorff J, Mincheva A, Lichter P, et al. Clustered arrangement of winged helix genes fkh-6 and MFH-1: Possible implications for mesoderm development. Development. 1996;122(6):1751–8.

388. Kaestner KH, Silberg DG, Traber PG, Schütz G. The mesenchymal winged helix transcription factor Fkh6 is required for the control of gastrointestinal proliferation and differentiation. Genes Dev. 1997;11(12):1583–95.

389. Aoki R, Shoshkes-Carmel M, Gao N, Shin S, May CL, Golson ML, et al. Foxl1-Expressing Mesenchymal Cells Constitute the Intestinal Stem Cell Niche. CMGH [Internet]. 2016 [cited 2019 Aug 28];2(2):175–88. Available from: http://dx.doi.org/10.1016/j.jcmgh.2015.12.004

390. Häcker U, Kaufmann E, Hartmann C, Jürgens G, Knöchel W, Jäckle H. The Drosophila fork head domain protein crocodile is required for the establishment of head structures. EMBO J [Internet]. 1995 Nov 1 [cited 2019 Aug 28];14(21):5306–17. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7489720

391. Topczewska JM, Topczewski J, Solnica-Krezel L, Hogan BL. Sequence and expression of zebrafish foxc1a and foxc1b, encoding conserved forkhead/winged helix transcription factors. Mech Dev [Internet]. 2001 Feb [cited 2019 Aug 28];100(2):343–7. Available from: http://

www.ncbi.nlm.nih.gov/pubmed/11165495

392. Topczewska JM, Topczewski J, Shostak A, Kume T, Solnica-Krezel L, Hogan BLM. The winged helix transcription factor Foxc1a is essential for somitogenesis in zebrafish. Genes Dev [Internet]. 2001 Sep 15 [cited 2019 Aug 28];15(18):2483–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11562356

393. Scheucher M, Dege P, Lef J, Hille S, Knöchel W. Transcription patterns of four different fork head/HNF-3 related genes (XFD-4, 6, 9 and 10) in Xenopus laevis embryos. Roux's Arch Dev Biol [Internet]. 1995 Jan [cited 2019 Aug 28];204(3):203–11. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28305961

394. Köster M, Dillinger K, Knöchel W. Activin A signaling directly activates Xenopus winged helix factors XFD-4/4', the orthologues to mammalian MFH-1. Dev Genes Evol [Internet]. 2000 May 19 [cited 2019 Aug 28];210(6):320–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11180837

395. Buchberger A, Schwarzer M, Brand T, Pabst O, Seidl K, Arnold HH. Chicken winged-helix transcription factor cFKH-1 prefigures axial and appendicular skeletal structures during chicken embryogenesis. Dev Dyn [Internet]. 1998 May 1 [cited 2019 Aug 28];212(1):94–101.

396. Wilm B, James RG, Schultheiss TM, Hogan BLM. The forkhead genes, Foxc1 and Foxc2, regulate paraxial versus intermediate mesoderm cell fate. Dev Biol [Internet]. 2004 Jul 1 [cited 2019 Aug 28];271(1):176–89. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15196959

397. Kume T, Jiang H, Topczewska JM, Hogan BLM. The murine winged helix transcription factors, Foxc1 and Foxc2, are both required for cardiovascular development and somitogenesis. Genes Dev [Internet]. 2001 Sep 15 [cited 2019 Aug 28];15(18):2470–82. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11562355

398. Hope IA, Mounsey A, Bauer P, Aslam S. The forkhead gene family of Caenorhabditis elegans. Gene [Internet]. 2003 Jan 30 [cited 2019 Aug 28];304(1–2):43–55. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12568714

399. Pérez Sánchez C, Casas-Tintó S, Sánchez L, Rey-Campos J, Granadino B. DmFoxF, a novel Drosophila fork head factor expressed in visceral mesoderm. Mech Dev [Internet]. 2002 Feb [cited 2019 Aug 28];111(1–2):163–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11804790

400. Huch M, Knoblich JA, Lutolf MP, Martinez-Arias A. The hope and the hype of organoid research. Dev [Internet]. 2017 [cited 2019 Aug 30];144(6):938–41. Available from: https://dev.biologists.org/content/develop/144/6/938.full.pdf

401. Fernandéz-Taboada E, Moritz S, Zeuschner D, Stehling M, Schöler HR, Saló E, et al. Smed-SmB, a member of the LSm protein superfamily, is essential for chromatoid body organization and planarian stem cell proliferation. Development [Internet]. 2010 Apr 1 [cited 2014 Aug 27];137(7):1055–65. Available from: http://dev.biologists.org/content/137/7/1055.long

References

402. Currie KW, Brown DDR, Zhu S, Xu CJ, Voisin V, Bader GD, et al. HOX gene complement and expression in the planarian Schmidtea mediterranea. Evodevo [Internet]. 2016 [cited 2019 Aug 19];7(1):7. Available from: www.geneious.

403. Cardona A, Fernández J, Solana J, Romero R. An in situ hybridization protocol for planarian embryos: Monitoring myosin heavy chain gene expression. Dev Genes Evol. 2005;215(9):482–8.

404. Ross KG, Omuro KC, Taylor MR, Munday RK, Hubert A, King RS, et al. Novel monoclonal antibodies to study tissue regeneration in planarians. BMC Dev Biol [Internet]. 2015;15. Available from: http://www.biomedcentral.com/1471-213X/15/2

405. González-Estévez C, Felix DA, Aboobaker AA, Saló E. Gtdap-1 promotes autophagy and is required for planarian remodeling during regeneration and starvation. Proc Natl Acad Sci U S A [Internet]. 2007 Aug 14 [cited 2014 Sep 4];104(33):13373–8. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1948951&tool=pmcentrez&rendertype=abstract

406. Eckelt K. Multi-approach analysis for identification and functional characterization of eye regeneration related genes of Schmidtea mediterranea. University of Barcelona, Barcelona, Spain; 2011.

407. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol [Internet]. 2019 Apr 18 [cited 2019 Jul 22];37(4):420–3. Available from: http://www.nature.com/articles/s41587-019-0036-z

408. Lupas A, Van Dyke M, Stock J. Predicting coiled coils from protein sequences. Science (80-) [Internet]. 1991 May 24 [cited 2019 Jul 22];252(5009):1162–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2031185

409. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics [Internet]. 2009 May 1 [cited 2019 Jul 19];25(9):1189–91. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp033

410. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol [Internet]. 2009 Mar 4 [cited 2019 Jul 19];10(3):R25. Available from: http://genomebiology.biomedcentral.com/articles/10.1186/gb-2009-10-3-r25

411. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics [Internet]. 2009 Aug 15 [cited 2019 Jul 19];25(16):2078–9. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp352

412. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat Biotechnol [Internet]. 2011 Jan 10 [cited 2019 Jul 19];29(1):24–6.

Available from: http://www.nature.com/articles/nbt.1754

413. Trifinopoulos J, Nguyen L-T, von Haeseler A, Minh BQ. W-IQ-TREE: a fast online phyloge-
      netic tool for maximum likelihood analysis. Nucleic Acids Res [Internet]. 2016 Jul 8 [cited
      2019 Jul 19];44(W1):W232–5. Available from: https://academic.oup.com/nar/article-lookup/
      doi/10.1093/nar/gkw256

414. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast uni-
      versal RNA-seq aligner. Bioinformatics [Internet]. 2013 Jan 1 [cited 2019 Sep 2];29(1):15–
      21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23104886

415. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential ex-
      pression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res [Internet].
      2015 Apr 20 [cited 2019 Sep 2];43(7):e47. Available from: http://academic.oup.com/nar/arti-
      cle/43/7/e47/2414268/limma-powers-differential-expression-analyses-for

416. Alexa A RJ. topGO: Enrichment Analysis for Gene Ontology. R package version 2.36.0.
      2019.

417. Wu M GL. TCseq: Time course sequencing data analysis. R package version 1.8.0. 2019.

418. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A method for assaying chroma-
      tin accessibility genome-wide. Curr Protoc Mol Biol [Internet]. 2015 Jan 5 [cited 2019 Sep
      17];2015:21.29.1-21.29.9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25559105

419. Lukoseviciute M, Gavriouchkina D, Williams RM, Hochgreb-Hagele T, Senanayake U,
      Chong-Morrison V, et al. From Pioneer to Repressor: Bimodal foxd3 Activity Dynamically
      Remodels Neural Crest Regulatory Landscape In Vivo. Dev Cell [Internet]. 2018 Dec 3
      [cited 2019 Sep 17];47(5):608-628.e6. Available from: http://www.ncbi.nlm.nih.gov/pub-
      med/30513303

420. Schmidl C, Rendeiro AF, Sheffield NC, Bock C. ChIPmentation: Fast, robust, low-input ChIP-
      seq for histones and transcription factors. Nat Methods [Internet]. 2015 Oct 17 [cited 2019
      Sep 17];12(10):963–5. Available from: http://www.nature.com/articles/nmeth.3542

421. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native
      chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding
      proteins and nucleosome position. Nat Methods [Internet]. 2013 Dec 6 [cited 2019 Sep
      17];10(12):1213–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24097267

422. Stark R BG. Bioconductor - DiffBind [Internet]. 2011 [cited 2019 Sep 17]. Available from:
      https://bioconductor.org/packages/release/bioc/html/DiffBind.html

423. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, et al. The Pfam pro-
      tein families database in 2019. Nucleic Acids Res [Internet]. 2019 Jan 8 [cited 2019
      Sep 15];47(D1):D427–32. Available from: https://academic.oup.com/nar/article/47/D1/
      D427/5144153

424. Kosugi S, Hasebe M, Tomita M, Yanagawa H. Systematic identification of cell cycle-depen-

dent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. Proc Natl Acad Sci [Internet]. 2009 Jun 23 [cited 2019 Aug 5];106(25):10171–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19520826

425. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform [Internet]. 2017 Sep 6 [cited 2019 Sep 15]; Available from: http://academic.oup.com/bib/article/doi/10.1093/bib/bbx108/4106928/MAFFT-online-service-multiple-sequence-alignment

426. Hemmati-Brivanlou A, Melton D. Vertebrate Embryonic Cells Will Become Nerve Cells Unless Told Otherwise. Cell [Internet]. 1997;88(1):13–7. Available from: http://linkinghub.elsevier.com/retrieve/pii/S009286740081853X

427. Gehart H, Clevers H. Tales from the crypt: new insights into intestinal stem cells. Nat Rev Gastroenterol Hepatol [Internet]. 2019;16(1):19–34. Available from: http://dx.doi.org/10.1038/s41575-018-0081-y

428. Telford MJ, Budd GE, Philippe H. Phylogenomic insights into animal evolution. Curr Biol. 2015;25(19):R876–87.

429. Ivankovic M, Haneckova R, Thommen A, Grohme MA, Vila-Farré M, Werner S, et al. Model systems for regeneration: planarians. Development [Internet]. 2019 Sep 1;146(17):dev167684. Available from: http://dev.biologists.org/lookup/doi/10.1242/dev.167700

430. Klemm SL, Shipony Z, Greenleaf WJ. Chromatin accessibility and the regulatory epigenome. Nat Rev Genet [Internet]. 2019;20(April):29–35. Available from: http://dx.doi.org/10.1038/s41576-018-0089-8

431. Dattani A, Sridhar D, Aziz Aboobaker A. Planarian flatworms as a new model system for understanding the epigenetic regulation of stem cell pluripotency and differentiation. Semin Cell Dev Biol [Internet]. 2019;87:79–94. Available from: https://doi.org/10.1016/j.semcdb.2018.04.007

432. Abcam.com. Epigenetics applications - ATAC-seq | Abcam [Internet]. [cited 2019 Sep 15]. Available from: https://www.abcam.com/epigenetics/epigenetics-application-spotlight-atac-seq

433. Furey TS. ChIP-seq and beyond: New and improved methodologies to detect and characterize protein-DNA interactions [Internet]. Vol. 13, Nature Reviews Genetics. Nature Publishing Group; 2012 [cited 2019 Aug 6]. p. 840–52. Available from: http://www.nature.com/articles/nrg3306

434. Paps J, Baguñà J, Riutort M. Lophotrochozoa internal phylogeny: New insights from an up-to-date analysis of nuclear ribosomal genes. Proc R Soc B Biol Sci. 2009;276(1660):1245–54.

2.   Müller WA, Müller WA. Model Organisms in Developmental Biology. In: Developmental Biology [Internet]. New York, NY: Springer New York; 1997 [cited 2019 Sep 13]. p. 21–121. Available from: http://link.springer.com/10.1007/978-1-4612-2248-4_3

3.   Jameson JL, DeGroot LJ, De Kretser DM (David M., Giudice L, Grossman A, Melmed S, et al. Endocrinology : adult and pediatric [Internet]. [cited 2019 Sep 10]. Available from: https://www.sciencedirect.com/book/9780323189071/endocrinology-adult-and-pediatric

4.   Galliot B, Ghila L. Cell plasticity in homeostasis and regeneration. Mol Reprod Dev. 2010;77(10):837–55.

5.   Pellettieri J, Alvarado AS. Cell Turnover and Adult Tissue Homeostasis: From Humans to Planarians. Annu Rev Genet [Internet]. 2007;41(1):83–105. Available from: http://www.annualreviews.org/doi/abs/10.1146/annurev.genet.41.110306.130244

6.   Rosenfeld CS. The Epigenome and Developmental Origins of Health and Disease. The Epigenome and Developmental Origins of Health and Disease. 2015. 1–542 p.

7.   Cook SF, Bies RR. Disease Progression Modeling: Key Concepts and Recent Developments [Internet]. Vol. 2, Current Pharmacology Reports. Springer International Publishing; 2016 [cited 2019 Sep 9]. p. 221–30. Available from: http://link.springer.com/10.1007/s40495-016-0066-x

8.   Sims-Lucas S, Good M, Vainio SJ. Editorial: Organogenesis: From development to disease. Frontiers in Cell and Developmental Biology [Internet]. 2017 Sep 20 [cited 2019 Sep 9];5(SEP):85. Available from: http://journal.frontiersin.org/article/10.3389/fcell.2017.00085/full

9.   Miller CJ, Davidson LA. The interplay between cell signalling and mechanics in developmental processes [Internet]. Vol. 14, Nature Reviews Genetics. 2013 [cited 2019 Sep 10]. p. 733–44. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24045690

10.   Wilson MR, Close TW, Trosko JE. Cell population dynamics (apoptosis, mitosis, and cell-cell communication) during disruption of homeostasis. Exp Cell Res [Internet]. 2000 Feb 1 [cited 2019 Sep 8];254(2):257–68. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10640424

11.   Charras G, Sahai E. Physical influences of the extracellular environment on cell migration. Nat Rev Mol Cell Biol [Internet]. 2014 Oct 30 [cited 2014 Oct 30];(Box 1). Available from: http://www.nature.com/doifinder/10.1038/nrm3897

12.   Sluys R, Riutort M. Planarian Diversity and Phylogeny. In 2018. p. 1–56. Available from: http://link.springer.com/10.1007/978-1-4939-7802-1_1

13.   Miyaoka Y, Ebato K, Kato H, Arakawa S, Shimizu S, Miyajima A. Hypertrophy and Unconventional Cell Division of Hepatocytes Underlie Liver Regeneration. Curr Biol [Internet]. 2012;22(13):1166–75. Available from: http://dx.doi.org/10.1016/j.cub.2012.05.016

ANNEXES

Annex I – Nucleotide and amioacidic sequences of bls genes.

Schmidtea mediterranea (*Smed*) *bls* members GenGank IDs and their scaffol distribution.

ORF of the *bls* genes in *Schmidtea mediterranea*. Amino acid sequence analysis of *Smed* BLS protein. The presence of the signal peptide (SP) and coiled coild (CC) were indicated

Species	Gene	Scaffold	Accession Number
Smed	*bls1a*	54	BK010973
Smed	*bls1b*	54	BK010974
Smed	*bls2a*	54	BK010975
Smed	*bls2b*	54	BK010976
Smed	*bls3a*	54	BK010977
Smed	*bls3b*	54	BK010978
Smed	*bls3c*	54	BK010979
Smed	*bls3d*	54	BK010980
Smed	*bls3e*	54	BK010981
Smed	*bls3f*	54	BK010982
Smed	*bls3g*	54	BK010983
Smed	*bls4a*	49	BK010984
Smed	*bls4b*	49	BK010985

>BLS2A

MNLKLSLLILSCFACMYVNGILGLRLLTGLKLNVDGLIKADLGLRLGLYLGAGNYRSVLEI-PAINNFLLGLRARVAGPIYAEVKAHLEEIGQISSGMYGLGIDRETVDNLVHHIRQLERRRYE-LEALRRKYFRSQILIDYLVKLLKIKM

>BLS2B

MNLKLSLLILSCFACMYVNGILGLRLLTGLKLNVDGLIKADLGLRLGLYLGAGNYRSVLEI-PAINNFLLGLRARVAGPIYAEVKAHLEEIGQISSGMYGLGIDRETVDNLVHHIRQLERRRYE-LEALRRKYFRSQNIDRLFSEAIKN

>BLS3A

MNLKLSLLILSCFTCMYVNGFVDLRLLTGLKLNVAGLIKTDLGLRLGLYLGAGYYRSVLEI-PAINHFLIGLRARVSGPIYANVQAHLEDIGKISSGIYGLGIDREKINRLVYQIRKLERRRYE-LEALRRTHFPDQDVNKVFEKCCVEVVEESS

>BLS3B

MNLKLSLLILSCFTCMYVNGFVDLRLLTGLKLNVAGLIKADLGLRLGLYLGAGYYRSVLEI-PAINHFLIGLRARVSGPIYANVQAHLEDIGKISSGIYGLGIDREKINRLVYQIRKLERRRYE-LEALRRTHFPDQDVNKVFEKMLCGSCRRIE

>BLS3C

MNLKLSLLILSCFTCMYVNGFVDLRLLTGLKLNVAGLIKADLGLRLGLYLGAGYYRSVLEI-PAINHFLIGLRARVSGPIYANVQAHLEDIGKISSGIYGLALTEKRSTD

>BLS3D

MHVCEWFCGSQTFDGTEIECGGLIKADLGLRLGLYLGAGYYRSVLEIPAINHFLIGLRAR-VSGPIYANVQAHLEDIGKISSGIYGLGIDREKINRLVYQIRKLERRRYELEALRRTHFPDQD-VNKVFEKCCVEVVEESSNAK

>BLS3E

MNLKLSLLILSCFTCMYVNGFVDLRLLTGLKLNVAGLIRLTWD

>BLS3F

MNLKLSLLILSCFTCMYVNGFVDLRLLTGLKLNVAGLIKADLGLRLGLYLGAGYYRSVLEI-
PAINHFLIGLRARVSGPIYANVQAHLEDIGKISSGIYGLGIDRERSTD

>BLS3G

MNLKLSLLILSCFTCMYVNGFVDLRLLTGLKLNVAGLIKADLGLRLGLYLGAGYYRSVLEI-
PAINHFLIGLRARVSGPIYANVQAHLEDIGKISSGIYGLGIDREKINRLVYQIRKLERRRYE-
LEALRRTHFPDQDVNKVFENVVWKLSKNRVMQNDLVNKCEIFHIYSK

>BLS5A

MNLKLSLLILSCFACMYLNGVVGLGLLTGLRVNVDGLIDVDLGLALGLKLGAGNYRSVLEI-
PAINHFLLGLRARLPGAIYAKVQARLEYMGRISSRIYGLGINQETVNNLVYYIRQLERRRYE-
LEALRRRYFRNQNIDEIFRAVNKK

>BLS5B

MNLKLSLLILSCFACMYLNGVVGLGLLTGLRVQARLEYMGRISSRIYGLGINQETVNNL-
VYYIRQLERRRYELEALRRRYFRNQNIDEIFRAVNKK

Nucleotide sequences of *bls* homologs in Tricladida species. Djap, *Dugesia japonica*. Spol, *Schmidtea polychroa*. Smes, *Schmidtea mediterranea* sexual strain.

>dd_Djap_v4_77219_2_1

TCGTAAGGCTTCTAATTCGTATCGTCTTCGTTCAAGTTTTCTTATTTGATAAACTAATCT-
GTTGATCTTTTCTCTGTCAATGCCTAAACCGTAAATTCCACTTGATATTTTTCCAATG-
TCTTCCAAATGTGCTTGTACATTTGCATAGATTGGACCCGATACACGGGCCCTGAGTC-
CAAGTAGAAAATGATTGATTGCCGGGATCTCTAATACTGATCGATAATTGCCGGCTCC-
CAAGTACAGACCCAATCTTAATCCCAAGTCAGCCTTAATGAGGCCCTCCACATTCA-
ATTTCAGTCCCGTCAAAAGTCTGAGATCCACAAAACCATTCACATACATGCATGTGAAG-
CAAGAAAGTATTAAAAGAG

>dd_Djap_v4_77219_1_1

GTCATCGTCAAGGTTTAACATTTTTCATTTTCACGTTTTACTTTTTATTAACTGCAC-
GAAATATTTCATCAATATTTTGATTTCTGAAATATCTTCTTCGTAAGGCTTCTAATTCG-
TATCTTCTTCGTTCAAGTTGTCTAATGTAATTAACTAGGTTGTTGACCGTTTCTTGG-
TTAATTCCTAAACCGTAAATCCTACTTGATATTCGTCCCATGAGTTCCAAACGGGCTT-
GTACCTTTGCATAGATTGCACCCGGTAGACGGGCTCTGAGTCCAAGTAAAAAATGATT-
GATTGCCGGGATCTCTAATACTGATCGATAATTGCCGGCTCCCAATTTCAGACCCAAT-
GCTAATCCCAAGTCAACATCAATGAGGCCGTTCACATTCACTCTTAATCCCGTCAAAAG-
TCCGAGACCCACAACACTATTCAGGTACATGCATGCGAAGCAAGAAAGTATTAAAAGA-
GATAGTTTCAAATTCATGCATGTACCTGA

>dd_Spol_v4_8725_2_1

TTTTAATACTTTCATGCTTCGCATGCATGTACGTGAATGGTATTCTGGGTCTCAGA-
CTTTTGACGGGACTGAAATTGAATGTGGAGGGCCTCATTAAGGCTGACTTGG-
GATTAAGATTGGGTCTGTACTTGGGAGCCGGCAATTATCGATCAGTATTAGAGATCC-
CGGCAATCAATCATTTTCTACTTGGACTCAGGGCCCGTGTATCGGGTCCAATCTATGCA-
AATGTACAAGCACATTTGGAAGACATTGGAAAAATATCAAGTGGAATTTACGGTTTAGG-
CATTGACAGAGAAAGATCAACAGATTAGTTTATCAAATAAGAAAACTTGAACGAAGAC-
GATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAAGATGTAAACAAGTG-
TTTGAAAAATGTTGTGTGGAAGTTGTCGAAGAATCGAGTAATGAAAAATGATTTAGTGA-
ATAAATGTGAAATTTTTC

>dd_Spol_v4_7617_2_1
CTTTGCTAACAAGATCATTTATTGATGATAGATATGCTAAATTTCACAATTCATTCACT-
CAATCATTTTGCATTATTTGATTCTTCAACTACTTCCACACAACATTTTTCAAAAACTT-
CATTTACATCTTTTCTCGGGAAATGAGTTTTTTGTAAGGCTTCTAATTCGTATCG-
TCTTCGTTCAAGTCTTCTAATGTAATCCACTAAGATGTTGATGATTTTTCTATCAATACC-
GAAACTGTAAATTCTTCTTGATAGTCGTCCCATGTCTTTGAATATGTTTTTGACTTTAG-
CATATATTCTAATCGGTAGTTGGGCTCTGATTCCAAGTAGAAAATTTTTGATTTCCGG-
GATCTGTAATACGGATCCAAAATGGCCTGCTCCCAATTTCAAACCCAATCCTAATCCTA-
TATCAGCTTTAACAATATCACATACATTCAGTTTCAGTCCAGCCAAAAGATTCAAAAGAC-
CATTCACATACAAGCATGTGAAGCAAGAAAGTATTAAAAGAGATAATTTCAAATTCATAT-
CACATACATTCAGTTTCAG

>dd_Spol_v4_7617_1_1
CTTTGCTAACAAGATCATTTATTGATGATAGATATGCTAAATTTCACAATTCATTCACT-
CAATCATTTTGCATTATTTGATTCTTCAACTACTTCCACACAACATTTTTCAAAAACTT-
CATTTACATCTTTTCTCGGGAAATGAGTTTTTTGTAAGGCTTCTAATTCGTATCG-
TCTTCGTTCAAGTCTTCTAATGTAATCCACTAAGATGTTGATGATTTTTCTATCAATACC-
GAAACTGTAAATTCTTCTTGATAATCGTCCCATGTCTTTGAATATTGTTTTGACTTTAG-
CATATATTCTAATCGGTAGTTGGGCTCTGATTCCAAGTAGAAAATTTTTGATTTCCGG-
GATCTGTAATACGGATCCAAAATGGCCTGCTCCCAATTTCAAACCCAATCCTAATCCTA-
TATCAGCTTTAACAATATCACATACATTCAGTTTCAGTCCAGCCAAAAGATTCAAAAGAC-
CATTCACATACAAGCATGTGAAGCAAGAAAGTATTAAAAGAGATAATTTCAAATTCATAT-
CACATACATTCAGTTTCAG

>dd_Spol_v4_8725_3_1
CTTGGACTCAGAGCCCGTCTACCGGGTGCAATCTATGCAAAGGTACAAGCCCGTTT-
GGAACTCATGGGACGAATATCAAGTAGGATTTACGGTTTAGGAATTAACCAAGA-
AACGGTCAACAACCTAGTTAATTACATTAGACAACTTGAACGAAGAAGATACGAATTA-
GAAGCCTTACGAAGAAGATATTTCAGAAATCAAAATATTGATGAAATATTTCGTGCAG-
TTAATAAAAAGTAAAACGTGAAAATAAAAAATGTTAAACCTTGACGATCACTTTCA-
AATAAAATATA

>dd_Spol_v4_8725_1_1
CCGGTAGACGGGCTCTGAGTCCAAGTAAAAAATGATTGATTGCCGGGATCTCAGA-
GCCCGTCTACCGGGTGCAATCTATGCAAAGGTACAAGCCCGTTTGGAACTCATGGGA-
CGAATATCAAGTAGGATTTACGGTTTAGGAATTAACCAAGAAACGGTCAACAACCTAG-
TTAATTACATTAGACAACTTGAACGAAGAAGATACGAATTAGAAGCCTTACGAAGAAGA-
TATTTCAGAAATCAAAATATTGATGAAATATTTCGTGCAGTTAATAAAAAGTAAAACGT-
GAAAATAAAAAATGTTAAACCTTGACGATCACTTTCAAATAAAATA

>dd_Smes_v1_35718_1_1
T T T T T T T T T A T T A A T G C A T A T A T T T T A T T T G A A A G T G A T C G T C A A G G -
TTTAACATTTTTTATTTTCACGTTTTACTTTTTATTAACTGCACGAAATATTTCATCA-
ATATTTTGATTTCTGAAATATCTTCTTCGTAAGGCTTCTAATTCGTATCTTCTTCGTTCA-
AGTTGTCTAATGTAATAAACTAGGTTGTTGACCGTTTCTTGGTTAATTCCTAAACCG-
TAAATCCTACTTGATATTCGTCCCATGTATTCCAAACGGGCTTGTACCTTTGCATAGATT-
GCACCCGGTAGACGGGCTCTGAGTCCAAGTAAAAAATGATTGATTGCCGGGATCTC-
TAATACTGATCGATAATTGCCGGCTCCCAATTTCAGACCCAATGCTAATCCCAAGTCA-
ACATCAATGAGGCCGTCCACATTCACTCTTAATCCCGTCAAAAGTCCGAGACCCACAA-
CACCATTCAGGTACATGCATGC

>dd_Smes_v1_35718_1_2

T T T T T T T T T T T A A T T A A A G C A T A T T T T A T T T G A T A G T G A T C G T C A-
AGATTAAACTTTTTTAATTTTCACATTTTAATTTTTAATAGCTTCACTAAATAATCTATCA-
ATATTTTGACTTCTGAAATATTTTCTTCGTAAGGCTTCTAATTCGTATCTTCTTCGTTCA-
AGTTGTCTAATGTAATAAACTAGGTTGTTGACCGTTTCTTGGTTAATTCCTAAACCG-
TAAATCCTACTTGATATTCGTCCCATGTATTCCAAACGGGCTTGTACCTTTGCATAGATT-
GCACCCGGTAGACGGGCTCTGAGTCCAAGTAAAAATGATTGATTGCCGGGATCTC-
TAATACTGATCGATAATTGCCGGCTCCCAAGTACAGACCCAATCTTAATCCCAAGTCA-
GCCTTAATGAGGCCATCCACATTCAATTTCAGTCCCGTCAAAAGTCTGAGATCCACA-
AAACCATTCACATACATGCATGT

ORF of the *bls* homologs in Tricladida species. Amino acid sequence analysis of Tricladida BLS protein homologs. The presence of the signal peptide (SP) and coiled coild (CC) were indicated. Djap, *Dugesia japonica*. Spol, *Schmidtea polychroa*. Smes, *Schmidtea mediterranea* sexual strain.

>dd_Djap_v4_77219_2_1
MYVNGFVDLRLLTGLKLNVEGLIKADLGLRLGLYLGAGNYRSVLEIPAINHFLLGLRARVSG-
PIYANVQAHLEDIGKISSGIYGLGIDREKINRLVYQIRKLERRRYELEALR
>dd_Djap_v4_77219_1_1
MNLKLSLLILSCFACMYLNSVVGLGLLTGLRVNVNGLIDVDLGLALGLKLGAGNYRSVLEI-
PAINHFLLGLRARLPGAIYAKVQARLELMGRISSRIYGLGINQETVNNLVNYIRQLERRRYE-
LEALRRRYFRNQNIDEIFRAVNKK
>dd_Spol_v4_8725_2_1
MYVNGILGLRLLTGLKLNVEGLIKADLGLRLGLYLGAGNYRSVLEIPAINHFLLGLRARVS-
GPIYANVQAHLEDIGKISSGIYGLGIDREKINRLVYQIRKLERRRYELEALRRTHFPDQD-
VNKVFEKCCVEVVEESSNEK
>dd_Spol_v4_7617_2_1
MNLKLSLLILSCFTCLYVNGLLNLLAGLKLNVCDIVKADIGLGLGLKLGAGHFGSVLQI-
PEIKNFLLGIRAQLPIRIYAKVKNIFKDMGRLSRRIYSFGIDRKIINILVDYIRRLERRRYELEAL-
QKTHFPRKDVNEVFEKCCVEVVEESNNAK
>dd_Spol_v4_7617_1_1
MNLKLSLLILSCFTCLYVNGLLNLLAGLKLNVCDIVKADIGLGLGLKLGAGHFGSVLQI-
PEIKNFLLGIRAQLPIRIYAKVKTIFKDMGRLSRRIYSFGIDRKIINILVDYIRRLERRRYELEAL-
QKTHFPRKDVNEVFEKCCVEVVEESNNAK
>dd_Spol_v4_8725_3_1
MGRISSRIYGLGINQETVNNLVNYIRQLERRRYELEALRRRYFRNQNIDEIFRAVNKK
>dd_Spol_v4_8725_1_1
MIDCRDLRARLPGAIYAKVQARLELMGRISSRIYGLGINQETVNNLVNYIRQLERRRYE-
LEALRRRYFRNQNIDEIFRAVNKK
>dd_Smes_v1_35718_1_1
MYLNGVVGLGLLTGLRVNVDGLIDVDLGLALGLKLGAGNYRSVLEIPAINHFLLGLRARLP-
GAIYAKVQARLEYMGRISSRIYGLGINQETVNNLVYYIRQLERRRYELEALRRRYFRNQNI-
DEIFRAVNKK
>dd_Smes_v1_35718_1_2
MYVNGFVDLRLLTGLKLNVDGLIKADLGLRLGLYLGAGNYRSVLEIPAINHFLLGLRARLP-
GAIYAKVQARLEYMGRISSRIYGLGINQETVNNLVYYIRQLERRRYELEALRRKYFRSQNI-
DRLFSEAIKN

