UNIVERSITAT DE BARCELONA

Characterization of the Hox patterning genes in acoel flatworms

Eduardo Moreno González

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tesisenxarxa.net) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tesisenred.net) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tesisenxarxa.net) service has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized neither its spreading and availability from a site foreign to the TDX service. Introducing its content in a window or frame foreign to the TDX service is not authorized (framing). This rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

UNIVERSITY OF BARCELONA
SCHOOL OF BIOLOGY
DEPARTMENT OF GENETICS

CHARACTERIZATION OF THE HOX PATTERNING GENES IN ACOEL FLATWORMS

PhD THESIS

Eduardo Moreno González

THESIS DIRECTOR:
Dr. Pedro Martínez Serra
ICREA Researcher
Department of Genetics
School of Biology
University of Barcelona
Barcelona, Spain

THESIS CO-DIRECTOR:
Dr. Jaume Baguñà i Monjo
Cathedratic Professor
Department of Genetics
School of Biology
University of Barcelona
Barcelona, Spain

Memoria presentada por Eduardo Moreno González para acceder al título de Doctor por la Universidad de Barcelona, bajo la dirección de los Doctores Pedro Martínez Serra y Jaume Baguñà Monjo.

Barcelona, a 10 de Marzo del 2010

Preface

This thesis is the result of a four-year PhD project carried out from January 2006 to December 2009. Most of my work has been carried out in the "Evolutionary Developmental Biology" Research Group of the Genetics Department, University of Barcelona, Spain. The Director of this thesis has been the ICREA Researcher Pedro Martinez and the Co-Director has been the Professor of Genetics Jaume Baguñà. The thesis has been possible thanks to the financial support of the Spanish Ministry of Education, which has provided me with a F.P.U. PhD scholarship.

This thesis contains six sections. In the first place, I start with a general introduction to the topics relevant to my research project. Second, I describe the list of major aims of this work. In the third place I include a general discussion of the results obtained and emphasize its scientific relevance. Next, I list the major conclusions from this project. This section is followed by a complete list of references. At the end I include some appendices with the publications derived from this thesis plus supplementary information.

Eduardo Moreno González

Barcelona, March 2010

Acknowledgments

To my family;
for the support, affection, understanding, patience, advice and encouragement.

To my friends;
to everybody who has accompanied me along this way,
to the ones who arrived and stayed,
to the ones who arrived and went but left me with their imprint,
to the ones that will come.

To mi advisor Pedro Martínez;
for his attention, interest, energy and support,
and for having done so rewarding this stage of my life and training.

To my co-advisor Jaume Baguñà;
for his contribution to the birth and maturation on this thesis.

To the people who have collaborated with me and made this thesis possible;
to Marga, Katrien, Andi and Peter, for their help, enthusiasm and encouragement.

To my lab mates;
for having always been ready to help me,
and for having created a really enriching multicultural environment.

To Barcelona.

Table of Contents

Preface i
Acknowledgments iii
Table of Contents v
List of Figures and Tables viii
List of Abbreviations xi
Informe del Director sobre los Artículos Publicados xv

General Introduction 1
1. On the origins of Bilateria 3
1.1. Radiata and Bilateria 3
1.1.1. The Bilateria 3
1.1.2. The Radiata 4
1.2. Radial or bilateral ancestor? Hypothesis on the type of symmetry borne by the last common ancestor of cnidarians and bilaterians 6
1.2.1. The Radial-Bilateral Transition hypothesis 6
1.2.2. The Turbellarian theory 7
1.3. How did the Urbilateria look like? Hypotheses on the nature and evolution of the last common bilaterian ancestor 8
1.3.1. The Planuloid-Acoeloid hypothesis 9
1.3.2. The Archicoelomate hypothesis 11
2. The approach from the Evolutionary Developmental Biology to the origin of the bilateral animals 13
2.1. Hox and ParaHox genes 15
2.1.1. Hox and ParaHox genes in bilaterians 16
2.1.2. Hox and ParaHox genes in non-bilaterians 20

3. The phylogenetic position of the acoelomorph flatworms within the Platyhelminthes and within the Bilateria as seen from morphological/embryological data and phylogenetic analyses 23
   3.1. Morphological and developmental features that separate Acoelomorpha from the rest of the Platyhelminthes 25
   3.2. The advent of molecular phylogenies. Are the Platyhelminthes basal bilaterians and monophyletic? 27
   3.3. Phylogenetic, genomic and phylogenomic data support the polyphyly of the Platyhelminthes and the basal position of Acoelomorpha within the Bilateria 28
   3.4. A cautionary note from phylogenomics and other molecular data 30
4. The acoel species *Symsagittifera roscoffensis* and *Isodiametra pulchra*, new animal models in Evolution and Development 31
   4.1. *Symsagittifera roscoffensis* Kostenko and Mamkaev, 1990 (Graff, 1891) 32
   4.2. *Isodiametra pulchra* Hooge and Tyler, 2005 (Smith and Bush, 1991) 35

Aims of This Study 37

General Discussion 41
1. The Hox-ParaHox complement in acoel flatworms, and its genomic arrangement and structure in the acoel *Symsagittifera roscoffensis* 43
   1.1. The Hox-ParaHox complement in acoels 43
   1.2. The arrangement of the genomic Hox gene containing regions in *S. roscoffensis* 45
   1.3. The structure of the genomic regions containing Hox genes in *S. roscoffensis* 46
2. Hox gene expression patterns in juvenile specimens of *Symsagittifera roscoffensis* 54
3. Nucleotide sequence and expression pattern of the Cdx ParaHox gene in *Symsagittifera roscoffensis* 56

4. Number, types, arrangement, and expression of Hox and ParaHox gene complements in Cnidaria 59
5. Scenarios for HOX and ParaHox cluster origin and evolution within the Metazoa 67
6. Functional analysis of the *IpHoxPost* gene in the acoel *Isodiametra pulchra* and the ancestral function of the posterior Hox gene in the LCBA 72
7. Axial body pattern mechanism before the advent of Hox genes: the Wnt genes and the co-option of Hox genes 77
8. On the LCBA and the Cambrian "explosion". A speculative scenario 79
9. Perspectives 84

Conclusions 85

References 89

Appendices 113

Appendix A. Publications 115

Appendix B. Primers used for RACE 117

Appendix C. Sequences 119

Appendix D. Papers not included in this thesis 129

List of Figures and Tables

FIGURES

Cover

General anatomy of an acoel (dorsal view). Scheme extracted from http://devbio.umesci.maine.edu/styler/globalworming

General Introduction

1. Representation of a bilaterally symmetrical animal. 4
2. Representation of a radially symmetrical animal. 5
3. Hypotheses on the origin of bilateral symmetry in metazoans. 8
4. Diagrammatic representation of the Planuloid–Acoeloid hypothesis. 9
5. Two possible models for the origin of the Bilateria according to the Planuloid-Acoeloid hypothesis. 10
6. Diagrammatic representation of the Archicoelomate hypothesis. 11
7. Origin of the Bilateria from a cnidarian polyp ancestor according to the Archicoelomate hypotheses. 12
8. Hox homeodomain consensus sequence represented as a sequence logo. 16
9. Three dimensional structure of the homeodomain with the three alpha helices binding to the major groove of the DNA chain. 16
10. Diagrammatic representation of HOX clusters and their corresponding expression patterns in deuterostomes and protostomes plotted onto a consensus phylogeny of the represented taxa. 18
11. Morphology of the Acoela. 26
12. Phylogenetic position of the phylum Acoelomorpha as a sister group to the Nephrozoa (protostomes + deuterostomes) bilaterians. 30
13. Dorsal view of an adult specimen of *Symsagittifera roscoffensis*. 33
14. Dorsal view of an adult specimen of *Isodiametra pulchra*. 35