Alignments using *bls* gene members together with riboprobes sequences used in this study.
Forward primers are in orange and reverse are in blue.

```
bls2a/1-504     AGTTT-TTTTCAGTCATGAATTTGAAATTATCTCTTTTAATACTTTCATGCTTCGCATGC
bls2b/1-472     AGTTT-TTTTCAGTCATGAATTTGAAATTATCTCTTTTAATACTTTCATGCTTCGCATGC
bls2_p1/1-204   ------------------------------------------------------------C
bls2_p2/1-116   ------------------------------------------------------------
bls3a/1-475     AGTTGATTTTCAGTCATGAATTTGAAATTATCTCTTTTAATACTTTCTTGCTTCACATGC
bls3b/1-487     AGTTGATTTTCAGTCATGAATTTGAAATTATCTCTTTTAATACTTTCTTGCTTCACATGC
bls3c/1-487     AGTTGATTTTCAGTCATGAATTTGAAATTATCTCTTTTAATACTTTCTTGCTTCACATGC
bls3d/1-484     AGTTGATTTTCAGTCATGAATTTGAAATTATCTC-TTAATACTTTCTTGCTTCACATGC
bls3e/1-484     AGTTGATTTTCAGTCATGAATTTGAAATTATCTCTTTTAATACTTTCTTGCTTCACATGC
bls3f/1-485     AGTTGATTTTCAGTCATGAATTTGAAATTATCTCTTTTAATACTTTCTTGCTTCACATGC
bls3g/1-485     AGTTGATTTTCAGTCATGAATTTGAAATTATCTCTTTTAATACTTTCTTGCTTCACATGC
bls3_p2/1-108   ------------------------------------------------------------
bls5a/1-505     AGTTT-TTTTCAGTCATGAATTTGAAACTATCTCTTTTAATACTTTCTTGCTTCGCATGC
bls5b/1-487     AGTTT-TTTTCAGTCATGAATTTGAAACTATCTCTTTTAATACTTTCTTGCTTCGCATGC
bls5_p1/1-186   ------------------------------------------------------------

bls2a/1-504     ATGTACGTGAATGGTATTCTGGGTCTCAGACTTTTGACGGGACTGAAATTGAATGTGGAT
bls2b/1-472     ATGTACGTGAATGGTATTCTGGGTCTCAGACTTTTGACGGGACTGAAATTGAATGTGGAT
bls2_p1/1-204   ATGTACGTGAATGGTATTCTGGGTCTCAGACTTTTGACGGGACTGAAATTGAATGTGGAT
bls2_p2/1-116   ------------------------------------------------------------
bls3a/1-475     ATGTATGTGAATGGTTTTTGTGGATCTCAGACTTTTGACGGGACTGAAATTGAATGTGGCG
bls3b/1-487     ATGTATGTGAATGGTTTTTGTGGATCTCAGACTTTTGACGGGACTGAAATTGAATGTGGCG
bls3c/1-487     ATGTATGTGAATGGTTTTTGTGGATCTCAGACTTTTGACGGGACTGAAATTGAATGTGGCG
bls3d/1-484     ATGTATGTGAATGGTTTTTGTGGATCTCAGACTTTTGACGGGACTGAAATTGAATGTGG-C
bls3e/1-484     ATGTATGTGAATGGTTTTTGTGGATCTCAGACTTTTGACGGGACTGAAATTGAATGTGGCG
bls3f/1-485     ATGTATGTGAATGGTTTTTGTGGATCTCAGACTTTTGACGGGACTGAAATTGAATGTGGCG
bls3g/1-485     ATGTATGTGAATGGTTTTTGTGGATCTCAGACTTTTGACGGGACTGAAATTGAATGTGGCG
bls3_p2/1-108   ------------------------------------------------------------
bls5a/1-505     ATGTACCTGAATGGTGTTGTGGGTCTCGGACTTTTGACGGGATTAAGAGTGAATGTGGAC
bls5b/1-487     ATGTACCTGAATGGTGTTGTGGGTCTCGGACTTTTGACGGGATTAAGAGTGAATGTGGAC
bls5_p1/1-186   -----------------------------------CGGGATTAAGAGTGAATGTGGAC

bls2a/1-504     GGCCTCATTAAGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGCAATTAT
bls2b/1-472     GGCCTCATTAAGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGCAATTAT
bls2_p1/1-204   GGCCTCATTAAGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGCAATTAT
bls2_p2/1-116   ------------------------------------------------------------
bls3a/1-475     GGCCTCATTAAGACTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGATATTAT
bls3b/1-487     GGCCTCATTAAGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGATATTAT
bls3c/1-487     GGCCTCATTAAGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGATATTAT
bls3d/1-484     GGGCTCATTAAGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGATATTAT
bls3e/1-484     GGCCTCATT-AGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGATATTAT
bls3f/1-485     GGCCTCATTAAGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGATATTAT
bls3g/1-485     GGCCTCATTAAGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGATATTAT
bls3_p2/1-108   ------------------------------------------------------------
bls5a/1-505     GGCCTCATTGATGTTGACTTGGGATTAGCATTGGGTCTGAAATTGGGAGCCGGCAATTAT
bls5b/1-487     GGCCTCATTGATGTTGACTT-GGATTAGCATTGGGTCTGAAATTGGGAGCCGGCAATTAT
bls5_p1/1-186   GGCCTCATTGATGTTGACTTGGGATTAGCATTGGGTCTGAAATTGGGAGCCGGCAATTAT

bls2a/1-504     CGATCAGTATTAGAGATCCCGGCAATCAATAATTTTCTACTTGGACTCAGGGCCCGTGTA
bls2b/1-472     CGATCAGTATTAGAGATCCCGGCAATCAATAATTTTCTACTTGGACTCAGGGCCCGTGTA
bls2_p1/1-204   CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTCTACTTGGACTCAGGGCCCGTGTA
bls2_p2/1-116   ------------------------------------------------------------
bls3a/1-475     CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTCTAATTGGACTCAGGGCCCGTGTA
bls3b/1-487     CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTCTAATTGGACTCAGGGCCCGTGTA
bls3c/1-487     CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTCTAATTGGACTCAGGGCCCGTGTA
bls3d/1-484     CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTCTAATTGGACTCAGGGCCCGTGTA
```

```
bls3e/1-484      CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTCTAATTGGACTCAGGGCCCGTGTA
bls3f/1-485      CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTCTAATTGGACTCAGGGCCCGTGTA
bls3g/1-485      CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTCTAATTGGACTCAGGGCCCGTGTA
bls3_p2/1-108    ------------------------------------------------------------
bls5a/1-505      CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTTTACTTGGACTCAGAGCCCGTCTA
bls5b/1-487      CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTTTACTTGGACTCAGAGCCCGTCTA
bls5_p1/1-186    CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTTTACTTGGACTCAGAGCCCGTCTA


bls2a/1-504      GCAGGTCCAATTTATGCAGAAGTAAAAGCACATTTGGAAGAAATTGGACAAATATCAAGT
bls2b/1-472      GCAGGTCCAATTTATGCAGAAGTAAAAGCACATTTGGAAGAAATTGGACAAATATCAAGT
bls2_p1/1-204    GCAGGTCCAATTTATGCAGAAGT-------------------------------------
bls2_p2/1-116    ------------------------------------------------------------
bls3a/1-475      TCGGGTCCAATCTATGCAAATGTACAAGCACATTTGGAAGACATTGGAAAAATATCAAGT
bls3b/1-487      TCGGGTCCAATCTATGCAAATGTACAAGCACATTTGGAAGACATTGGAAAAATATCAAGT
bls3c/1-487      TCGGGTCCAATCTATGCAAATGTACAAGCACATTTGGAAGACATTGGAAAAATATCAAGT
bls3d/1-484      TCGGGTCCAATCTATGCAAATGTACAAGCACATTTGGAAGACATTGGAAAAATATCAAGT
bls3e/1-484      TCGGGTCCAATCTATGCAAATGTACAAGCACATTTGGAAGACATTGGAAAAATATCAAGT
bls3f/1-485      TCGGGTCCAATCTATGCAAATGTACAAGCACATTTGGAAGACATTGGAAAAATATCAAGT
bls3g/1-485      TCGGGTCCAATCTATGCAAATGTACAAGCACATTTGGAAGACATTGGAAAAATATCAAGT
bls3_p2/1-108    ------------------------------------------------------------
bls5a/1-505      CCGGGTGCAATCTATGCAAAGGTACAAGCCCGTTTGGAATACATGGGACGAATATCAAGT
bls5b/1-487      CCGGGTGCAATCTATGCAAAGGTACAAGCCCGTTTGGAATACATGGGACGAATATCAAGT
bls5_p1/1-186    CCGGGTGCAATCTATGCAAAGGTACAAGCCCGTTTGGAATACA-----------------


bls2a/1-504      GGGATGTACGGTTTAGGTATTGACAGAGAAACGGTCGACAACCTAGTTCATCACATTAGA
bls2b/1-472      GGGATGTACGGTTTAGGTATTGACAGAGAAACGGTCGACAACCTAGTTCATCACATTAGA
bls2_p1/1-204    ------------------------------------------------------------
bls2_p2/1-116    -------------------------GAAACGGTCGACAACCTAGTTCATCACATTAGA
bls3a/1-475      GGAATTTACGGTTTAGGCATTGACAGAGAAAGATCAACAGATTAGTTTATCAAATAAGA
bls3b/1-487      GGAATTTACGGTTTAGGCATTGACAGAGAAAGATCAACAGATTAGTTTATCAAATAAGA
bls3c/1-487      GGAATTTACGGTTTA-GCATTGACAGAGAAAGATCAACAGATTAGTTTATCAAATAAGA
bls3d/1-484      GGAATTTACGGTTTAGGCATTGACAGAGAAAGATCAACAGATTAGTTTATCAAATAAGA
bls3e/1-484      GGAATTTACGGTTTAGGCATTGACAGAGAAAGATCAACAGATTAGTTTATCAAATAAGA
bls3f/1-485      GGAATTTACGGTTTAGGCATTGACAGAGAAA-GATCAACAGATTAGTTTATCAAATAAGA
bls3g/1-485      GGAATTTACGGTTTAGGCATTGACAGAGAAAGATCAACAGATTAGTTTATCAAATAAGA
bls3_p2/1-108    -----------------CATTGACAGAGAAAGATCAACAGATTAGTTTATCAAATAAGA
bls5a/1-505      AGGATTTACGGTTTAGGAATTAACCAAGAAACGGTCAACAACCTAGTTTATTACATTAGA
bls5b/1-487      AGGATTTACGGTTTAGGAATTAACCAAGAAACGGTCAACAACCTAGTTTATTACATTAGA
bls5_p1/1-186    ------------------------------------------------------------


bls2a/1-504      CAACTTGAACGAAGACGATACGAATTAGAAGCCTTACGAAGAAAATATTTCAGAAGTCAA
bls2b/1-472      CAACTTGAACGAAGACGATACGAATTAGAAGCCTTACGAAGAAAATATTTCAGAAGTCAA
bls2_p1/1-204    ------------------------------------------------------------
bls2_p2/1-116    CAACTTGAACGAAGACGATACGAATTAGAAGCCTTACGAAGAAAATATTTCAGAAGTC-A
bls3a/1-475      AAACTTGAACGAAGACGATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAA
bls3b/1-487      AAACTTGAACGAAGACGATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAA
bls3c/1-487      AAACTTGAACGAAGACGATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAA
bls3d/1-484      AAACTTGAACGAAGACGATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAA
bls3e/1-484      AAACTTGAACGAAGACGATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAA
bls3f/1-485      AAACTTGAACGAAGACGATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAA
bls3g/1-485      AAACTTGAACGAAGACGATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAA
bls3_p2/1-108    AAACTTGAACGAAGACGATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAA
bls5a/1-505      CAACTTGAACGAAGAAGATACGAATTAGAAGCCTTACGAAGAAGATATTTCAGAAATCAA
bls5b/1-487      CAACTTGAACGAAGAAGATACGAATTAGAAGCCTTACGAAGAAGATATTTCAGAAATCAA
bls5_p1/1-186    ------------------------------------------------------------


bls2a/1-504      -ATATTGATAGATTATTTA------------GTGAAGCTATTAAAAATTAAAATGTGAA
bls2b/1-472      AATATTGATAGATTATTTA------------GTGAAGCTATTAAAAATTAAAATGTGAA
bls2_p1/1-204    ------------------------------------------------------------
bls2_p2/1-116    AATATTGATAGATTATTTA------------GTGAA---------------------
bls3a/1-475      GATGTAAACAAAGTGTTTG-AAAAATGTTGTGTGGAAGTTGTCGAAGAATCGAGTA----
```

```
bls3b/1-487    GATGTAAACAAAGTGTTTGAAAAAATGTTGTGTGGAAGTTGTCGAAGAATCGAGTAATGC
bls3c/1-487    GATGTAAACAAAGTGTTTG-AAAAATGTTGTGTGGAAGTTGTCGAAGAATCGAGTAATGC
bls3d/1-484    GATGTAAACAAAGTGTTTG-AAAAATGTTGTGTGGAAGTTGTCGAAGAATCGAGTAATGC
bls3e/1-484    GATGTAAACAAAGTGTTTG--AAAATGTTGTGTGGAAGTTGTCGAAGAATCGAGTAATGC
bls3f/1-485    GATGTAAACAAAGTGTTTG-AAAATGTTGTGTGGAAGTTGTCGAAGAATCGAGTAATGC
bls3g/1-485    GATGTAAACAAAGTGTTTG--AAAATGTTGTGTGGAAGTTGTCGAAGAATCGAGTAATGC
bls3_p2/1-108  GATGT-------------------------------------------------------
bls5a/1-505    AATATTGATGAAATATTTC------------GTGCAGTTAATAAAAAGTAAAACGTGAA
bls5b/1-487    AATATTGATGAAATATTTC------------GTGCAGTTAATAAAAAGTAAAACGTGAA
bls5_p1/1-186  ------------------------------------------------------------
```

Aligments using *bls* gene members together sequence used to perform RNAi experiments. Forward primer is in orange and reverse is in blue.

```
bls2a/1-504    AGTTT-TTTTCAGTCATGAATTTGAAATTATCTCTTTTAATACTTTCATGCTTCGCATGC
bls2b/1-472    AGTTT-TTTTCAGTCATGAATTTGAAATTATCTCTTTTAATACTTTCATGCTTCGCATGC
bls3a/1-475    AGTTGATTTTCAGTCATGAATTTGAAATTATCTCTTTTAATACTTTCTTGCTTCACATGC
bls3b/1-487    AGTTGATTTTCAGTCATGAATTTGAAATTATCTCTTTTAATACTTTCTTGCTTCACATGC
bls3c/1-487    AGTTGATTTTCAGTCATGAATTTGAAATTATCTCTTTTAATACTTTCTTGCTTCACATGC
bls3d/1-484    AGTTGATTTTCAGTCATGAATTTGAAATTATCTCTTT-AATACTTTCTTGCTTCACATGC
bls3e/1-484    AGTTGATTTTCAGTCATGAATTTGAAATTATCTCTTTTAATACTTTCTTGCTTCACATGC
bls3f/1-485    AGTTGATTTTCAGTCATGAATTTGAAATTATCTCTTTTAATACTTTCTTGCTTCACATGC
bls3g/1-485    AGTTGATTTTCAGTCATGAATTTGAAATTATCTCTTTTAATACTTTCTTGCTTCACATGC
bls3_p2/1-108  ------------------------------------------------------------
bls5a/1-505    AGTTT-TTTTCAGTCATGAATTTGAAACTATCTCTTTTAATACTTTCTTGCTTCGCATGC
bls5b/1-487    AGTTT-TTTTCAGTCATGAATTTGAAACTATCTCTTTTAATACTTTCTTGCTTCGCATGC

bls2a/1-504    ATGTACGTGAATGGTATTCTGGGTCTCAGACTTTTGACGGGACTGAAATTGAATGTGGAT
bls2b/1-472    ATGTACGTGAATGGTATTCTGGGTCTCAGACTTTTGACGGGACTGAAATTGAATGTGGAT
bls3a/1-475    ATGTATGTGAATGGTTTTTGTGGATCTCAGACTTTTGACGGGACTGAAATTGAATGTGGCG
bls3b/1-487    ATGTATGTGAATGGTTTTTGTGGATCTCAGACTTTTGACGGGACTGAAATTGAATGTGGCG
bls3c/1-487    ATGTATGTGAATGGTTTTTGTGGATCTCAGACTTTTGACGGGACTGAAATTGAATGTGGCG
bls3d/1-484    ATGTATGTGAATGGTTTTTGTGGATCTCAGACTTTTGACGGGACTGAAATTGAATGTGGCG
bls3e/1-484    ATGTATGTGAATGGTTTTTGTGGATCTCAGACTTTTGACGGGACTGAAATTGAATGTGGCG
bls3f/1-485    ATGTATGTGAATGGTTTTTGTGGATCTCAGACTTTTGACGGGACTGAAATTGAATGTGGCG
bls3g/1-485    ATGTATGTGAATGGTTTTTGTGGATCTCAGACTTTTGACGGGACTGAAATTGAATGTGGCG
bls3_p2/1-108  ------------------------------------------------------------
bls5a/1-505    ATGTACCTGAATGGTGTTGTGGGTCTCGGACTTTTGACGGGATTAAGAGTGAATGTGGAC
bls5b/1-487    ATGTACCTGAATGGTGTTGTGGGTCTCGGACTTTTGACGGGATTAAGAGTGAATGTGGAC

bls2a/1-504    GGCCTCATTAAGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGCAATTAT
bls2b/1-472    GGCCTCATTAAGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGCAATTAT
bls3a/1-475    GGCCTCATTAAGACTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGATATTAT
bls3b/1-487    GGCCTCATTAAGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGATATTAT
bls3c/1-487    GGCCTCATTAAGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGATATTAT
bls3d/1-484    GG-CTCATTAAGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGATATTAT
bls3e/1-484    GGCCTCATT-AGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGATATTAT
bls3f/1-485    GGCCTCATTAAGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGATATTAT
bls3g/1-485    GGCCTCATTAAGGCTGACTTGGGATTAAGATTGGGTCTGTACTTGGGAGCCGGATATTAT
bls3_p2/1-108  ------------------------------------------------------------
bls5a/1-505    GGCCTCATTGATGTTGACTTGGGATTAGCATTGGGTCTGAAATTGGGAGCCGGCAATTAT
bls5b/1-487    GGCCTCATTGATGTTGACTT-GGATTAGCATTGGGTCTGAAATTGGGAGCCGGCAATTAT

bls2a/1-504    CGATCAGTATTAGAGATCCCGGCAATCAATAATTTTCTACTTGGACTCAGGGCCCGTGTA
bls2b/1-472    CGATCAGTATTAGAGATCCCGGCAATCAATAATTTTCTACTTGGACTCAGGGCCCGTGTA
bls3a/1-475    CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTCTAATTGGACTCAGGGCCCGTGTA
bls3b/1-487    CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTCTAATTGGACTCAGGGCCCGTGTA
bls3c/1-487    CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTCTAATTGGACTCAGGGCCCGTGTA
bls3d/1-484    CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTCTAATTGGACTCAGGGCCCGTGTA
```

```
bls3e/1-484      CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTCTAATTGGACTCAGGGCCCGTGTA
bls3f/1-485      CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTCTAATTGGACTCAGGGCCCGTGTA
bls3g/1-485      CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTCTAATTGGACTCAGGGCCCGTGTA
bls3_p2/1-108    ------------------------------------------------------------
bls5a/1-505      CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTTTACTTGGACTCAGAGCCCGTCTA
bls5b/1-487      CGATCAGTATTAGAGATCCCGGCAATCAATCATTTTTTACTTGGACTCAGAGCCCGTCTA


bls2a/1-504      GCAGGTCCAATTTATGCAGAAGTAAAAGCACATTTGGAAGAAATTGGACAAATATCAAGT
bls2b/1-472      GCAGGTCCAATTTATGCAGAAGTAAAAGCACATTTGGAAGAAATTGGACAAATATCAAGT
bls3a/1-475      TCGGGTCCAATCTATGCAAATGTACAAGCACATTTGGAAGACATTGGAAAAATATCAAGT
bls3b/1-487      TCGGGTCCAATCTATGCAAATGTACAAGCACATTTGGAAGACATTGGAAAAATATCAAGT
bls3c/1-487      TCGGGTCCAATCTATGCAAATGTACAAGCACATTTGGAAGACATTGGAAAAATATCAAGT
bls3d/1-484      TCGGGTCCAATCTATGCAAATGTACAAGCACATTTGGAAGACATTGGAAAAATATCAAGT
bls3e/1-484      TCGGGTCCAATCTATGCAAATGTACAAGCACATTTGGAAGACATTGGAAAAATATCAAGT
bls3f/1-485      TCGGGTCCAATCTATGCAAATGTACAAGCACATTTGGAAGACATTGGAAAAATATCAAGT
bls3g/1-485      TCGGGTCCAATCTATGCAAATGTACAAGCACATTTGGAAGACATTGGAAAAATATCAAGT
bls3_p2/1-108    ------------------------------------------------------------
bls5a/1-505      CCGGGTGCAATCTATGCAAAGGTACAAGCCCGTTTGGAATACATGGGACGAATATCAAGT
bls5b/1-487      CCGGGTGCAATCTATGCAAAGGTACAAGCCCGTTTGGAATACATGGGACGAATATCAAGT


bls2a/1-504      GGGATGTACGGTTTAGGTATTGACAGAGAAACGGTCGACAACCTAGTTCATCACATTAGA
bls2b/1-472      GGGATGTACGGTTTAGGTATTGACAGAGAAACGGTCGACAACCTAGTTCATCACATTAGA
bls3a/1-475      GGAATTTACGGTTTAGGCATTGACAGAGAAAAGATCAACAGATTAGTTTATCAAATAAGA
bls3b/1-487      GGAATTTACGGTTTAGGCATTGACAGAGAAAAGATCAACAGATTAGTTTATCAAATAAGA
bls3c/1-487      GGAATTTACGGTTTAG-CATTGACAGAGAAAAGATCAACAGATTAGTTTATCAAATAAGA
bls3d/1-484      GGAATTTACGGTTTAGGCATTGACAGAGAAAAGATCAACAGATTAGTTTATCAAATAAGA
bls3e/1-484      GGAATTTACGGTTTAGGCATTGACAGAGAAAAGATCAACAGATTAGTTTATCAAATAAGA
bls3f/1-485      GGAATTTACGGTTTAGGCATTGACAGAGAAA-GATCAACAGATTAGTTTATCAAATAAGA
bls3g/1-485      GGAATTTACGGTTTAGGCATTGACAGAGAAAAGATCAACAGATTAGTTTATCAAATAAGA
bls3_p2/1-108    ----------------CATTGACAGAGAAAAGATCAACAGATTAGTTTATCAAATAAGA
bls5a/1-505      AGGATTTACGGTTTAGGAATTAACCAAGAAACGGTCAACAACCTAGTTTATTACATTAGA
bls5b/1-487      AGGATTTACGGTTTAGGAATTAACCAAGAAACGGTCAACAACCTAGTTTATTACATTAGA


bls2a/1-504      CAACTTGAACGAAGACGATACGAATTAGAAGCCTTACGAAGAAAATATTTCAGAAGTCAA
bls2b/1-472      CAACTTGAACGAAGACGATACGAATTAGAAGCCTTACGAAGAAAATATTTCAGAAGTCAA
bls3a/1-475      AAACTTGAACGAAGACGATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAA
bls3b/1-487      AAACTTGAACGAAGACGATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAA
bls3c/1-487      AAACTTGAACGAAGACGATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAA
bls3d/1-484      AAACTTGAACGAAGACGATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAA
bls3e/1-484      AAACTTGAACGAAGACGATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAA
bls3f/1-485      AAACTTGAACGAAGACGATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAA
bls3g/1-485      AAACTTGAACGAAGACGATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAA
bls3_p2/1-108    AAACTTGAACGAAGACGATACGAATTAGAAGCATTACGACGAACTCATTTCCCTGACCAA
bls5a/1-505      CAACTTGAACGAAGAAGATACGAATTAGAAGCCTTACGAAGAAGATATTTCAGAAATCAA
bls5b/1-487      CAACTTGAACGAAGAAGATACGAATTAGAAGCCTTACGAAGAAGATATTTCAGAAATCAA


bls2a/1-504      -ATATTGATAGATTATTTA------------GTGAAGCTATTAAAAATTAAAATGTGAA
bls2b/1-472      AATATTGATAGATTATTTA------------GTGAAGCTATTAAAAATTAAAATG-TGA
bls3a/1-475      GATGTAAACAAAGTGTTTG-AAAAATGTTGTGTGGAAGTTGTCGAAGAATCGAGTA----
bls3b/1-487      GATGTAAACAAAGTGTTTGAAAAAATGTTGTGTGGAAGTTGTCGAAGAATCGAGTAATGC
bls3c/1-487      GATGTAAACAAAGTGTTTG-AAAAATGTTGTGTGGAAGTTGTCGAAGAATCGAGTAATGC
bls3d/1-484      GATGTAAACAAAGTGTTTG-AAAAATGTTGTGTGGAAGTTGTCGAAGAATCGAGTAATGC
bls3e/1-484      GATGTAAACAAAGTGTTTG--AAAATGTTGTGTGGAAGTTGTCGAAGAATCGAGTAATGC
bls3f/1-485      GATGTAAACAAAGTGTTTG-AAAAATGTTGTGTGGAAGTTGTCGAAGAATCGAGTAATGC
bls3g/1-485      GATGTAAACAAAGTGTTTG--AAAATGTTGTGTGGAAGTTGTCGAAGAATCGAGTAATGC
bls3_p2/1-108    GATGT-------------------------------------------------------
bls5a/1-505      AATATTGATGAAATATTTC------------GTGCAGTTAATAAAAAGTAAAACGTGAA
bls5b/1-487      AATATTGATGAAATATTTC------------GTGCAGTTAATAAAAAGTAAAACGTGAA


bls2a/1-504      AATTAAAAAAGTTTAATCTTGACGATCACTATCAAATAA
bls2b/1-472      AAATTAA--------------------------------
```

```
bls3a/1-475      ----------------------------------------
bls3b/1-487      AAAATGA---------------------------------
bls3c/1-487      AAAATGATT-------------------------------
bls3d/1-484      AAAATGA---------------------------------
bls3e/1-484      AAAATGA---------------------------------
bls3f/1-485      AAAATGA---------------------------------
bls3g/1-485      AAAATGA---------------------------------
bls3_p2/1-108    ----------------------------------------
bls5a/1-505      AATAAAAAATGTTAAACCTTGACGATCACTTTCAAATAA
bls5b/1-487      AATAAAAAATGTTAAACCTTGA------------------
```

Annex II – DVL and FOXK alignments of *Schmidtea mediterranea* (*Smed*), *Xenopus laevis* (*Xlea*) and *Homo sapiens* (*Hsa*).