General Discussion

1. Structure of the coding regions of *SrHox5* and *SrHoxPost* genes. 48
2. Sequence alignment of anterior Hox orthologues. 50
3. Sequence alignment of central Hox orthologues. 50
4. Sequence alignment of posterior Hox orthologues. 51
5. Comparative pictures of *SrHox1* expression patterns and *SrPBX*. 52
6. Alignment of HB1 elements from mouse, human, two *Drosophila* species (*D. melanogaster* and *D. funebris*), and the *S. roscoffensis* gene *SrHoxPost*. 53
7. Diagrammatic dorsal view of the general morphology of one juvenile specimen of *S. roscoffensis* showing the nested domains of Hox gene expression. 55
8. Cdx homeobox alignment and the corresponding phylogenetic tree built with the software CLC Sequence Viewer 6.3 by UPGMA with 100 replicates for the bootstrap analysis. 57
9. *SrCdx* expression patterns in juvenile and adult specimens of *S. roscoffensis*. 58
10. Three possible scenarios for the evolution of the Hox gene axial patterning system and the bilateral symmetry in metazoans. 66
11. Origin of the ProtoHox gene within the metazoans. 67
12. Multigene (or Segmental) Duplication and Tandem Duplication scenarios to explain the origin and evolution of HOX and ParaHox gene clusters at the C-BLCA. 68
13. Evolution of HOX and ParaHox clusters in metazoans from a single ProtoHox cluster after the divergence of Porifera. 71
14. Main features of the acoeloid-like LCBA. 82

TABLES

General Introduction

1. Summary of the Hox and ParaHox related genes recovered from cnidarian species analyzed so far. 21

General Discussion

1. Hox and ParaHox genes orthologies between cnidarians and bilaterians according to different authors: 1. Kamm *et al*., 2006; 2. Chourrout *et al*., 2006; 3. Ryan *et al*., 2007; 4. Chiori *et al*., 2009; 5. Quiquand *et al*., 2009. 61

List of Abbreviations

General abbreviations

AP: antero-posterior (axis)

ASW: artificial sea water

b: bursa

BAC: bacterial artificial chromosome

bp: base pair(s)

BrdU: bromo-deoxy-uridine

bs: bursa sphincter

BSA: bovine serum albumin

cb: commissural brain

C-BLCA: cnidarian-bilaterian last common ancestor

cDNA: complementary Deoxyribonucleic Acid

cp: central parenchyma or gut

DAPI: 4',6-diamidino-2-phenylindole

dsRNA: double-stranded Ribonucleid Acid

DV: dorso-ventral (axis)

ec: egg cytoplasm

EDTA: ethylenediaminetetraacetic acid

en: egg nucleus

ep: epidermis

EST: expressed sequence tag

FISH: fluorescence *in situ* hybridization

FITC: fluorescein isothiocyanate

HB1: homeodomain binding sites

HP: hexapeptide (motif)

ISH: *in situ* hybridization

kb: kilo base pairs

LBA: long branch attraction

LCBA: last common bilaterian ancestor (referring to the last common ancestor of Acoelomorpha + Protostomata-Deuterostomata)

m: mouth

min: minutes

mo: mouth opening

Myr: millions of years

n: egg nucleolus

nc: nerve cords

OA: oral-aboral (axis)

ORF: open reading frame

PBS: phosphate buffered saline with bovine serum albumin

PCR: polymerase chain reaction

P-DLCA: protostome-deuterostome last common ancestor

PG: paralogues group

pp: peripheral parenchyma

RACE: rapid amplification of cDNA ends

RBT: Radial-Bilateral transition

RNAi: interference ribonucleid acid

RT: room temperature

SSC: sodium chloride sodium citric acid

SDS: sodium dodecyl sulfate

st: statocyst

sv: seminal vesicle

WISH: whole mount *in situ* hybridization

List of species abbreviations for alignments

Bl	*Branchiostoma lanceolatum* (cephalochordate)
Ci	*Ciona intestinales* (urochordate)
Ch	*Clytia hemisphaerica* (cnidarian, hydrozoan)
Cl	*Convolutriloba longifissura* (acoel flatworm)
Cr	*Convolutriloba retrogemma* (acoel flatworm)
Cq	*Culex quinquefasciatus* (arthropod, insect)
Df	*Drosophila funebris* (arthropod, insect)
Dm	*Drosophila melanogaster* (arthropod, insect)
Ed	*Eleutheria dichotoma* (cnidarian, hydrozoan)
Hs	*Homo sapiens* (vertebrate, mammal)
Ip	*Isodiametra pulchra* (acoel flatworm)
Mm	*Mus musculus* (vertebrate, mammal)
Nv	*Neanthes virens* (annelid, polychaete)
Pv	*Patella vulgata* (mollusc, gastropod)
Sk	*Saccoglossus kowalevskii* (hemichordate)
Sp	*Strongylocentrotus purpuratus* (echinoderm)
Sr	*Symsagittifera roscoffensis* (acoel flatworm)

1

GENERAL INTRODUCTION

1. ON THE ORIGINS OF BILATERIA

1. 1. RADIATA AND BILATERIA

The geometry of the body has been one of the first and main characters used for the classification of the organisms. Traditionally, two main geometrical criteria have been used: the symmetry of the body plan and the polarity of the body axes. In Biology, symmetry is defined as the regular arrangement of body structures relative to the axis of the body. Since the beginning of the 19th Century, symmetry was considered one of the most fundamental properties of organism form, and a prime criterion to arrange the diversity of living beings into hierarchies (Haeckel, 1866). On the other hand, polarity along an axis occurs when one end (or pole) of the axis is different to the opposite one. This concept is currently applied, for instance, to body and limb axes.

Within the metazoans, two major groups have classically been established based on their primary symmetry, the polarity of their body axes, and the number of embryonic layers. These groups are the Radiata and the Bilateria.

1. 1. 1. The Bilateria

The Bilateria (Hatschek, 1888) is a monophyletic clade within the Kingdom Metazoa, which includes all the groups with primary bilateral symmetry, that is, symmetrical about a plane running from their frontal end to their caudal end (for instance, head to tail), and with identical or nearly identical right and left halves. Within the Bilateria, though, some organisms bear radial symmetry, this being the result of secondary acquisitions within specific lineages and specific live cycle stages. That is, for instance, the case of the echinoderms where adults are characterized by pentaradial or other sorts of body symmetry.

A bilateral animal bears a single plane of symmetry, named sagittal plane, which results from the intersection of two orthogonal axes of polarity, the so-called antero-

posterior (AP) and the dorso-ventral (DV) axes (Figure 1). The AP axis generally corresponds to the direction of the main body length and to the direction of locomotion (with the sensory organs and brain located anteriorly), while the DV axis is usually oriented with respect to gravity, the ventral side facing the substrate (Manuel, 2009). Moreover, all Bilateria are triploblastic animals; that is, their bodies develop from three embryonic layers: endoderm, mesoderm and ectoderm.

Most metazoans (around the 99% of species) are bilaterians. This stresses the great evolutionary success of bilateral symmetry since it confers clear adaptive advantages to bilaterian animals over other organisms having other kinds of body symmetry, for instance facilitating directed locomotion or to improve the efficiency of internal circulation (Finnerty, 2005).

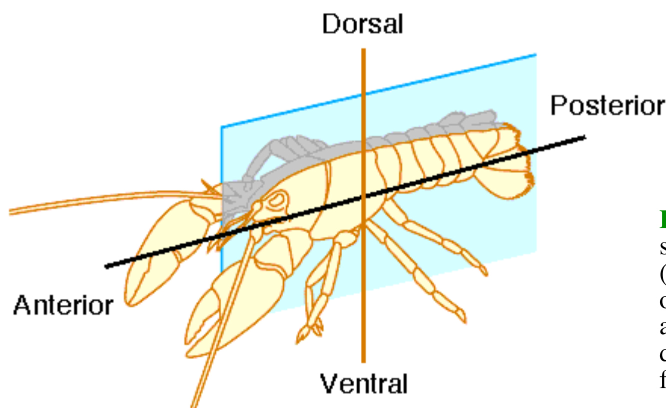


Figure 1. Representation of a bilaterally symmetrical animal. The symmetry plane (light blue) is determined by the two orthogonal polarity axes of the body: the antero-posterior axis (black line) and the dorso-ventral axis (orange line). Modified from http://io.uwinnipeg.ca/~simmons/

1. 1. 2. The Radiata

In the early 19th century, the famous zoologist George Cuvier proposed a superphyletic designation, the Radiata, to encompass the so-called radially symmetric animals (jellyfish, polyps, starfish, sea urchins, and some Protozoa) (Cuvier, 1817). Nowadays, however, the term Radiata describes a clade within the Kingdom Metazoa made up of only two phyla, Ctenophora (comb jellies) and Cnidaria (sea anemones, corals, sea pens, jellyfishes, box jellies and hydrozoan polyps) (Martindale *et al*., 2002).