Sequences and aligment of DVL proteins of *Smed*, *Xlae* and *Hsa*. *Xlae* and *Hsa*.

```
>Smed_dvl1
MEETRIIYYVDDEETPYLIKFHSPPEQITLGDFKNALNRPNYKFFFKSLDDDFGVVKEEITDDDAKLPYVNGRVVSWLV-
VSEGSTQSDNHSSSGKEVLLVDSDKSKDKGTVSDSSDPKSPSFRNYNKIPTKHSSSSKKQEESNKIHRQNHKFTGPE-
KITLDETDDAFDEIDSIYNEDKVPPLRKFSDFKHSVKLKKLRNAQGSHGNSNSSSNNNNSNSNNASNNAPAKQQPI-
YESSSSMMSSDLDTTSFFDSEDDSSRFSSATETTMSSKYGKQRRQLRRRRKMPHLSRASSFSSMTDSTVSLNIITVTL-
NMDTVPFLGISIVGQTNGNQENGDGGIYVGSIMKGGAVALDGRIEPGDMILEVNGISFENVSNEEAVRTLREQVQKLG-
PVTLVVAKSWDPNPTGYMLPQQDPVRPIDPRAWVLHTQAMGNMAPPNQPPVASGDQFIQSGKYLPYAGAMSTV-
ASTITTTSSSLKSSSVVEVQNPQNKFIGRTDETIPSPLTTNHEPSVIIRAMAQADSGLPIRDRLWLKITIYNAFIGS-
DLVDWLYSHVQGFTDRKDARKFATNLLKMGFIRHTVNKSSFSEQCYYVLNDMMGALSVMNLESEIDSVSVVAGQQQSK-
TRLVHQDPGCKESPILSSGFQSACSNKWSSYNIQSNPGPEMYSQAQIMTMNNNNAFKYQNFYTDNMQPQSQSLVK-
PLGLSSNLKSSRGPGSVNSGSVVTRNCVCLNSMTVLKAVKHMWESGANKSVMAATDLSGGSGGGSSGGGEGGGGST-
VREVEPVSTASMRNQDINSIVSVNSTTCRLCGGEQEESDLYSNNDEDDNPHRHHSMKMDAPATTSGSSASGSSGN-
VTRNIQMPVYHMQNPAPVIGSGIVLPDASQYPLHNSPPPSYQQSMAIGTILGKNQSIMPGDNSSNQFSLAISNGIL-
NENSSSGFFSDGVDKCD
```

```
>Smed_dvl2
MTNCATSGNVISDETRIIYHIDEEETPYLIKLSISPDKVTLGDLKNALNRPHYKYFFKSMDDDFGVVKEEITDDEAKLP-
CFKGRVISWLVTAEGSTVSDNVDSNGILDKNESRMLPFQESHFPLINNIKASGGTTTNESDTICDTCTDTDSVYSAA-
QDRVGPLRSFHDYKQAGRVAAHANRVNTNTPNGQNPIYETNSSMMSSDLESTSFFDSEDESSRFSTTTCTTMSS-
RYGRQKQQRRRRPPAISRASSFSSITDSTMSLNIVTVRLNMDTVKFLGISIVGQSNKGGDGGIYVGSIMKGGA-
VAQDGRIEPGDMILEVNDISFEDMSNDDAVRTLREQVQKPGPINLVVAKCWDPNPKGYFTIPRQEPVRPIDPRAWVLHT-
NAMTAGASEPPSSVNGVHPQVSNLVAPSMQSLLSGGTMLAGTSAATFNAAAFGYMPQPQNINQNTASVSTVGGPPGAS-
VGFFGYPMGMPGQFSQGAGSIVTTSSSLPESERYQEELHLTKNTDVGTILRVLSQPDSGLDIRDRLWLKITLPNAFIG-
SNLVDWLYRHIEGFSDRKEARKYAANLLKFGYIKHTVNKVTFSEQCYYVLGNTTLNMSRLSLDQVESVSEVGVNGPHH-
LAALPPPNFSSNKQPISSCINQPPLNINPQLTATSEPLPSNNANVATATASSNSQYSVVGPLPCSQPSQHASSNASA-
SAIKKSGSCNSLSGSSSSTSSSSSSNRSNTRINGNASSVSNMISKNPPPKIPPRTIASVSTNSTNPIISGFQNRGQSS-
VSQ
```

```
>Xle_dvl1
MAETKIIYHIDEEETPYLVKLPVPPEKVTLADFKNVLSNRPVHHYKFFFKSMDQDFGVVKEEISDDNAKLPCFN-
GRVVSWLVLAESSHSDGGSQSTESRTDLPLPIERTGGIGDSRPPSFHPNASSSRDGLDNETGTDSVVSHRRDRHRR-
KNRETHDDVPRINGHPKLDRIRDPGGYDSASTVMSSELESSSFVDSDEDENTSRLSSSTEQSTSSRLIRKHKRRRRKQK-
MRQIDRSSSFSSITDSTMSLNIITVTLNMEKYNFLGISIVGQSNDRGDGGIYIGSIMKGGAVAADGRIEPGDMLLQVND-
VNFENMSNDDAVRVLREIVSKPGPISLTVAKCWDPTPRSYFTIPRAEPVRPIDPAAWITHTSALTGAYPRYGMGPSMSI-
ITCTSSSLTSSKPESEKQEDSPLSVKSDMATIVKVMQVPDSGLEIRDRMWLKITISNAVIGADVVDWLYTHVEGFKER-
REARKYASSMLKHGYLRHTVNKITFSEQCYYVFGDLCGNVAALNLNDGSSGTSDQDTLAPLPHPAAPWPLGQGYSY-
QYPLAPPCFPPTYQEPGFSYGSGSAGSQHSEGSKSSGSSRSNREGKRSSGREKERRSTGSGSGSDRAPRSGGSKNER-
PLSQHSHHSHSSVSRSQRSHRSNSHHSHGPPGLPPLFSLPKIGSKVYGTSGPPGGPPVRELANVPPELTGSRQSFQKA-
MGNPCEFFVDIM
```

```
Xle_dvl2
MAETKVIYHLDEEETPYLVKVPVPATDIRLRDFKAALGRGHAKYFFKAMDQDFGVVKEEISDDNAKLPCFNDRV-
VSWLASSEGSQPDSAPPAPATEVRPEPPPPVPPPIPPPPAERTSGIGDSRPPSFHPNVSGSTEQLDQDNES-
VISMRRDRVRRRESSEQAGVGRGVNGRTERHLSGYESSSTLLTSEIETSICDSEEDDTMSRFSSSTEQSSAS-
RLLKRHRRRRKQRPPRLERTSSFSSVTDSTMSLNIITVTLNMEKYNFLGISIVGQSNERGDGGIYIGSIMKGGA-
VAADGRIEPGDMLLQVNDINFENMSNDDAVRVLRDIVHKPGPIVLTVAKCWDPSPQGYFTLPRNEPIHPIDPAAWVSH-
SAALSGSFPVYPGSASMSSMTSSTSVTETELSHALPPVSLFSLSVHTDLASVVKVMASPESGLEVRDRMWLKITIP-
NAFLGSDVVDWLYHHVEGFQDRREARKFASNLLKAGFIRHTVNKITFSEQCYYIFGDLTGCENYMTNLSLNDNDGSS-
GASDQDTLAPLPLPGASPWPLLPTFSYQYQAPHPYSTQPPAYHELSSYSYGMGSAGSQHSEGSRSSGSNRSDGGRGMQ-
KDDRSGVAGVGGGDSKSGSGSESEYSTRSSIRRVGGGEAGPPSERSTSSRLPPHHPPSVHSYAAPGVPLSYNPMMLM-
MMPPPPLPPPGVCPPNSSVPPGAPPLVRDLASVPPELTATRQSFHMAMGNPSEFFVDVM
```

```
>Xle_dvl3
MGETKVIYHLDEQETPYLVKLPVPAEKVTLGDFKNILNKPNYKFFFKSMDDDFGVVKEEISDDNAKLPCFNGRVVC-
WLVSADGSQSDAGSVCADIQSDLPPPIERTGGIGDSRPPSFHPNTRGSQENLDNETETDSVVSARRERPGRKETSEHA-
TRINGTSKMERRRDTGGYESSSTLMSSELDSTSFFDSDEDDSTSRFSNSTEQSSASRLMRRHKRRRRKPKAPQIERSSS-
FSSITDSTMSLNIITVTLNMEKYNFLGISIVGQSNERGDGGIYIGSIMKGGAVAADGRIEPGDMLLQVNDTNFENMSND-
DAVRVLRDIVHKPGPITLTVAKCWDPSPRNCFTLPRSEPIRPIDPAAWVSHTAAMTGTYPAYGMSPSMSTITSTSS-
SITSSIPETERFDDFQLSIHSDMVTIVKAMSSSESGLEVRDRMWLKITIPNAFIGSDVVDWLYHHVEGFTDRREAR-
KYASNLLKAGYIRHTVNKITFSEQCYYIFGDLCGNMANLSLNDHDGSSGTSDQDTLAPLPHPGAAPWPIAFQYQYPLPH-
PYSPHPGFPDPAYIYGGGSAGSQHSEGSRSSGSNRSSTEKRKDRETKGGDSKSGGSGSESDHTTRSSLRRDRAASER-
SVPASEHSHRSHHSIAHSIRSHHTHQSFGPPGIPPLYGAPMMMMPAPVSVMGPPGAPPSRDLASVPPELTASRQSFR-
MAMGNPSEFFVDVIKEFWGV

>Hsa_dvl1
MAETKIIYHMDEEETPYLVKLPVAPERVTLADFKNVLSNRPVHAYKFFFKSMDQDFGVVKEEIFDDNAKLPCFNGRV-
VSWLVLAEGAHSDAGSQGTDSHTDLPPPLERTGGIGDSRPPSFHPNVASSRDGMDNETGTESMVSHRRERARRRNREE-
AARTNGHPRGDRRRDVGLPPDSASTALSSELESSSFVDSDEDGSTSRLSSSTEQSTSSRLIRKHKRRRRKQRLRQA-
DRASSFSSITDSTMSLNIVTVTLNMERHHFLGISIVGQSNDRGDGGIYIGSIMKGGAVAADGRIEPGDMLLQVNDVN-
FENMSNDDAVRVLREIVSQTGPISLTVAKCWDPTPRSYFTVPRADPVRPIDPAAWLSHTAALTGALPRYGTSPCS-
SAVTRTSSSSLTSSVPGAPQLEEAPLTVKSDMSAVVRVMQLPDSGLEIRDRMWLKITIANAVIGADVVDWLYTHVEG-
FKERREARKYASSLLKHGFLRHTVNKITFSEQCYYVFGDLCSNLATLNLNSGSSGTSDQDTLAPLPHPAAPWPLGQ-
GYPYQYPGPPPCFPPAYQDPGFSYGSGSTGSQQSEGSKSSGSTRSSRRAPGREKERRAAGAGGSGSESDHTAPSGVGSS-
WRERPAGQLSRGSSPRSQASATAPGLPPPHPTTKAYTVVGGPPGGPPVRELAAVPPELTGSRQSFQKAMGNPCEFFVDIM

>Hsa_dvl2
MAGSSTGGGGVGETKVIYHLDEEETPYLVKIPVPAERITLGDFKSVLQRPAGAKYFFKSMDQDFGVVKEEISDDNARLP-
CFNGRVVSWLVSSDNPQPEMAPPVHEPRAELAPPAPPLPPLPPERTSGIGDSRPPSFHPNVSSSHENLEPETETES-
VVSLRRERPRRRDSSEHGAGGHRTGGPSRLERHLAGYESSSTLMTSELESTSLGDSDEEDTMSRFSSSTEQSSAS-
RLLKRHRRRRKQRPPRLERTSSFSSVTDSTMSLNIITVTLNMEKYNFLGISIVGQSNERGDGGIYIGSIMKGGA-
VAADGRIEPGDMLLQVNDMNFENMSNDDAVRVLRDIVHKPGPIVLTVAKCWDPSPQAYFTLPRNEPIQPIDPAAWVSH-
SAALTGTFPAYPGSSSMSTITSGSSLPDGCEGRGLSVHTDMASVTKAMAAPESGLEVRDRMWLKITIPNAFLGS-
DVVDWLYHHVEGFPERREARKYASGLLKAGLIRHTVNKITFSEQCYYVFGDLSGGCESYLVNLSLNDNDGSSGAS-
DQDTLAPLPGATPWPLLPTFSYQYPAPHPYSPQPPPYHELSSYTYGGGSASSQHSEGSRSSGSTRSDGGAGRT-
GRPEERAPESKSGSGSESEPSSRGGSLRRGGEASGTSDGGPPPSRGSTGGAPNLRAHPGLHPYGPPPGMALPYNPMMV-
VMMPPPPPPVPPAVQPPGAPPVRDLGSVPPELTASRQSFHMAMGNPSEFFVDVM

>Hsa_dvl3
MGETKIIYHLDGQETPYLVKLPLPAERVTLADFKGVLQRPSYKFFFKSMDDDFGVVKEEISDDNAKLPCFNGRV-
VSWLVSAEGSHPDPAPFCADNPSELPPPMERTGGIGDSRPPSFHPHAGGGSQENLDNDTETDSLVSAQRERPRRRDGPE-
HATRLNGTAKGERRREPGGYDSSSTLMSSELETTSFFDSDEDDSTSRFSSSTEQSSASRLMRRHKRRRRKQKVSRI-
ERSSSFSSITDSTMSLNIITVTLNMEKYNFLGISIVGQSNERGDGGIYIGSIMKGGAVAADGRIEPGDMLLQVNEIN-
FENMSNDDAVRVLREIVHKPGPITLTVAKCWDPSPRGCFTLPRSEPIRPIDPAAWVSHTAAMTGTFPAYGMSPSL-
STITSTSSSITSSIPDTERLDDFHLSIHSDMAAIVKAMASPESGLEVRDRMWLKITIPNAFIGSDVVDWLYHNVEGFT-
DRREARKYASNLLKAGFIRHTVNKITFSEQCYYIFGDLCGNMANLSLHDHDGSSGASDQDTLAPLPHPGAAPWPMAF-
PYQYPPPPHPYNPHPGFPELGYSYGGGSASSQHSEGSRSSGSNRSGSDRRKEKDPKAGDSKSGGSGSESDHTTRSSL-
RGPRERAPSERSGPAASEHSHRSHHSLASSLRSHHTHPSYGPPGVPPLYGPPMLMMPPPPAAMGPPGAPPGRDLASVP-
PELTASRQSFRMAMGNPSEFFVDVM


Smed_dvl1/1-935   -----------MEETRIIYYVDDEETPYLIKFHSPPEQITLGDFKNAL-NRP--NYKFFF
Smed_dvl2/1-775   MTNCATSGNVISDETRIIYHIDEEETPYLIKLSISPDKVTLGDLKNAL-NRP--HYKYFF
Xlae_dvl1/1-708   -----------MAETKIIYHIDEEETPYLVKLPVPPEKVTLADFKNVLSNRPVHHYKFFF
Xle_dvl2/1-736    -----------MAETKVIYHLDEEETPYLVKVPVPATDIRLRDFKAAL-GR--GHAKYFF
Xle_dvl3/1-717    -----------MGETKVIYHLDEQETPYLVKLPVPAEKVTLGDFKNIL-NKP--NYKFFF
Hsa_dvl1/1-695    -----------MAETKIIYHMDEEETPYLVKLPVAPERVTLADFKNVLSNRPVHAYKFFF
Hsa_dvl2/1-736    MAGSSTGGG-GVGETKVIYHLDEEETPYLVKIPVPAERITLGDFKSVL-QRP-AGAKYFF
Hsa_dvl3/1-716    -----------MGETKIIYHLDGQETPYLVKLPLPAERVTLADFKGVL-QRP--SYKFFF


Smed_dvl1/1-935   KSLDDDFGVVKEEITDDDAKLPYVNGRVVSWLVVSEGS--------TQSDNHSSSGKEV-
Smed_dvl2/1-775   KSMDDDFGVVKEEITDDEAKLPCFKGRVISWLVTAEGS--------TVSDN---------
Xlae_dvl1/1-708   KSMDQDFGVVKEEISDDNAKLPCFNGRVVSWLVLAESSHSDGGS--QSTESRTDLPLPI-
Xle_dvl2/1-736    KAMDQDFGVVKEEISDDNAKLPCFNDRVVSWLASSEGSQPDSAPPAPATEVRPEPPPPVP
Xle_dvl3/1-717    KSMDDDFGVVKEEISDDNAKLPCFNGRVVCWLVSADGSQSDAGS--VCADIQSDLPPPI-
```

```
Hsa_dvl1/1-695   KSMDQDFGVVKEEIFDDNAKLPCFNGRVVSWLVLAEGAHSDAGS--QGTDSHTDLPPPL-
Hsa_dvl2/1-736   KSMDQDFGVVKEEISDDNARLPCFNGRVVSWLVSSDNPQPEMAP--PVHEPRAELAPPA-
Hsa_dvl3/1-716   KSMDDDFGVVKEEISDDNAKLPCFNGRVVSWLVSAEGSHPDPAP--FCADNPSELPPPM-


Smed_dvl1/1-935  LLVDSDKSKDKGTVSDSSDPKSPSFRNYNK--IPTK-HS-SSSKKQEESNKIHRQNHKFT
Smed_dvl2/1-775  -------VDSNGILDK---NESRMLPFQESHFPLINNI-KASGGTTTNES---------
Xlae_dvl1/1-708  -------ERTGGIGDS---RPPSFHP----------NA-SSSRDGLDNE----------
Xle_dvl2/1-736   PPIPPPPAERTSGIGDS---RPPSFHP----------NV-SGSTEQL--D----------
Xle_dvl3/1-717   -------ERTGGIGDS---RPPSFHP----------NT-RGSQENLDNE----------
Hsa_dvl1/1-695   -------ERTGGIGDS---RPPSFHP----------NV-ASSRDGMDNE----------
Hsa_dvl2/1-736   PPLPPLPPERTSGIGDS---RPPSFHP----------NV-SSSHENLEPE----------
Hsa_dvl3/1-716   -------ERTGGIGDS---RPPSFHP----------HAGGGSQENLDND----------


Smed_dvl1/1-935  GPEKITLDETDDAFDEIDSIYN--EDKVPPLRKFSD-FKHSVKLKKLRNAQGSHGNSNSS
Smed_dvl2/1-775  -------DTICDTCTDTDSVYSAAQDRVGPLRSFHD-YKQAG-----------------
Xlae_dvl1/1-708  -------------TGTDSVVSHRRDR--HRRKNRETHDDVP-----------------
Xle_dvl2/1-736   -------------QDNESVISMRRDR--VRRRESS--EQAGV----------------
Xle_dvl3/1-717   -------------TETDSVVSARRER--PGRKETS--EHAT----------------
Hsa_dvl1/1-695   -------------TGTESMVSHRRER--ARRRNRE---EAA----------------
Hsa_dvl2/1-736   -------------TETESVVSLRRER--PRRRDSS--EHGAG---------------
Hsa_dvl3/1-716   -------------TETDSLVSAQRER--PRRRDGP--EHAT----------------


Smed_dvl1/1-935  SNNNNSNSNNASNNAPA-KQQPIYESSSSMMSSDLDTTSFFDS--EDDSSRFSSATETTM
Smed_dvl2/1-775  --RVAAHANRVNTNTPN-GQNPIYETNSSMMSSDLESTSFFDS--EDESSRFSTTTCTTM
Xlae_dvl1/1-708  --RINGHPKLDRIRDPG-G----YDSASTVMSSELESSSFVDSDEDENTSRLSSSTEQST
Xle_dvl2/1-736   GRGVNG--RTERHLS---G----YESSSTLLTSEIE-TSICDSEEDDTMSRFSSSTEQSS
Xle_dvl3/1-717   --RINGTSKMERRRDTG-G----YESSSTLMSSELDSTSFFDSDEDDSTSRFSNSTEQSS
Hsa_dvl1/1-695   --RTNGHPRGDRRRDVGLP----PDSASTALSSELESSSFVDSDEDGSTSRLSSSTEQST
Hsa_dvl2/1-736   GHRTGGPSRLERHLA---G----YESSSTLMTSELESTSLGDSDEEDTMSRFSSSTEQSS
Hsa_dvl3/1-716   --RLNGTAKGERRREPG-G----YDSSSTLMSSELETTSFFDSDEDDSTSRFSSSTEQSS


Smed_dvl1/1-935  SSKYGKQRRQLRRRRKMPHLSRASSFSSMTDSTVSLNIITVTLNMDTVPFLGISIVGQTN
Smed_dvl2/1-775  SSRYGRQK-QQRRRRPPAISRASSFSSITDSTMSLNIVTVRLNMDTVKFLGISIVGQSN
Xlae_dvl1/1-708  SSRLIRKHKRRRRKQKMRQIDRSSSFSSITDSTMSLNIITVTLNMEKYNFLGISIVGQSN
Xle_dvl2/1-736   ASRLLKRH-RRRRKQRPPRLERTSSFSSVTDSTMSLNIITVTLNMEKYNFLGISIVGQSN
Xle_dvl3/1-717   ASRLMRRHKRRRRKPKAPQIERSSSFSSITDSTMSLNIITVTLNMEKYNFLGISIVGQSN
Hsa_dvl1/1-695   SSRLIRKHKRRRRKQRLRQADRASSFSSITDSTMSLNIVTVTLNMERHHFLGISIVGQSN
Hsa_dvl2/1-736   ASRLLKRH-RRRRKQRPPRLERTSSFSSVTDSTMSLNIITVTLNMEKYNFLGISIVGQSN
Hsa_dvl3/1-716   ASRLMRRHKRRRRKQKVSRIERSSSFSSITDSTMSLNIITVTLNMEKYNFLGISIVGQSN


Smed_dvl1/1-935  GNQENGDGGIYVGSIMKGGAVALDGRIEPGDMILEVNGISFENVSNEEAVRTLREQVQKL
Smed_dvl2/1-775  ---KGGDGGIYVGSIMKGGAVAQDGRIEPGDMILEVNDISFEDMSNDDAVRTLREQVQKP
Xlae_dvl1/1-708  ---DRGDGGIYIGSIMKGGAVAADGRIEPGDMLLQVNDVNFENMSNDDAVRVLREIVSKP
Xle_dvl2/1-736   ---ERGDGGIYIGSIMKGGAVAADGRIEPGDMLLQVNDINFENMSNDDAVRVLRDIVHKP
Xle_dvl3/1-717   ---ERGDGGIYIGSIMKGGAVAADGRIEPGDMLLQVNDTNFENMSNDDAVRVLRDIVHKP
Hsa_dvl1/1-695   ---DRGDGGIYIGSIMKGGAVAADGRIEPGDMLLQVNDVNFENMSNDDAVRVLREIVSQT
Hsa_dvl2/1-736   ---ERGDGGIYIGSIMKGGAVAADGRIEPGDMLLQVNDMNFENMSNDDAVRVLRDIVHKP
Hsa_dvl3/1-716   ---ERGDGGIYIGSIMKGGAVAADGRIEPGDMLLQVNEINFENMSNDDAVRVLREIVHKP


Smed_dvl1/1-935  GPVTLVVAKSWDPNPTGYM-LPQQDPVRPIDPRAWVLHTQAMGNMAPPNQPPVASGDQFI
Smed_dvl2/1-775  GPINLVVAKCWDPNPKGYFTIPRQEPVRPIDPRAWVLHTNAMTAGA--SEPPSSV-----
Xlae_dvl1/1-708  GPISLTVAKCWDPTPRSYFTIPRAEPVRPIDPAAWITHTSALT---------------
Xle_dvl2/1-736   GPIVLTVAKCWDPSPQGYFTLPRNEPIHPIDPAAWVSHSAALS---------------
Xle_dvl3/1-717   GPITLTVAKCWDPSPRNCFTLPRSEPIRPIDPAAWVSHTAAMT---------------
Hsa_dvl1/1-695   GPISLTVAKCWDPTPRSYFTVPRADPVRPIDPAAWLSHTAALT---------------
Hsa_dvl2/1-736   GPIVLTVAKCWDPSPQAYFTLPRNEPIQPIDPAAWVSHSAALT---------------
Hsa_dvl3/1-716   GPITLTVAKCWDPSPRGCFTLPRSEPIRPIDPAAWVSHTAAMT---------------


Smed_dvl1/1-935  QSGKYLPYAGAMSTVASTITTT--------------------------SSSLK----
Smed_dvl2/1-775  -NGVHPQVSNLVAPSMQSLLSGGTMLAGTSAATFNAAAFGYMPQPQNINQNTASVSTVGG
Xlae_dvl1/1-708  --GAYPRYG--MGPSMSIITCT--------------------------SSSL-----
```

```
Xle_dvl2/1-736    --GSFPVYP--GSASMSSMTSS--------------------------------------
Xle_dvl3/1-717    --GTYPAYG--MSPSMSTITST----------------------------SSSI-----
Hsa_dvl1/1-695    --GALPRYG--TSPCSSAVTRTS----------------------------SSSL-----
Hsa_dvl2/1-736    --GTFPAYP--GSSSMSTITSG--------------------------------------
Hsa_dvl3/1-716    --GTFPAYG--MSPSLSTITST----------------------------SSSI-----


Smed_dvl1/1-935   -------------------------SSSVVEVQNPQNKFIGRTDETIPSPLTTNHE
Smed_dvl2/1-775   PPGASVGFFGYPMGMPGQFSQGAGSIVTTSSSLPESERYQ---------EELHLTKNTD
Xlae_dvl1/1-708   -------------------------TSSKPESEKQE-----------DSPLSVKSD
Xle_dvl2/1-736    -------------------------TSVTETELSHA-----LPPVSLFSLSVHTD
Xle_dvl3/1-717    -------------------------TSSIPETERFD----------DFQLSIHSD
Hsa_dvl1/1-695    -------------------------TSSVPGAPQLE----------EAPLTVKSD
Hsa_dvl2/1-736    -------------------------SSLPDGCEGR------------GLSVHTD
Hsa_dvl3/1-716    -------------------------TSSIPDTERLD----------DFHLSIHSD


Smed_dvl1/1-935   PSVIIRAMAQADSGLPIRDRLWLKITIYNAFIGSDLVDWLYSHVQGFTDRKDARKFATNL
Smed_dvl2/1-775   VGTILRVLSQPDSGLDIRDRLWLKITLPNAFIGSNLVDWLYRHIEGFSDRKEARKYAANL
Xlae_dvl1/1-708   MATIVKVMQVPDSGLEIRDRMWLKITISNAVIGADVVDWLYTHVEGFKERREARKYASSM
Xle_dvl2/1-736    LASVVKVMASPESGLEVRDRMWLKITIPNAFLGSDVVDWLYHHVEGFQDRREARKFASNL
Xle_dvl3/1-717    MVTIVKAMSSSESGLEVRDRMWLKITIPNAFIGSDVVDWLYHHVEGFTDRREARKYASNL
Hsa_dvl1/1-695    MSAVVRVMQLPDSGLEIRDRMWLKITIANAVIGADVVDWLYTHVEGFKERREARKYASSL
Hsa_dvl2/1-736    MASVTKAMAAPESGLEVRDRMWLKITIPNAFLGSDVVDWLYHHVEGFPERREARKYASGL
Hsa_dvl3/1-716    MAAIVKAMASPESGLEVRDRMWLKITIPNAFIGSDVVDWLYHNVEGFTDRREARKYASNL


Smed_dvl1/1-935   LKMGFIRHTVNKSSFSEQCYYVLNDMM----GALSVMNLESEIDSVSVVAGQQQSKTRLV
Smed_dvl2/1-775   LKFGYIKHTVNKVTFSEQCYYVLGNTT----LNMSRLSLD-QVESVSE-VGVNGPHHLAA
Xlae_dvl1/1-708   LKHGYLRHTVNKITFSEQCYYVFGDLC----GNVAALNLN---DGSS---GTSDQDTLAP
Xle_dvl2/1-736    LKAGFIRHTVNKITFSEQCYYIFGDLT-GCENYMTNLSLN-DNDGSS---GASDQDTLAP
Xle_dvl3/1-717    LKAGYIRHTVNKITFSEQCYYIFGDLC----GNMANLSLN-DHDGSS---GTSDQDTLAP
Hsa_dvl1/1-695    LKHGFLRHTVNKITFSEQCYYVFGDLC----SNLATLNLN---SGSS---GTSDQDTLAP
Hsa_dvl2/1-736    LKAGLIRHTVNKITFSEQCYYVFGDLSGGCESYLVNLSLN-DNDGSS---GASDQDTLA-
Hsa_dvl3/1-716    LKAGFIRHTVNKITFSEQCYYIFGDLC----GNMANLSLH-DHDGSS---GASDQDTLAP


Smed_dvl1/1-935   HQDP--GCKESPILSSGFQSACSNKWSSYNIQSNPGPEMYSQAQIMTMNNNNAFKYQNFY
Smed_dvl2/1-775   LPPPNFSSNKQPIS------------SCIN----QPPLNIN------------------
Xlae_dvl1/1-708   LPHP---AAPWPLG------------QGYSYQY-PLAPPCF------------------
Xle_dvl2/1-736    LPLP--GASPWPLL------------PTFSYQY-QAPHPYS------------------
Xle_dvl3/1-717    LPHP--GAAPWPI-------------AFQYQY-PLPHPYS------------------
Hsa_dvl1/1-695    LPHP---AAPWPLG-----------QGYPYQY-PGPPPCF------------------
Hsa_dvl2/1-736    -PLP--GATPWPLL-----------PTFSYQY-PAPHPYS------------------
Hsa_dvl3/1-716    LPHP--GAAPWPM-------------AFPYQYPPPPHPYN------------------


Smed_dvl1/1-935   TDNMQPQSQSLVKPLGLSSNLKSSRGPGSVNSGSVVTRNCVCLNSMTVLKAVKHMWESGA
Smed_dvl2/1-775   -----PQLTATSEPLP-SNNANVATATASSNSQYSV---------VGPLPCSQPSQHASS
Xlae_dvl1/1-708   -----P--PTYQEP-------GFSYGSGSAGSQH----------------SEGSKSSGS
Xle_dvl2/1-736    -----TQPPAYHEL------SSYSYGMGSAGSQH---------------SEGSRSSGS
Xle_dvl3/1-717    -----PH-PGFPDP-------AYIYGGGSAGSQH---------------SEGSRSSGS
Hsa_dvl1/1-695    -----P--PAYQDP-------GFSYGSGSTGSQQ---------------SEGSKSSGS
Hsa_dvl2/1-736    -----PQPPPYHEL------SSYTYGGGSASSQH---------------SEGSRSSGS
Hsa_dvl3/1-716    -----PH-PGFPEL-------GYSYGGGSASSQH---------------SEGSRSSGS


Smed_dvl1/1-935   NKSVMAAT-DLSGGSGGGSSGGGEGGGGSTVREVEPVST-ASMRNQDINSIVSVNSTTCR
Smed_dvl2/1-775   NASASAIK-----KSGSCNSLSGSSSSTSSSSSSNRSNT-------------------R
Xlae_dvl1/1-708   SRSNREGK-RSSGREKERRST-------GSGSGSDRAPR-S------------------
Xle_dvl2/1-736    NRSDGGRGMQKDDRSGVAGVGGGDSKS-GSGSESEYSTR-SSIR--------------R
Xle_dvl3/1-717    NRSSTE-------KRKDRETKGGDSKSGGSGSESDHTTR-SSLR--------------R
Hsa_dvl1/1-695    TRSSR----RAPGREKERRAAG----AGGSGSESDHTAP-SG-----------------
Hsa_dvl2/1-736    TRSDGGAG-----RTGRPEERAPESKS-GSGSESEPSSRGGSLR--------------R
Hsa_dvl3/1-716    NRSGSD------RRKEKDPKAGDSKSGGSGSGSESDHTTR-SSLRG-----------PR


Smed_dvl1/1-935   LCGGEQEESDLYSNNDEDDNP----HRHHSMKMDAPATTSGSS----ASGSSGNVTRNIQ
```

```
Smed_dvl2/1-775 I--------------------------NGNASSVSNMISKNP----------------
Xlae_dvl1/1-708 ------GGS-------KNERP-LSQHSHHSHSSVSRSQRSHRSNSHHSHG-PPGLPPLFS
Xle_dvl2/1-736  VGGGEA------------GP-PSE---RSTSSRLP---PHHPPSVHSYA-APGVPLSYN
Xle_dvl3/1-717  ----DRAAS-------ERSVP-ASEHSHRSHHSIAHSIRSHHT--HQSFG-PPGIPPLYG
Hsa_dvl1/1-695  -----VGSS-------WRERP-AGQ---LSRGSSPRSQASATA---------PGLPP---
Hsa_dvl2/1-736  --GGEASGT-------SDGGP-PPS---RGSTGGAPNLRAH--PGLHPYGPPPGMALPYN
Hsa_dvl3/1-716  ----ERAPS-------ERSGPAASEHSHRSHHSLASSLRSHHT--HPSYG-PPGVPPLYG


Smed_dvl1/1-935 MPVYHMQNPAPVIGSGIVLPDASQYPLHNSPP--PSYQQSMAIGTILGKNQSIM-----P
Smed_dvl2/1-775 --------------------------PPKIPP--RTIA---------SVSTNSTNP
Xlae_dvl1/1-708 LPKIGS----------KVYG-T-----SGPPGGPP-VRELA---------NVP-----P
Xle_dvl2/1-736  -PMMLMMMPPP----PLP-PPGVCPPNSSVPPGAPPLVRDLA---------SVP-----P
Xle_dvl3/1-717  APMMMM----------PAPVSV-----MGPPGAPP-SRDLA---------SVP-----P
Hsa_dvl1/1-695  -PHPTT----------KAYTVV-----GGPPGGPP-VRELA---------AVP-----P
Hsa_dvl2/1-736  -PMMVVMMPPP----PPPVPPAV------QPPGAPP-VRDLG--------SVP-----P
Hsa_dvl3/1-716  PPMLMM----------PPPPAA-----MGPPGAPP-GRDLA---------SVP-----P


Smed_dvl1/1-935 GDNSSNQ-FSLAISNGILNENSSSGFFSDGVDKCD--
Smed_dvl2/1-775 IISGFQNRGQSSVSQ--------------------
Xlae_dvl1/1-708 ELTGSRQSFQKAMGNP-------CEFFVDIM------
Xle_dvl2/1-736  ELTATRQSFHMAMGNP-------SEFFVDVM------
Xle_dvl3/1-717  ELTASRQSFRMAMGNP-------SEFFVDVIKEFWGV
Hsa_dvl1/1-695  ELTGSRQSFQKAMGNP-------CEFFVDIM------
Hsa_dvl2/1-736  ELTASRQSFHMAMGNP-------SEFFVDVM------
Hsa_dvl3/1-716  ELTASRQSFRMAMGNP-------SEFFVDVM------
```

Sequences and aligment of FOXK proteins of *Smed*, *Xlae* and *Hsa*. *Xlae* and *Hsa.*

```
>Smed_foxK1-2.1
MSGDYYDDSLDSDNNLPQYARITFFGQVPYIMQKERVIIGRNSAAGSVDIDVGAVTFVSRKHLELTYSYQKLKVK-
CLGKNGIFIDNIFKSHSFIPYELPSKCTLRFPSTDVQFCVEQLVGIKSSDRGRSYGRMKNSLRYVTDNDESPEYKRM-
KIQSRSDQDSVSNDNQEGAFNSLREVICNLSDEVEEYIHNDEDDDNHKINASECDGLMNDTNEENEGVLIDNIGAY-
SLDTNGSITTLTKEYTLENMKSNDMNSTNILRGIDNQSNISNRNVAKLFLSNLRCSNEQYYTNNGQITVLDSDDL-
TIEHGSDTQKPPYSYAQLIVQAVSTARDRQLTLNGIYNYISKNYPYFKAHDKGWQNSVRHNLSLNRYFIKVPRGQDEP-
GKGSFWRVDPAYENKLIAQAFRKRRLRNNASGIVNEDGLVSQVLNVSNIDANSQVSNAKVFGINKSGQIASRTPNFGN-
VITLKRACDINTQNGGTTYVLNSSSALSSNNEPQKYIVSNRNNSNFESLSNISSNKTGTIFKCQSSKTVYLDGNNI-
KISGTSGVVSKNRGLQIRSQGNNPGTLISLGGKLFSTANIKSLQIQQQSHPRILTSNSSFDPSNSNKSSITNTINPK-
GNNFYIISPTGIKNQKTLTSSFDDNIVHNSLQSPITISNSNRRIINLQTQSPNIRKIDLSNGSMNKETDLKSPIILK-
KVKINNNFLTLKSEPKINNIRQNQDSVNGDVNMRYSIQSCDKNFNDNLSDEDIKLQDESDSQVFISDRDIDNSEMDH-
FLISHSDPSVSGLPLQEHSEPGSSPDLYKHDDDMWPEDEVHKMESLKYSMHDGIVDTECD

>Smed_foxK1-2.2
MSVDFEDETLKSDSDIIYSDYIEYARITFGNENAYFMKTEKITIGRNSADGTVDIDVGPHTFVSRKHLEMMYSYK-
KLKIKCLGKNGIFIDNYFKAHSAIPYELPYECTLRFPSTNVEVNVKQLVGRKSGKCNSMESDNDANDFIDTRTLV-
NSRKRKATQTINENFSYDIFESVNITNKEITINNFDVVPMRRPMEMKRDISDSNMKSDLDRTMTESVIHNESSTINS-
SPNYNNHNNNSNFVTHHRKPMNSAPVISSGTCTTGLNPKLEQQIPSIINSSHNTTTQTSTISTNKKFQNVCIEKSN-
TIPQIVLNARPYRSYSYKGEVIYTHNSQVFSSIYDSKDLAIEYGTDTQKPPYSYAQLIVQAVSSSRDRQLTLNGI-
YNYISKHYPYFKSHDKGWQNSVRHNLSLNRYFIKVPRGQDEPGKGSFWRVDGAYESKLIAQAFRKRRLRSNNCGSIC-
SENIPSASRVFSLSSGRAGSLCAVSDRNAKLNNIFTLRRAGDQKAPTTYVFRKMSDTQQLTENVFGSAAKLISKQSD-
VNNSSNTISPDNVFHTNHNHNNNINISVNSNSKSNTSSKLGQKRIFQPPISMINKPPPSNANGNGNANSNSTIFSIGG-
KVFSTNKLKPIQLLSTKNSASFVSVKQQSPQQQHHQQTGNKYILISTSQIDNQSSQNSHNNNIHSNQMISITGSGTL-
NRHQEPDHHSDHQTRASNKPPRINPIILSDNQLEMAHHYKSQQPISPLSVSETCEFPLDEPSQLDSSSTELGPLSP-
GIKFTQSLQQEDEDDDLDLEMDQFLDSYSILDHSNDIDSPADLYKEAATEIDDYCVWQDECLGLQDVEIEIH

>Smed_foxK1-1
MDNEFIHARISGINILYLMKNNTCVIGRDVSSKVDLTITSSPCISRMHLKLIANDNRLFLKCFGKNGIFINES-
FQVYTLDEVPLPALSTIRFPSTNIELQIESRNYLLENSKKKFPLKKRLYMTNKDCDSLGIDENYLSESSVDAI-
ELLSQRNHRMHLQNSSDMIVHSPTDQGTETIKPPYSYAQLIIQAIISEESQQMTLSEIYRYISKNFPYYKMNQKG-
WQNSIRHNLSLNRYFIRIPRSHNNCGKSAFWKLDKSQEAQLIKQAFWKRRMKNSFVIVNKPANNNNDNDNNNNVNNNNN-
```

NNFKRNASQEHCQSSSSPLLAPIKTIASNSSMIIDQSTLSSDSIRNTEVPMYFTPFDI

>Xlae_foxK1
MAVAVCGAVVPVVARLEGREFEYLMKKRSVTIGRNSSQGCVDVSMGHSSFISRRHLEIFIGGSGDGDDADVGDFYL-
RCLGKNGVFVDGVFQRRGAPPLQLPRVCTFRFPSTNIKITFTALAIDKKQKLEAPESPVKPVQQISPLTIHIPDNI-
AHLISPLPSPTGTISAANSCPSSPRGAGSSGFKFGRVIPPDLIAEAAQSENDKDASGGDSPKDDSKPPYSYAQLIVQA-
ITMAPDKQLTLNGIYTHITKNYPYYRTADKGWQNSIRHNLSLNRYFIKVPRSQEEPGKGSFWRIDPASESKLVEQA-
FRKRRPRGVPCFRTPLGPLSSRSAPASPNHSGVFSAHSSGVQTPESLSREGSPIPLEPDASVIHPKLAVIQEAR-
FAQSAPGSPLSSQPVLITVQRQLPQTIKPVTYTVAAPVTTATSQQAVMQTVHVVHQIPAVSVTNVTGLTPINTYTVG-
GQTMVAQAAVMAQPKLEHQENGDHKEVKVKVEAIPAIGHPALTTASRIIQTSSSAPLQTVTIVQTPLGQHQLPIKA-
VTQNGTHVVPITTAIQGQVTTANSSYSLIESPWQWRGNGTRAASPLHMLATHASASASLPTKRQNGDQSEQPDIKRGKT-
DEREVLAMTGLDAQSEMAMAASNEQENQK

>Hsap_foxK1
MAEVGEDSGARALLALRSAPCSPVLCAAAAAAAFPAAAPPPAPAQPQPPPGPPPPPPPPLPPGAIAGAGSSGGSS-
GVSGDSAVAGAAPALVAAAAASVRQSPGPALARLEGREFEFLMRQPSVTIGRNSSQGSVDLSMGLSSFISRRHLQLS-
FQEPHFYLRCLGKNGVFVDGAFQRRGAPALQLPKQCTFRFPSTAIKIQFTSLYHKEEAPASPLRPLYPQISPLKIHI-
PEPDLRSMVSPVPSPTGTISVPNSCPASPRGAGSSSYRFVQNVTSDLQLAAEFAAKAASEQQADTSGGDSPKDESKP-
PFSYAQLIVQAISSAQDRQLTLSGIYAHITKHYPYYRTADKGWQNSIRHNLSLNRYFIKVPRSQEEPGKGSFWRID-
PASEAKLVEQAFRKRRQRGVSCFRTPFGPLSSRSAPASPTHPGLMSPRSGGLQTPECLSREGSPIPHDPEFGSKLAS-
VPEYRYSQSAPGSPVSAQPVIMAVPPRPSSLVAKPVAYMPASIVTSQQPAGHAIHVVQQAPTVTMVRVVTTSANSAN-
GYILTSQGAAGGSHDAAGAAVLDLGSEARGLEEKPTIAFATIPAAGGVIQTVASQMAPGVPGHTVTILQPATPVTL-
GQHHLPVRAVTQNGKHAVPTNSLAGNAYALTSPLQLLATQASSSAPVVVTRVCEVGPKEPAAAVAATATTTPATAT-
TASASASSTGEPEVKRSRVEEPSGAVTTPAGVIAAAGPQGPGTGE

> Hsap_foxK2
MAAAAAALSGAGTPPAGGGAGGGGAGGGGSPPGGWAVARLEGREFEYLMKKRSVTIGRNSSQGSVDVSMGHSSFISR-
RHLEIFTPPGGGGHGGAAPELPPAQPRPDAGGDFYLRCLGKNGVFVDGVFQRRGAPPLQLPRVCTFRFPSTNIKIT-
FTALSSEKREKQEASESPVKAVQPHISPLTINIPDTMAHLISPLPSPTGTISAANSCPSSPRGAGSSGYKVGRVMPS-
DLNLMADNSQPENEKEASGGDSPKDDSKPPYSYAQLIVQAITMAPDKQLTLNGIYTHITKNYPYYRTADKGWQNSI-
RHNLSLNRYFIKVPRSQEEPGKGSFWRIDPASESKLIEQAFRKRRPRGVPCFRTPLGPLSSRSAPASPNHAGVLSAHSS-
GAQTPESLSREGSPAPLEPEPGAAQPKLAVIQEARFAQSAPGSPLSSQPVLITVQRQLPQAIKPVTYTVATPVTTST-
SQPPVVQTVHVVHQIPAVSVTSVAGLAPANTYTVSGQAVVTPAAVLAPPKAEAQENGDHREVKVKVEPIPAIGHATL-
GTASRIIQTAQTTPVQTVTIVQQAPLGQHQLPIKTVTQNGTHVASVPTAVHGQVNNAAASPLHMLATHASASASLPT-
KRHNGDQPEQPELKRIKTEDGEGIVIALSVDTPPAAVREK

```
Smed-foxK1-2.1/1-826  MSGDYYDDSLD-----------------------------------------------
Smed-foxK1-2.2/1-834  MSVDFEDETLK-----------------------------------------------
Smed-foxK1-1/1-358    ----------------------------------------------------------
Xlae-foxk1/1-641      --------------------------MAVAVCGAVVP---------------------
Hsap-foxK1/1-733      MAEVGEDSGARALLALRSAPCSPVLCAAAAAAAFPAAAPPPAPAQPQPPPGPPPPPPPP
Hsap-foxK2/1-656      ----------------------MAAAAAALSGAGTPP---------------------


Smed-foxK1-2.1/1-826  ----------------SDNNL-----P----------------QYARITFFGQVPYIM
Smed-foxK1-2.2/1-834  ----------------SDSDIIYSDYI----------------EYARITFGNENAYFM
Smed-foxK1-1/1-358    ---------------MDNEFI----------------------HARIS-GINILYLM
Xlae-foxk1/1-641      ----------------------------------------VVARLE-GREFEYLM
Hsap-foxK1/1-733      PPGAIAGAGSSGGSSGVSGDSAVAGAAPALVAAAAASVRQSPGP-ALARLE-
                      GREFEFLM
Hsap-foxK2/1-656      ----AGGGAGGGGAG-------GGGSP------------PGGWAVARLE-GREFEYLM


Smed-foxK1-2.1/1-826  QKERVIIGRNSAAGSVDIDVGAVTFVSRKHLELTY---------------SYQK----
Smed-foxK1-2.2/1-834  KTEKITIGRNSADGTVDIDVGPHTFVSRKHLEMMY---------------SYKK----
Smed-foxK1-1/1-358    KNNTCVIGRD-VSSKVDLTITSSPCISRMHLKLIA---------------NDNR----
Xlae-foxk1/1-641      KKRSVTIGRNSSQGCVDVSMGHSSFISRRHLEIFI--GGSGDGDDA---------DV-
Hsap-foxK1/1-733      RQPSVTIGRNSSQGSVDLSMGLSSFISRRHLQLSF---------------QEPH----
Hsap-foxK2/1-656      KKRSVTIGRNSSQGSVDVSMGHSSFISRRHLEIFTPPGGGGHGGAAPELPPAQPRPDAG


Smed-foxK1-2.1/1-826  -LKVKCLGKNGIFIDNIFKSHSFIPYELPSKCTLRFPSTDVQFCVEQLVGIKSSDRGRS
```

```
Smed-foxK1-2.2/1-834  -LKIKCLGKNGIFIDNYFKAHSAIPYELPYECTLRFPSTNVEVNVKQLVGRKS---GKC
Smed-foxK1-1/1-358    -LFLKCFGKNGIFINESFQVYTLDEVPLPALSTIRFPSTNIELQI--------------
Xlae-foxk1/1-641      DFYLRCLGKNGVFVDGVFQRRGAPPLQLPRVCTFRFPSTNIKITFTAL-----------
Hsap-foxK1/1-733      -FYLRCLGKNGVFVDGAFQRRGAPALQLPKQCTFRFPSTAIKIQFTSL-----------
Hsap-foxK2/1-656      DFYLRCLGKNGVFVDGVFQRRGAPPLQLPRVCTFRFPSTNIKITFTAL-----------


Smed-foxK1-2.1/1-826  GRMKNSLRYVTDNDESP---------EYKRMKIQSRSDQ-----DSVSNDNQEGAFNSL
Smed-foxK1-2.2/1-834  ----NSME--SDNDANDFIDTRTLVNSRKRKATQTINENFSYDIFESVNITNKEITINF
Smed-foxK1-1/1-358    ----ESRNYLLENSKKKF--------PLKKRLYMTNKDC------SLGIDENYLSESSV
Xlae-foxk1/1-641      -----------------------------------------------AIDKKQKLEAPESPV
Hsap-foxK1/1-733      ------------------------------------------------YHKEEAPASPL
Hsap-foxK2/1-656      -----------------------------------------------SSEKREKQEASESPV


Smed-foxK1-2.1/1-826  REVICNLSDEVEEYIHNDEDDDNHKINASECDGLMNDTNEENEGVLIDNIGAYSLDTN-
Smed-foxK1-2.2/1-834  DVVPMRRPMEMKR----DISDSNMK---SDLDRTMTESVIHNESSTINSSPNYNNHNNN
Smed-foxK1-1/1-358    DAI-------------ELLSQRNHRM---------------------------------
Xlae-foxk1/1-641      KPV-----------Q-QISPLTIHI---------------------------------
Hsap-foxK1/1-733      RPL-----------YPQISPLKIHI---------------------------------
Hsap-foxK2/1-656      KAV-----------QPHISPLTINI---------------------------------


Smed-foxK1-2.1/1-826  -----------------GSITTLTK---EYTLENMKSNDMNSTNILRGIDNQSNISN--
Smed-foxK1-2.2/1-834  FVTHHRKPMNSAPVISSGTCTTGLNPKLEQQIPSIINSSHNTTTQTSTISTNKKFQNVC
Smed-foxK1-1/1-358    ----------------------------------------------------HLQN--
Xlae-foxk1/1-641      -----------P--------------D-NIAHLISPLPSPTGTI-------SAANSC
Hsap-foxK1/1-733      -----------P--------------EPDLRSMVSPVPSPTGTI-------SVPNSC
Hsap-foxK2/1-656      -----------P--------------D-TMAHLISPLPSPTGTI-------SAANSC


Smed-foxK1-2.1/1-826  --RNVAKLFLSNLR------CSNEQYYTNNGQI--TVLDSDDLTI--------------
Smed-foxK1-2.2/1-834  EKSNTIPQIVLNARPYRSYSYKGEVIYTHNSQVFSSIYDSKDLAI--------------
Smed-foxK1-1/1-358    -----------------------------------------SSDMIVHSP-----------
Xlae-foxk1/1-641      ----------SSPR------GAGSSGFKFGRVI------PPD--LIAE---AAQSENDK
Hsap-foxK1/1-733      ----------ASPR------GAGSSSYRFVQNV------TSDLQLAAEFAAKAASEQQA
Hsap-foxK2/1-656      ----------SSPR------GAGSSGYKVGRVM------PSDLNLMAD---NSQPENEK


Smed-foxK1-2.1/1-826  v---EHGSDTQKPPYSYAQLIVQAVSTARDRQLTLNGIYNYISKNYPYFKAHDKGWQNS
Smed-foxK1-2.2/1-834  ----EYGTDTQKPPYSYAQLIVQAVSSSRDRQLTLNGIYNYISKHYPYFKSHDKGWQNS
Smed-foxK1-1/1-358    ---TDQGTETIKPPYSYAQLIIQAIISEESQQMTLSEIYRYISKNFPYYKMNQKGWQNS
Xlae-foxk1/1-641      ASGGDSPKDDSKPPYSYAQLIVQAITMAPDKQLTLNGIYTHITKNYPYYRTADKGWQNS
Hsap-foxK1/1-733      TSGGDSPKDESKPPFSYAQLIVQAISSAQDRQLTLSGIYAHITKHYPYYRTADKGWQNS
Hsap-foxK2/1-656      ASGGDSPKDDSKPPYSYAQLIVQAITMAPDKQLTLNGIYTHITKNYPYYRTADKGWQNS


Smed-foxK1-2.1/1-826  RHNLSLNRYFIKVPRGQDEPGKGSFWRVDPAYENKLIAQAFRKRRLRNNASGIVNEDGL
Smed-foxK1-2.2/1-834  RHNLSLNRYFIKVPRGQDEPGKGSFWRVDGAYESKLIAQAFRKRRLRSNNCGSICSENI
Smed-foxK1-1/1-358    IRHNLSLNRYFIRIPRSHNNCGKSAFWKLDKSQEAQLIKQAFWKRRMKN----------
Xlae-foxk1/1-641      IRHNLSLNRYFIKVPRSQEEPGKGSFWRIDPASESKLVEQAFRKRRPRGVPC-------
Hsap-foxK1/1-733      IRHNLSLNRYFIKVPRSQEEPGKGSFWRIDPASEAKLVEQAFRKRRQRGVSC-------
Hsap-foxK2/1-656      IRHNLSLNRYFIKVPRSQEEPGKGSFWRIDPASESKLIEQAFRKRRPRGVPC-------


Smed-foxK1-2.1/1-826  -VSQVLNVSNIDANSQVSNAKVFGINKSGQIASRTPNFGNVITLKRACDINTQNGGTTY
Smed-foxK1-2.2/1-834  PSASRVFSLSSGRAGSLCA------------VSDRNAKLNNIFTLRRAGD--QKAPTTY
Smed-foxK1-1/1-358    ---------------------------------------------------------SF
Xlae-foxk1/1-641      --------------------F---------RTP--------------------
Hsap-foxK1/1-733      --------------------F---------RTP--------------------
Hsap-foxK2/1-656      --------------------F---------RTP--------------------


Smed-foxK1-2.1/1-826  VL-------------------------NSSSALSSNNEPQKYIVSNRNNSNFESLSN
Smed-foxK1-2.2/1-834  VFRKMSDTQQLTENVFGSAAKLISKQSDVNNSSNTIS----PDNVFHTNHNHNNI---N
Smed-foxK1-1/1-358    VI------------------------------------VNKPANNNNDNDNNNNV----
Xlae-foxk1/1-641      --------------LGP----------------LSSRSAPA-----SPNHSGV----
Hsap-foxK1/1-733      --------------FGP----------------LSSRSAPA-----SPTHPGL----
Hsap-foxK2/1-656      --------------LGP----------------LSSRSAPA-----SPNHAGV----
```

```
Smed-foxK1-2.1/1-826   ISSNKTGTIFKCQSSKTVYLDGNNIKISGTSGVVSKNRGLQI--------------RSQ
Smed-foxK1-2.2/1-834   ISVN--------SNSK-----------SNTSSKLGQKRIFQPPISMINKPPPSNANGNN
Smed-foxK1-1/1-358     ----------------------------------------------------------
Xlae-foxk1/1-641       FSAHSSGV----QTPESLSREGSPIPLEPDASVI-----HPKLAVIQEARFAQSAPGSP
Hsap-foxK1/1-733       MSPRSGGL----QTPECLSREGSPIPHDPEFGS--------KLASVPEYRYSQSAPGSP
Hsap-foxK2/1-656       LSAHSSGA----QTPESLSREGSPAPLEPEPGAA-----QPKLAVIQEARFAQSAPGSP


Smed-foxK1-2.1/1-826   GNNPGTLISLGGKLFSTANIKSLQIQQQSHPRILTSNSSF--DPSNSNKSSITNTINP-
Smed-foxK1-2.2/1-834   ANSNSTIFSIGGKVFSTNKLKPIQL-------LSTKNSAS--FVSVKQQSPQQQHHQQ-
Smed-foxK1-1/1-358     ----------------------------------------------------------
Xlae-foxk1/1-641       LSSQPVLITVQRQLPQTI--KPVTYTVAAPVTTATSQQAVMQTVHVVHQIPAVSVTNV-
Hsap-foxK1/1-733       VSAQPVIMAVPPR-PSSLVAKPVAY--MPASIVTSQQPAGHAIHVVQQAPTVTMVRVVT
Hsap-foxK2/1-656       LSSQPVLITVQRQLPQAI--KPVTYTVATPVTTSTSQPPVVQTVHVVHQIPAVSVTSV-


Smed-foxK1-2.1/1-826   KG----NNFYI-----------ISPTGIKNQKTLTSSFDDNIVHNSLQSPITISNS---
Smed-foxK1-2.2/1-834   TG----NKYIL-----------ISTSQIDNQSS-QNSHNNNIHSNQM---ISITGSG--
Smed-foxK1-1/1-358     -------------------------------NNNNNNNFKRNAS------------
Xlae-foxk1/1-641       TGLTPINTYTVGGQTM-----VAQAAVMAQPK-LEHQENGDHKEVKVKVEAIPAIGHPA
Hsap-foxK1/1-733       TSANSANGYILTSQGAAGGSHDAAGAAVL--D-LGSEARGLEEKPTIAFATIPAAGGVI
Hsap-foxK2/1-656       AGLAPANTYTVSGQAV-----VTPAAVLAPPK-AEAQENGDHREVKVKVEPIPAIGHAT


Smed-foxK1-2.1/1-826   ----NRRIINLQTQSPNIRKIDLSNGSMNKETDL--KSPIILKKVKINNNFLTL---KS
Smed-foxK1-2.2/1-834   --TLNRH------QEPD-HHSDHQTRASNKPPRI---NPIIL----SDNQLEMAHHYKS
Smed-foxK1-1/1-358     ----QEH-----CQSSS------------------SPLLA------------------
Xlae-foxk1/1-641       LTTASRI-----IQTSS-------SAPLQTVTIV--QTP--LG-------------QH
Hsap-foxK1/1-733       QTVAS--------QMAP-------GVPGHTVTILQPATPVTLG-------------QH
Hsap-foxK2/1-656       LGTASRI-----IQTAQ-------TTPVQTVTIVQ-QAP--LG-------------QH


Smed-foxK1-2.1/1-826   EPKINNIRQNQ-----------DSVNGDVN-MRYSIQSCDKNFNDNLSDEDIKLQDESD
Smed-foxK1-2.2/1-834   QQPISPLSVSETCEFPLDEPSQLDSSSTELGPLSPGIK-----FTQSLQQED-------
Smed-foxK1-1/1-358     --PIKTIASNS--------SMIIDQST----------------------------
Xlae-foxk1/1-641       QLPIKAVTQNGTHVVPITTAIQGQVTTAN---SSYSLI-------------------E
Hsap-foxK1/1-733       HLPVRAVTQNGKHAVP---------TNS----------------------------
Hsap-foxK2/1-656       QLPIKTVTQNGTHVASVPTAVHGQVNNA----------------------------


Smed-foxK1-2.1/1-826   SQVFISDRDIDNSEMDHFLISHSDPS------------------------VSGLP
Smed-foxK1-2.2/1-834   -----EDDDLD-LEMDQFLDSYS------------------------------I
Smed-foxK1-1/1-358     ----LSSDSIRNTEVPMYF----------------------------------
Xlae-foxk1/1-641       SPWQWRGNGTRAASPLHMLATHASAS-------------------------ASLPT
Hsap-foxK1/1-733       ----LAGNAYALTSPLQLLATQASSSAPVVVTRVCEVGPKEPAAAVAATATTTPATATT
Hsap-foxK2/1-656       ----------AASPLHMLATHASAS------------------------ASLPTK


Smed-foxK1-2.1/1-826   QEHSEPGSSP-DLYKHD----DD--MWPEDEVHKMESLKYSMHDGIVDTECD----
Smed-foxK1-2.2/1-834   LDHSNDIDSPADLYKEAATEIDDYCVW-QDECLGLQDVEIEIH------------
Smed-foxK1-1/1-358     --------TPFDI-----------------------------------------
Xlae-foxk1/1-641       RQNGDQSEQP-DIKRGK---TDER----E--VLAMTGLDAQSEMAMAASNEQENQK
Hsap-foxK1/1-733       SASASSTGEP-EVKRSR---VEEP----SGAVTTPAGVIAAAGPQGPGTGE-----
Hsap-foxK2/1-656       RHNGDQPEQP-ELKRIK---TEDG----EGIVIALS-----VDTPPAAVREK----
```

Annex III - Fox Family analysis in *Schmidtea meditteranea*

Nucleotide sequences of all *Smed fox* genes.

```
>foxA1-1
atgcttggaaaaaatccttatgaaactgcaatgagcaacgtgtattctctacctccgggaggttctatttacaatat-
gaacccgatgagtatatcatcagctggctacaactctcaacaagtatcaacactatcgttgaacttgaccggaatcg-
gacctcattcattaagcccaatgagtgcaagcatgtcgggtatagctgcaatggccggtggaatgagacaaggtctt-
gagttgggtcttggtagaagtgatagtccaagagataaaaattcaatttccaataacaaccgaccatatcaaagaagt-
tacactcatgccaagcctccatacagttatataagtttgataacaatggcgattcaaaattctccagtaaacatgtg-
cactctatcggagatctatcaattcattatggatcattttccatactatcgtcaaaatcaacagcgatggcagaattc-
gattcgacattctttgtccttcaacgattgctttgttaaggttagtagaagcccagaaaaaccaggtaaaggct-
catattggaccttgcatcctcaatcaggtaacatgtttgaaaacggttgttatctcagaagacaaaagcgattcaaa-
gatccacacagagaaatcggcagacagagtcaaagagctgccactggtcctggatcaaatgtcacagaaaacaat-
cacgacaacgcatcgcaagaagctagtgataacgcagaaagtgatacgaaacccaacatcaagcaacttgatttat-
caagcgatctcttaactaatcaaggtcataatattaaaaatactaatccaacttctgttagtcagagttgttcgat-
gtttcatcggaaaaaggaaaactgctcaccagtagaaatgaaattgaataaccaaaaccaacaatcaaaccagcaagaa-
catccacaaatccattacaatcccaatcagcaattctactcaaatcagcaaaacattttccaacaaagttctcttgat-
cattacagtctattagcatccgatgatcctcttggtcaaggtatgcacttgccaccaggtgcaaatagtgttttcg-
gactttacggggcacataacttaccaaacgatgatcaaatttcagtgtcattaccatcgatatccttatccggacatc-
cgtatgacaatttatcaacagcaatggcatatcaatatgaagcatctcaacacaattcttcattactaacgacaag-
taatccgttctcaatagatcgtttgatgcatccaagactagtcgctgcagcgatgggggtcagtccccatgatactc-
tatacgcaggagctaccggcccatcagttgatctagaacacatgaaatactactcaaactacaacaatgtgcctcct-
tattcctctgcaatgtctgactactacaaatatgtacaaatcctcagccgggcaacagcgacatgagtctt
```

```
>foxA1-2
atgagtgagttttctccccagtcagtgatacatcaattcgaatgaatcgacagtttcatcaagaatctttttcccaa-
caagtcaccattggacgcatcattagatggcaattcacataaacgaatctcatgttcgaaacctccatacagctacat-
cagcctcattactatggcgattcaaaattctccccagaaaatgtgtacattgtcggaaatctatcaatttatcatg-
gactcttttccatattatcagcaaaaccaacagcgctggcaaaactcaatccgacattctctttccttttaacgatt-
gtttcatcaaagctagccggtgcacagagaaaccgggaaaaggctcattttgggttctccatccggattcgggaaatat-
gtttgaaaatggttgtttttttgagacgacaaaaacgatttaaagatccaagaaaacacaaattgattttattaggaaaa-
gattcagaaaaacaaatcaacactcaccggatttccgagttgagaattcatccaattgataatccatcgaattggatc-
catgcggaaaaatcaacggaattgattgctttggaaaaatccaataattcatcgatattttcaattccaccaatag-
tagacaaatcgatttcgattccaatatcaccatgccgtcgatttattcagagaatttcctcccaaatctttcgaat-
gatcgattcctcattgacaaaattatgcaccataaaatgacgat
```

```
>foxAt
atggcaatccgatcgtcgcctaatcaaaaacaaactttgaatggaatttatgattggattaggaaacattttg-
catatttcaaattggatactcaaagatggcagaactcaattcgtcatgctctgtcatttaatgattgcttcataaaa-
ttggcgagaccaattggtgaatctggaaaaggttgctattgggctattcatcctgaagcaaaagatcattttcaattc-
ggatcattattacgtagatataaaaagttcactcagtcagatagaaatcacaaaatttggaattatttcccatattct-
caccaggcttcagcaatcgtcacaaaaaatattgcttttccttataataccaatgatcgatttttttattataaa
```

```
>foxC2-1
atgcagctttgcaaattcgaagttaattcatcaaatagttcaataacagcttccggatcaaatggaccattattttt-
gccgtctttagcaacggtagctaacaactgtcatttacctctaaatgtcagtaacagcaataacaaccaaaatgat-
gaacctccttttatttcatggaatccaactaacatggctgcttgctacaatcaaggacttgcatctccaaaaccag-
catcattttgggctgcttcaaacaatggatcgcttaattcaaacaattttatgtcttgcctagctaatcctggt-
caaggtttaattcaacagacatatgaacacatgtccagttattatgggaattctagtcttatcccaccatacagcac-
gagttattcaatgggcaatagtccaattccaaatgcaaggtttctgacacactacggaagccattcaaatcatc-
caaatatgtcctcatcaagaaatccatcaagagatcttgtaaaacctccctattcatatatagctttgattaccatg-
gctgttcatgcgcatccagaaaaaaaagtcaccttaagtgggatatatcaatttataatggaaaaattcccatac-
tatcgtgaaaataaacaaggctggcaaaattcaattcgacacaatttatcgttaaatgaatgttttataaaaatac-
caagagatgacaagaaacccggaaaaggttcatattgggcactgcaccctcaagcgtacaatatgtttgaaaacg-
gttcttttttaagaaggcgacgacgatttaaagcaaagatgttctccaggatcgtgaagataggaaaagaaaacaa-
caggaagatgaatgtaacatttttacatctacaagcaaagatgacgatgaaaaaaatgaaaaactttcagttatatc-
taaaaatggagatttaaacaaaagttattcagaaagtgtcgaacattcagatcagagtcaaagtataatagagaat-
gaattatcacaagaaacaaaacttggtcatgattcatcaattaacaacgtgaaccatacaaacaatgaaacattc-
```

caaaatagaacatcacaattcgaaacatctttaggtggatcagaggataacccttttcatttcgacgataacaagaatt-
taacttcatttaacaattttggcaatggattttttcatgaaaaatgaatcatcagattcaataccatttccattaactct-
gacaaacacaaatgaagggctctcagtaaataataatacgaattcagactactttcaattcatgggtagatatgcaggt-
caagaatcaactgacgtttccacggcgtttagttctgtcacaaatcacagtggagatttagaagatcatattcaaaat-
tatcagcaattgtctgcatttcaaaattacactaacccacattctagtctgacagcagccgctgcagctgcttggtat-
gcagcagctcatgataccaattcagtcctaggatatccaatttctaataactacaatgcatcatctagcatattag-
gctgcaatccgaatatgcctttttctggttccttgtctcctcattcagcacaggtatcaaataccagttcatcgat-
gaccactagtcaaagcaatacaatatcatcattttcaataacactgatttggtgggctcattaaatttcggtgctg-
gaatactctcttcaaatagtcctaattatattccaacacatcatcccagtttaacaaattattacttaatt

>foxC2-2
atgaatgttgatcacatctcggtattaactgcggcaaattcagcgccacaaattccgcgccatccaatttccgat-
catgcgtcgttgttgtattattcatcagcttatcaatttcccggatgtgaatcttcgagcttcttgtggccaaat-
tcttcatacaattctaattcggattcctggaaagctgcctacttaaattcagatccccatcatctaaattcttactatc-
catcgttgaattctcaagatgttcaacaagatttgcttttcacaaatcaattcattcagattttattccatgt-
cagcctcgtgttacaataattatcaaaaccaattcccaatcattaataccgattcaaaatatcctaattcccctatc-
gataatcagaataatattaaaatggagaataaaaacgataacagatctttagttaagcccccatattcatatatag-
cactgataacaatggcaataacatctcaaccagaccatcgaataactttgaatggcatctatcagtttatttcagaaa-
gatttccttattatcgtgagaataagcaaggctggcagaattctattaggcataatttatcgctcaatgagtgtttcat-
taaggtatctagaggtgaccaaagatctggaaaaggcagttattggactttgcacccgaatgcatataatatgtttt-
gaaaatggatcttttcttcgtagaagacgccgtttttaagtataaggacgatacgatatcaactgaagcaactgcta-
caaactcagatataataatttcaaattctgatgatcttgaaaaggttcatacatccaaagttgaaaatcctcattc-
tataacaaatgacttcaaaggtttgaattttttcattgaagaaaaatagtacttttggcgatgaatgtcaagaactttt-
gactgtcagtcgaaagaatcatattcaaaaattcatcaacattaagccgaatgttataaaagaaatagagaatttcga-
caaacaagcttttgatttcaatcctaatgggttatacaaccaatttccttattttttcaaatattagcccccagaac-
gaatatcctccgtcaaataaccacctcagacctccattgccttacagcatgaatttttatgcagaaaccgaatccatatt-
tattgagcaactatttatctcatcctcaagaaatttcc

>foxD3-1
atgaacaaattttcattatattactctactttaaatcatttatccaaagaagaggaatccaattccatggcattg-
gaaatgaatcatctctattggtcactgttcggtcagcgctctgatgcgaaattaatgaagaattttggagatgactct-
tacttgatgccatatcacagatctaattctatagagtatttaagagcgtctcaaagaataaacatgaaatatgaacaa-
gatatgccccaggcgttcaaagattctcagactattgaacaaataaaaaacttacttcgagatgatgatgtggatg-
tagacgaattgaaagaatgcgaaagtgatgataatatacactcagaaagtgatttagatcgaatatgttcagaaga-
caaaaaagatgttgataataaaataccgatgtctcaagaaagatgcaaatctaaatctcataatgtaaaacctc-
catattcatatatagctctaattaccatggcaattttaagatctcctcaaagaaaattgacattgagtggaatat-
gtgaatttataatgggacgttttccatattataaagatagatttccagcctggcaaaatagtattcgtcataatt-
tatctttgaatgattgcttcataaaaataccaagagaaccgggaaatcctggcaaaggaaattactggacattagatc-
caagatcagaagatatgtttgataatggcagcttcttacgaagaagaaaacgatataagcgtcagttaccatcagaaat-
gttcaatcacaatcaatctcatttgataattcctccaacttcattcagaatgactattccaccaaatccaataa-
cagttcaacaaaaccttgtaaatcagttgttatttcatcaaaatattactaagtcaattattcccaatccgaataatat-
cattcctccattgaatgctattcgttatccaagaccgtttgagaataatttcagaggacatgagttcataggatcgc-
tagttccccccttatccacacgatgggagaatagttattccttcgacatcaaaaactgaatacaacgagccgataaata-
aaagatgcaaactcaattcatcagaagacgaactgtcgacatttcaaaagttttccatttctcacataatttcagac-
gattcaacagataataaaactaaacaatctgaagaatcttatgctaatcagttcataccgcaaacgcttcgactttggc-
catgcaatcctttttccagtggcaaaaccagtgatcaaatcgacttggatacctcctaatttcatcaac

>foxD2
atgaaatccgacagcaagcatagcaaaccccccttactcctatattgccctcattacaatgtcaattctccact-
caaaagaaaagagattaactctgaatgaaatttgttcatttattatattaaatttcccttactatcgtgagaga-
tttccatcgtggcagaattccatcaggcacaatctctcattgaatgattgttttgtcaaacagcctcgggaacccg-
gaagttccggcaaaggaaattattggaatttagaccctgcagcggtcagcatgtttgagaacggaagttttctga-
gacgtcgaacaagatacaaatcatcgaaacaagtgagaaaaaatcaaagtgaagaagaattcaacaaatttcacaaaa-
caataaatgatgaaaatcgtgttcatcgacagatagtaaaaaatatttggaattttccattgaagctttattagag

>foxD3-2
atggtatataaaatcagtaattcaagtaccaaggaaagccggcaactttcaaaatctgaagtaataaaaaatgcttc-
caaacctccatattcgtatatagctctaatagcaatggctatttctagttctccttcaaaaaaacttacattaagt-
gaaatatgtgatttcattatgaagaaattctcttattatcgagatagatttccagtatggcaaaatagtatccgaca-
caatctttctttaaatgattgtttcataaagattcccagagattcaaataatccgggtaaaggcaacttttggagtt-
tagatccccaatctgaaggaatgtttgataatggaagttttctaagacgcagaaggagattcaaaagtagattgccaga-

gataatgagagcaaatattagacggataaatatcaacaatgagattctatcaaaaattctaaactcaacaatttac-
gaatatttgaatccgtgcacaaaacttttatttatgaacttagtgccatatgaggcttttcaaaatgtgttcttccgt-
gacaatatttcaggatcgtctgaaatcagtttgaaatcacaatcttcccataaatctttaaacaataaattttttgattg-
caaacataataaataaagtg

>foxF1-1
atgaaaatttctaacagccacgaaaacctcgctgattcccaagcggatgacactgtggatcaaaatattaaaatt-
gcactcactaaaggtacagaaatatatttaaaacctcccttttcatacatctcgttgattgcaatggcaatc-
caagcttctccctatcaacagtgcacattaaacgagatatacgaatatttaagcgaacatttttaaattcttcagag-
gcgagtatcaaggatggaaaaaattcaattagacataatttatctctcaacgattgtttcatcaaaattccaaaagga-
tttggaagaataggggaaaggacattattggacaatgaatgaaaatgcaagtttttttatttaatgaaaagtgcaatcg-
cagaagaccaagaggatttcgacaaaaattgcagtccatcgaaagaagctcaaattttacattattatccacacca-
cataataatggtgtaacgttttataattattgtgatattccggtccatagtgatgcacaaactaaccacatttattc-
gaatagttttttatttatcaaaaaatgagtgggaaaattattcgaattgcatgtttttgcacaaattcgtgttgtaat-
caattt

>foxF1-2
atgaacaatgcagattccaatcgttcgggagtttatcttcaatcacaattctctcatcaacttctaggatcattt-
gaaagtcctatttccagcacacagcttaaagaaacagctgcagtaaatggccaggaaataaaatcagaaattcat-
tattctccagaattttccacagattatatgccaattaattcagagctttatcaatcaagtttcttgaacaattct-
tacgaccatcaacaatttgtcaattcctcatcattaaattatttcaattccaataaaaagttatgtgggttttt-
tatcatcaacaatgaataactttacggatctagacagaatagattacaatagaattaatgatattaattttt-
gaacaatgctatcagcctttttcaaaactctaaaaatgaacaccatagagaaggaaatgcgttaaatcatttt-
taattcattatctcgacaaattgaattagtgaatggagattccaatagtctttctcaatacattcacttaact-
cagcctttcaataatcaacttactaataacacagatcctcaaagcagtccaaattatgaatctaacgaaatta-
aacctaatgaattacctgtcataaattcatctatcgttactaaagattgtgaaattaacaatacttcatgttta-
aaaaaatgtgaagaaaatgattttgtcagtaatgatagcaattctccaacaaagctcaatgaagaaaatactacc-
gacattgaaaattgtagtaataaagaagatgatattaaaaaagaaaaatctattgaaaaagaaactgatactatt-
gaaaaaatgaaatcagaaaaatctttaaaatcatcaactgcaaaacgtaaatccgaaaaacctgaacaatcatatatt-
gcattaatagttatggctattcaagcttcaccaataagaaaatgtacattaagtgaaatacaattatttgcag-
caaaattttccctttttttcaaggtccttatcaaggatggaaaaattctgttcgacacaatttatctttaaatgaat-
gtttttatcaaacttcccaaaggtatgggcagacctggtaaaggacattactggacaatcgatcctactgctgaatt-
tatgtttgaagacggaggaaatagaagaagaccaagaggtttcagaagaaaagtatctacacgaagcaactctggttc-
cgctttctcgaagtctccatacacagtcacttcagtttcatataataatagcagtccaatgatgttttcacctactaat-
gtacgatgcaatcaattaggaaatagcgcatcatcttgtaattccatgaataatcttaatggaccagttactgacaat-
atgccacctcttttcgatatgtatcatcaaagctttggtctttttactcagcaggaaacaaatgatccaatgaaaact-
gtaatgccaaattatcctcagcacgattccggatttttttacagattctttctatccaaattcatcatcagattc-
cagttttaataaacaagagattccaacagcttcatcattatttcatcctaattttatgtctaatacaattagtcaaac-
gtcaaattcttcaagtatgtctcatccttccatatctccttatgtctcaggaaattctagaaatagtcctttatt-
tacttcgctttcttcaactttgcaagctcagaatatatctgaattaaaaatttctgatcagtcatcatttcattctcct-
caatcatcaacttcttccactgtagcagctgttgccgctgcgtcgctttattcaaattggtttcccaatggacct-
gttgaggcagctaaatatttatagaccattcctcagagtcagagagaatcgaaaatttatatgcatcgagcgctt-
tacctagacttctcaaacatacattcagaaaatgatgaaattcttgaaagaaattcaatgtatcaaccatacatgaca-
gattgtttattaccatacttaccttcgagtcaaatgtttaatatagcaaacatgacagacgtacgttccagtaacata-
aacattgataatgcggacagatccatctgggttcacaacaaattt

>foxG
atgagaaattttccttattataaaaataataaacaaggatggcagaattctattcgacataatttgtctttgaataaat-
gctttatcaaagttcccagaggatatgatgatccaggtaaaggtaactattggatggtagatccggcttgtgaagac-
gtgtatattggtggaataactggcaagttgcgaagacgatcaagttctgtccaacgaatgcaacgtctcggcttgagat-
gcggaccttggccaacttcactcgaccactgtttacgacttggaccaccggctatcagcaacattttcccccatccat-
cagatccactgcaaagcgattcatctaattgttgttctccagtgaaaaatactttctcagattttcagaataaaaatc-
taccttggaatcctatcaatcaaaattgcttttttcccggtgttctttctcatcctcctcttaccccaaacctt-
gacttctacagacatttacgattgcagaatatcattccaaaaatgtggtattcaaatcccgttactttgatgaaggaga-
aaaagattttt

>foxL1t
atggaaaattttgctcgttgcattaacaatcagcttgaagacaattcgaattcgtgctcaattggatcgccatcga-
catcatcgaaaatgatgatggaattaatgaatttaatgccacctcgatttcttggattgaatttcggtttactgaatc-
cgtttccatggaatccaacaatgatgcaccttgagaagccaccgtattcatatatagccctaattgctatggccattc-

gcaatacccccgataaaaagatcacactgagcggaatatacaaattcatcacagacaactttccgttttatcac-
cgaaacaaacaaggatggcagaactcaattaggcataatctcagtctcaatgattgctttatcaagattgcccgt-
gacaaaataatcctgggaaaggcaattactggactctgaatgagaacttcgaagaaatgtttgatcatggaattttc-
gtcgacggagacgaaagaccaaatcactgaccaacaagcaagattgccaagacagtgaaagccatcaaacccgaagtt-
gttcatctgttaattccgactcaattgaaatgagaaataaaacccaagaatcaactccgaatttcttcattgaccagtt-
gcttaaaagtcagtctctgaatccttcggtcattgtcaagcagtcaaccgtcgaaccgaaatcagaattctccta-
aatctaaagaattcattcaataactgttcgaacatgtggccaatttggcagagagacttgatcctgtcaaagtcccta-
caataatttgctactaaactcagacaatccactatcaaacaattcacagtcgttattattcaaaaagccttcacat

>foxQ/D
atggaaaattctttatcacgaaatcaacgcaaaataatttcttttcaatcaacaattccaatgtcatttcaaa-
gatgatccagtcctcaatgcctatttcggttatgccagttggaccacagaacaatccattctctcaaaactcaatg-
atttctcttcaaccattactaaaagcctttcatccggtggtcggggatttctgtaagaaaacaacacaagatgatc-
cgaaaccgaactatagttatattggtctcatttcaatggcaattttgagtacgaaagaaaagaaatggtttttatc-
cgatatttatcaatggatccaagatcattattcatactttcaaacgagaggacctggttggagaaatagcatccga-
cacaatctttctttaaatgattgttttgtcaaagttggtagatcatccaatggaaaaggtcactattggggaatc-
caccctgctaatattgaagatttcaaacgaggtgattttcgtagaagaagagcacaaaggaaagtacggcgagctt-
tagggttgacttgtccagatgaagacgatactccatctccatctccaacacattcgccaaaagcatttgattggc-
caataactgcgcacaatgaaacaccgaatttccagatttctgataattccattcataaaatgaattctattattca-
caacaaaaccgaaaatgacattccactaatttccaattccggatacaatccgtatttctggccgcataacattgaa-
gatgaggacaagttaaaattaatcaaatcaagtcaaaattctagaacatttgacatagaaaatattctcaacccag-
taacaaaaaaagttcagttgaaagaagtgtttataatcagttttttttcacatttaaacttcctaaatatcctcaatg-
gaatacgttttgcccgacagttgccatccactttttttcagagagaatcaagttcaccaatttcttgaattcaaacacc-
gagccattgcaaatagatcagtaaatttaatgttgaaaagtgtaacaaatattgaacaaact

>foxJ1-1
atgttgcagcactgtctcagttatggtgaacgcactagagaggaattggaacaaaatgtctacagcaagtatcaat-
gtcctgatcgtcgatcctcagcgtcctcgaccacttcaacaggatccaatgaatccaatagaagcattaatcga-
caaaagtattcctttttcgcaaatgagttgccagatcttggcgaatacttttcataagaattggtcagaacaatgtg-
gaaatgagtccgggtttgatgattctctcaccaatctcaactggcttcataatgtgaatccaatgaagattaatgag-
caaatgtcccatgttcacgttggatatccgagagaccggctggtgtcgaacagctaccaaccaaggaactatcaca-
catttaaccatcagaatcaggcatccactcaacaggattcctctgcattcacgtttccaatgatttctgccgat-
cagttttcatttgacaaacctgattataaaaataatccaaacttgaagccaacatttacttttgcagttttgatttc-
gtgggccatgaaagaactcgggaaaccaaaaatcactttgtcggacatttacggatggatatcggacaattttgcatac-
taccgctactcagattctagttggcagaattcagttcgccataacctatcgatgaataaatattttcaaaaggttc-
caaggagaaaagatgaacctggaaagggtggattttggcgaatgaatccagataatatcaaagaattggatcacaatt-
tatcaaagatcaaatataacccttcgagtcattcaatcattcctatcaaaacggagaaactgccaatcatttccagt-
tatcgacaaatgaccagctacccccctacagtgtctaccaacagcaatagcaaccatataaataccttcactatc-
caagagcactcgcgaccaatgattagcaatcacaatgcagtcaaacgacctagaaagtctccagtaataactattc-
cagaaaactgggatgcattggaatgtttaaaggagacggacacaaaacaagtaaacaacatcaaaatcaatgaaaa-
caatcatttccctcagtatgaagaactattgttatttgatggcccttctgcttacaaaaatgatcttgacttctcttc-
gatggataattcatctactgattcatttccatatctttctcagtgtccgccaaatagtaatgagatttctctcag-
caatttcgaggatgacagtctgggtgatattcttgctagtccaatgcttctcgatggagaagagatcatggacaacgaa-
gatttcaagtccatggtctaccaaaccgaatatgacttatttccaacctcaaaagatgatctacaagaactcaaca-
cattgttgggcatatat

>foxJ1-2
atgaatctgaataaattgacctcccaggctttgcctatttcaccaatatcggcttcttcaatgagttgtggaaagtat-
gttgacctaagacaaccgtatgcgtcagacccgagtgtcagaagcgatattatctgcagctgggaccaaatcgac-
gattctacgagaatattctataagacccatactacaggacgaccacctttctcccatataagcttgatatgcatggc-
tattcaagacataggtcaatctagaattacctcgactcagatttgcgagtggattataataaatttcccatactacca-
gattctggacaattcctggcaaaactccgttagaaatcttctctcagttagtaagtgttttcaaaaggtaccccgccg-
taaagatgagccagaaaaaggtggctttgg