The Radiata are radially symmetrical animals. That is, they are characterized by having a single axis of symmetry and a discrete number of symmetry planes (Figure 2).

The axis of symmetry corresponds to the polarity axis of the body, extending from the centre of the oral surface, which contains the mouth, to the centre of the opposite, or aboral, end. This axis is called oral-aboral axis (OA). The number of symmetry planes varies from two (biradially symmetrical bodies as Ctenophores), four (tetraradially symmetrical bodies as most cnidarian medusae) to several.

Besides, the Radiata are diploblastic animals, since they only develop from two embryonic layers, endoderm and ectoderm, lacking the third embryonic layer, the mesoderm. However, whether some radial animals (e.g. anthozoan and hydrozoan cnidarians) do actually have a true mesoderm or not is still a matter of dispute (Seipel and Schmid, 2006).

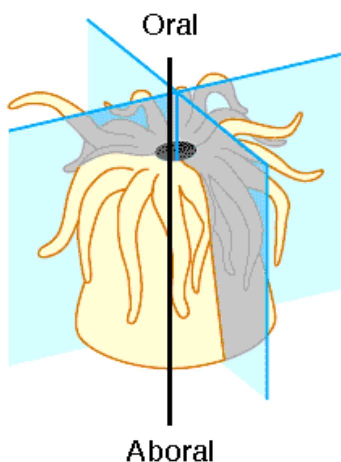


Figure 2. Representation of a radially symmetrical animal. It bears several planes of symmetry (light blue) laid out around a longitudinal axis with oral and aboral ends (black line). Modified from http://io.uwinnipeg.ca/~simmons/

A mere 0.5-1% of the extant metazoan species belong to the Radiata clade, all of them living in aquatic environments. This is likely so because radial symmetry is particularly suited to sessile and free-floating life styles, which allows them to face their environment from all sides. Even so, two structures are often bilaterally symmetrical in polyps of the anthozoan order Octocorallia. These are: 1) the pharynx (due to its oblong and elongated form, and to the presence of a single ciliated groove or siphonoglyph), and 2) the spatial distribution of the eight retractor muscles of the mesenteries. These structures, however, are evolutionary acquisitions, or synapomorphies, of the Anthozoa class with no counterparts in bilaterian body plans (Marques and Collins, 2004).

Quite unexpectedly, recent molecular studies have shown the presence and expression in cnidarians of many of the genes involved in the dorsoventral patterning of bilaterians, suggesting the presence of a second, or directive, axis (specifically in anthozoans), perpendicular to its OA axis (Finnerty *et al*., 2004). In consequence Cnidaria are considered, by some, as true bilaterians, having acquired secondarily its radial symmetry. Under this view, the clade Bilateria should include Cnidaria plus the "classical" Bilateria. The authors defending the presence of mesoderm in Cnidaria suggest that the name Bilateria should be changed to Triploblastica, which is referring to the presence of mesoderm (Baguñà *et al*., 2008).

1. 2. RADIAL OR BILATERAL ANCESTOR? HYPOTHESIS ON THE TYPE OF SYMMETRY BORNE BY THE LAST COMMON ANCESTOR OF CNIDARIANS AND BILATERIANS

1. 2. 1. The Radial-Bilateral Transition hypothesis

The traditional view of animal evolution postulates a gradual increase in complexity from the earliest "lower" metazoans to the present-day "higher" metazoans. Under this vision, the whole Bilateria clade would derive from a single common ancestor bearing radial symmetry. This process is called the Radial-Bilateral Transition (RBT). The RBT hypothesis assumes deep morphologic changes in the body of the radial ancestor. Thus, animals with a unique axis (OA) and radial symmetry had to be transformed into animals with two orthogonal axes (AP and DV) and bilateral symmetry. Moreover, other innovations were also associated with the appearance of bilateral symmetry, altogether considered as synapomorphies of the Bilateria. First of all, a new embryonic layer, the mesoderm, which turned the diploblastic ancestor into a triploblastic organism. Second, the loosely arranged nerve cells of diploblasts aggregated into condensed axial nerve cords with an anterior concentration of neural cells ('early brain ganglia') that favoured the establishment of directional forward movements. Finally, a through gut with distinct mouth, anus, and intermediate regions evolving from a sac-like gut (Holland, 1998).

Because most recent phylogenies place the Cnidaria as the sister group of the Bilateria, the RBT hypothesis assumes the appearance and evolution of radial symmetry sometime before the divergence of cnidarians and bilaterians. Therefore, the last common ancestor of cnidarians and bilaterians (C-BLCA) is modelled as a radially symmetrical organism. Later on, bilateral symmetry independently evolved, as suggested by Hyman (Hyman, 1940) and Beklemishev (Beklemishev, 1969), in bilaterians as well as in some structures within the radial Cnidaria.

1. 2. 2. The Turbellarian theory

Developed by Hadzi (Hadzi, 1944) and advocated by both Hand (Hand, 1959) and Willmer (Willmer, 1990), the Turbellarian theory proposes that radial symmetry is a derived feature within the Cnidaria, and suggests that bilateral symmetry evolved prior to the cnidarian-bilaterian last common ancestor. For the supporters of this theory, the siphonoglyph and the arrangement of the retractor muscles that attach to the mesenteries in many species within the Anthozoa are clear signs of ancestral bilateral symmetry in Cnidarians. Indeed, as has recently been proposed, the radial symmetry in cnidarian polyps could be the result of a simplification process from a bilateral symmetry free-living ancestor and due to the adaptation to sedentary life (Kaufman, 2008).

According to the Turbellarian theory, the C-BLCA would have been a bilaterally symmetrical animal, from which the radial symmetry would have evolved subsequently within the Cnidaria (Figure 3) (Finnerty, 2003). The Planulozoa' hypothesis (Wallberg *et al*., 2004) represents a new version of this theory, and suggests that both cnidarians and bilaterians evolved from a planulomorph ancestor that showed traits of bilateral symmetry.

Besides these two classical hypotheses, a new hypothesis has recently been proposed that leaves out the traditional vision of a gradual increase in complexity during evolution suggesting instead that asymmetry, radial symmetry and bilateral symmetry are derived features from an ancestral cylindrical symmetrical state, determined by a single symmetry axis and an infinite number of planes of symmetry (Manuel, 2009).

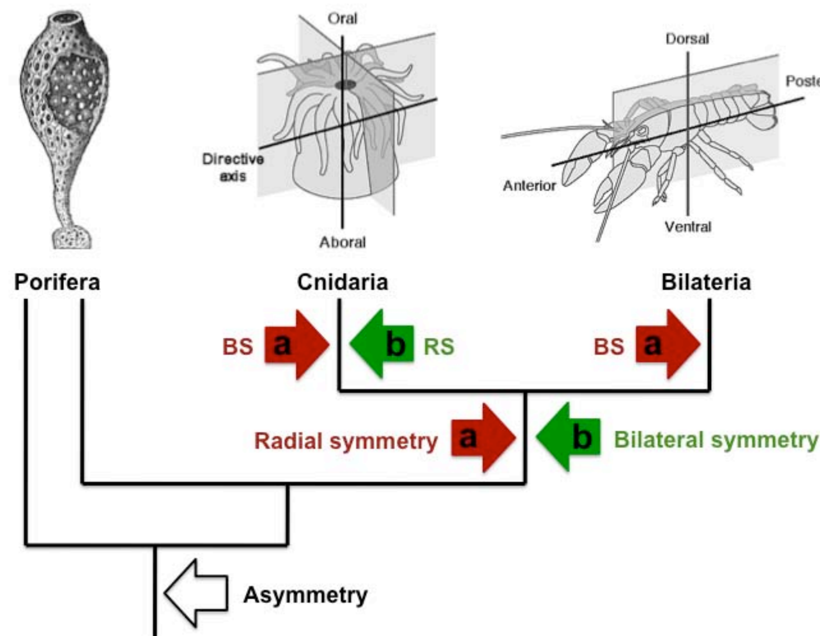


Figure 3. Hypotheses on the origin of bilateral symmetry in metazoans. Hypothesis a (red) corresponds to the Radial-Bilateral Transition hypothesis: absence of symmetry in the metazoan ancestor was followed by the appearance of radial symmetry sometime before the divergence of cnidarians and bilaterians. Hence, bilateral symmetry (BS) evolved independently in bilaterians and cnidarians. Hypothesis b (green) represents the Turbellarian theory: the C-BLCA was a bilaterally symmetrical animal; hence, radial symmetry (RS) evolved subsequently within the Cnidaria. Modified from Finnerty, 2003.

1. 3. HOW DID THE URBILATERIA LOOK LIKE? HYPOTHESES ON THE NATURE AND EVOLUTION OF THE LAST COMMON BILATERIAN ANCESTOR

The nature of the last common bilaterian ancestor (LCBA), commonly known as the Urbilateria, has been under discussion for decades. Obviously the chosen models depend very much on the assumed phylogenetic relationship of the metazoan groups. Firstly, it is important to highlight that if the Turbellarian/Planulozoa hypothesis is correct the Urbilateria does not really correspond to what is considered today as LCBA but to the C-BLCA. On the other hand, if the RBT hypothesis is right, all bilaterians came from radial ancestors, and therefore the Urbilateria corresponds to the current LCBA. Under the last scenario, two main hypotheses on the origin and morphological features of the Urbilateria are currently being contemplated: the Planuloid-Acoeloid hypothesis (Figure 4) and the Archicoelomate hypothesis (Figure 6). These hypotheses lead to radically opposed views about the main morphological and embryological features of the Urbilateria and the ulterior evolution of bilaterian characters and groups.

1. 3. 1. The Planuloid-Acoeloid hypothesis

The Planuloid-Acoeloid hypothesis (Figure 4) was proposed by von Graff in 1882 (von Graff, 1882) and later supported by Hyman (Hyman, 1951) and Salvini-Plawen (Salvini-Plawen, 1978). It features a gradual, step-by-step evolution of the Bilateria starting with a diploblastic planuloid or planula-like organism, with a complexity similar to the planula larvae of a modern cnidarian. By neoteny or progenesis, these organisms became sexually mature at the larval stage and lost the following life cycle stages of sessile polyp and/or medusa. However, the contrary scenario is also possible: cnidarians could have been also derived from the planuloid ancestor. In brief, planuloid organisms presumably swimming in the water column sunk to the bottom and shifted to crawl over the substrate. This move led to the development of dorso-ventrally flattened body shapes and ultimately to the appearance of bilateral symmetry and subsequently to the first triploblastic bilateral organisms: the acoeloid. Under such scenario, the LCBA would have looked like a simple unsegmented and acoelomate "worm-like" organism, (to some extent) to the present-day acoel flatworms. From this ancestor, pseudocoelomate protostomates and, later on, protostomian and deuterostomian coelomates derived (Baguñà and Riutort, 2004).

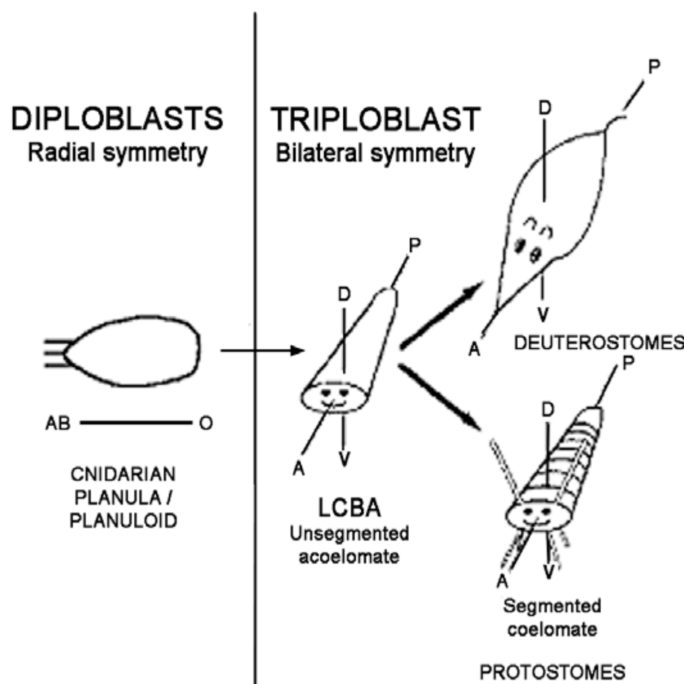


Figure 4. Diagrammatic representation of the Planuloid–Acoeloid hypothesis. From a planuloid diploblastic organism, similar to the planula larva of modern cnidarians but able to reproduce, originated the bilaterian ancestor: a simple unsegmented and acoelomate acoeloid similar to present-day acoel flatworms. From this ancestor, pseudocoelomate and coelomate protostomes and deuterostomes derived. The OA axis of cnidarians and the AP axis of bilaterians would be homologous, and the DV axis would be an innovation of the Bilateria. A, anterior; AB, aboral; D, dorsal: O, oral; P, posterior; V, ventral. Modified from Baguñà and Riutort, 2004.

To explain the transformation of a radially symmetrical planuloid with a single body axis, into a two axial, bilaterally symmetrical acoeloid, two scenarios have been traditionally contemplated. First, once established onto the substrate, the planuloid gradually shifted the posteriorly placed blastopore (mouth/anus) to a central (ventral) position. In this way the OA axis of cnidarians would correspond to the AP axis of the Bilateria (OA and AP axes would be homologous) and the DV axis would be a key innovation of the Bilateria (Figure 5 A). Moreover, the anus would appear later on, and independently, in different bilaterian clades. A second scenario sees the planuloid settling with the blastopore down, facing the substrate. This necessitates of the subsequent compression of the OA axis and elongation of its orthogonal axis, which will give rise to the AP axis of Bilateria. The OA axis would then be homologous to the DV axis, while the AP axis will form de novo with the displacement of the apical organ leading to the formation of a new brain anteriorly (Figure 5 B).

It is necessary to stress that, however interesting and appealing, these scenarios and hypothesis are just so stories with, so far, and very scant experimental support.

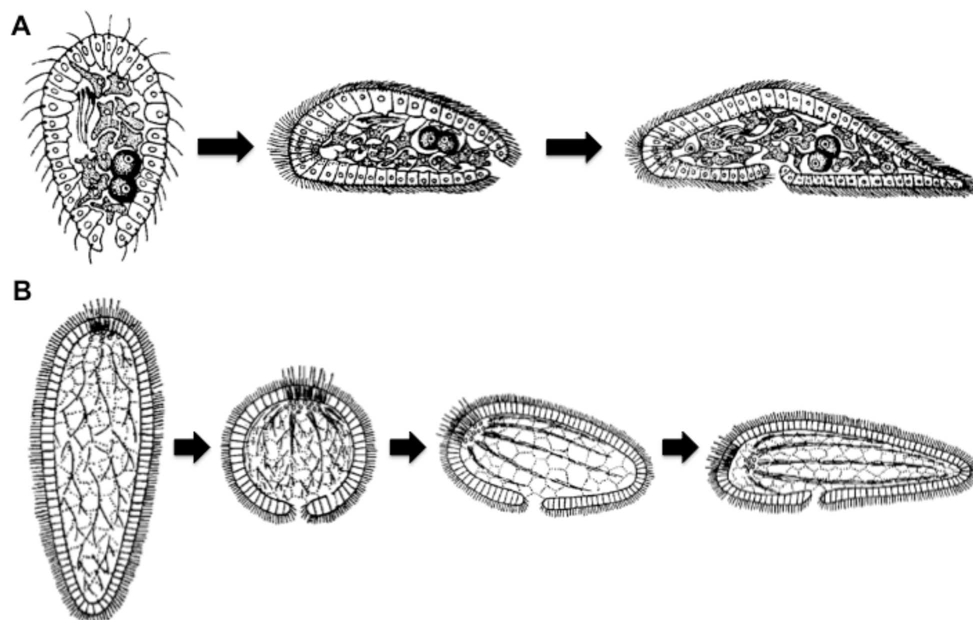


Figure 5. Two possible models for the origin of the Bilateria according to the Planuloid-Acoeloid hypothesis. **A.** A radially symmetrical ancestor settles on an antimere, which turns into the ventral side of the Bilateria. The now posteriorly placed blastopore (mouth/anus) moves to a central (ventral) position. **B.** A radially symmetrical ancestor settles down on its oral pole, which turns into the ventral side of Bilateria. The ancestor stretched its body and the apical brain area moved forward giving rise to a new AP axis. Modified from Hyman, 1951.