>foxJ1-3
atggcaatgctagacttaggtaagccgaaaatagctctcaatgaaatctatgaatggattcgtgaaaactttctttac-
tatcgaaaatccgactcgagttggcagaattcagttcgacataatttgtcattgaagaaatgttttgaaaaagttcccc-
ggaaaaaagacgagcgaggcaaaggaggatttttggaggataaatccgaatttcacggagaatctcgggaataatttcat-
caaatatcgtcggcaatttcatgtttatgcaacaccgccgccacttccaccatcgcaaccattggtatcatcgcaaccg-
caaccgtcggcacagacccaatccaatggttaccatttgagcaattccgtgtacaaatatgccggaaataacaataat-

gttgccagtaaaggctggagaggaaattcgaatgccaattgcgaaattgtgaattttaaaccgctggcaaattaccac-
caacaccaacaacagcagcagcagcagtcccaccagcatttgaaactagtgaacattcccagtgtcagatcagtcaa-
caagccatataaaagactgcaaacaactcatatgaaacccatattcaaaaagccatatgcattgtccagtctcaatcag-
tatgaaaccattgaggaaaacgatcgcgtggttgcccaaatgcgcctttatggctctccaagacccaaatctcccac-
gctgttctcattgtgtgaaaacgaaccgttgctatcagaaatatcctccagcaactccaattattcggcgttcggattc-
gaggatttcgattcgacaacttcctcttcattcgatacagaattcagctgctcatcgaatttccacacgtcctcagat-
gacaattactccttgaatattgccctggatatcgatgagttcatcccacagttc

>foxJ1-4
atgctcacccagcatgatagtacatgtctaccaaaaattcatataaagtattcacaacctgaacaacggctttc-
caagaattatctaacgacttcaaaattgaattccattcagatttatacgacagaagttttcaatggaaaaaaactt-
tcagtccttgctgaaaaatttcagaaaaattggttagataaaaaaggaatcgactcaaaagaaatggacgattcact-
tactaatctcaattggttacataatgtaccaagtattttgggaccagaatctcctttgaattcaccagtccatttt-
gctcaatcaggctaatcatcaattattgccgaatctcgaacgaccagctaataatatcattgctaataattttgta-
aataatgagataaatatttccaagcaatcggagtttttccaataacttttcctccaatccttccttcactcatcatctc-
gacagtgctctgaaagatggtggactattagatcctttagtacggcatgattacaagaccaattggtctggcaaac-
cgcctttttcttatgcaacactgatttgtatgtcaatgagagaattgaacaaatcaaaaatcactctgtcggatatt-
tacggatggatcacagataatttcatctactatcgattgtcagattcgagttggcagaattccgtcaggcataatt-
tatctttgaacaaatgtttcgaaaaagtgccaagaagcaaagatgaacgcggtaaggggggattttggagaattaatc-
caaaatatgctgacagattggaatcaaacctcataaaatatcgacggcaatttcctttgtataacatgcctaaaata-
acaaatggtttaccggccgcctcgactttgttatcaagaacaacaaattatagttacatgctctgtcagtcatccg-
taaatactgataatacttggtcaaattcaatgacaactaacaatcaaaactcctttgcatttcgtgatcctcaacc-
gattttgagcgaaaattcgttgccttttaaaagtaaaagaggatgggataattctggtgcaacttatctgccgaatct-
gaccaattataaccaaggaaaattgatctcttctagtgccttttccaacttcaataacgatgtgacaagcaacgaa-
gattattgtcaaaatcgaaaacacaattctaaaacatatcacactcaatcgtatcgaaatcacaacatgacaaagttg-
caaagtcttgatgagtttcgaacatcagaaaccggtcaaaaggtcgacacgcaaatgagaacggaaatcactccgaat-
cattctccgacaatgtattcattgttgaatcatttgaactcacctccgtcccatgtgtctatggacgatatagattct-
gcagcctcatcatcatttgatgcagaattaactgccaatactggatcctgtgccaccctccattctcattccgat-
gatcccgatgaagctcttcggtttggcacaagtcttggagatgatttcatggtgggttccgatgattacttaaattgc-
gattcgattcgtagtcatgatcagttattcgattcggttaatagtcaatttatt

>foxJ1-5
cgccggttactaagacttttataaatatctactaaattagttttgcgcataatttcagttataacttttgagttaat-
taatcagttttctttacaaattgtgccaataaatttacgatatataattccattatacatttcccaattatcgtct-
gacattttagcgttgctaataatcggattatgatttctgaaacgagactcggtttggctgagaaatttcgccaaaattg-
gcaaacacaatatccttgttcgatgattgaaaaatctgaaacgagaaatttccatcaaaactttcttaaacgaaacaa-
cagaaattttaatcaatccttttttcggaggtaaccataatgataatgaagatgattcactgacaaatctgaattg-
gttgcagaacctaaatttattcaatatgtctcaaactcaaccgcctctaagtccttctccgactcatcagatgacaat-
gccttctccccagtttcctttatccatgaacccaatgaaccaattaaaccgaataatacctcaatattttgaaagtag-
tagcctttacgccgaagtcgccatgaatttggagttcatgaatgatgaagccagtgttcattttttttagaatgaacaac-
caacttaggccaccattttcagatataaacttaatatcaatggccatccacagctacaacaaaatgaaagtcagttttc-
cagaaatttgtcagtggatagaagaaacctttccatattacaaatcactttcagatgattggaagacgggaata-
aaaaaccatttgacattgagcaaatgctttcaaaaggttcccagacgaagagatgaaacggaaaaaggcggatttttg-
gagatttagtaatgagttttaaattacctggaagaaatcgaatacaaacattatcgggtgtaaattttgataatc-
gaatcaatcacgacaacgcaaatccctccaatgatatcacttacaagttcggcaaaatcatcaatcggaagcggacaat-
gaataaatttgtcacacaggacgagaatgaaaactcggcgtttcggattccaaaaaagaaatggttttccaatgaga-
tgattgatttccagaacggtaaatcaattttgatccataacaaagccaataataattccaacaacatggacaatatc-
caccacggaatcagagagtgcaaaccaagggagtccttattgctacaatatcaagatggctccttgtgtaacatta-
accaaatgaatctggcaataaacagttgcgcgaatgaatctatgtctaatattggaatttcagtattagctaattcg-
gatctttcacggcttcaagagtcaaatgaacttttccatcagaacataatgtgtaatcaaataaatcccatagttacc-
gcttcccaatctccatgtccagaacaagtcacggaagaggaattcctttctggtcaagcattcactaacaattatata-
acattgaataccaatgctgacagtctaccttccatgcttgacactcaaaccaactacagtcacaatgatttcaatat-
cagcgacctaagcatcactggaatcggactcaagccgcctgaatggtggaatcaaagcttttccgccagtcaaat-
cagtttatccaaattcaattgctcattccgatctgtgaccgtggcgaataatggagaacagatttcccctttctg-
cattggtattgatgaagatcatcacggactttggtcatcgaatgatatttcaaatctttcagaattggacgtgct-
gtttggattgagc

>foxK1-2.1
atgtctggggattattacgacgattcccttgattctgataacaatttacctcaatatgctcgtattacttttttc-
gggcaagttccatatattatgcaaaaagaacgagttattattggtcgtaatagtgctgctggatctgtagacattgat-

```
gttggtgctgtaacttttgtttctcgaaagcatctagaattgacatattcatatcaaaagttgaaagttaagtgccttg-
gaaaaaatggaatttttatcgataatatttttaaatctcattcttttattccttatgaattaccatccaaatgtactc-
tacgctttccaagtaccgatgtgcaattctgtgtcgaacagttagtagggataaaatcatcagacagaggacgttct-
tatggaaggatgaaaaattctttaaggtatgtcactgacaatgatgagtcaccagaatataaacgtatgaaaattcaat-
cacgaagcgatcaagatagtgtttcgaatgataatcaagaaggagcatttaatagtttgagagaggttatatgta-
atctgagtgatgaggttgaggaatatattcacaatgatgaagatgatgataatcacaaaattaatgcgtctgaat-
gtgatggactgatgaacgatactaatgaagaaaacgaaggagttttgattgataatattggggcctattcactt-
gacacaaatgggtcaattacgactttaaccaaagaatacacactagaaaacatgaagtcaaatgacatgaattcca-
caaatattttgaggggaattgataatcagtcaaatatttctaatagaaatgtagccaaactgtttctcagtaatcttc-
gttgctcaaatgaacaatattataccaataatggacagataactgttttggattcagatgatcttactatcgag-
catggatcggacactcaaaaacctccatattcatacgctcaattaatagtccaagcagttagtacggctagggacc-
ggcaattaacacttaatggcatttacaattacatcagtaaaaactatccttatttttaaagctcacgataaaggctg-
gcaaaattcagttcgtcacaatttatctttaaatcggtactttattaaagttccaagaggtcaagatgaaccgggta-
aaggttcattttggcgagttgatccagcttatgaaaacaaattaatagcccaagcatttcgaaaaagaagattacg-
caataatgcatcaggtatcgttaatgaagatggcttagtgtctcaggtcctaaatgtatcgaatattgatgctaacagc-
caagtttccaacgctaaggtatttggcattaataaatctggtcaaattgctagtagaactccaaattttggaaat-
gttattactctgaaacgtgcatgtgatatcaatacacaaaacggtggcacaacatatgtgcttaatagttcatct-
gctcttttcatcaaataacgaacctcagaaatacatagtttctaatcgtaataattccaattttgaatcccttttc-
caatatatcctccaataaaactggtacaattttcaagtgtcagtcatcaaaaacagtttaccttgatgggaataatat-
caaaatatcaggaacttcgggcgtagtttctaaaaaccgaggtttacaaatcagatcacaaggaaataatcctg-
gaacacttattagcctaggaggcaaattgttttcaacagcaaatatcaaatcacttcaaattcagcagcaatct-
catcccagaattctaacatctaattcctcatttgatccttctaattctaataaaagttcaattaccaatacaat-
caatcccaaaggaaataatttctatataatctctcctacaggtattaaaaatcaaaagacgcttacttcatcatttgat-
gacaacattgtccataattctttgcaatctcctatcacaatatcgaattcaaatcgtcgtatcatcaatcttcaga-
cacaaagtccaaacataagaaaaattgatttatcaaatggttctatgaacaaagaaaccgatctgaaaagtcctat-
tatactgaaaaaggttaaaataaacaataattttctcacgttgaaatcggaaccaaaaatcaataacatccgccaaaat-
caagattccgttaatggcgatgtgaatatgagatattctatccaaagctgcgataaaaactttaatgataatctttc-
cgacgaagacatcaaattacaagacgagtcagatagtcaggttttttatatccgatcgtgacattgacaattctgaaatg-
gatcacttttttgatatcacactctgacccctctgtatcaggtctcccgctccaggaacacagtgagccgggatcttcgc-
cagatttgtacaaacatgatgatgatatgtggccagaagatgaagttcacaaaatggagtccttgaagtatagtatg-
cacgacggtattgtcgatactgagtgtgat
```

>foxK1-2.2
```
atgtctgtggattttgaagatgaaactttgaaatcagattcagatatcatttattctgattatattgaatacg-
caagaataacatttggaaatgaaaatgcatatttttatgaaaaccgaaaaaatcaccatcggtagaaatagtgcagat-
ggaacggtagatattgatgttggtcctcacacatttgtttcacgcaaacatttggaaatgatgtattcatataagaaat-
taaaaattaaatgcttgggaaagaatggaatttttatcgataactattttaaagctcattctgcaattccctac-
gaacttccttatgagtgtactctgcgatttcccagtacaaacgttgaagtgaatgttaaacaacttgtcggtagaaagt-
caggcaaatgtaattctatggaatctgataatgacgcgaatgatttttatcgatacccgaacattagtaaacagcc-
gcaagagaaaagcaacacaaacaatcaacgaaaactttagttacgatatatttgaaagtgtgaatatcacaaata-
aggaaattacaataaataatttcgatgtcgtcccgatgcgaaggccgatggaaatgaaaagagacataagtgacag-
caacatgaaatcagatctcgatcggacaatgactgagtcggtaatccacaatgaaagcagcaccatcaacagcagccc-
gaattataacaatcacaataataatagtaacttcgtcacgcatcatcggaaacctatgaattccgccccagtcatat-
catctggtacctgcacaacaggactgaatccgaaactcgaacagcagattccatcaataatcaacagcagtcacaatac-
gacaacgcaaacatcaacgatatcaacgaataagaaattccaaaatgtttgcatcgagaaatccaatacaattcca-
caaatcgttcttaacgcgcgcccatatcggtcatattcatacaaaggtgaagtgatttacactcacaattctcagg-
tattcagctcgatatacgactcgaaggacctggccattgagtacgggacagacactcagaaacctccgtattcatac-
gcccagctgattgtacaagctgtcagctcgtcccgggacagacagctgaccctgaatggcatatacaattacat-
cagcaagcattatccctatttcaaatcacacgataaaggatggcaaaattccgttcgtcacaacttgtcgctgaac-
cgctactttatcaaggttcctagaggccaggatgagcctggcaaaggttccttttggcgtgttgacggtgcctac-
gagagcaaattgatcgcacaagcgttcagaaagcggcgattacgcagcaacaattgcggtagcatctgtagtgaaaa-
cattcccagcgcttccagagtattcagcctcagcagtgggcgagcaggcagtctttgtgcagtttctgatcggaatgc-
caaattgaataacattttcacgctgaggagagctggtgatcagaaggcgcccacgacatacgtgtttagaaagatgagc-
gatactcaacaattgacagaaaatgttttcggatcagcagccaagctgattagcaagcaatcagatgtcaataattc-
tagcaacacaatttctcccgacaatgtctttcataccaatcacaatcacaataatatcaatatcagtgtcaacagcaa-
cagcaagtccaacacaagctcgaaacttggacagaaacgcattttccagccgccaatcagcatgatcaacaagccgc-
caccttccaatgccaacggcaacggcaatgccaatagcaacagcacgatattcagcatcggtggcaaagtcttctcaac-
gaataaattgaaacctatacagctactttcaaccaagaattccgcctcatttgtatcagtaaagcagcagtctccacag-
caacagcaccatcagcaaaccggaaacaaatatatcctcatttccacatcccaaatcgacaatcaaagttcccaaaa-
cagccacaataacaatatccacagcaaccaaatgatcagtattaccggttccggcacccctaaaccgtcaccaggaac-
cagatcacccatagtgaccatcaaactcgggcatcaaataagcccccccgcatcaatccgatcatcctatcagataat-
```

caactggaaatggctcatcattacaagtcccagcaaccgatctcaccattgagcgtatcagaaacatgtgaatttc-
cgctcgatgaaccgagccagctcgattccagttccacggagcttgggcccctctcgcctggcattaagttcact-
cagtccttgcaacaggaggacgaagacgatgaccttgatctggaaatggatcaatttctcgactcatactcgatcctt-
gaccattccaacgatatcgattctcccgctgacctttacaaggaggccgctacagaaattgatgattattgcgtgtg-
gcaagatgaatgcctgggcctgcaggacgtcgagattgagatccattgagct

>foxK1-1

atggataatgaatttattcacgctagaatttctggaattaatattctttatttaatgaaaaataatacatgtgttattg-
gtcgcgatgtcagttctaaagtcgatttaacaattacctcatctccttgtatatcaagaatgcacttgaaattaatagc-
gaatgataatagacttttcttaaaatgctttggaaaaaatggaatatttataaacgagtcttttcaagtgtacacttta-
gatgaagttcccttgccagcattgtcgacgattaggtttccaagcacaaatattgaacttcaaatagagtctaggaat-
tacttattagaaaacagtaaaaagaaatttcctctcaaaaaacgattgtatatgacaaataaagattgtgattcacttg-
gaatcgatgagaattatttatcagaatcatcggttgatgcgatagaattattgtcacaaagaaatcacagaatgcatc-
tacaaaattcaagtgacatgattgtacattctccgactgatcaaggcactgaaacgatcaaacctccatattcgtac-
gcacaactcattattcaggcaataataagtgaggaatctcagcaaatgactttgagtgaaatataccgatatatcag-
taaaaactttccatattacaaaatgaatcaaaaggatggcagaattcaattcgtcacaatttatcattgaatag-
atactttattcgaattccaagatcacataataattgcggtaaaagtgcattttggaagcttgataagtcacaggag-
gcccaattaattaagcaagcattttggaagcgaagaatgaaaaattcattcgttatagtaaacaaaccagccaataa-
caataatgacaatgacaataataacaatgtcaacaacaataacaacaacaatttaaacgaaatgcgagccaagaa-
cactgtcaatccagctcatcgcctctgcttgctcctattaaaaccatagcgtcaaactcgtcaatgataatagatcagt-
caactttgagtagcgattctattcgaaatactgaagttcccatgtactttacaccgttcgacatct

>foxN2-1

atgattgattctaaccagtctatttatgatttgccataccgtgcaagaattatgatgaatgcatgtattgacaat-
caaaagtctggctcaaatacaaaaattaatcatattagcaatattgaaaatcttaaaccgatgaactcaaaatctgt-
gttaaaaactaacaataattacacttattcttctggagatcttcaacctcttcagtggcttcagcaagataccattct-
tagtattactccattagataatgaagaagaatatgttcctggtaaagatgacataatatctataacaacttctc-
cagtgaaagtcgatgtaatttctgaattaaaatcagaagaatcattgtaccaaaattcaacaatcgcagattat-
tatcggtatcaagcgatgcgagccgcaatgaccaacggtcatcaaaatggaggctacagacaacaacacggatatt-
tatcctactcaaactcttcaaccggcgttcgaccaaacaattcaggcagccagcataatctttgtcattcagtcaat-
gaatatcaaaagcctccgttcagttatactcacataattttcatggcaatcgagacaagtcccaataaagccatga-
cagtgaatgatatatacgtttggtgtgaaacgcatttcccatattatacacaagctggagttggatggaaaaatagtct-
gagacacaatttgtcgattaacaaatcattcaagaaatcagtcgagatggaaggggcggacctggtcgaggtgct-
tattgggttgttgaacctcgggaacgcaataatttgattgatgcaattagacgcagcccttgcagccttggtagttttc-
catttctcaacaatacatatccatggggcactggaaattgttactcaaacaacatttttcaatctgcaaatactct-
caacaacaacttagcctcgctaacggcgacacgttctgttgttaacaataagctcaattcttctatcgatcta-
caatcctcatcacgtgggcctttattgcgtattggtgggcaacttactgatgcgtctaatttgattcgatcaactc-
cagttattcagaggctgaacgattccagcaacatcatcgaattatctaaaactcctcaaatcatatccaattccat-
tactttatctacagttcctttaattcagtttgatcaaatcaaatttgaatctgatatgacaacattcaatgaatctata-
gattctaatgatctctgttctgctcaacaagatgaagaagactgtcgtggaaaggtatttattatcgatgactcgac-
cgagtcgaacccagaattttatcagacatatctggacattctaagaggtctgatttataatgacattgaaaaatcg-
gattcaattagatcagaaaataccaaagattgcacagaacacttaaaaaatcggaaacaattttttttctaaaataact-
gtaaatcaagcgatgatattctaccaatagaagatatcgctcctatatctagaaaaataatcaagaaacattgctcct-
taaaacaaaaagaaatgcaaatgcgtaaaatccgtttctgattctggtgggtctgattcatcactatcaaata-
aagaattctctgaattgtctgacacggatttcaattctcgtataagcatgggtttacgaaatcaaaaaaaagaga-
aaaaaaggctatatttcagtaaattaaagtcattaaccaacagtagaaaatcatcgtataatcgaaatcgtatta-
ataacatagtcagagcaccacattcagatcatatttacgcaattctgcaggaatttctcccacccgccaataaatcg-
gaagaacactctccaattaatcaaatgattgatgacaaagaatcccctttcttccacgatggccgatgataaatctc-
tatccctcgatttgtcacaaacatccttatcaaaaagatccatgcacaaattgcagagtcttaaaaaatcaat-
gaaaaagttcaaatttttcagattccgacgaagaaatctttaatactactaattatgagaatatttctaacaaact-
cataagaaattctttgtcattcatcacaagaaacggaaattaaacaagaacatgaattcatcgaaaattatta-
aatataattcccgattgttaaaggaaaattcagaatacataaacaaaaaaccgtaccggaaatataactacgat-
gacgatgaagaacgtgaaatcgttgttatgttaaaagcgaaattacagatgacacctttttctagtccaatatata-
aatctttgaaattcgaggatttcaataatgcatttcagggaactcttgaaatctgagcgataataaatctaatagc-
tacgattctcagtttattcctttcaaaaaaatccgacctaacaaagatacacacgtgaaacggaaagttggcagac-
caaagaaatctgagtgcaattcggtgacttttcgatactcaaatgatcttcgtaacaataaacagtcaaatatta-
aatcgaaaaagaaaacaaacaccgaaacgagatttaaaatacctcaaagaatgcaattcaaacaaaaaagttgcca-
cagaaatcagcaataaagatgagcaagtggcaattagcgtcttagcaaacttacagaaaaagaaccaacaagatcggat-
caaaaatatgtacaatttatcgaattatatccaaagaaacagttcagattcaaacatgagttagcaatattcccccat-
gaagccggattctcatgcatcgccataattgcaattggaaaactgcaaaattacgaattatttttttttaccagaaattc-
tacaattgttttttatgacagtttgataaattcatatctgattttactagttgctaataataatggtgctccttcagt-

caatgcgtttatttatcatctcttcatctctgtataaattccatcttctctgtgtctattggtgt

>foxN2-2
atgaaacctactatactaaaggtatattcaaattttcctgatattccttatgatattaaaaagtgtgaagataat-
catttgattaaaatttcaaatgttgaaaatcctggaattccacaatttaaatcagaaatgcttacaaattcattctgc-
gaaaatgcatatcctgttgaagacagttatattgattccaataaagatttagtgtttttagaagattttttatcg-
gtttcaagaacgtagaaatgctttcattgattcgaaaaaatctcaaacacaaattaatcagtcaaacttgaaacgcttt-
tacgaaagaccgagttttagttatacgcaattaatatatatggctatagaaaattcgccgggtaaagcaatgacagt-
gaacgatatttataattggtgtgaattaaattttccttattatgctaacgcgattcacggttggaaaaattctttaaga-
cacaatttgtccattaacaaatgttttaaaaaaattactaaaatgggtcaaggaggcagagggtctctatggattgta-
gatcctcgtgagaaacaacatttagtttcaaccatgtgtcgggcaaatgtggcttcctttaatgctccatcgtttata-
attcaaaccgctccagcatcactcgaccatcatcgaattgttattggtgatgatcactcattagatattccacaggat-
gatacaattgttttctctgacgattctacatcgtctcaaatacagtctgatttatcattaattcttgaggaaaagcct-
gtgtatagctggacagatacgcgcttaccttgtgactataccgttcctttgtcacctgaaaattgttccgatgcat-
cagagaccgattacaacgattggaatagagcaatgatttctttagaacctattttccctgtcaaacagcaccagat-
caaatcaaaatttcctttaaatcccatcgtgacacctccgagttcttttgaacataattattcaatccagaaccta-
cagataattatcgacgatccgtgcgccattgtacggcatatatttcacccagtaatcaactcaaacgcctcattttg-
tacatccagtgattataatgacgaagaatggcctcgaacaccttctccagtgcctcacgttgctcaaatcgaaccgc-
cacaatctaaactgaaaatcgtcaaatcgggattaaagcgcggacggcctaaaaactcgacaaaacgagaattggaatt-
gaaaaaatctcaaaagcctcagaaaaaactcaatcacaacgaccggatcaataatatgtacaatttgaatagtttcat-
caactcgttg

>foxN2-3
atgactacaaacagctctttcatatcaacaagagttgcaaattacaaactttcatctagctatcctagtatta-
caccattatattcgaaagatttaaattataatggttctttatcatttacttcatatactaaatcatggagtagt-
tatcagaatatggcactaaaagcaacagataatttgcgcaatcaggatttaaccaataattctgatctacaacctt-
taaactggcttcaaaaagaaagtattattgacattgatcctcttgatatcgaagacgaattaaaagatgattcta-
atactcctggatatgatgatttaattgatccccatatttcacatgccaaagttgaagttgaatcacaagaaaattc-
taacatacatgattttgtaagattccaagcaatgcgaaatagcatgaacaacagtaatcgtgttgcgttgagaaacaa-
caactcatttacagactgcaatcaaaattacagtcaccaggctagtggattcgcgaagcctaatttaagttacaca-
caattaatatttatggctatcgaaacctgcaaagacagagctatgacagttaatgatatttaccagtggtgtgaagt-
caattatccttattacaagcatattggaccttcatggaagaattcactccggcacaatctttctatcaataaat-
catttcgcaggatgcctcgtgatggacaaacaactatcggcaattatgagatgaacagtggcgacgacatgagt-
gaagaaacggtatcccaacgagataataagcaaggaggccctggcagaggtgcatattgggtggtagatgacaaggaac-
gaaagaatctgctggattcaatttatcgggcacatcaaactatacccctgagtgatcgaaacggtcagcggattgatc-
caccgtattctcttattgtcaatccaccgaaaactgggcctatgttgcgcattgatggccgttttacagacgcgtc-
gacgtatctgcagtcgaaacgcagcgatgatgaagtagcaaatgagaactacaacatccaatcagattatgcaaat-
gccgagaacctgtcaatgcagaaaatcgagaacacagaagcagccgagtacacatcaatgatgtttaatcttgccac-
taccagcagtccagaaactggtggaggagacacatcggttaaattatcagcgagctacaacagtaactgctcgaga-
atggaaggtaaatgggatgctcattcaatgcctagcaacgatatttattctcatttagaggcagactcagacttgtgt-
caaccgaaattaaaaaatcgacgtaaatcaaagcatgaaaaggttcgcaattacagtgaaatggaatcagctacaaaa-
gatctttttccaaatcgcaaaatcagaaatccgatctcattaaaacaaataatggcatgacaagtgtaaatgtaatt-
gttgctcctcatttcgatcacatctatgctcgttatcaacaattcctaccgccggcaattgccagtgacatggata-
aaggggtattcgggcaagagcatgatataaacccggatgccgatttcctcaagaacggatatcgtgatgatgacgat-
gaagacaatgaaaatgaaggagagcaaggagagattgagaatgaagatgaacaatcgaatttgaactcgttggaagc-
catcaaacaagagaaactcgaaaccgataccgcaatgttggcaacaaaactgaatggctcccgacggatcaagctat-
gctggaatgggaaattgaattctgcagctgtagcattcactcgacagagaaaggttcgcagggtttttaaatcggc-
gagctcggcgtagtctaagaattaggaaaccgagggagccgattgaaccgaatgagacggcaatcaagaaagagat-
gattgaaccgaacgagcagacagtgatagaagatgccagtgatggctacgccacagattccactatgcccatgagcaa-
cagaaaatcgatcacaagaaccaatcggatagttgccaacaaaacccacccggtccgaatccttccagtgaagcgta-
aagcaagtctcagtaaatcaggtcctcccgaaaagcgataccgttcaaaacggaagcgaacgcgtccattgagaatc-
gacgacacctgct

>foxNt
atgtgtgcgtttgatgatagtttgaccgaaatcgattggctaaattcgttatgtattcagaaaaaaagtgaaat-
taaggaagggaatgtcatttctactacaatatcaccatatagccttattacaacacatcataattctccaaat-
gacaattgtaatacttcacgaattcatggatatgataaaccttcatatagctatgcatatttgataaagctggc-
tattgaaagttcaaatagtaaaagatgaagttaaatgaaatttattcgtggattagtgaaacatttccgtattata-
aatatatgcaaaaagataataacaaaggttggaagaatgcaatacggcacaatctctctttgaataaaattttta-
aacgttgtgctaaagattcctcccaattgggaaagggctctttttggacgatcgactctagtctcgacgatgatgtc-

```
caatacccaccggtcaagaaaatgaaattgatgcttaattcaacagaaacagttagatcttcttccatcgaaaaactg-
gatcttatccattcgactgaaatcactcaaaatatctcaaaagaaatgtcaattccggtttcgatagtttacacttat-
tatcttcaagcagcaataatgagtataatttaatcgaatcaagtgctacaccgatgtcgaacatcagcacgtcaagtt-
tagaaatgatgcccgagaaaacccaaataacatttgccaccctcgagtcacgcgtcctaccgaacaatagtgcattac-
ggtgtattccaatgagattgaaaatcataatttctcaaaagaaccttgtaatgacctcggatattcttgtcaaattagc-
tattctaaccaaccgatctatagaaactgtctatttaaatcaaatgtctcaaccacggaattaaacccgatattgt-
caacagtggaatgctcaagagaaatacttcatccaaaggtactaaacaaacatttaaattcagacccaatttcatc-
caagaatgaccttttcacagacaagctcgaggttgatgaagacattatgtctgttaaaagtgccaacaatgatttt-
gtagtaaaaactccagtgctcagccaaatcaattctcagcctgtagagttgggagggaacaattcggaaat-
caaagaagtttatgaggaaatgaattgactaaaggttcattgcgagattcacatttaatttcaagcctaccccaa-
gatacgattgataatcttaatagttttgataatttgagtttttatgattcttttgcttctatccctcagccttt-
gtttgactctgaatcacgtagttttgaactttcccaaagcacaagaaaatatctatataatttatctggtgtta-
aagaaaatctgctacaccgatctcagtggttagatctgcaggaagaagactatcaaaatcaatctctgagtgaga-
catttttttgaaactctgagatctgaagaaatgcgacgaaattataaatcagaaacctatccagatagtcctaatc-
gcgacgtttctcctaccaaagaatatttggaactgaaatcctctctcatcagctacagatacaaaagcgacgccg-
caaactccacttccatgagtctcagcacttcctcaacggaaaagtcatcaaatactctcagatctaaatcagatatc-
tataaaaagtcaaattcattcgacagcacttctgaaattgaacattttaaaagaagcatatccgaaaatcaaacttgc-
gatggagatgaagaaattacagatgtatttccctgggattctatcatctgagattcgtcggacagccttatgtatt-
tatttgtatttgttctgtgttatttttattattggaaatatttgatactttcattgtataaatcagatcatttattat-
caatttta
```

>foxO
```
atgaataacatgctagacgattgccattctccagaattccgagcccgatcgcagacctggggaggaaatgataactct-
tatcgatcgagaatcattgaggcagatctttctcagaaatacacgactttgccagcgattcctacacatctttc-
gaggagtcgactgttcaagaaaacatttccccgttgaaggattctctgccgaaaattcttccagcaaaaaaatcatc-
gagaaaaacccatggggtcaggaaacgtattctgacttgattgaagccgctatcaattcccatcccaatcaaatg-
gctactttgcaacaaatatgaatttatttccaagaacaataaatatttgcagaacgagttgatgctacttcctcg-
gctggttggaaaaattcgatccggcataatttgtcactccatgacaaattcgtgaaatgcccgaaaaaaaatgagaa-
catgaaaagctcgttgtggtcaattaacacaaagtgctacaaacgagagcgctccaacagcatggattgcaagaggtc-
gggagtcgatctaatagccagacgcaaattgctaaaggaacaacggcgtcacatttcaagcaattccaacattgtatc-
gaattctaagatatcacccaattccgacaagtacactgctgccactgacgataatagcatccacaagttgtcccta-
acaaccagtcctccttcacttatgatagcgctgagtcggccatggaaacgatttcagatttcatctataaagacaa-
caagctcaaatatagatctccacaaacaattcatgaatcaaacaaggaaatgcaacttgatagtagctataacaaatat-
gattgtcatctccaccaaacccttcatgaagacctacaacgactcgatgtcaattcaatggaacattaccattata-
aaaaagatcagaattttaattacaacaaatcacagtgtcttgatttccatcgggctgctcttgaatactccagatt-
tattgccgacatatttattgatgatgttgaa
```

>foxP
```
atgataaatcaggctctaaatttcagtttgttacaaccgaatcccagcagcaaaccattaaaaatatccccgt-
caattcaatatccagatgttcagaaattcaatgaaaataaaatgtcaaatccgaatcctttcttgattagtgagtat-
cagaaaatcaagcaattgattgataaatcaggaatgttattgccgaatgagaagcttgtaacaaaaaatcggggaga-
atgttactgggctaattgttttttttcaaagtgatattgaagagaatttcaaaaaacatctcaacgaatgtcatcatat-
gacgacgtctgcacttgcacaattagaaatttcattcgaaaaagctatgcaattatcaaaactgctattaaatgaa-
caaaatcatctttataacatgttcttgcatcttgatcaaaaactgaaaagtcttcctaatgacaattgtttattacta-
aatgatagtcaatatcttgcagctcttttgaataaatttaaacaaatgcctaatattttcgacttcttcgctgattc-
taatcgattcaatggagacctaaactcctttcatccatcacatagactagacaatcatcagtttacaataagtaatgc-
tagccatcatactcaacagccttcacctcatcaacaacaatctcccttgactgttccagcattatcgagtattcct-
caacaattaattgataatctttcatcagttttaatgagacaacaacaacaacagcagccaccacaattgataccat-
tatcatcacaaaactctaatcaacaacccaattttcaacagtcaattgcttctcctcctcaactaatcagtcaa-
caaactcattcggcgtttacaaataatattctaccactcataaacgccaataacaatattaatagtctactcgtgaa-
cagttcagatccaaatcttatcaatatcagtcccaataatcaactcaacaacattgacaacaatctcaaagatattt-
gtctcaatccagatcctcagattccaattatccaaggtctgatccggaatcgaatcacgtaaaaaatgaaaaaat-
caaacctatcaaaagggcgtcctcaaacagtcagttatcatttttcccaaaaggaatgacagaccgcgagtattaccg-
gcagtttcctgtccgacccagtgccagttatgtcaacttaatcaaaactgcaattttggaatcaccaagacgtgaatt-
gtcactgaatgaaatttatgtttggatgcaaactgaattcgcttatttccgtgacaaggaacagaagtggaaaaatgc-
catcagacataacttgtcgcttcataaatgctttcaaagaaaacatgggaaattatggactttcaatgaaaatgaata-
caacatgaaaaagtcgagaaatcgatttcagtacaatgttccaattccaggagctgagaatcagacaaatgatccgga-
cagttacgaagatgaagaagatccccaagaacagttggatagttctgaatacatgataaagaaggaaatggaattcaat-
gatgaaaatatgcctgaagacaaacatttagagattgatgagtggtccaattatcttcacaaaacctctttgat-
cagttcaagtgacgataacatgaaaccaccagatatcagtgtcgat
```

Amio acid sequences of all *Smed* FOX proteins. FKD, FHA and NLS domains are colored

>foxA1-1
MLGKNPYETAMSNVYSLPPGGSIYNMNPMSISSAGYNSQQVSTLSLNLTGIGPHSLSPMSASMSGIAAMAGGMRQGLEL-
GLGRSDSPRDKNSISNNNRPYQRSYTHAKPPYSYISLITMAIQNSPVNMCTLSEIYQFIMDHFPYYRQNQQRWQNSI-
RHSLSFNDCFVKVSRSPEKPGKGSYWTLHPQSGNMFENGCYLRRQKRFKDPHREIGRQSQRAATGPGSNVTENNHD-
NASQEASDNAESDTKPNIKQLDLSSDLLTNQGHNIKNTNPTSVSQSCSMFHRKKENCSPVEMKLNNQNQQSNQQEHPQI-
HYNPNQQFYSNQQNIFQQSSLDHYSLLASDDPLGQGMHLPPGANSVFGLYGAHNLPNDDQISVSLPSISLSGHPYDNL-
STAMAYQYEASQHNSSLLTTSNPFSIDRLMHPRLVAAAMGVSPHDTLYAGATGPSVDLEHMKYYSNYNNVPPYSSAMS-
DYYKYVQNPQPGNSDMSL

>foxA1-2
MSEFFSPVSDTSIRMNRQFHQESFPNKSPLDASLDGNSHKRISCSKPPYSYISLITMAIQNSPQKMCTLSEIYQFIMDS-
FPYYQQNQQRWQNSIRHSLSFNDCFIKASRCTEKPGKGSFWVLHPDSGNMFENGCFLRRQKRFKDPRKHKLILLGKD-
SEKQINTHRISELRIHPIDNPSNWIHAEKSTELIALEKSNNSSIFFNSTNSRQIDFDSNITMPSIYSENFLPNLSN-
DRFLIDKIMHHKMTISLDQSYDGSSQNLDYFL