1. 3. 2. The Archicoelomate hypothesis

The Archicoelomate hypothesis (Sedgwick, 1884) posits that the Urbilateria was a rather large and complex organism derived from either the larval or the polyp stage of a Tetracorallia cnidarians (whose gastral cavity was divided by septa into four chambers). From this complex Urbilateria evolved the even more complex protostomate and deuterostomate groups. As a consequence, acoelomate and pseudocoelomate unsegmented bilaterians had to originate, at early and/or late stages of bilaterian evolution, by morphological simplification of the complex coelomate ancestor (Figure 6).

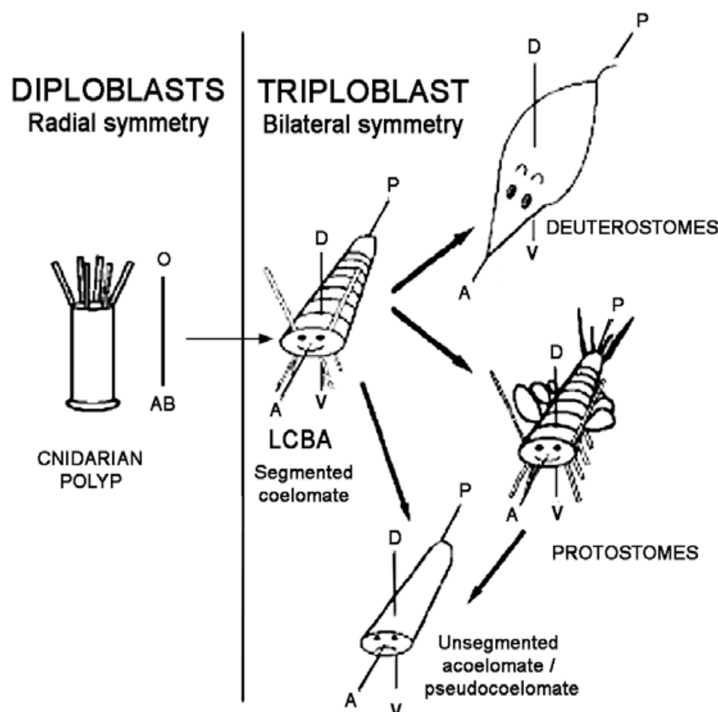


Figure 6. Diagrammatic representation of the Archicoelomate hypothesis. It features a swift transition from either a larval or an adult cnidarian to a complex or very complex LCBA, bearing through gut, eyes, coelom and, likely, segments and appendages. From this ancestor evolved the more complex protostomates and deuterostomates. In turn, acoelomate and pseudocoelomate non-segmented bilaterians derived at early and/or late stages of bilaterian evolution by morphological simplification from complex coelomate ancestors. A, anterior; AB, aboral; D, dorsal: O, oral; P, posterior; V, ventral. Modified from Baguñà and Riutort, 2004.

The developmental scenario of such transformation starts by considering the homology between the gastral pouches of cnidarians and the gastral evaginations (enterocoels) that later will form the coelomic cavities of deuterostomes. Thus, it is postulated that the coelomic cavities of the bilaterian ancestor arose by closure of the gastral pouches in a polyp or a late larvae of a cnidarian. In both cases, the OA and DV axes would be homologous, and the AP would originate de novo by orthogonal elongation of the body.

If an adult polyp was the actual ancestor (Figure 7), and however difficult to imagine it, it had to start crawling on their oral surface, which will became the ventral side. Subsequently, the oral opening (actually oval) elongated and sealed centrally except at both ends (amphistomy) giving rise simultaneously to the mouth and the anus (Sedgwick, 1884) and to a bilaterian organism with a through gut (Arendt *et al*., 2001; Holland, 2000). If, instead, a larva were the ancestor, the OA axis would be homologous with the DV axis. Then the apical pole (containing a nervous organ, the so-called apical organ) would move forward, towards one side of the larvae, giving rise to the future anterior region. In that case, the OA and DV axes are homologous, with the AP axis originating de novo.

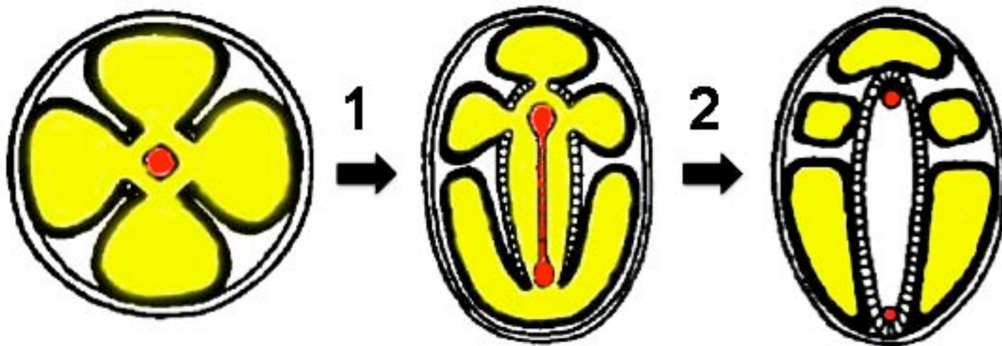


Figure 7. Origin of the Bilateria from a cnidarian polyp ancestor according to the Archicoelomate hypothesis. (1) A tetracorallian polyp elongates its primary mouth (red) and starts to close the gastral pockets. (2) The new mouth and anus forms at the ends of the primary mouth (red) as the result of its central closure (amphistomy). In turn, the gastral pockets are already closed and separate from the gut originating the coelomic cavities (yellow). Modified from Remane, 1950.

Recent molecular genetic studies have revealed striking similarities between genes and genetic networks specifying segmentation in such distant groups as vertebrates, annelids and arthropods (Holland *et al*., 1997; Tautz, 2004). Besides, a parallel set of similarities have also been described between the same groups of organisms for the spatial organization of their Central Nervous System (CNS, see, for instance: Denes *et al*., 2007). These data have been taken as solid evidence backing the Archicoelomate theory, and, therefore, for a complex macroscopic Urbilateria already bearing a through-gut, eyes, a complex CNS, coelom and, very likely, segments and appendages (Adoutte *et al*., 1999, 2000; Balavoine and Adoutte, 2003; De Robertis, 2008).

As it was the case with the Planuloid-Acoeloid hypothesis, the morphological and embryological evidence currently supporting the Archicoelomate hypothesis is very scant and most of its scenarios should be considered, like in the Planuloid-Acoeloid hypothesis, as mere just so stories.

2. THE APPROACH FROM THE EVOLUTIONARY DEVELOPMENTAL BIOLOGY TO THE ORIGIN OF THE BILATERAL ANIMALS

To properly assess the homologies between the body axes of cnidarians and bilaterians, we should go beyond comparative anatomy and embryology. Analyses of molecular mechanisms involved in the establishment of body axes and in the maintenance of axial polarities seems a powerful alternative and, at first view, a suitable tool to reveal homologies. Conservation of gene orthologies, gene expression patterns, and functional interactions, should reflect shared ancestry and homology (Nielsen and Martinez, 2003). This premise has been adopted as the base of the Evolutionary Developmental Biology research program since the early 1990's (Gilbert *et al*., 1996).

Briefly, Evolutionary Developmental Biology (informally, Evo-Devo) is interested in the comparison of the developmental processes, and the genes involved in them, among different organisms with the aim of understanding the phyletic relationships between organisms and to infer the changes in developmental processes that brought about changes in form during evolutionary time. Historically, evolutionary biologists (after the Synthesis) tried to understand the evolution of form and function looking at the changes in allele frequencies of specific genes thought to be involved in the adaptation of organisms to their environment. Instead, developmental biologists, and later on developmental geneticists, have focused on how alterations in gene expression and function underlie the changes in body shape during development, independently of their actual adaptative value. Bringing together both approaches promised to unravel how changes in gene regulatory networks controlling specific morphologies during development could generate enough phenotypic variation for natural selection to act upon.

The unexpected finding that the developmental-genetic toolkit is highly conserved across animal phyla and that only a small fraction of genes is actually involved in the regulation of embryonic development brought up two important consequences. First, a small set of developmental genes is enough to generate the huge diversity of body plans among the Metazoa. Second, small variations in the number of these genes and in the developmental gene networks that control them are enough to produce both small morphological changes and those bringing the so-called morphological novelties (Raff, 2000). In other words, the huge biodiversity of body plans doesn't seem to be due to drastic differences in gene number and/or gene types, as postulated by the classical neodarwinian Modern Synthesis, but rather is the consequence of changes in the gene regulatory networks that control embryonic developmental processes (Carroll *et al.*, 2005).

Hence, we assume that the thorough study of changes and innovations in the regulatory gene networks determining the axial patterns and germ layer formation of radial ancestors and bilaterian descendants will represent one of the most insightful approaches to analyze the origin of bilaterian animals. Consequently, and despite the time elapsed (600-550 Myr from the Ediacaran-Cambrian boundary), a register of these changes should still be traceable within the genomes of extant organisms. Accordingly, a comparative analysis of developmental genes between cnidarians and bilaterians must represent an ideal intellectual tool to establish their evolutionary relationships and to infer the mechanisms that produced the necessary morphological changes underlying the origin of new body plans.

Because bilateral symmetric animals are characterized by the presence of orthogonal AP and DV axes, with the concurrent presence of mesoderm and a centralized nervous system, developmental genes involved in patterning the AP body axis (Hox and ParaHox) and the DV body axis (Decapentaplegic/Bone Morphogenetic Protein and Short Gastrulation/Chordin) plus those involved with the development of the mesodermal layer (Forkhead, Brachyury, Goosecoid, Hedgehog, Twist, Mef2, Snail, etc) should prove to be the best targets to study the origin of Bilateria.

Among these genes, in this Thesis I have focused on the Hox and ParaHox homeotic genes. Reasons to choose them are discussed in the following section.

2. 1. HOX AND PARAHOX GENES

Hox genes encode for regulatory transcription factors involved in the regionalization of the antero-posterior axis during early embryonic development of bilateral animals (McGinnis and Krumlauf, 1992). In the 1980s, genes of the Hox family were for the first time cloned and sequenced (Gehring, 1985, 1998; Lewis, 1978) and soon after found to be present in all bilaterian organisms tested. Developmental biologists realized that animals with different body plans, e.g. insects and vertebrates, shared this same set of regulatory genes specifying position along the major body axis of the embryo. A link between genes and development was soon established, fuelling the development of the Evo-Devo as a new discipline (Raff, 2000).

Hox genes are functionally defined as homeotic genes because its lack of function lead to homeotic transformations of body segments; producing major changes in the morphology of the individual. In homeotic mutants, different body parts are often lost or, more often, one part of the body is transformed into a replica of another part. In *Drosophila* for instance, loss-of-function mutations of the gene Ultrabithorax (Ubx) cause the partial or whole transformation of the third thoracic segment into a duplicate of the second thoracic one, giving rise to a four-winged fly (Lewis, 1978).

In terms of DNA sequence and protein structure, the Hox are defined as homeobox genes because they possess a homeobox DNA sequence, which is a stretch of DNA about 180 nucleotides long that encodes a protein domain known as the homeodomain fold (Bürglin, 1996). The homeodomain fold is a structural domain of a 60-amino acid helix-turn-helix structure in which short loop regions connect three alpha helices (Figures 8 and 9). The homeodomain binds to DNA and regulate the expression of many targets (Bürglin, 1996; Gehring *et al*., 1994a).

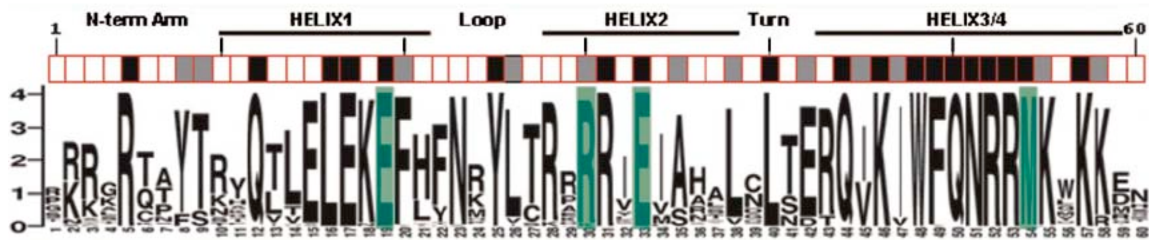


Figure 8. Hox homeodomain consensus sequence represented as a sequence logo (http://weblogo.berkeley.edu/logo.cgi). Black- and grey-filled boxes above the sequence logo indicate conserved positions, including invariant residues or conservative changes, respectively. Positions highlighted in green indicate residues specifically conserved within the Hox subclass, termed Hox generic signatures. Extracted from Merabet *et al*., 2009.

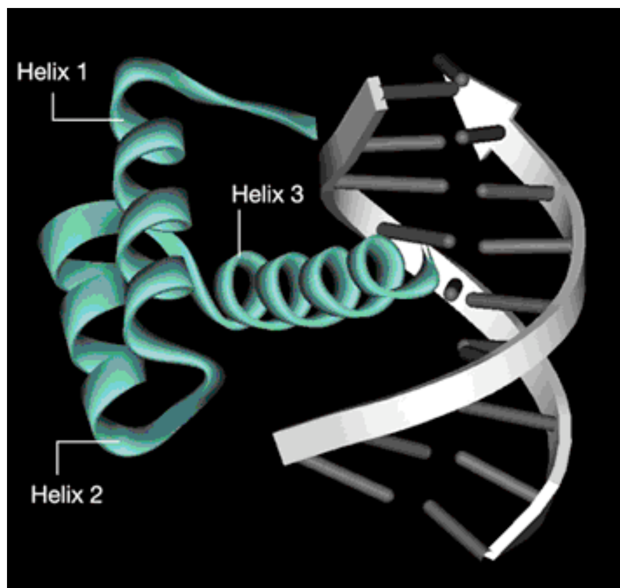


Figure 9. Three dimensional structure of the homeodomain with the three alpha helices binding to the major groove of the DNA chain. Extracted from Abate-Shen, 2002.

2. 1. 1. Hox and ParaHox genes in bilaterians

As mentioned, Hox genes play an evolutionary conserved role in patterning the antero-posterior axis of all bilaterians (Gilbert, 2006; Pearson *et al*., 2005). In fact, different Hox expression domains specify the regional identity over most of the AP body axis. However, Hox genes are not involved in the early establishment of the AP axis, a process requiring several maternally localised determinants and signalling molecules. Instead, the main function of Hox genes is to control the correct specification of structures along an already established axis (Carroll *et al*., 1988; Nüsslein-Volhard *et al*., 1987).

In distantly related taxa, such as vertebrates and arthropods, Hox genes were found to be arranged in evolutionarily conserved genomic clusters. However, the level of cluster conservation varies among the bilaterian groups. Duboule (2007) has recently proposed four types of differentially arranged HOX clusters. First, the canonical, contiguous, cluster of vertebrates, very compact, and with all its genes in the same transcriptional orientation. Second, the disorganized cluster with non-Hox genes interspersed among the Hox genes and with Hox genes transcribed in both orientations (e.g. in the sea urchin *Strongylocentrotus purpuratus*). Third, the so-called split clusters, resulting from the break of the original single cluster into two or more pieces (e.g. the *Drosophila'* HOM 'cluster'). Finally he recognizes, the atomized clusters, where the original cluster is completely disintegrated (e.g. the urochordate *Oikopleura'* 'cluster'), and the Hox genes dispersed in the genome.

In the organized cluster, the relative spatial and temporal expression of the different Hox genes correlates with their relative position within the cluster. Hence, genes located towards the 3' end of the cluster are expressed in more anterior body regions and earlier during embryonic development than genes located towards the 5' end. These properties are called spatial and temporal collinearity, respectively (Figure 10) (Duboule, 1994; McGinnis and Krumlauf, 1992). In non-organized clusters, the sequence of expression domains along the major axis of the different Hox genes, with respect to their *Drosophila* or mice orthologues, is usually kept. But in that case, the proper use of the term spatial collinearity is unsuitable since genes are not located in a single cluster. According to Duboule (2007) such phenomenon should be denominated trans-collinearity, to distinguish it from cis-collinearity: the correspondence between the physical order of Hox genes in an organized cluster and their domains of expression along the major body axis. Finally, it is important to point out that clear examples of temporal trans-collinearity have not been reported. This would indicate that temporal collinearity is more dependent than spatial collinearity on the presence of an intact cluster (Duboule, 2007).