>foxAt
MAIRSSPNQKQTLNGIYDWIRKHFAYFKLDTQRWQNSIRHALSFNDCFIKLARPIGESGKGCYWAIHPEAKDHFQFG-
SLLRRYKKFTQSDRNHKIWNYFPYSHQASAIVTKNIAFPYNTNDRFFYYK

>foxC2-1
MQLCKFEVNSSNSSITASGSNGPLFLPSLATVANNCHLPLNVSNSNNNQNDEPPFISWNPTNMAACYNQGLASPKPASF-
WAASNNGSLNSNNFMSCLANPGQGLIQQTYEHMSSYYGNSSLIPPYSTSYSMGNSPIPNARFLTHYGSHSNHPNMSSS-
RNPSRDLVKPPYSYIALITMAVHAHPEKKVTLSGIYQFIMEKFPYYRENKQGWQNSIRHNLSLNECFIKIPRDDKKPGK-
GSYWALHPQAYNMFENGSFLRRRRRFKAKDVLQDREDRKRKQQEDECNIFTSTSKDDDEKNEKLSVISKNGDLNKSYS-
ESVEHSDQSQSIIENELSQETKLGHDSSINNVNHTNNETFQNRTSQFETSLGGSEDNPFHFDDNKNLTSFNNFGNGFFM-
KNESSDSIPFPLTLTNTNEGLSVNNNTNSDYFQFMGRYAGQESTDVSTAFSSVTNHSGDLEDHIQNYQQLSAFQNYT-
NPHSSLTAAAAAWYAAAHDTNSVLGYPISNNYNASSSILGCNPNMPFSGSLSPHSAQVSNTSSSMTTSQSNTISSFF-
NNTDLVGSLNFGAGILSSNSPNYIPTHHPSLTNYYLI

>foxC2-2
MNVDHISVLTAANSAPQIPRHPISDHASLLYYSSAYQFPGCESSSFLWPNSSYNSNSDSWKAAYLNSDPHHLN-
SYYPSLNSQDVQQDLLFHKSIHSDFYSMSASCYNNYQNQFPIINTDSKYPNSPIDNQNNIKMENKNDNRSLVKP-
PYSYIALITMAITSQPDHRITLNGIYQFISERFPYYRENKQGWQNSIRHNLSLNECFIKVSRGDQRSGKGSYWTLHP-
NAYNMFENGSFLRRRRRFKYKDDTISTEATATNSDIIISNSDDLEKVHTSKVENPHSITNDFKGLNFSLKKNSTFGDEC-
QELFDCQSKESYSKFINIKPNVIKEIENFDKQAFDFNPNGLYNQFPYFSNISPQNEYPPSNNHLRPPLPYSMNFMQK-
PNPYLLSNYLSHPQEIS

>foxD3-1
MNKFSLYYSTLNHLSKEEESNSMALEMNHLYWSLFGQRSDAKLMKNFGDDSYLMPYHRSNSIEYLRASQRINMKYEQD-
MPQAFKDSQTIEQIKNLLRDDDVDVDELKECESDDNIHSESDLDRICSEDKKDVDNKIPMSQERCKSKSHNVKPPYSY-
IALITMAILRSPQRKLTLSGICEFIMGRFPYYKDRFPAWQNSIRHNLSLNDCFIKIPREPGNPGKGNYWTLDPRSED-
MFDNGSFLRRRKRYKRQLPSEMFNHNQSHLIIPPTSFRMTIPPNPITVQQNLVNQLLFHQNITKSIIPNPNNIIP-
PLNAIRYPRPFENNFRGHEFIGSLVPPYPHDGRIVIPSTSKTEYNEPINKRCKLNSSEDELSTFQKFSISHIIS-
DDSTDNKTKQSEESYANQFIPQTLRLWPCNPFPVAKPVIKSTWIPPNFIN

>foxD2
MKSDSKHSKPPYSYIALITMSILHSKEKRLTLNEICSFIILNFPYYRERFPSWQNSIRHNLSLNDCFVKQPREPGSSGK-
GNYWNLDPAAVSMFENGSFLRRRTRYKSSKQVRKNQSEEEFNKFHKTINDEKSCSSTDSKKYLEFSIEALLE

>foxD3-2
MVYKISNSSTKESRQLSKSEVIKNASKPPYSYIALIAMAISSSPSKKLTLSEICDFIMKKFSYYRDRFPVWQNSI-
RHNLSLNDCFIKIPRDSNNPGKGNFWSLDPQSEGMFDNGSFLRRRRRFKSRLPEIMRANIRRININNEILSKILNSTIY-
EYLNPCTKLLFMNLVPYEAFQNVFFRDNISGSSEISLKSQSSHKSLNNKFLIANIINKV

>foxF1-1
MKISNSHENLADSQADDTVDQNIKIALTKGTEIYLKPPFSYISLIAMAIQASPYQQCTLNEIYEYLSEHFKFFRGEYQG-
WKNSIRHNLSLNDCFIKIPKGFGRIGKGHYWTMNENASFLFNEKCNRRRPRGFRQKLQSIERSSNFTLLSTPHNNGVT-
FYNYCDIPVHSDAQTNHIYSNSFYLSKNEWENYSNCMFCTNSCCNQF

>foxF1-2
MNNADSNRSGVYLQSQFSHQLLGSFESPISSTQLKETAAVNGQEIKSEIHYSPEFSTDYMPINSELYQSSFLNN-
SYDHQQFVNSSSLNYFNSNKKLCGFLSSTMNNFTDLDRIDYNRINDINFEQCYQPFQNSKNEHHREGNALNHF-
NSLSRQIELVNGDSNSLSQYIHLTQPFNNQLTNNTDPQSSPNYESNEIKPNELPVINSSIVTKDCEINNTSCLK-
KCEENDFVSNDSNSPTKLNEENTTDIENCSNKEDDIKKEKSIEKETDTIEKMKSEKSLKSSTAKRKSEKPEQSY-
IALIVMAIQASPIRKCTLSEIYNYLQQNFPFFQGPYQGWKNSVRHNLSLNECFIKLPKGMGRPGKGHYWTIDPTAE-
FMFEDGGNRRRPRGFRRKVSTRSNSGSAFSKSPYTVTSVSYNNSSPMMFSPTNVRCNQLGNSASSCNSMNNLNG-
PVTDNMPPLFDMYHQSFGLFTQQETNDPMKTVMPNYPQHDSGFFTDSFYPNSSSDSSFNKQEIPTASSLFHPNFMSN-
TISQTSNSSSMSHPSISPYVSGNSRNSPLFTSLSSTLQAQNISELKISDQSSFHSPQSSTSSTVAAVAAASLYSNWFP-
NGPVEAAKYFIDHSSESERIENLYASSALPRLSNIHSENDEILERNSMYQPYMTDCLLPYLPSSQMFNIANMTDVRSS-
NINIDNADRSIWVHNKF

>foxG
MRNFPYYKNNKQGWQNSIRHNLSLNKCFIKVPRGYDDPGKGNYWMVDPACEDVYIGGITGKLRRRSSSVQRMQRLGLR-
CGPWPTSLDHCLRLGPPAISNIFPHPSDPLQSDSSNCCSPVKNTFSDFQNKNLPWNPINQNCFFPGVLSHPPLTPNLD-
FYRHLRLQNIIPKMWYSNPVTLMKEKKIF

>foxL1t
MENFARCINNQLEDNSNSCSIGSPSTSSKMMMELMNLMPPRFLGLNFGLLNPFPWNPTMMHLEKPPYSYIALIA-
MAIRNTPDKKITLSGIYKFITDNFPFYHRNKQGWQNSIRHNLSLNDCFIKIARDKNNPGKGNYWTLNENFEEMFDHGN-
FRRRRRKTKSLTNKQDCQDSESHQTRSCSSVNSDSIEMRNKTQESTPNFFIDQLLKSQSLNPSVIVKQSTVEPKSEF-
SLNLKNSFNNCSNMWPIWQRDLILSKSYNNLLLNSDNPLSNNSQSLLFKKPSH

>foxQ/D
MENSFITKSTQNNFFSINNSNVISKMIQSSMPISVMPVGPQNNPFSQNSMISLQPLLKAFHPVVGDFCKKTTQDDPK-
PNYSYIGLISMAILSTKEKKMVLSDIYQWIQDHYSYFQTRGPGWRNSIRHNLSLNDCFVKVGRSSNGKGHYWGIHPA-
NIEDFKRGDFRRRRAQRKVRRALGLTCPDEDDTPSPSPTHSPKAFDWPITAHNETPNFQISDNSIHKMNSIIHNKTEN-
DIPLISNSGYNPYFWPHNIEDEDKLKLIKSSQNSRTFDIENILNPVTKKVQLKEVFYNQFFSHLNFLNILNGIRFAR-
QLPSTFFRENQVHQFLEFKHRAIANRSVNLMLKSVTNIEQT

>foxJ1-1
MLQHCLSYGERTREELEQNVYSKYQCPDRRSSASSTTSTGSNESNRSINRQKYSFSQMSCQILANTFHKNWSEQCGNES-
GFDDSLTNLNWLHNVNPMKINEQMSHVHVGYPRDRLVSNSYQPRNYHTFNHQNQASTQQDSSAFTFPMISADQFSFDK-
PDYKNNPNLKPTFTFAVLISWAMKELGKPKITLSDIYGWISDNFAYYRYSDSSWQNSVRHNLSMNKYFQKVPRRKDEP-
GKGGFWRMNPDNIKELDHNLSKIKYNPSSHSIIPIKTEKLPIISSYRQMTSYPPTVSTNSNSNHINTFTIQEHSRP-
MISNHNAVKRPRKSPVITIPENWDALECLKETDTKQVNNIKINENNHFPQYEELLLFDGPSAYKNDLDFSSMDNSSTDS-
FPYLSQCPPNSNEISLSNFEDDSLGDILASPMLLDGEEIMDNEDFKSMVYQTEYDLFPTSKDDLQELNTLLGIY

>foxJ1-2
MNLNKLTSQALPISPISASSMSCGKYVDLRQPYASDPSVRSDIICSWDQIDDSTRIFYKTHTTGRPPFSHISLICMAIQ-
DIGQSRITSTQICEWIIINFPYYQILDNSWQNSVRNLLSVSKCFQKVPRRKDEPEKGGFW

>foxJ1-3
MAMLDLGKPKIALNEIYEWIRENFLYYRKSDSSWQNSVRHNLSLKKCFEKVPRKKDERGKGGFWRINPNFTENL-
GNNFIKYRRQFHVYATPPPLPPSQPLVSSQPQPSAQTQSNGYHLSNSVYKYAGNNNNVASKGWRGNSNANCEIVNF-
KPLANYHQHQQQQQQQSHQHLKLVNIPSVRSVNKPYKRLQTTHMKPIFKKPYALSSLNQYETIEENDRVVAQMRLYG-
SPRPKSPTLFSLCENEPLLSEISSSNSNYSAFGFEDFDSTTSSSFDTEFSCSSNFHTSSDDNYSLNIALDIDEFIPQF

>foxJ1-4
MLTQHDSTCLPKIHIKYSQPEQRLSKNYLTTSKLNSIQIYTTEVFNGKKLSVLAEKFQKNWLDKKGIDSKEMDDSLTNL-
NWLHNVPSILGPESPLNSPVHLLNQANHQLLPNLERPANNIIANNFVNNEINISKQSEFSNNFSSNPSFTHHLDSALK-
DGGLLDPLVRHDYKTNWSGKPPFSYATLICMSMRELNKSKITLSDIYGWITDNFIYYRLSDSSWQNSVRHNLSLNK-
CFEKVPRSKDERGKGGFWRINPKYADRLESNLIKYRRQFPLYNMPKITNGLPAASTLLSRTTNYSYMLCQSSVNTDNT-
WSNSMTTNNQNSFAFRDPQPILSENSLPFKSKRGWDNSGATYLPNLTNYNQGKLISSSAFSNFNNDVTSNEDYC-
QNRKHNSKTYHTQSYRNHNMTKLQSLDEFRTSETGQKVDTQMRTEITPNHSPTMYSLLNHLNSPPSHVSMDDIDSAAS-
SSFDAELTANTGSCATLHSHSDDPDEALRFGTSLGDDFMVGSDDYLNCDSIRSHDQLFDSVNSQFI

```
>fxJ1-5
MISETRLGLAEKFRQNWQTQYPCSMIEKSETRNFHQNFLKRNNRNFNQSFFGGNHNDNEDDSLTNLNWLQNLNLF-
NMSQTQPPLSPSPTHQMTMPSPQFPLSMNPMNQLNRIIPQYFESSSLYAEVAMNLEFMNDEASVHFFRMNNQLRPPFS-
DINLISMAIHSYNKMKVSFPEICQWIEETFPYYKSLSDDWKTGIKNHLTLSKCFQKVPRRRDETEKGGFWRFSNEFLNY-
LERNRIQTLSGVNFDNRINHDNANPSNDITYKFGKIINRKRTMNKFVTQDENENSAFRIPKKKWFSNEMIDFQNGKSIL-
IHNKANNNSNNMDNIHHGIRECKPRESLLLQYQDGSLCNINQMNLAINSCANESMSNIGISVLANSDLSRLQESNELF-
HQNIMCNQINPIVTASQSPCPEQVTEEEFLSGQAFTNNYITLNTNADSLPSMLDTQTNYSHNDFNISDLSITGIGLKP-
PEWWNQSFSASQISLSKFNCSFRSVTVANNGEQISPFCIGIDEDHHGLWSSNDISNLSELDVLFGLS


>foxK1-2.1
MSGDYYDDSLDSDNNLPQYARITFFGQVPYIMQKERVIIGRNSAAGSVDIDVGAVTFVSRKHLELTYSYQKLKVK-
CLGKNGIFIDNIFKSHSFIPYELPSKCTLRFPSTDVQFCVEQLVGIKSSDRGRSYGRMKNSLRYVTDNDESPEYKRM-
KIQSRSDQDSVSNDNQEGAFNSLREVICNLSDEVEEYIHNDEDDDNHKINASECDGLMNDTNEENEGVLIDNIGAY-
SLDTNGSITTLTKEYTLENMKSNDMNSTNILRGIDNQSNISNRNVAKLFLSNLRCSNEQYYTNNGQITVLDSDDL-
TIEHGSDTQKPPYSYAQLIVQAVSTARDRQLTLNGIYNYISKNYPYFKAHDKGWQNSVRHNLSLNRYFIKVPRGQDEP-
GKGSFWRVDPAYENKLIAQAFRKRRLRNNASGIVNEDGLVSQVLNVSNIDANSQVSNAKVFGINKSGQIASRTPNFGN-
VITLKRACDINTQNGGTTYVLNSSSALSSNNEPQKYIVSNRNNSNFESLSNISSNKTGTIFKCQSSKTVYLDGNNI-
KISGTSGVVSKNRGLQIRSQGNNPGTLISLGGKLFSTANIKSLQIQQQSHPRILTSNSSFDPSNSNKSSITNTINPK-
GNNFYIISPTGIKNQKTLTSSFDDNIVHNSLQSPITISNSNRRIINLQTQSPNIRKIDLSNGSMNKETDLKSPIILK-
KVKINNNFLTLKSEPKINNIRQNQDSVNGDVNMRYSIQSCDKNFNDNLSDEDIKLQDESDSQVFISDRDIDNSEMDH-
FLISHSDPSVSGLPLQEHSEPGSSPDLYKHDDDMWPEDEVHKMESLKYSMHDGIVDTECD


>foxK1-2.2
MSVDFEDETLKSDSDIIYSDYIEYARITFGNENAYFMKTEKITIGRNSADGTVDIDVGPHTFVSRKHLEMMYSYK-
KLKIKCLGKNGIFIDNYFKAHSAIPYELPYECTLRFPSTNVEVNVKQLVGRKSGKCNSMESDNDANDFIDTRTLV-
NSRKRKATQTINENFSYDIFESVNITNKEITINNFDVVPMRRPMEMKRDISDSNMKSDLDRTMTESVIHNESSTINS-
SPNYNNHNNNSNFVTHHRKPMNSAPVISSGTCTTGLNPKLEQQIPSIINSSHNTTTQTSTISTNKKFQNVCIEKSN-
TIPQIVLNARPYRSYSYKGEVIYTHNSQVFSSIYDSKDLAIEYGTDTQKPPYSYAQLIVQAVSSSRDRQLTLNGI-
YNYISKHYPYFKSHDKGWQNSVRHNLSLNRYFIKVPRGQDEPGKGSFWRVDGAYESKLIAQAFRKRRLRSNNCGSIC-
SENIPSASRVFSLSSGRAGSLCAVSDRNAKLNNIFTLRRAGDQKAPTTYVFRKMSDTQQLTENVFGSAAKLISKQSD-
VNNSSNTISPDNVFHTNHNHNNINISVNSNSKSNTSSKLGQKRIFQPPISMINKPPPSNANGNGNANSNSTIFSIGG-
KVFSTNKLKPIQLLSTKNSASFVSVKQQSPQQQHHQQTGNKYILISTSQIDNQSSQNSHNNNIHSNQMISITGSGTL-
NRHQEPDHHSDHQTRASNKPPRINPIILSDNQLEMAHHYKSQQPISPLSVSETCEFPLDEPSQLDSSSTELGPLSP-
GIKFTQSLQQEDEDDDLDLEMDQFLDSYSILDHSNDIDSPADLYKEAATEIDDYCVWQDECLGLQDVEIEIH


>foxK1-1
MDNEFIHARISGINILYLMKNNTCVIGRDVSSKVDLTITSSPCISRMHLKLIANDNRLFLKCFGKNGIFINES-
FQVYTLDEVPLPALSTIRFPSTNIELQIESRNYLLENSKKKFPLKKRLYMTNKDCDSLGIDENYLSESSVDAI-
ELLSQRNHRMHLQNSSDMIVHSPTDQGTETIKPPYSYAQLIIQAIISEESQQMTLSEIYRYISKNFPYYKMNQKG-
WQNSIRHNLSLNRYFIRIPRSHNNCGKSAFWKLDKSQEAQLIKQAFWKRRMKNSFVIVNKPANNNNDNDNNNNVNNNNN-
NNFKRNASQEHCQSSSSPLLAPIKTIASNSSMIIDQSTLSSDSIRNTEVPMYFTPFDI


>foxN2-1
MIDSNQSIYDLPYRARIMMNACIDNQKSGSNTKINHISNIENLKPMNSKSVLKTNNNYTYSSGDLQPLQWLQQDTIL-
SITPLDNEEEYVPGKDDIISITTSPVKVDVISELKSEESLYQNSTIADYYRYQAMRAAMTNGHQNGGYRQQHGYL-
SYSNSSTGVRPNNSGSQHNLCHSVNEYQKPPFSYTHIIFMAIETSPNKAMTVNDIYVWCETHFPYYTQAGVGW-
KNSLRHNLSINKSFKKISRDGRGGPGRGAYWVVEPRERNNLIDAIRRSPCSLGSFPFLNNTYPWGTGNCYSNNIF-
QSANTLNNNLASLTATRSVVNNKLNSSIDLQSSSRGPLLRIGGQLTDASNLIRSTPVIQRLNDSSNIIELSKTPQIISN-
SITLSTVPLIQFDQIKFESDMTTFNESIDSNDLCSAQQDEEDCRGKVFIIDDSTESNPEFYQTYLDILRGLIYNDIEKS-
DSIRSENTKDCTEHLKNRKQFFSKNNCKSSDDILPIEDIAPISRKIIKKHCSLKQKRKCKCVKSVSDSGGSDSSL-
SNKEFSELSDTDFNSRISMGLRNQKKEKKRLYFSKLKSLTNSRKSSYNRNRINNIVRAPHSDHIYAILQEFLPPANK-
SEEHSPINQMIDDKESLSSTMADDKSLSLDLSQTSLSKRSMHKLQSLKKSMKKFKFSDSDEEIFNTTNYENISNKLIRN-
SLSFITRKRKLNKNMNSSKIIKYNSRLLKENSEYINKKPYRKYNYDDDEERENRCYVKSEITDDTFSSPIYKSLKFED-
FNNAFQGTLENLSDNKSNSYDSQFIPFKKIRPNKDTHVKRKVGRPKKSECNSVTFRYSNDLRNNKQSNIKSKKKNKTP-
KRDLKYLKECNSNKKVATEISNKDEQVAISVLANLQKKNQQDRIKNMYNLSNYIQRNSSDSNMS
```

>foxN2-2
MKPTILKVYSNFPDIPYDIKKCEDNHLIKISNVENPGIPQFKSEMLTNSFCENAYPVEDSYIDSNKDFSVLEDFYR-
FQERRNAFIDSKKSQTQINQSNLKRFYERPSFSYTQLIYMAIENSPGKAMTVNDIYNWCELNFPYYANAIHGWKNSL-
RHNLSINKCFKKITKMGQGGRGSLWIVDPREKQHLVSTMCRANVASFNAPSFIIQTAPASLDHHRIVIGDDHSLDIPQD-
DTIVFSDDSTSSQIQSDLSLILEEKPVYSWTDTRLPCDYTVPLSPENCSDASETDYNDWNRAMISLEPIFPVKQHQ-
IKSKFPLNPIVTPPSSFEHNYSIQNLQIIIDDPCAIVRHIFHPVINSNASFCTSSDYNDEEWPRTPSPVPHVAQIEP-
PQSKLKIVKSGLKRGRPKNSTKRELELKKSQKPQKKLNHNDRINNMYNLNSFINSL

>foxN2-3
MTTNSSFISTRVANYKLSSSYPSITPLYSKDLNYNGSLSFTSYTKSWSSYQNMALKATDNLRNQDLTNNSDLQPLN-
WLQKESIIDIDPLDIEDELKDDSNTPGYDDLIDPHISHAKVEVESQENSNIHDFVRFQAMRNSMNNSNRVALRNNNS-
FTDCNQNYSHQASGFAKPNLSYTQLIFMAIETCKDRAMTVNDIYQWCEVNYPYYKHIGPSWKNSLRHNLSINKSFR-
RMPRDGQTTIGNYEMNSGDDMSEETVSQRDNKQGGPGRGAYWVVDDKERKNLLDSIYRAHQTIPLSDRNGQRIDP-
PYSLIVNPPKTGPMLRIDGRFTDASTYLQSKRSDDEVANENYNIQSDYANAENLSMQKIENTEAAEYTSMMFNLATTSS-
PETGGGDTSVKLSASYNSNCSRMEGKWDAHSMPSNDIYSHLEADSDLCQPKLKNRRKSKHEKVRNYSEMESATKDLF-
PNRKNQKSDLIKTNNGMTSVNVIVAPHFDHIYARYQQFLPPAIASDMDKGVFGQEHDINPDADFLKNGYRDDDDEDNE-
NEGEQGEIENEDEQSNLNSLEAIKQEKLETDTAMLATKLNGSRRIKLCWNGKLNSAAVAFTRQRKVRRVLNRRARRSL-
RIRKPREPIEPNETAIKKEMIEPNEQTVIEDASDGYATDSTMPMSNRKSITRTNRIVANKTHPVRILPVKRKASLSKS-
GPPEKRYRSKRKRTRPLRIDDTC

>foxNt
MCAFDDSLTEIDWLNSLCIQKKSEIKEGNVISTTISPYSLITTHHNSPNDNCNTSRIHGYDKPSYSYAYLIKLAIESSN-
SKKMKLNEIYSWISETFPYYKYMQKDNNKGWKNAIRHNLSLNKIFKRCAKDSSQLGKGSFWTIDSSLDDDVQYP-
PVKKMKLMLNSTETVRSSSIEKLDLIHSTEITQNISKEMSIPVSIVYTLLSSSSNNEYNLIESSATPMSNISTSS-
LEMMPEKTQITFATLESRVLPNNSAFTVYSNEIENHNFSKEPCNDLGYSCQISYSNQPIYRNCLFKSNVSTTELN-
PILSTVECSREILHPKVLNKHLNSDPISSKNDLFTDKLEVDEDIMSVKSANNDFVVKTPVLSQINSQPVELGGNN-
SEIKEVYEENELTKGSLRDSHLISSLPQDTIDNLNSFDNLSFYDSFASIPQPLFDSESRSFELSQSTRKYLYNLS-
GVKENLLHRSQWLDLQEEDYQNQSLSETFFETLRSEEMRRNYKSETYPDSPNRDVSPTKEYLELKSSLISYRYKS-
DAANSTSMSLSTSSTEKSSNTLRSKSDIYKKSNSFDSTSEIEHFKRSISENQTCDGDEEITDVFPWDSII

>foxO
MNNMLDDCHSPEFRARSQTWGGNDNSYRSRIIEADLSQKYTTFASDSYTSFEESTVQENISPLKDSLPKILPAK-
KSSRKNPWGQETYSDLIEAAINSHPNQMATLQQIYEFISKNNKYFAERVDATSSAGWKNSIRHNLSLHDKFVKCPK-
KNENMKSSLWSINTKCYKRERSNSMDCKRSGVDLIARRKLLKEQRRHISSNSNIVSNSKISPNSDKYTAATDDNSIH-
KLSPNNQSSFTYDSAESAMETISDFIYKDNKLKYRSPQTIHESNKEMQLDSSYNKYDCHLHQTLHEDLQRLDVNSMEHY-
HYKKDQNFNYNKSQCLDFHRAALEYSRFIADIFIDDVE

>foxP
MINQALNFSLLQPNPSSKPLKISPSIQYPDVQKFNENKMSNPNPFLISEYQKIKQLIDKSGMLLPNEKLVTKNRGE-
CYWANCFFQSDIEENFKKHLNECHHMTTSALAQLEISFEKAMQLSKLLLNEQNHLYNMFLHLDQKLKSLPNDNC-
LLLNDSQYLAALLNKFKQMPNIFDFFADSNRFNGDLNSFHPSHRLDNHQFTISNASHHTQQPSPHQQQSPLTVPALS-
SIPQQLIDNLSSVLMRQQQQQQQPPQLIPLSSQNSNQQPNFQQSIASPPQLISQQTHSAFTNNILPLINANNNINSLL-
VNSSDPNLINISPNNQLNNIDNNLKDILSQSRSSDSNYPRSDPESNHVKNEKIKPIKRASSNSQLSFFPKGMT-
DREYYRQFPVRPSASYVNLIKTAILESPRRELSLNEIYVWMQTEFAYFRDKEQKWKNAIRHNLSLHKCFQRKHGKLWTF-
NENEYNMKKSRNRFQYNVPIPGAENQTNDPDSYEDEEDPQEQLDSSEYMIKKEMEFNDENMPEDKHLEIDEWSNYLHKT-
SLISSSDDNMKPPDISVD

Graphical representation of *fox* genes expression changes during cell differentiation.
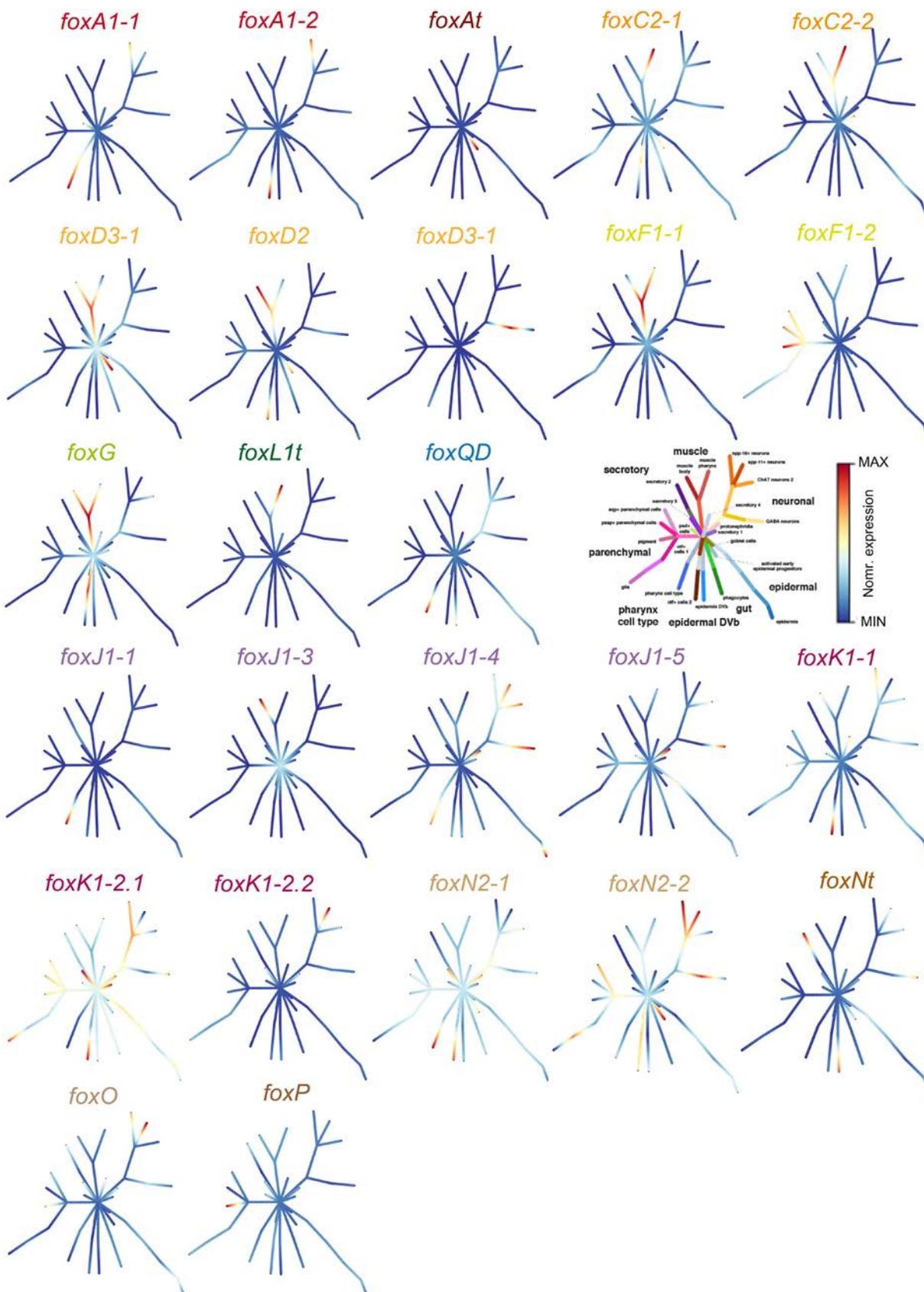
Annex IV - Table of primers used for clonig main genes for each chapter.

Chapter I

primer name	technique	5' to 3'		vector/s
		Fw	Rv	
bls3 repeat	PCR	AACCCTGACTATAGTGTGATCG	TCAGGGAAATGAGTTCGTCGT	pSPARK
bls3	dsRNA	TACTTTCTTGCTTCACATGCATGTATG	ACATCTTGGTCAGGGAAATGAG	pCRII
bls2 p1	qPCR/ISH	CATGTACGTGAATGGTATTCTGGG	ACTTCTGCATAAATTGGACCTGC	pSPARK
bls2 p2	ISH	GAAACGGTCGACAACCTAGTTC	AATATTTCAGAAGTCAAATATTGATAGATTATTTAGTGAA	pSPARK
bls3 p1	qPCR	TACTTTCTTGCTTCACATGCATGTATG	GGCCCTGAGTCCAATTAGAAAAT	pSPARK
bls3 p2	ISH/dsRNA	CATTGACAGAGAAAAGATCAACAGAT	ACATCTTGGTCAGGGAAATGAG	pSPARK/ pCRII
bls5 p1	qPCR/ISH	CGGGATTAAGAGTGAATGTGGAC	TGTATTCCAAACGGGCTTGTACC	pSPARK
pitx	qPCR	CCTTTTGGAACTCTAATGTCACC	TTGAATGACCAAGGGAAAGG	
th	qPCR	GATTGGCAACCCTGTATTG	TCACCGGATGAAGATAGAAG	
ura4	qPCR	TTCACGTTGTCGATCTAGCC	CGAATATCCTCTGCCAGTGC	
dd4277	ISH	ATCAGCAGAGAAAGCCCAAC	ACAGGCCACCATCACAAGTT	pJC53

Chapter II

primer name	technique	5' to 3'		vector/s
		Fw	**Rv**	
wnt1	dsRNA	CTGCAACCGACCTTTCCTAA	GAAGCCCTGATAAAACAAGCAA	pCRII
wnt1	ISH	AACACCAGATGGTGGCATTT	CCATTCGGGTTTTGAATCAC	pSPARK
notum	dsRNA	ATCAAAACCGGCAAGTCTCC	ACCCAACGATTTCGCAATTA	pCRII
notum	ISH	ATTGAATGATCCGCAATCCA	TACAAACGTTCGCTGCAATG	pSPARK
foxG	dsRNA	GATGGTAGATCCGGCTTGTG	AATCGCTTTGCAGTGGATCT	pCRII
foxG	ISH	GGATGGCAGAATTCTATTCGAC	GGTTTGGGGTAAGAGGAGGA	pSPARK
pitx	ISH	GTCATTCTCCATCGGCTCAT	TGACAACATTGGCTGTCGAT	pJC53
foxK1-2.1	dsRNA/ ISH	TCAATCACGAAGCGATCAAG	CTGGATCAACTCGCCAAAAT	pJC53
foxK1-2.2	dsRNA/ ISH	GAGAGCAAATTGATCGCACA	GGCTGTTTTGGGAACTTTGA	pJC53
foxK1-1	dsRNA/ ISH	TGTGTTATTGGTCGCGATGT	TTCTTCGCTTCCAAAATGCT	pCRII/ pSPARK

Chapter III

primer name	technique	5' to 3'		vector/s
		Fw	Rv	
foxA1-1	dsRNA/ISH	CCGGTAGCTCCTGCGTATAG	CACAGAGAAATCGGCAGACA	pCRII/pSPARK
foxA1-2	dsRNA/ISH	CACCATTGGACGCATCATTA	CTTTTCCCGGTTTCTCTGTG	pCRII/pSPARK
foxAt	dsRNA/ISH	TCGTCGCCTAATCAAAAACA	CCCAATAGCAACCTTTTCCA	pCRII/pSPARK
foxC2-1	dsRNA/ISH	AACAGCTTCCGGATCAAATG	TTAAATCGTCGTCGCCTTCT	pCRII/pSPARK
foxC2-2	dsRNA/ISH	AAATTCAGCGCCACAAATTC	TTCTGCCAGCCTTGCTTATT	pCRII/pSPARK
foxD3-1	dsRNA/ISH	CCAGGCGTTCAAAGATTCTC	GATTTCCCGGTTCTCTTGGT	pCRII/pSPARK
foxD3-2	dsRNA/ISH	AAAATGCTTCCAAACCTCCA	AAGTTGCCTTTACCCGGATT	pCRII/pSPARK
foxD2	dsRNA/ISH	CGACAGCAAGCATAGCAAAC	ATTCTGCCACGATGGAAATC	pCRII/pSPARK
foxF1-1	dsRNA/ISH	GGCGAGTATCAAGGATGGAA	GCATCACTATGGACCGGAAT	pJC53
foxF1-2	dsRNA/ISH	CAGATCCTCAAAGCAGTCCA	GCCCATACCTTTGGGAAGTT	pJC53
foxG	ISH	GGATGGCAGAATTCTATTCGAC	GGTTTGGGGTAAGAGGAGGA	pCRII/pSPARK
foxL1t	dsRNA/ISH	AATGCCACCTCGATTTCTTG	AACAACTTCGGGTTTGATGG	pCRII/pSPARK
foxQ/D	dsRNA/ISH	TGGTCGGGGATTTCTGTAAG	ACCCTAAAGCTCGCCGTACT	pCRII/pSPARK
foxJ1-1	dsRNA/ISH	AAATGAGTCCGGGTTTGATG	GATTGGCAGTTTCTCCGTTT	pCRII/pSPARK
foxJ1-2	dsRNA/ISH	AAATTGACCTCCCAGGCTTT	TATGGGAGAAAGGTGGTCGT	pCRII/pSPARK
foxJ1-3	dsRNA/ISH	GGTAAGCCGAAAATAGCTCTCA	CGGTTCGTTTTCACACAAT	pCRII/pSPARK
foxJ1-4	dsRNA/ISH	TGGGACCAGAATCTCCTTTG	CTGACGGAATTCTGCCAACT	pCRII/pSPARK
foxJ1-5	dsRNA/ISH	AGACTCGGTTTGGCTGAGAA	GCCGTGAAAGATCCGAATTA	pCRII/pSPARK
foxK1-2.1	dsRNA/ISH	TCAATCACGAAGCGATCAAG	CTGGATCAACTCGCCAAAAT	pJC53
foxK1-2.2	dsRNA/ISH	GAGAGCAAATTGATCGCACA	GGCTGTTTTGGGAACTTTGA	pJC53
foxK1-1	dsRNA/ISH	TGTGTTATTGGTCGCGATGT	TTCTTCGCTTCCAAAATGCT	pCRII/pSPARK
foxN2-1	dsRNA/ISH	GGCGACACGTTCTGTTGTTA	TCTGCAATTTGTGCATGGAT	pCRII/pSPARK
foxN2-2	dsRNA/ISH	CCGGGTAAAGCAATGACAGT	AGGTAAGCGCGTATCTGTCC	pCRII/pSPARK
foxN2-3	dsRNA/ISH	TCTTGCCCGAATACCCCTTT	GCCTCGTGATGGACAAACAA	pCRII/pSPARK
foxJ2/3	dsRNA/ISH	ACCGAAATCGATTGGCTAAA	ACTTGACGTGCTGATGTTCG	pCRII/pSPARK
foxO	dsRNA/ISH	CTTTTGCCAGCGATTCCTAC	AGCAATTTGCGTCTGGCTAT	pCRII/pSPARK
foxP	dsRNA/ISH	GACGTCTGCACTTGCACAAT	AGGACGCCCTTTTGATAGGT	pCRII/pSPARK

Annex V – "Posterior wnts have distinct roles in specification and patterning of the planarian posterior region"

Posterior wnts have distinct roles in specification and patterning of the planarian posterior region

Sureda-Gómez, Miquel, **Pascual-Carreras, Eudald** and Adell, Teresa

International Journal of Molecular Sciences. 2015, 16, 26543–26554

Impact factor (2015): 4.01

International Journal of
Molecular Sciences

MDPI

Article

Posterior Wnts Have Distinct Roles in Specification and Patterning of the Planarian Posterior Region

Miquel Sureda-Gómez, Eudald Pascual-Carreras and Teresa Adell *

Department of Genetics and Institute of Biomedicine, University of Barcelona, Barcelona E-08028, Catalonia, Spain; msureda6@gmail.com (M.S.-G.); eudald.pascual@gmail.com (E.P.-C.)
*   Correspondence: tadellc@ub.edu; Tel.: +34-93-402-15-00

Abstract: The wnt signaling pathway is an intercellular communication mechanism essential in cell-fate specification, tissue patterning and regional-identity specification. A βcatenin-dependent signal specifies the AP (Anteroposterior) axis of planarians, both during regeneration of new tissues and during normal homeostasis. Accordingly, four *wnts* (posterior *wnts*) are expressed in a nested manner in central and posterior regions of planarians. We have analyzed the specific role of each posterior *wnt* and the possible cooperation between them in specifying and patterning planarian central and posterior regions. We show that each posterior *wnt* exerts a distinct role during re-specification and maintenance of the central and posterior planarian regions, and that the integration of the different wnt signals (βcatenin dependent and independent) underlies the patterning of the AP axis from the central region to the tip of the tail. Based on these findings and data from the literature, we propose a model for patterning the planarian AP axis.

Keywords: patterning; identity specification; wnt signaling; planarians

1. Introduction

The wnt signaling pathway is an intercellular communication mechanism with essential roles in cell-fate specification, tissue patterning and specification of regional identity [1]. Wnts, the secreted elements of the pathway, interact with the membrane receptors Frizzleds and their co-receptors (LRP4/5, Ror, Ryk) to transduce different signals that branch mainly in three pathways: the canonical or βcatenin-dependent wnt signaling, and two non-canonical or βcatenin-independent signals, which regulate either JNK (c-Jun N-terminal kinase) or PKC (Protein kinase C) pathways [1–3]. The βcatenin-dependent pathway exerts its function by regulating the nuclear translocation of βcatenin, and is mainly involved in cell-fate specification. One of the most conserved roles of βcatenin-dependent Wnt signaling is the specification of the AP (Anteroposterior) axis, where βcatenin is required to confer posterior features in most developmental models studied [4]. βcatenin-independent pathways are mainly involved in the control of cell shape and movements [5,6].

Planarians are an ideal model for the study of cell-fate specification and patterning, since they are extremely plastic. They are bilateral animals with a complex cephalized nervous system and a three-branched gut which converges into a pharynx, which takes in food and expulses debris through a ventral mouth [7]. Planarians can regenerate any amputated part, even the head, in a few days, and they continuously remodel their tissues while they grow and shrink according to food availability and temperature. Those capabilities are due to the presence of a population of totipotent stem cells all around their body, the neoblasts, which are able to differentiate to any cell type [7–9]. Because of its astonishing regenerative abilities, planarians have been established as a unique model to understand stem cell biology and the molecular mechanisms underlying patterning and

regional identity specification. Specifically, the function of the Wnt signaling pathway has been extensively studied in planarians [10–19]. Due to its plasticity, even in the adult stage, the phenotypes generated when silencing Wnt pathway elements had no precedent in the field of developmental biology and were extremely informative. RNAi experiments demonstrate that *Smed-βcatenin1* is essential to pattern the AP axis in planarians, since its inhibition generates anteriorized phenotypes ranging from "tailless" planarians to "two-headed" planarians and, most strikingly, "radial-like hypercephalized" planarians [12,20]. Moreover, the study of several elements of the pathway confirms this function, since inhibition of *APC* and *axin*, elements of the βcatenin destruction complex, lead to posteriorization of planarians [11,18]. Interestingly, the *Smed-βcatenin1*-dependent Wnt signal is required to specify AP identities both during planarian regeneration and during homeostasis [12,17,18].

Consistent with the role of the βcatenin-dependent Wnt signal in AP axial specification, 4 *wnts* are expressed in the posterior part of planarians in a nested manner, which we name in this study posterior *wnts* (*Smed-wnt1*, *Smed-wnt11-1*, *Smed-wnt11-2* and *Smed-wnt11-5*) [15,17–19]. Since planarians such as *S. mediterranea* typically measure at least 1–2 mm in length, the field is too large to be patterned by a single morphogen. It has therefore been proposed that cooperation between posterior *wnts* could be required to pattern the AP axis [20]. Out of the four posterior *wnts*, however, only *Smed-wnt1* and *Smed-wnt11-2* have been studied functionally. During regeneration of the tail, *Smed-wnt1* inhibition leads to "tailless" or "two-headed" planarians, and *Smed-wnt11-2* inhibition leads to "tailless" planarians [14,15,19]. Although those two *wnts* seem to be regulators of βcatenin activity, because its silencing produces an anteriorized phenotype, the strong anteriorization of planarians produced after *Smed-βcatenin1* silencing has never been phenocopied by the inhibition of any *wnt*. The purpose of the present study is to analyze the specific role of each posterior *wnt* and the possible cooperation among them both during regeneration and maintenance of the AP axis. Our data demonstrates that each posterior *wnt* exerts a distinct function during posterior regeneration, and that the inhibition of all of them generates a stronger anteriorization than the inhibition of any of them alone. During homeostasis, simultaneous silencing of the four posterior *wnts* also generates a stronger phenotype than silencing any *wnt* alone, although a shift of posterior to anterior identity is never achieved. We conclude that the integration of the different Wnt signals (βcatenin dependent and independent) underlies the patterning of the AP axis from the central region to the tip of the tail.

2. Results

2.1. Individual Posterior Wnts Exert Specific Roles during Posterior Regeneration

To study the role of each posterior *wnt* during posterior regeneration, we first analyzed their expression pattern by *in situ* hybridization. In agreement with previous reports, the four posterior *wnts* are found to be expressed in a graded manner along the AP axis in intact planarians (Figure S1A) [19]. *Smed-wnt1* expression is restricted to few cells in the posterior midline; *Smed-wnt11-1* and *Smed-wnt11-2* are expressed from the mouth to the tip of the tail, and *Smed-wnt11-1* also in the mouth itself; and *Smed-wnt11-5* is expressed from the pre-pharyngeal region to the tip of the tail. Interestingly, all of them are expressed as a gradient, higher in the most posterior tip. Moreover, posterior *wnts* are also expressed in a temporal manner during posterior regeneration, being *Smed-wnt1* the first one, expressed few hours after cutting (Figure S1B) [14,19], followed by *Smed-wnt11-1* and *Smed-wnt11-2*, which are detected 2 days after cutting (Figure S1B) [19]. *Smed-wnt11-5* is expressed at all regeneration stages, since its expression is not lost after cutting the tail but just re-patterned (Figure S1B) [19]. Those expression patterns suggest that each posterior *wnt* could exert a specific role during posterior specification and patterning, and that the cooperation between them could enable a correct and complete posterior pattern.

To test the specific role of each posterior *wnt*, we analyzed the morphology and pattern of the tail regenerated by planarians in which each posterior *wnt* alone was silenced. Phenotypes

were analyzed by morphological observation and by immnohistochemistry with anti-synapsin and anti-βcatenin2 antibodies, to visualize the nervous and the digestive system, respectively (Figure 1 and Figure S2). As expected, inhibition of *Smed-wnt1* led to "tailless" and "two-headed" planarians (Figure 1A and Figure S1). Immunohistochemical analysis showed that "two-headed" planarians always differentiate a second pharynx in the opposite direction to the original one, according to the new axis generated in the posterior tip (Figure 1A(D′)). "Tailless" planarians showed a rounded closure of ventral nerve cords (VNCs) and an undefined posterior tip (Figure 1A(B′,C′)) [15]. Among "tailless" planarians two different phenotypes could be distinguished: animals in which only one pharynx was observed (sometimes in an opposite orientation) (Figure 1A(B′)) and animals in which two pharynges in opposite orientation could be observed (Figure 1A(C′)). Silencing of *Smed-wnt11-1* lead to the regeneration of shorter tails, in which the distance from the pharynx to the posterior tip was clearly shorter (Figure 1 and Figure S2). Immunohistochemical analysis showed that those animals close properly the VNCs in the posterior tip, and no signal of anteriorization can be observed (Figure 1A(E′)). Again, two different phenotypes could be distinguished when analyzing the central region, since in some animals a second pharynx appeared in parallel and very close to the pre-existing one (Figure 1A(F′)). Interestingly, two-pharynged *Smed-wnt11-1* RNAi animals never showed two mouths (Figure 1A(F″)). Silencing of *Smed-wnt11-2* always lead to the regeneration of "tailless" planarians, as had been already reported (Figure 1 and Figure S1) [15,19]. Immunohistochemical analysis demonstrated that although *Smed-wnt11-2* RNAi animals only show the normal pre-existing pharynx, a second mouth appears in half of the animals (Figure 1A(G′,G″)). A second pharynx associated to the second mouth has not been observed, although in few cases a pharynx primordium could be guessed (Figure 1A(H′)). *Smed-wnt11-5* RNAi animals apparently regenerated a perfect tail (Figure S2). Immunohistochemical analysis corroborates that VNCs close normally in the posterior tip. However, in most of the cases a second pharynx, oriented either in the same or in opposite direction with respect to the original one, can be observed (Figure 1A(I′,J′)). A second mouth always differentiates associated to the second pharynx (Figure 1A(I″)). Thus, in *Smed-wnt11-5* RNAi animals, posterior identity appears normal but the central region appears duplicated. The quantification of the different phenotypes observed after silencing each posterior *wnt* alone allows the visualization of the different degrees of anteriorization generated (Figure 1B).

We then analyzed posterior identity specification of planarians in which posterior *wnts* were silenced. Posterior *wnts* and *fz4* were used as markers [18]. Results show that after *Smed-wnt1* RNAi the rest of posterior *wnts* and *fz4* disappear or decrease significantly, demonstrating the loss of posterior identity in "tailless" and "two-headed" phenotypes (Figure 1C). In contrast, after *Smed-wnt11-1* and *Smed-wnt11-5* RNAi, all posterior markers were expressed in the same pattern and levels as in controls, in agreement with the normal posterior closure of the VNCs in the posterior tip (Figure 1C). Thus, *Smed-wnt11-1* and *Smed-wnt11-5* RNAi animals have normal posterior identity. *Smed-wnt11-2* RNAi animals displayed a significant decrease in the expression of posterior markers, according to the "tailless" phenotype observed. Expression of *Smed-wnt1* appeared not only diminished but totally disorganized (Figure 1C) [19]. Interestingly, expression of *Smed-wnt11-1* and *Smed-wnt11-2* was found to be dependent on *Smed-wnt1*, although it remains unclear whether this is a direct regulation or a consequence of the loss of posterior identity.

Taken together, these results suggest that *Smed-wnt1* and *Smed-wnt11-2* specify posterior identity, although only *Smed-wnt1* RNAi animals exhibit a shift in polarity. Moreover, *Smed-wnt11-2* exerts a role in patterning or specifying central identity, since its inhibition duplicates the mouth. *Smed-wnt11-1* and *Smed-wnt11-5* are not required to specify the identity of the posterior tip. However, they have a role in patterning or specifying the central region, since ectopic pharynges differentiate when they are silenced. *Smed-wnt11-1* would be also required to properly elongate the tail.
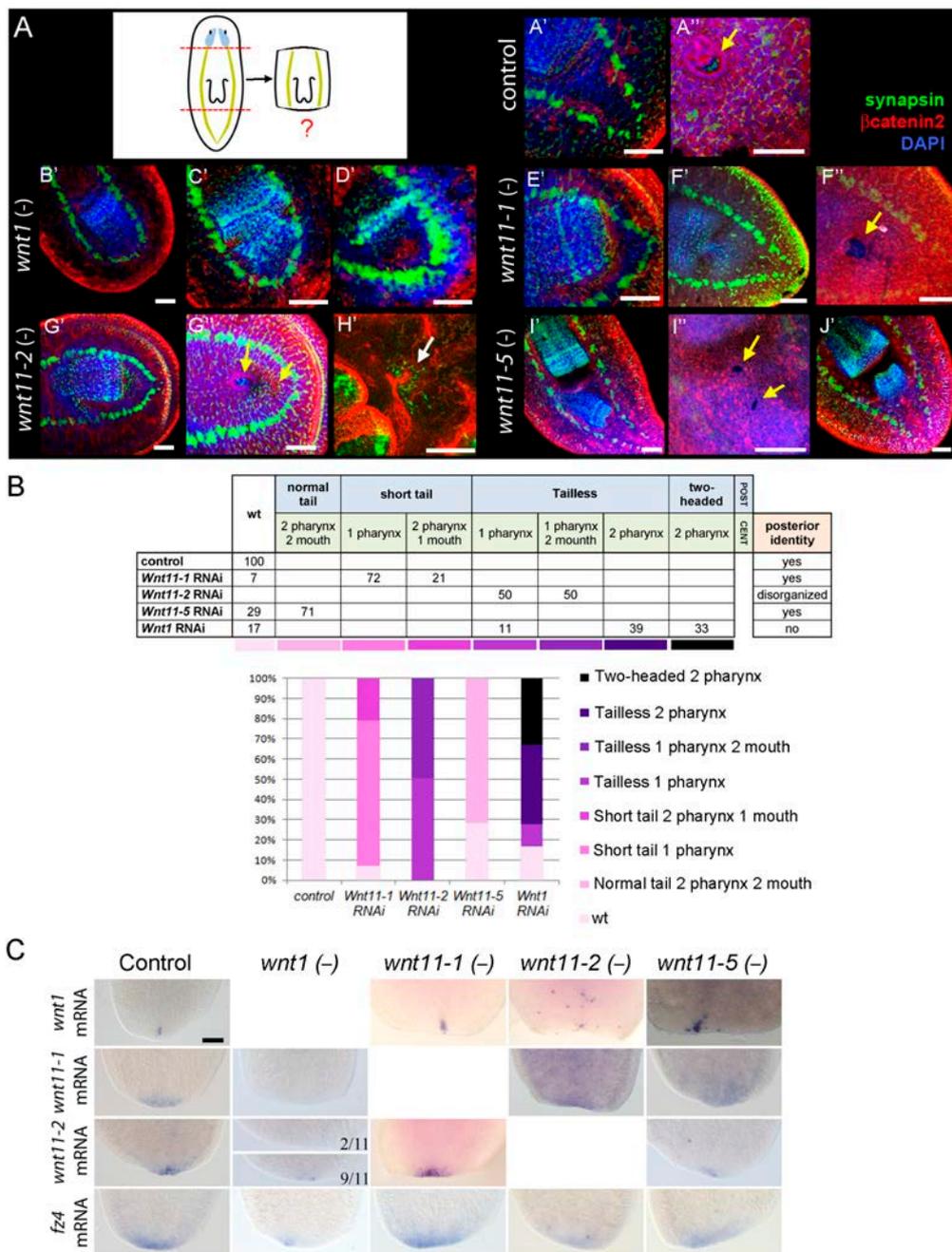
Figure 1. Each posterior *wnt* exerts a distinct function during planarian posterior regeneration. (**A**) Immunohistochemical analysis of planarian tail after silencing of *Smed-wnt1* (**B′**–**D′**), *Smed-wnt11-1* (**E**–**F″**), *Smed-wnt11-2* (**G′**–**H′**) and *Smed-wnt11-5* (**I′**–**J′**). Anti-synapsin labels the nervous system (green), anti-βcatenin1 labels adherent junctions (red), and nuclei are stained with DAPI (4′,6-diamidino-2-phenylindole)(blue). **A″**, **F″**, **G″** and **I″** show a magnification of the plane corresponding to the mouth opening of **A′**, **F′**, **G′** and **I′**, respectively (mouth openings are indicated with yellow arrows). A primordium of a second pharynx in *Smed-wnt11-2* RNAi animals is shown in **H′** (white arrow). Animals were fixed at 20 days of regeneration. All images correspond to confocal z-projections; (**B**) Quantification of the different phenotypes observed after silencing each posterior *wnt*. (The number of animals analyzed for each condition was at least *n* = 14.). wt; wild type; (**C**) *In situ* hybridization analysis of the expression of posterior markers in 3-day regenerating posterior *wnt* RNAi. (The number of animals analyzed for each condition was at least *n* = 11.). Anterior is left/up, posterior is right/down in (**A**); anterior is up, posterior is down in **C**. Scale bar: 100 μm (**A**,**C**).

2.2. Smed-wnt11-2 and Smed-wnt11-5 but not Smed-wnt11-1 Cooperate with Smed-wnt1 in Specifying Posterior Identity

To study whether posterior *wnts* play a cooperative role in posterior specification and patterning, we silenced *Smed-wnt1* (the only *wnt* that leads to shift of posterior to anterior identity upon silencing) simultaneously with *Smed-wnt11-1*, *Smed-wnt11-2* or *Smed-wnt11-5*. The resulting phenotypes were analyzed by immunohistochemistry with anti-synapsin and anti-βcatenin2 antibodies to visualize the nervous and digestive systems (Figure 2A). The phenotypes obtained after double inhibition were quantified and compared with those obtained after single inhibition of each posterior *wnt* (Figure 2A), allowing visualization of the degree of cooperation between the different posterior *wnts* in central and posterior specification. In these experiments, the penetrance of the phenotypes of single *wnt* RNAi was milder than in the experiments shown above, since half the amount of dsRNA was injected for each gene in order to maintain the total amount of dsRNA injected per animal (see Section 4.2). Quantification of the different phenotypes shows that simultaneous silencing of *Smed-wnt1* together with *Smed-wnt11-2* or *Smed-wnt11-5* increased the number of "two-headed" planarians from 20% in *Smed-wnt1* RNAi planarians to 70% in the doubles [14]. In contrast, simultaneous silencing of *Smed-wnt1* together with *Smed-wnt11-1* decreased the frequency of "two-headed" planarians from 20% to 8%. Interestingly, two new phenotypes not observed in the single inhibition experiment appeared in these experiments. Firstly, we observed "tailless" planarians with two pharynges in parallel, which is the addition of the suppression of the posterior identity after *Smed-wnt1* silencing together with the appearance of an ectopic pharynx after *Smed-wnt11-1*. In addition, "tailless" planarians were observed with two pharynges in tandem and in the same orientation, which is the addition of the suppression of the posterior identity after *Smed-wnt1* silencing and the duplication of the central identity produced by *Smed-wnt11-5* silencing (Figure 2A(A′,B′)). According to the phenotypes observed, analysis of the posterior marker *fz4* in the double RNAi planarians revealed a loss or reduction in *Smed-wnt1/Smed-wnt11-2* and *Smed-wnt1/Smed-wnt11-5* RNAi planarians (Figure 2B). *Smed-wnt1/Smed-wnt11-1* RNAi animals also displayed a mild reduction of *fz4* expression, possibly due to the inhibition of *Smed-wnt1*. Taken together, these results demonstrate that *Smed-wnt11-2* and *Smed-wnt11-5*, but not *Smed-wnt11-1*, cooperate with *Smed-wnt1* in specifying posterior identity. The contribution of *Smed-wnt11-2* in posterior specification could be predicted according to its requirement in single RNAi experiments. However, the contribution of *Smed-wnt11-5* in posterior specification should be in cooperation with *Smed-wnt1*, since its inhibition alone never induces posterior defects. The possible cooperation between *Smed-wnt11-1*, *Smed-wnt11-2* and *Smed-wnt11-5* in the specification and patterning of the central region requires further attention.
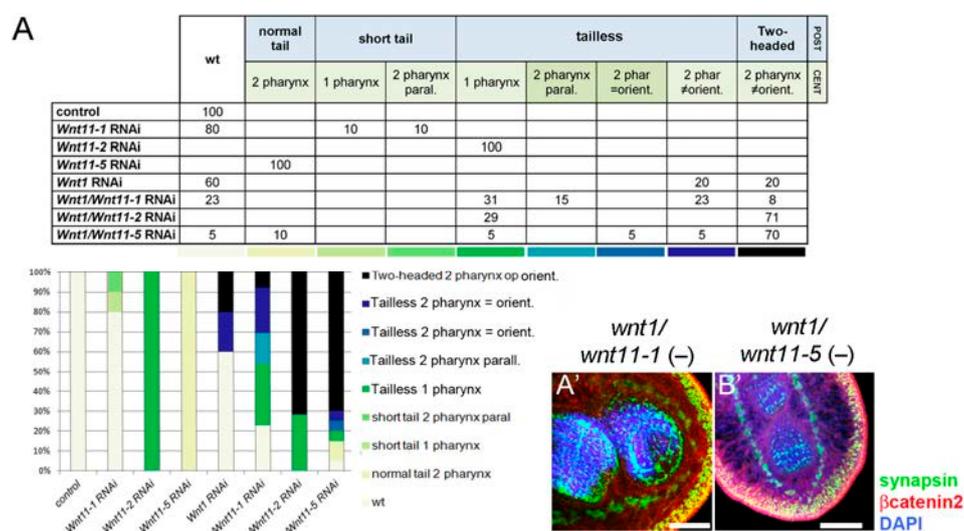

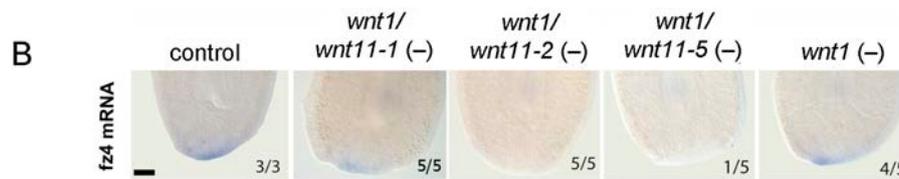|	wt	normal tail	short tail		tailless				Two-headed	
		2 pharynx	1 pharynx	2 pharynx paral.	1 pharynx	2 pharynx paral.	2 phar =orient.	2 phar ≠orient.	2 pharynx ≠orient.	POST / CENT
control	100									
Wnt11-1 RNAi	80		10	10						
Wnt11-2 RNAi					100					
Wnt11-5 RNAi		100								
Wnt1 RNAi	60							20	20	
Wnt1/Wnt11-1 RNAi	23				31	15		23	8	
Wnt1/Wnt11-2 RNAi					29				71	
Wnt1/Wnt11-5 RNAi	5	10			5		5	5	70	

Figure 2. *Cont.*

Figure 2. Cooperation of posterior *wnts* to specify posterior identity during regeneration. (**A**) Quantification of the different phenotypes observed after silencing each posterior *wnt* alone and *Smed-wnt1* in combination with the other posterior *wnts*. Two new phenotypes appeared after silencing *Smed-wnt1/Smed-wnt11-1* (**A′**) and *Smed-wnt1/Smed-wnt11-5* (**B′**), both of which show a "tailless" morphology next to the differentiation of a second pharynx alongside the original one. Animals were fixed at 20 days of regeneration. (The number of animals analyzed for single RNAi was *n* = 4–10 and for double *n* = 7–20.). **A′**,**B′** images correspond to confocal z-projections; and (**B**) *In situ* hybridization analysis of the expression of posterior markers in 3-day regenerating animals in which *Smed-wnt1* was silenced together with the other posterior *wnts*. (The number of animals analyzed for each condition was *n* = 3–5.). Anterior is left/up, posterior is right/down in **A′**,**B′**; anterior is up, posterior is down in (**B**). Scale bar: 100 μm (**A**,**B**).

2.3. Silencing of All Posterior Wnts Together Is Insufficient to Transform Posterior Identity into Anterior during Homeostasis

βcatenin-dependent Wnt signaling is also required for the maintenance of posterior identity and pattern during planarian homeostasis, since *Smed-βcatenin1* inhibition in intact planarians produces the appearance of ectopic eyes and brain in the posterior tip [12,17,18]. To analyze the possible cooperation between posterior *wnts* in the maintenance and pattern of the AP axis during homeostasis, we silenced them simultaneously and analyzed the resulting phenotypes after 6 weeks of inhibition. As a previous step, we silenced each *wnt* alone and showed that posterior eyes were not induced in any case (Figure S3). However, *Smed-wnt11-1* and *Smed-wnt11-2* RNAi planarians did show "tailless" phenotypes. RNAi of the four posterior *wnts* simultaneously produced an evident reduction of the tail, generating a strong "tailless" phenotype. In those animals, the pattern of the central region was also affected, and 3 types of central phenotypes could be distinguished: animals with two pharynges in opposite orientation, animals with a disorganized pharynx, and animals without a pharynx, due to its expulsion (Figure 3A,B). Despite the strong phenotype observed in posterior *wnt* RNAi planarians, the differentiation of ectopic anterior structures never occurred. The analysis of anterior and posterior identity markers corroborates the "tailless" phenotype, since RNAi planarians completely lost the expression of the posterior marker *fz4*, and the anterior markers *sFRP* [18] and *notum* [16] never appear in the posterior region (Figure 3C). Moreover, *sFRP* staining also revealed disorganization of the pharynx (Figure 3C).
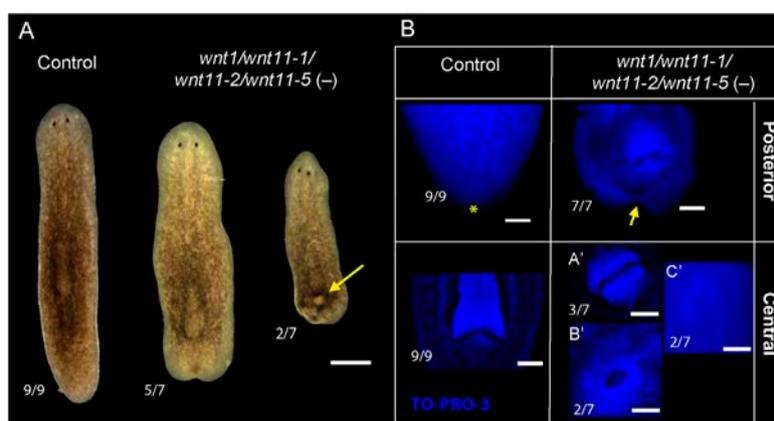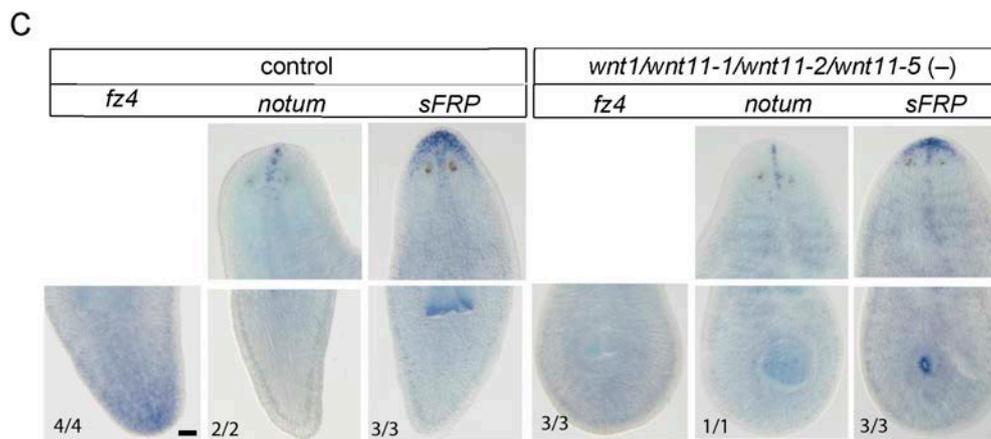

Figure 3. *Cont.*

Figure 3. Silencing of all posterior *wnts* during homeostasis generates a strong "tailless" phenotype, without neither posterior nor anterior identity (**A**) After 6 rounds of *Smed-wnt1/Smed-wnt11-1/ Smed-wnt11-2/Smed-wnt11-5* inhibition, planarians show a "tailless" phenotype in which the central region is also affected, since the pharynx cannot be maintained (yellow arrow points to a hole generated after the expulsion of the pharynx). (The number of animals analyzed for each condition was *n* = 7–9); (**B**) TO-PRO-3 staining of the nucleus shows the "tailless" shape of RNAi planarians (yellow arrow) compared to controls (yellow asterisk), and the disorganization of the central region (**A′**, two pharynges; **B′**, disorganized pharynx; **C′**, no pharynx, after expulsion). All images correspond to confocal z-projections; and (**C**) "Tailless" *Smed-wnt1/wnt11-1/wnt11-2/wnt11-5* RNAi animals do not show expression of either posterior (*Fz4*) or anterior (*sFRP*, *notum*) markers in the posterior region. Anterior markers are normally expressed in the anterior region. Anterior is up, posterior is down in **all images**. Scale bar: 500 μm (**A**), 100 μm (**B**, **A′**, **B′** and **C′**) and 100 μm (**C**).

Taken together, these results show that disruption of the central and posterior regions in intact planarians is much stronger when silencing all posterior *wnts* simultaneously than when they are silenced individually, providing evidence of cooperation in the patterning of these regions. However, in contrast to the results reported for *Smed-βcatenin1* silencing [12], a shift of posterior identity to anterior was not observed under homeostatic conditions.

3. Discussion

Depending on the dose and time of inhibition, *Smed-βcatenin1* RNAi induces a gradual anteriorization of planarians, from "tailless" to "radial-like hypercephalized" animals. Consequently, it has been proposed that the graded activation of Smed-βcatenin1 from posterior to anterior is responsible for specifying the whole AP axis in planarians [12,20]. However, the *wnts* responsible for the nuclear localization of Smed-βcatenin1 in such a broad domain remained mainly elusive. Until now, only the involvement of *Smed-wnt1* in Smed-βcatenin1 nuclearization had been suggested, since it is the only *wnt* for which inhibition induces the appearance of a posterior head during posterior regeneration [14,15]. However, the strong anteriorization observed after *Smed-βcatenin1* silencing has never been observed following inhibition of any *wnt*. In this study, we analyzed the function of the four *wnts* which are expressed in the posterior part of planarians (posterior *wnts*) and explored the possibility that they cooperate to pattern planarian AP axis (Figure 4). Our results confirm that *Smed-wnt1* is the only *wnt* for which inhibition leads to a shift in posterior polarity during regeneration, when posterior identity must be re-specified. Moreover, we reproduce the "tailless" phenotypes after inhibition of *Smed-wnt11-2* [15,19], which also must exert a role in posterior specification, according to the decreased and disorganized pattern of posterior markers. In contrast, our results demonstrate that *Smed-wnt11-1* and *Smed-wnt11-5* are not required for posterior specification, since the tip of the tail in those RNAi animals regenerates normally and posterior markers are normally expressed. Interestingly, our data point to a role for *Smed-wnt11-5* in the

specification or patterning central identity, since *Smed-wnt11-5* RNAi animals regenerate a second pharynx and mouth posteriorly to the pre-existing one. The shorter tail of *Smed-wnt11-1* RNAi planarians could indicate a role for this *wnt* in the extension of the tail. Moreover, our data suggests that *Smed-wnt11-1* could exert a direct role in the formation of the mouth, since it is expressed in this organ, and *Smed-wnt11-1* RNAi planarians never duplicate the mouth despite the presence of two pharynges.
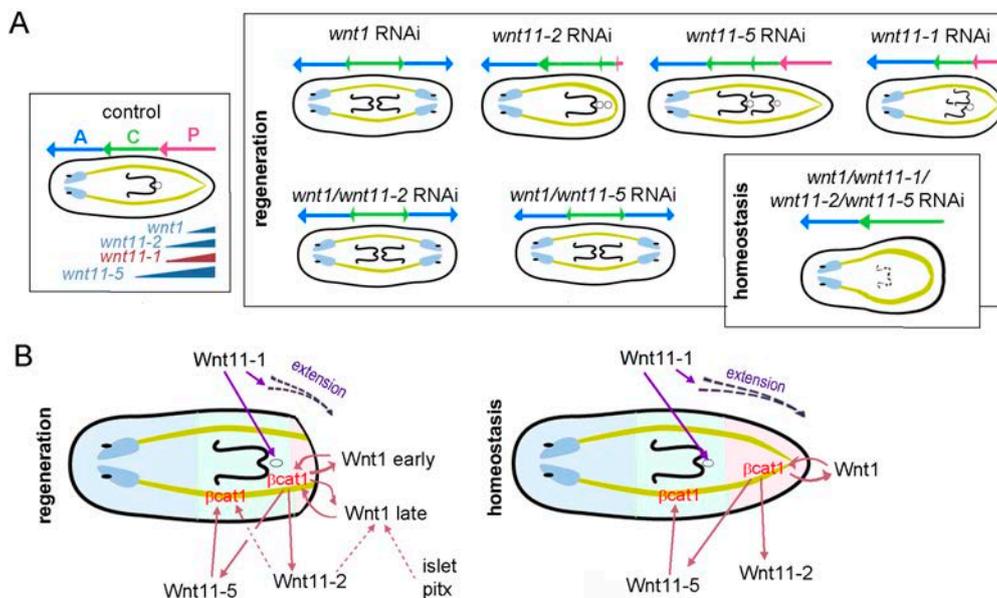

Figure 4. Summary and working model. (**A**) Scheme of the phenotypes generated after silencing the different posterior *wnts* (blue A, anterior; green C, central; pink P, posterior). The strongest phenotype is represented; (**B**) Proposed model of the function of posterior *wnts* in central and posterior specification and patterning (in the planarian: blue is anterior; green is central and pink is posterior).