Phylogenetic reconstructions have led to classify Hox genes into four classes of paralogues groups (PGs): anterior class (PG1-2, placed at the 3' end of the cluster),

group 3 (PG3), central class (PG4-8, placed in the middle of the cluster) and posterior class (PG9-13, placed at the 5' end of the cluster). PGs refers, thus, to their position in the cluster and their domain of expression along the AP body axis (McGinnis and Krumlauf, 1992; Pearson *et al*., 2005).

The total number of Hox genes within clusters varies among the bilaterian lineages, spanning from the 8 genes in the split Antennapedia-Ultrabithorax complex of *Drosophila melanogaster* to 13 genes in chordates and 14 in cephalochordates. The number of clusters within a genome is also variable due to partial or total genomic duplications (Figure 10). At least 7 clusters are present in the genome of the zebra fish (Amores *et al*., 1998) and 8 clusters in other teleost fishes (Meyer and Malaga-Trillo, 1999), whereas tetrapod vertebrates possess 4 clusters (Wada *et al*., 1999).

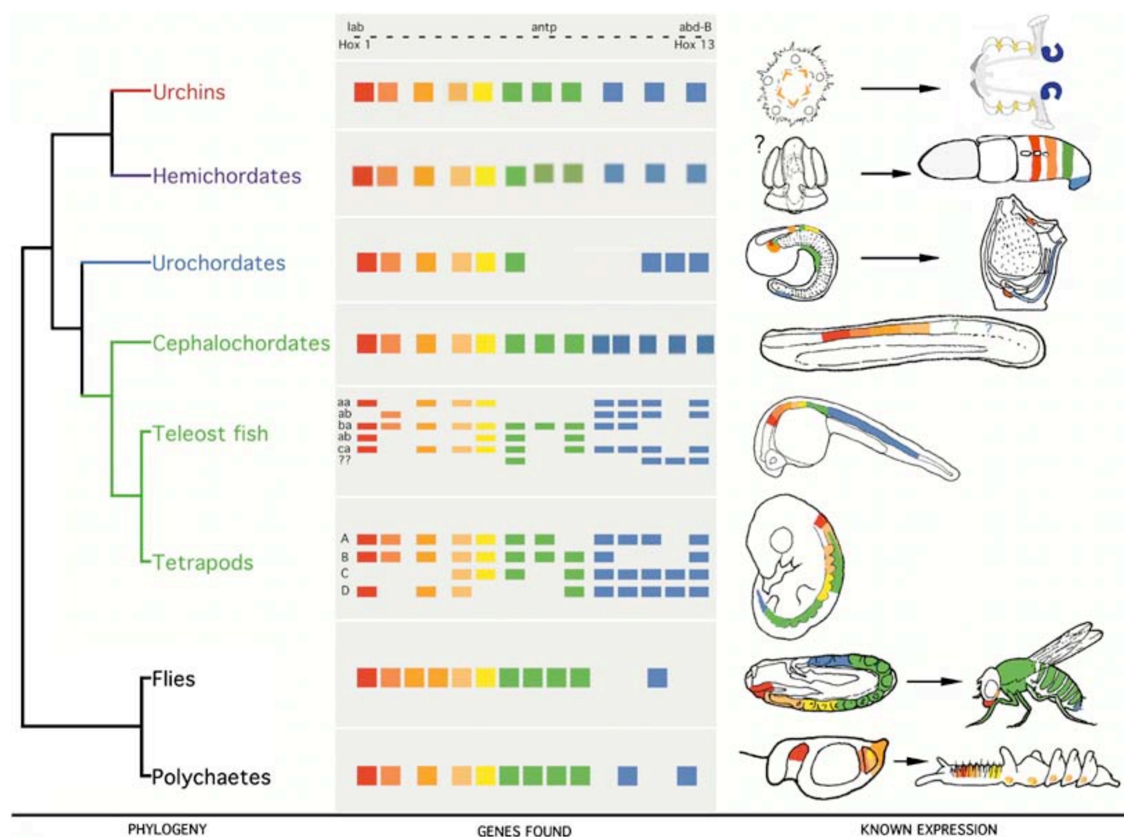


Figure 10. Diagrammatic representation of HOX clusters and their corresponding expression patterns in deuterostomes and protostomes plotted onto a consensus phylogeny of the represented taxa. Each group of paralogue genes is represented in a different colour. Extracted from Swalla, 2006.

Comparative analyses between Hox gene complements in Deuterostomia, Ecdysozoa, and Lophotrochozoa has led to the suggestion that the LCBA bore a large HOX cluster made by 7 or 8 genes (Balavoine *et al*., 2002; de Rosa *et al*., 1999). According to these authors, the LCBA and the protostome-deuterostome last common ancestor (P-DLCA) would be equivalent.

The ParaHox group is made of three genes phylogenetically related to Hox genes, but placed outside the HOX cluster. These genes are Cdx (Caudal-type homeobox), Xlox/Ipf1/Pdx (*Xenopus laevis* homeobox 8 / Insulin promoter factor 1 / Pancreatic and duodenal homeobox gene-1) and Gsx (Genomic screened homeobox).

In all protostome and most deuterostome species, ParaHox genes are dispersed in the genome. Although ParaHox cluster degeneration and break up occurred within vertebrates (e.g. hagfishes and teleost; Furlong *et al*., 2007; Mulley *et al*., 2006), ParaHox genes in the amphioxus *Branchiostoma floridae*, humans, mice, basal ray-finned fishes and *Xenopus*, are arranged in genomic clusters (Brooke *et al*., 1998; Ferrier *et al*., 2005; Furlong *et al*., 2007; Mulley *et al*., 2006).

Since molecular phylogenetic analysis suggests that Gsx is most closely related to Hox1-2, Xlox to Hox3, and Cdx to Hox9-14 (Brooke *et al*., 1998), the ParaHox gene cluster would be an ancient paralogue of the HOX one; and the two gene clusters would have arisen by duplication of a ProtoHox gene cluster (Brooke *et al*., 1998). However, this hypothesis have been challenged by recent phylogenies including sequences from cnidarian species (Chiori *et al*., 2009; Ryan *et al*., 2007).

In some organisms, the ParaHox genes are also involved in patterning some structures along the major body axis (mainly the endoderm), and show both spatial and temporal collinearity. That is the case of the gut in the sea urchin *Strongylocentrotus purpuratus* (Arnone *et al*., 2006). On the contrary, ParaHox genes are expressed in a non-collinear manner in the urochordate *Ciona intestinalis* (Ferrier and Holland, 2002) and in the polychaete annelid *Capitella* sp. I (Fröbius and Seaver, 2006).

To summarize, the general features of Hox and ParaHox genes suggest a causal relationship between the functional diversification of these genes within bilaterian lineages and the major metazoan transitions, such as the origin of symmetry, the origin of bilaterians or the origin of vertebrates.

2. 1. 2. Hox and ParaHox genes in non-bilaterians

The Hox-ParaHox genes are not innovations of the Bilateria. Indeed, clear homologues of the anterior Hox class genes and the anterior ParaHox gene Gsx have been reported from cnidarians (Finnerty *et al*., 2004; Gauchat *et al*., 2000; Ryan *et al*., 2007). An Xlox homologue has also recently been found in the hydrozoans *Turritopsis dohrnii* and *Clytia hemisphaerica* (Quiquand *et al*., 2009). Moreover, in some cnidarian species, posterior-like Hox genes and a putative Cdx like gene have been cloned, though their orthologies to bilaterian genes are uncertain, and they might have originated as cnidarian specific duplications (Chourrout *et al*., 2006). Paralogues belonging to the central class have so far not been recovered from cnidarians, suggesting that cnidarians never have had this group of genes.