Based on the results obtained in this study, we hypothesize that *Smed-wnt1*, *Smed-wnt11-2* and *Smed-wnt11-5* could act in a *βcatenin*-dependent manner, nuclearizing Smed-βcatenin1 in different domains along the AP axis (Figure 4B). Whereas *Smed-wnt1* and *Smed-wnt11-5* could be direct regulators of the βcatenin destruction complex in the posterior and central region, respectively, *Smed-wnt11-2* could be modulating Smed-βcatenin1 activity indirectly, at least in the posterior region. This possibility is supported by the observation that *Smed-wnt1* does not disappear but shows a disorganized pattern in *Smed-wnt11-2* RNAi planarians (Figure 4B). At this point, it should be noted that two different stages of *Smed-wnt1* expression occur during regeneration: an early *Smed-wnt1* expression, which occurs during wounding and is stem-cell independent, and a late *Smed-wnt1* expression, localized in the most posterior tip (the area which would correspond to the posterior organizer), that is stem-cell dependent [19]. We hypothesize that posterior identity is established by early *Smed-wnt1* expression, which triggers the sustained activation of *Smed-βcatenin1* in posterior regions through the subsequent activation of the late *Smed-wnt1* expression (Figure 4B). *Smed-wnt11-2*, for which inhibition leads to "tailless" planarians, would be required for the proper expression pattern of the late *Smed-wnt1*. The concentration of *Smed-wnt1* in the posterior tip would be essential for the establishment of the organizing region, which is responsible for growth and pattern rather than for identity specification. Additional factors, such as *Smed-pitx* or *Smed-islet*, could cooperate with *Smed-wnt11-2*, since their inhibition leads to suppression of late *Smed-wnt1* expression and regeneration of "tailless" planarians [21,22]. It has been proposed that the "tailless" phenotype could also be the result of *Smed-wnt11-2* acting in the establishment of the posterior midline [19]. In our view, the abolishment of the posterior midline goes together

with the disruption of the posterior organizer. *Smed-wnt11-1* RNAi animals regenerate a shorter tail showing a proper terminal identity. Moreover, their occasionally duplicated pharynx never locates in tandem, like in *Smed-wnt11-5* or *Smed-wnt1* RNAi planarians. For that reason, we hypothesize that *Smed-wnt11-1* would not function in a *βcatenin*-dependent manner but it would be involved in the non-canonical/βcatenin-independent Wnt signaling, a well known mechanism to regulate migration and cell movement, which are the main morphogenetic processes required for tissue extension and epithelial rearrangements [6]. The possible non-canonical function of *Smed-wnt11-1* and *Smed-wnt11-2* compared to the βcatenin-dependent function of *Smed-wnt11-5* is further supported by their evolutionary origin, since phylogenetic analysis shows that *Smed-wnt11-5* does not branch with Wnt11 but with the Wnt4 family [23]. Moreover, a *wnt4* has been suggested to act in a βcatenin-dependent manner in the platyhelminth *Schistosoma* [24]. Altogether, our results suggest that posterior *wnts* act in cooperation to provide a precise spatiotemporal control of the AP axis, from the pre-pharyngeal region to the tip of the tail.

The cooperation and integration of βcatenin-dependent and -independent Wnt signaling has been demonstrated to be essential also in the patterning of the AP neuroectoderm axis in sea urchin [25]. In cnidarians, it has been suggested that the patterning of the oral-aboral axis could be established by the cooperation between different Wnts, a "Wnt code", which would exert the function of the Hox code in bilatelians [26]. If the cooperation of posterior *wnts* is also required for maintenance of the AP pattern during homeostasis in planarians, then we expect that inhibition of the whole posterior *wnt* complement would lead to the abolishment of the identities from the pre-pharynx to the tail. Our results show that inhibition of posterior *wnts* during homeostasis one by one never induces the appearance of ectopic anterior structures but only generates mild "tailless" phenotypes. In contrast, inhibition of all posterior *wnts* together leads to a strong "tailless" phenotype, in which posterior markers disappear and also the central region is affected, since the pharynx cannot be maintained, which in fact is a feature of *Smed-βcatenin1* RNAi animals. This result confirms the hypothesis that posterior *wnts* cooperate to pattern the AP axis, including central and posterior regions. However, inhibition of the whole posterior *wnt* complement never induces the appearance of ectopic anterior structures, as occurs after *Smed-βcatenin1* silencing. One reason could be that silencing all posterior *wnts* simultaneously affects not only the bcatenin-dependent but also the bcatenin-independent Wnt signaling, which could prevent cell tip specification. Further RNAi analysis with different combinations of posterior *wnts* should be performed. A second reason could be that RNAi inhibition of the secreted elements of the pathway is less efficient than inhibition of the intracellular element, particularly considering that we are silencing four genes simultaneously. However, it must be noted that silencing of *Smed-wnt1* alone produces a strong anteriorization of planarians during regeneration but has no apparent phenotype during homeostasis. This observation could indicate that the signals which trigger posterior identity are different in the context of regeneration, when the posterior organizer must be re-specified, compared with the context of homeostasis, when the posterior organizer must be only maintained. A robust signaling network could underlie the maintenance of the posterior organizer (high levels of Smed-βcatenin1). Only the inhibition of *Smed-βcatenin1* itself or downstream elements, like *Smed-teashirt* [27], or the removal of the organizer after a posterior amputation, enables its re-specification towards a different fate.

4. Experimental Section

4.1. Planarian Culture

Planarians used in the presented experiments correspond to the clonal strain of *S. mediterranea* known as BCN-10 biotype. They were maintained as previously described [28]. Planarians used in these experiments were 4–6 mm length and were starved for 1 week before used for experiments.

4.2. RNAi Analysis

Double-stranded RNAs (dsRNAs) used in these experiments were synthesized by *in vitro* transcription (Roche) as previously described [29]. dsRNA microinjections were performed in the digestive system of planarians following the standard protocol of a 3 × 32 nL/injection of double-stranded (ds) RNA for three consecutive days before being amputated [29]. In regeneration experiments, 2 consecutive rounds of dsRNA injections were performed (1 round corresponds to 1 week, in which animals are injected on the first 3 days and amputated on the fourth). Animals were amputated transversally in 3 parts (heads, trunks and tails). In homeostasis experiments, 1 round of injection corresponds to 1 week in which dsRNA is injected on the first 3 days. Control animals were injected with dsRNA for the green fluorescent protein (GFP) sequence. In simultaneous gene-silencing experiments, the total amount of dsRNA injected for each gene and also the total amount of dsRNA injected in each animal was maintained constant by injecting the amount of GFP required.

4.3. Whole-Mount in Situ Hybridization

The RNA probes used in the present experiments were synthesized *in vitro* using Sp6 or T7 polymerase (Roche, Sant Cugat del Vallès, CAT, Spain) and DIG-modified ribonucleotides (Roche). Afterwards they were purified by ethanol precipitation and 7.5 M ammonium acetate addition. For *in situ* hybridization, animals were killed with HCl 2%, and fixed in Carnoy. An *in situ* Pro hybridization robot (Abimed/Intavis, Tübinguen, BW, Germany) was used for the *in situ* protocol, as previously described [30,31]. The temperature used for hybridizations was 56 °C, and were carried out for 16 h. A Leica MZ16F microscope (Leica Microsystems, Mannhiem, BW, Germany) was used to observe the samples. Images were captured with a ProgRes C3 camera from Jenoptik (Jena, TH, Germany).

4.4. Immunostaining

Immunostaining was carried out as described in previous studies [32]. The antibodies used in these experiments were: anti-synapsin (anti-SYNORF1,1:50, Developmental Studies Hybridoma Bank, Iowa City, IA, USA), anti-Smed-β-catenin2 (1:1000) [33] and anti-α-tubulin (AA4, 1:20, Developmental Studies Hybridoma Bank). Alexa 488-conjugated goat anti-mouse (1:400, Molecular Probes, Waltham, MA, USA) and Alexa 568-conjugated goat anti-rabbit (1:1000, Molecular Probes) were used as a secondary antibodies. Nuclei were stained with DAPI (1:5000) or TO-PRO®-3 (1:3000, Thermo Fisher Scientific, Waltham, MA, USA). A Leica TCS-SP2 (Leica Lasertechnik, Heidelberg, BW, Germany) adapted for an inverted microscope (Leitz DMIRB, Leica Lasertechnik, Heidelberg, BW, Germany) and a Leica TCS SPE (Leica Microsystems, Mannhiem, BW, Germany) were used to obtain confocal images. Representative confocal stacks for each experimental condition are shown.

References

1. Komiya, Y.; Habas, R. Wnt signal transduction pathways. *Organogenesis* **2008**, *4*, 68–75. [CrossRef] [PubMed]

2. Van Amerongen, R.; Nusse, R. Towards an integrated view of Wnt signaling in development. *Development* **2009**, *136*, 3205–3214. [CrossRef] [PubMed]

3. Mikels, A.J.; Nusse, R. Wnts as ligands: Processing, secretion and reception. *Oncogene* **2006**, *25*, 7461–7468. [CrossRef] [PubMed]

4. Petersen, C.P.; Reddien, P.W. Wnt signaling and the polarity of the primary body axis. *Cell* **2009**, *139*, 1056–1068. [CrossRef] [PubMed]

5. Mlodzik, M. Planar cell polarization: Do the same mechanisms regulate *Drosophila* tissue polarity and vertebrate gastrulation? *Trends Genet.* **2002**, *18*, 564–571. [CrossRef]

6. Wallingford, J.B.; Fraser, S.E.; Harland, R.M. Convergent extension: The molecular control of polarized cell movement during embryonic development. *Dev. Cell* **2002**, *2*, 695–706. [CrossRef]

7. Saló, E. The power of regeneration and the stem-cell kingdom: Freshwater planarians (*Platyhelminthes*). *Bioessays* **2006**, *28*, 546–559. [CrossRef] [PubMed]

8. Wagner, D.E.; Wang, I.E.; Reddien, P.W. Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. *Science* **2011**, *332*, 811–816. [CrossRef] [PubMed]

9. Baguna, J.; Saló, E.; Auladell, C. Regeneration and pattern formation in planarians. III. That neoblasts are totipotent stem cells and the cells. *Development* **1989**, *107*, 77–86.

10. Almuedo-Castillo, M.; Saló, E.; Adell, T. Dishevelled is essential for neural connectivity and planar cell polarity in planarians. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 2813–2818. [CrossRef] [PubMed]

11. Iglesias, M.; Almuedo-Castillo, M.; Aboobaker, A.A.; Saló, E. Early planarian brain regeneration is independent of blastema polarity mediated by the Wnt/β-catenin pathway. *Dev. Biol.* **2011**, *358*, 68–78. [CrossRef] [PubMed]

12. Iglesias, M.; Gomez-Skarmeta, J.L.; Saló, E.; Adell, T. Silencing of *Smed-βcatenin1* generates radial-like hypercephalized planarians. *Development* **2008**, *135*, 1215–1221. [CrossRef] [PubMed]

13. Almuedo-Castillo, M.; Sureda-Gómez, M.; Adell, T. Wnt signaling in planarians: New answers to old questions. *Int. J. Dev. Biol.* **2012**, *56*, 53–65. [CrossRef] [PubMed]

14. Petersen, C.P.; Reddien, P.W. A wound-induced Wnt expression program controls planarian regeneration polarity. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 17061–17066. [CrossRef] [PubMed]

15. Adell, T.; Salò, E.; Boutros, M.; Bartscherer, K. Smed-Evi/Wntless is required for β-catenin-dependent and -independent processes during planarian regeneration. *Development* **2009**, *136*, 905–910. [CrossRef] [PubMed]

16. Petersen, C.P.; Reddien, P.W. Polarized *notum* activation at wounds inhibits Wnt function to promote planarian head regeneration. *Science* **2011**, *332*, 852–855. [CrossRef] [PubMed]

17. Petersen, C.P.; Reddien, P.W. *Smed-βcatenin-1* is required for anteroposterior blastema polarity in planarian regeneration. *Science* **2008**, *319*, 327–330. [CrossRef] [PubMed]

18. Gurley, K.A.; Rink, J.C.; Sánchez Alvarado, A. β-Catenin defines head *versus* tail identity during planarian regeneration and homeostasis. *Science* **2008**, *319*, 323–327. [CrossRef] [PubMed]

19. Gurley, K.A.; Elliott, S.A.; Simakov, O.; Schmidt, H.A.; Holstein, T.W.; Alvarado, A.S. Expression of secreted Wnt pathway components reveals unexpected complexity of the planarian amputation response. *Dev. Biol.* **2010**, *347*, 24–39. [CrossRef] [PubMed]

20. Salo, E.; Adell, T.; Cebria, F. Gradients in planarian regeneration and homeostasis. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, 1–13.

21. März, M.; Seebeck, F.; Bartscherer, K. A Pitx transcription factor controls the establishment and maintenance of the serotonergic lineage in planarians. *Development* **2013**, *140*, 4499–4509. [CrossRef] [PubMed]

22. Hayashi, T.; Motoishi, M.; Yazawa, S.; Itomi, K.; Tanegashima, C.; Nishimura, O.; Agata, K.; Tarui, H. A LIM-homeobox gene is required for differentiation of Wnt-expressing cells at the posterior end of the planarian body. *Development* **2011**, *138*, 3679–3688. [CrossRef] [PubMed]

23. Riddiford, N.; Olson, P.D. Wnt gene loss in flatworms. *Dev. Genes Evol.* **2011**, *221*, 187–197. [CrossRef] [PubMed]

24. Li, H.F.; Wang, X.B.; Jin, Y.P.; Xia, Y.X.; Feng, X.G.; Yang, J.M.; Qi, X.Y.; Yuan, C.X.; Lin, J.J. Wnt4, the first member of the Wnt family identified in *Schistosoma japonicum*, regulates worm development by the canonical pathway. *Parasitol. Res.* **2010**, *107*, 795–805. [CrossRef] [PubMed]

25. Range, R.C.; Angerer, R.C.; Angerer, L.M. Integration of canonical and noncanonical wnt signaling pathways patterns the neuroectoderm along the anterior-posterior axis of sea urchin embryos. *PLoS Biol.* **2013**, *11*, 1001467. [CrossRef] [PubMed]

26. Guder, C.; Philipp, I.; Lengfeld, T.; Watanabe, H.; Hobmayer, B.; Holstein, T.W. The Wnt code: Cnidarians signal the way. *Oncogene* **2006**, *25*, 7450–7460. [CrossRef] [PubMed]

27. Reuter, H.; März, M.; Vogg, M.C.; Eccles, D.; Grífol-Boldú, L.; Wehner, D.; Owlarn, S.; Adell, T.; Weidinger, G.; Bartscherer, K. β-Catenin-dependent control of positional information along the ap body axis in planarians involves a teashirt family member. *Cell Rep.* **2015**, *10*, 253–265. [CrossRef] [PubMed]

28. Fernandéz-Taboada, E.; Moritz, S.; Zeuschner, D.; Stehling, M.; Schöler, H.R.; Saló, E.; Gentile, L. Smed-SmB, a member of the LSm protein superfamily, is essential for chromatoid body organization and planarian stem cell proliferation. *Development* **2010**, *137*, 1055–1065. [CrossRef] [PubMed]

29. Sánchez Alvarado, A.; Newmark, P. Double-stranded RNA specifically disrupts gene expression during planarian regeneration. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 5049–5054. [CrossRef] [PubMed]

30. Umesono, Y.; Watanabe, K.; Agata, K. A planarian orthopedia homolog is specifically expressed in the branch region of both the mature and regenerating brain. *Dev. Growth Differ.* **1997**, *39*, 723–727. [CrossRef] [PubMed]

31. Molina, M.D.; Saló, E.; Cebrià, F. The BMP pathway is essential for re-specification and maintenance of the dorsoventral axis in regenerating and intact planarians. *Dev. Biol.* **2007**, *311*, 79–94. [CrossRef] [PubMed]

32. Ross, K.G.; Omuro, K.C.; Taylor, M.R.; Munday, R.K.; Hubert, A.; King, R.S.; Zayas, R.M. Novel monoclonal antibodies to study tissue regeneration in planarians. *BMC Dev. Biol.* **2015**, *15*, 2. [CrossRef] [PubMed]

33. Chai, G.; Ma, C.; Bao, K.; Zheng, L.; Wang, X.; Sun, Z.; Salò, E.; Adell, T.; Wu, W. Complete functional segregation of planarian β-catenin-1 and -2 in mediating Wnt signaling and cell adhesion. *J. Biol. Chem.* **2010**, *285*, 24120–24130. [CrossRef] [PubMed]
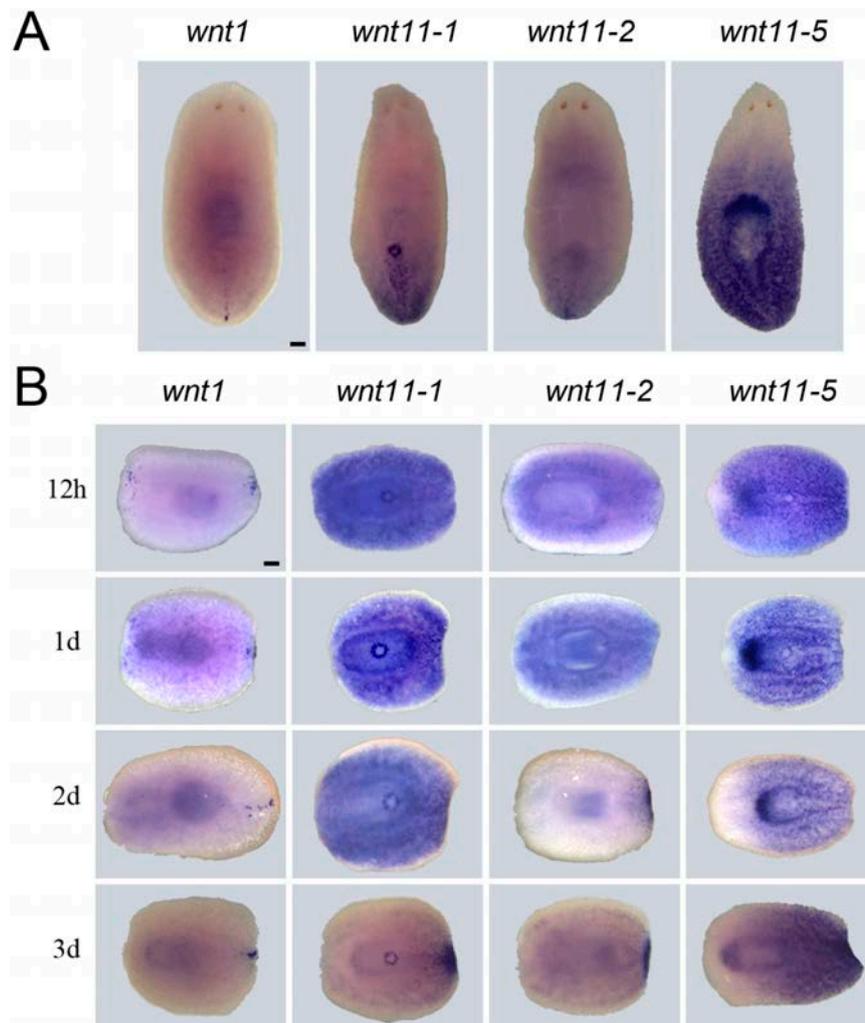
26554

Supplementary Information

Figure S1. Expression pattern of posterior *wnts* in intact and regenerating animals. (**A**) *In situ* hybridization of posterior *wnts* in intact animals. *Smed-wnt1* is expressed as a stripe of cells in the posterior dorsal midline; *Smed-wnt11-1* is expressed in the mouth and as a posterior gradient from the mouth to the tail; *Smed-wnt11-2* is expressed as a posterior gradient, concentrated in the posterior midline; and *Smed-wnt11-5* is expressed in the esophagus and as a gradient from the prepharynx to the tail; and (**B**) *In situ* hybridization of posterior *wnts* in regenerating trunks at 12 h, 1 day, 2 days and 3 days post-amputation. Anterior blastemas are shown on the left and posterior blastemas on the right. The first wnt to be expressed in the regenerating region is *Smed-wnt1*, at 12 h. At 1 day, *Smed-wnt1* decreases its expression in anterior blastemas and concentrates in the posterior. At 2 days, it recovers the expression pattern observed in intact animals (posterior dorsal midline) and disappears in anterior blastemas. *Smed-wnt11-1* expression is only maintained in the mouth during early regeneration stages, and appears in the regenerating region at day 2, at the same time that *Smed-wnt11-2*. *Smed-wnt11-5* keeps the expression observed in intact animals. At 2 days of regeneration, it starts to re-scale from anterior to posterior to recover the gradient seen in intact animals. At 3 days of regeneration, the expression of all posterior *wnts* resembles the one observed in intact planarians. (The number of animals analyzed for each condition was at least *n* = 5.) Scale bar: 100 μm (**A**,**B**).
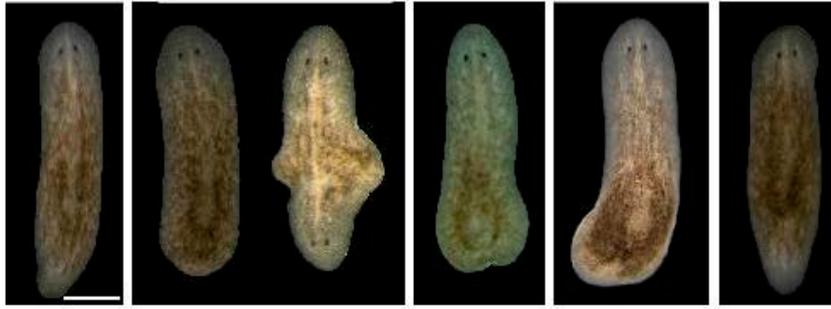
Figure S2. Phenotypes of posterior *wnts* RNAi during regeneration. Stereomicroscope views of regenerating trunk pieces showing the different phenotypes. From left to right: control animals have a wild-type appearance; *Smed-wnt1* RNAi planarians exhibit 2 different phenotypes, "tailless" (50% of animals) and "two-headed" (33% of animals); *Smed-wnt11-1* RNAi generates "short tail" planarians (93% of animals); *Smed-wnt11-2* RNAi generate "tailless" planarians (100% of animals); and *Smed-wnt11-5* RNAi planarians have a wild type tail. Images correspond to 20 days regenerating animals. (The number of animals analyzed for each condition was at least *n* = 14.) Scale bar: 500 μm.
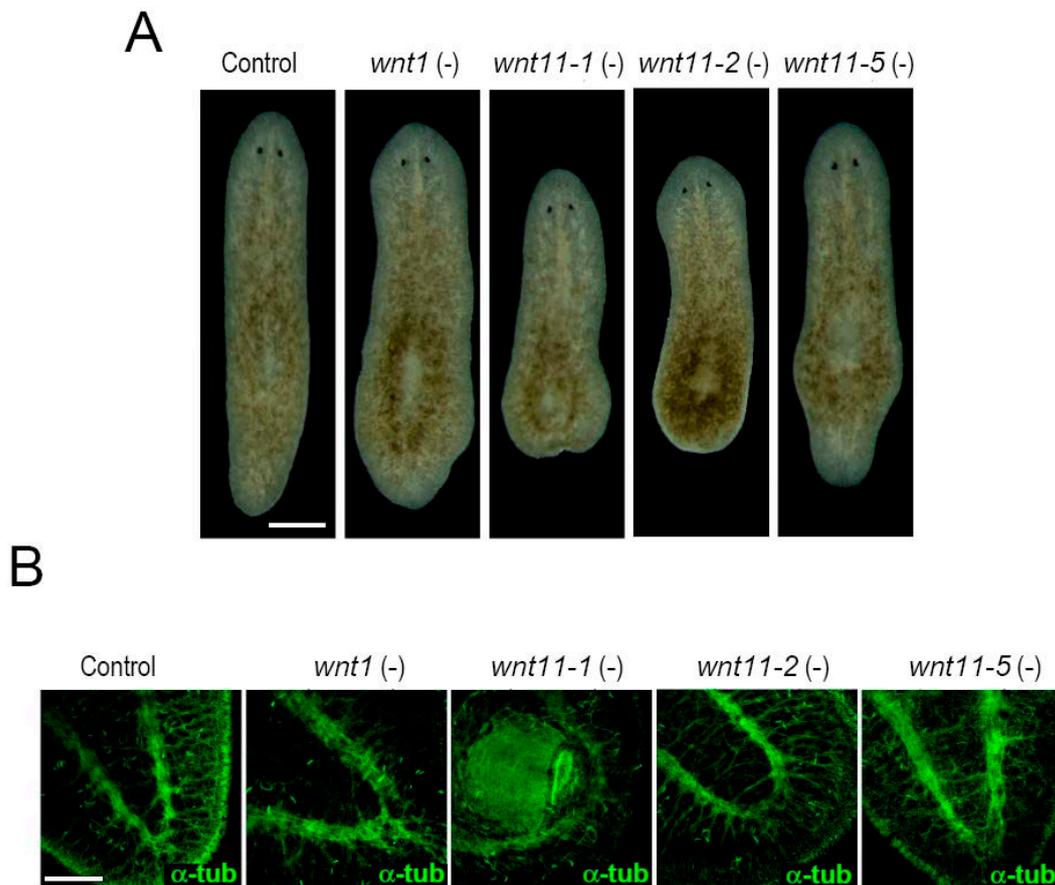

Figure S3. Phenotypes of posterior *wnt* RNAi during homeostasis. (**A**) Stereomicroscope view of the different phenotypes following RNAi in intact planarians. From left to right: controls resemble wild type; *Smed-wnt1* RNAi planarians also resemble wild type; *Smed-wnt11-1* RNAi generates "short tail" planarians; *Smed-wnt11-2* RNAi generates "tailless" planarians (100% of animals); and *Smed-wnt11-5* RNAi planarians have a tail that resembles wild type. (The number of animals analyzed for each condition was at least *n* = 10.); and (**B**) α-Tubulin immunostaining showing the morphology of the ventral nerve cords in the posterior tip of planarians after silencing posterior *wnts*. All images correspond to 20 days regenerating animals. Scale bar: 500 μm (**A**) and 100 μm (**B**).

ACKNOWLEDGEMENTS

I amb aquesta ultima pagina s'acaba aquesta tesi. Personalment també acabo un projecte que m'ha canviat per complert com a persona. Aquest viatge que va començar fa més de 5 anys, m'ha fet conèixer una mica més com sóc; i m'ha permès créixer personalment, cien-tíficament i emocionalment.

Quan vaig començar la tesi no estava gens preparat per poder assumir tot el que ha supo-sat per mi. Realment, es complicat transmetre amb paraules el que he sentit durant aquests anys. Encara que sembli ambivalent, durant la tesis et sents molt sol, però a la vegada molt acompanyat. T'enfrontes a un repte diari intel•lectual i experimental, que has de fer tu sol. Evidentment que et recolzes amb companys, directors i amics. Però al final hi ets tu. Sol.

A la vegada són els anys que més m'he sentit acompanyat. No només al departament, sinó que fora també. He tingut la sort de poder fer créixer les meves amistats. I no vull dir incrementar-ne el nombre (que també), sinó de fer que les que tenia fossin més fortes. Al laboratori he tingut la sort de compartir aquests anys amb gent meravellosa. L'Alex que em va ensenyar els primers mesos, i em va guiar en aquell nou mon. El Jose, que a més de també guiar-me en el món científic, ho va fer en el noble art de beure cervesa (ara ell ja no en beu, perquè ja se l'ha beure tota la durant la seva tesi). El Sisco i en els seus mons DC de supermans i legos. La Bea, la Sara i l'Aina amb les quals sempre m'han fet sentir d'allò més còmode al lab i a tot arreu on hem anat i ens hem trobat. La Maria, que quan era un "canijo" sempre em va donar bons consells, i ara serà tribunal de la meva tesi (flipa!!). El Kike, amb el qual sempre he trobat que m'hi he entès, dins i fora del lab. EL Miquel que ha estat com un germà a dins de lab. Moltes gracies! Haig de mencionar a l'Ari que tot i que en el màster, jo no hi vaig anar a fer amics i m'asseia al final de la classe, em va saber aguantar i sempre ens ho hem sabut passar bé allà on hem anat. Al màster també, hi vaig conèixer a la Nieves! Gràcies per la teva amistat i ser-hi tots aquests anys!

Quasi a la vegada que vaig entrar jo, va entrar també una noia portuguesa. Qui m'havia de dir que es convertiria en una de les meves millors amigues. Gràcies Nidia, per escoltar-me tant i fer que els dies al lab fossin tan bons. Gràcies.

Al Francescs (Cebrià i Mestres), els hi haig d'agrair les converses al passadís. Cada dilluns, rigorosament, hem passat revista al nostre estimat barça: GrisMan, #valverdeout, ... tots fora i plantilla nova!

A tu Susanna, que ja ens ho hem dit tot, només et em quedar donar-te les gràcies per haver construït aquesta amistat amb mi. Big sista!!

Carlos. Gracias por decir que sí a todos mis proyectos. Gracias por estos años. Lo empeza-mos juntos y lo hemos terminado junto! Joder! Estos últimos meses han sido jodidos, pero contigo al lado han sido muchísimo mejor y más divertidos.

Evidentment haig de donar les gràcies al estudiants que he tingut aquests anys. Ramon, Raquel i Pablo. Transmetre el que he après a vosaltres, m'ha fet millor científic. En general voldria agrair tots els estudiants que han passat aquests anys pel grup i pel lab: Elena, Co-ral, Sheila, Jordi, ...

Especialment, haig de donar les gracies al Dani i a la Maria, perquè estic segur que conviu-re amb mi aquesta última part de la tesi no ha estat fàcil. Gràcies per entendre-ho. I molta sort amb els vostres projectes.

Es acollonant que les amistat que vaig crear al començar la carrera, les segueixi mantenint. Gràcies Roger, Pol, Clàudia, Rosa i Marta. La vostra amistat ha estat imprescindible. I les festes encara més. Gràcies Youssef. Em compartit moltíssim fora del lab. I tot va començar també aquella classe del T2, i mira ara. Les voltes que dona la vida que vas tornar a fer la tesi allà on l'he fet jo. I hem pogut compartir UB+, Samus, Canadà,... la puta bida joder.

UB+ i Samurais. No entenc la tesi sense els partits de futbol amb vosaltres. Un dels millors remeis quan no sortia la *in situ*, era venir jugar amb vosaltres.

Durant aquests anys he après que el nostre Departament (ara secció de Genètica) té un encant especial. No se que té, que atreu gent bonica a treballar-hi. Gràcies a totes les mosques pels dinars, els congressos, les festes i birres. Paula, Elena, Haritz, Qi, José, Carlos, Martí, Giacomo. Gràcies a la Mariona, a qui torbo que és una de es persones més dolces del món, i que sempre m'ha animat amb un somriure a la boca. A la tercera planta també hi he trobat gent increïble. Isaac, Bàrbara, NoeS, Edgar, Guillem, Judit, Laura, Neus, Ester, Aldo. Moltes gracies! Gràcies també al Manel Bosch per totes les hores de confo! En el Departament, he pogut organitzar sopars i calçotades (Rafa la salsa sempre ha estat exquisida), amb les seves proves, novatades, cançons, juervezas. Joder i lo be que ens ho hem passat! A més, vem engegar un projecte genial, PhD DAY. Allà hi vaig conèixer al Jose EFE. Muchas gracias por hacerlo todo tan fácil.

Arran de preparar tot aquests merders al Departament, he trobat dues de les persones més importants de la meva vida. El Victor i la Núria. Victor, moltes gràcies per sempre treure el millor de mi, i sempre fer-me veure el meu costat bo. A més amb tu vaig descobrir la nit barcelonesa, que no té preu. Nuria (mi amiga!). Gracias por todas las meriendas, cafés y aires que hemos tomado. El uno con el otro nos hacemos bien y somos mejores como personas. Gracias por sacar-me de mis casillas de vez en cuando. Es necesario. Us estimo moltíssim a tots dos.

En aquest tesi, hi ha hagut dos grans projectes que sense dues persones no haguessin estat possibles. Gràcies Marta i Sergio (i Pep!), per fer-los possible. M'heu ensenyat moltíssim. Gràcies per escoltar i respondre els meus emails, i reanalitzar les dades cada vegada que teníem idees noves. Heu estat molt claus.

I would like to thank Dr. Phil Newmark. He gave the opportunity to be at his lab. It was one of the best experiences of my life. Umir and Rosa, thank you so much to be so kind to me. FISH works much better after my stay at Madison. Oooh Madison, what a city...no, you do not need to go there. In addition, in Madison I met Nicola!! Thanks for saving me. I am so thankful to have met you.

Gràcies a deu, fora del dia a dia de la tesi també he tingut una mica de vida personal. I aquests buits, els he omplert amb la gent que més estimo. Gràcies Paula i Carles, perquè sempre us he sentit molt a prop tot i ser tan lluny. Al Sherpa, que sempre t'ha flipat això dels cucs immortals. Els que heu estat més a prop mai us ho podré agrair prou que quasi mai em preguntéssiu per la tesi!! Un dels millors remeis per no pensar ens els mal de caps, és estar amb vosaltres. Totes les prèvies del Barça, festes de Rising, concerts de Penguis, caps de setmana, sopars i dinars... i sobretot totes les converses de whatsapp! Gràcies Navarro, Marcelo, Lolo, Adri, Rius (i Sònia), Perrier, Elena, Pau (i Marta), Edu i Roberto (i Ari). La veritat és que no ser que faria sense vosaltres.

A l'iniciï de la tesi (i abans) vaig compartir part de la vida amb l'Alba. Al final no va sortir bé. Però t'estic molt agraït pels anys que vam passar junts. Una de les raons que m'animaven atirar endavant, eres tu. Sempre em vas ajudar a créixer i a deixar els neguits del lab, al lab. Moltes gràcies per ser-hi aquells anys.

Haig de donar les gracies al Josep Marí, el meu professor de biologia de secundaria i batxillerat. Moltes gràcies per fer-me apassionar per la biologia. Si he arribat fins aquí i estic escrivint aquestes línies, és en part gràcies a tu.

Gràcies mare i pare. Gràcies pels viatges amunt i avall. Pels respirs que he pogut tenir a Corbera. Per aquests últims mesos de cuidar-me i sobretot dels tuppers. Gràcies per preguntar-me sempre per la feina i voler saber una mica més de les planaries. Gràcies avis, perquè encara m'hagi fet gran, quan ens veiem encara segueixo sent un nen, i el vostre nét. A la resta de la família: Jordi, Lluïsa, Abril, Carles i Min. Moltes gràcies per sempre preguntar com m'anava i ser-hi sempre. Us estimo molt.

A l'Arnau. Que s'ha convertit en un dels meus millors amics i amb qui ho puc compartir tot sense por. Gracias tambien a ti Rocío! Arnau, gràcies per la dibuixar la planaria (després de molt demanar), de fer la portada, i de tots els sopar i dinars mutus. Estic molt content de que ens haguem retrobat. T'estimo.

Finalment. Haig d'agrair a la Teresa i l'Emili. M'heu guiat en aquest camí, i ser que ha estat difícil. A vegades he estat una mica caòtic escrivint i intentat decidir cap a on tirar els projectes. Però sempre heu tingut el punt per marcar-me la bona direcció i deixar fer el que m'ha vingut de gust. Heu estat els meus pare i mare científics. Si tinc esperit i autonomia científica, es per parlar i discutir amb vosaltres. Me'n sento molt orgullós.

He tingut molta sort de poder fer una tesi. I d'acabar-la.

M'he adonat que la sort no és que estar ni el lloc adequat ni en el moment adequat. La sort es conèixer la gent adient i envoltar-me de persones com vosaltres. M'heu fet millor i mai oblidaré aquesta etapa.