It is important to highlight that these genes are not linked in single genomic clusters in any cnidarian species studied to date. Thus, *Nematostella vectensis* shows the anterior Hox genes linked in one region of the genome, while Gsx and Cdx are placed together in another region, and the "posterior Hox" appear in a different place of the genome (Ryan *et al*., 2007). In the hydromedusae *Eleutheria dichotoma* Hox genes are also physically dispersed in the genome (Kamm *et al*., 2006). As for the rest of cnidarians, no information is currently available on gene assemblies. Table 1 summarizes the list of Hox and ParaHox related genes recovered from different cnidarians.

	Hox 1	Hox 2	Post A	Post B	Post C	Gsx	Xlox	Cdx
Nematostella vectensis	✓	✓		✓	✓	✓	✓	✓
Metridium senile	✓							
Acropora millepora	✓					✓		
Cassiopeia xamachana				✓	✓	✓		
Clytia hemisphaerica	✓		✓	✓	✓	✓	✓	✓
Eleutheria dichotoma	✓		✓		✓	✓		✓
Hydra sp.	✓		✓	✓	✓	✓		
Hydractinia symbiolongicarpus						✓		
Podocoryne carnea	✓			✓		✓		
Turritopsis dohrnii							✓	

Table 1. Summary of the Hox and ParaHox related genes recovered from all cnidarian species analyzed so far. Genes have been classified according to phylogenetic analyses (taken from Chiori *et al*., 2009, and Quiquand *et al*., 2009). Note that posterior genes correspond to three groups specific of cnidarians, paraphyletic to those of bilaterian Hox9-14 (Chiori *et al*., 2009). Colour code: orange, Anthozoa; green, Scyphozoa; blue, Hydrozoa.

As regards Hox and ParaHox expression patterns in cnidarians, its strong variability among species has precluded the establishment of sound parallelisms with their bilaterian counterparts. Their role, if any, in OA axial positional information is equally obscure. Indeed, both Hox and ParaHox show different, and often completely opposite, expression patterns along the OA axis (Finnerty *et al*., 2003; Kamm *et al*., 2006; Yanze *et al*., 2001).

In ctenophores, evidence for the presence of Hox-ParaHox genes is extremely flimsy. Whereas Finnerty and colleagues reported one homeobox gene related to members of the Antennapedia Hox genes in the species *Beroë ovata* (Finnerty *et al*., 1996), further attempts to clone more genes yielded no results. Recently, and using degenerate polymerase chain reaction (PCR)-based methodologies, Pang and Martindale (2008) recovered four homeoboxes in *Mnemiopsis leidyi*, all belonging to the ANTP class. Phylogenetic analyses, however, could not reliably place them into any specific families. Altogether, it is unlikely that ctenophores bear Hox or ParaHox genes.

As far as it goes, evidence for Hox and ParaHox genes in the enigmatic phylum Placozoa (actually made by *Trichoplax adherens* and other undescribed species or varieties) is reduced to only one putative ParaHox gene (*Trox2*) assignable by phylogenetic analyses to the Gsx family. No Hox genes have been reported (Jakob *et al*., 2004; Monteiro *et al*., 2006). If Hox and ParaHox genes are considered sister groups of genes (García-Fernández, 2005a) the absence of Hox genes in *Trichoplax* is best explained to result from secondary losses.

Last but not least, the most basal metazoan phylum, the Porifera or sponges, do not seem to hold either Hox or ParaHox genes. Only genes belonging to the NK homeobox gene family have been reported from the complete genome of *Amphimedon queenslandica* (demosponge), strongly suggesting that Hox-ParaHox genes in bilaterians and cnidarians arose from the expansion of one ProtoHox gene that perhaps appeared in their common ancestor from within the NK group and after the split of the sponges (Larroux *et al*., 2007).

Altogether, and on the basis of data reported so far, the origin of the Bilateria correlates with the appearance of new classes of Hox genes, the group 3 and central Hox classes, followed by a general increase in the number of Hox genes (from 2 to 7 or more). Nevertheless, the huge difference between the small set of Hox-ParaHox genes in cnidarians and the more diverse and complex set of Hox genes in the P-DLCA makes very difficult to understand how this expansion occurred in the bilaterian lineage, after they split from cnidarians. An alternative scenario would imply that a cnidarian/bilaterian radial (or even bilateral) ancestor had a basic set of 3 or 4 Hox genes plus 3 ParaHox genes, from which bilaterians evolved with the expansion of the HOX cluster while cnidarians, and maybe placozoans and even ctenophores and poriferans, diversified and somehow by secondary simplification reduced their number of Hox-ParaHox genes. However interesting, such scenario is highly unparsimonious and not backed by current phylogenetic and phylogenomic results.

Recently, molecular phylogenetic studies have suggested that a group of turbellarian flatworms, the Acoelomorpha (acoels + nemertodertmatids), do no actually

belong to the phylum Platyhelminthes but form the first bilaterian offshoot previous to the protostome-deuterostome split (evidence backing this conclusion will be presented and discussed later on in this Introduction). As a result, the Acoelomorpha are nowadays considered the sister group of all remaining Bilateria, the so-called Nephrozoa (Ecdysozoa, Lophotrochozoa and Deuterostomia). The name Nephrozoa refers to the presence in them of an excretory system, and should be considered a synapomorphy for this group (Jondelius *et al*., 2002). This phylogenetic position of the Acoelomorpha suggests that the LCBA do not actually corresponds to the P-DLCA but to a simpler ancestor, in which the complement of Hox and ParaHox genes should be intermediate in number between the small set of cnidarians and the large set of the P-DLCA. Testing this important hypothesis is the main aim of this thesis. To fulfil it, I have analyzed the Hox and ParaHox complements in acoel flatworms, and determined whether or not they are compatible with such an "intermediate" position given to the acoelomorphs.

3. THE PHYLOGENETIC POSITION OF THE ACOELOMORPH FLATWORMS WITHIN THE PLATYHELMINTHES AND THE BILATERIA AS SEEN FROM MORPHOLOGICAL / EMBRYOLOGICAL DATA AND PHYLOGENETIC ANALYSES

In most "classical" zoological textbooks the phylum Platyhelminthes (commonly known as flatworms) has traditionally been considered the first lineage of bilaterians and a good proxy for the bilaterian ancestor. In fact, the monophyletic nature of Platyhelminthes has been accepted for almost 150 years (Ehlers, 1985) based on a set of shared characters such as acoelomate body structure, a ciliated monolayered epidermis, a frontal organ, the presence of stem-cells called neoblasts, their hermaphroditism, and the lack of hindgut and anus, being the mouth the only opening to the digestive tissue. However, besides the presence of neoblasts (still a matter of discussion), all these features are plesiomorphies; that is, characters shared with non other bilaterian organisms and, hence, of no use to define the phylum as monophyletic. The lack of

shared characters (synapomorphies) linking the main clades of the Platyhelminthes led Smith *et al*. (1986) to question its monophyly early on.

The Acoelomorpha, formed by the orders Acoela and Nemertodermatida, were classified within the free-living Platyhelminthes (class Turbellaria; Gegenbaur, 1859). In most phylogenetic proposals, the acoelomorphs are considered the most basal plathelminth clade. As in other orders of plathelminthes (e.g. Catenulida and Proseriata), the Acoelomorpha members bear a gravity-sensitive organ called statocyst placed in the anterior area of the body and embedded within the central nervous system. In acoels the statocyst consists of a round capsule surrounding a hemispherical concretion, the statolith. In nemertodermatids there are, instead, two statoliths per statocyst. In catenulids and proseriates, processes of outer neurons penetrate the capsule of the statocyst. Such innervations have not been found in acoelomorphs. The different structure of this organ among plathelminthes indicates that the different types of statocysts have likely evolved independently (Ehlers, 1991). As in most plathelminthes, the digestive system of acoelomorphs bears a single opening (the mouth), a pharynx, and a single layered gut epithelium. Whereas in nemertodermatids the gut is fully epithelial, most acoels, with the exception of some with truly cellular guts (Pedersen, 1964), lack a gut lumen and the gut epithelium has been transformed into a syncitial tissue mass.

In summary, despite no morphological/embryological synapomorphies are known linking the three main clades of plathelminthes, the phylum Platyhelminthes is considered monophyletic. Acoelomorphs are, therefore, considered plathelminthes, and very likely one of its most basal clades. As regards to the phylogenetic position of the Platyhelminthes within the Bilateria, the prevalent idea, based on its lack of coelom and presence of a closed gut (no anus), is to consider them one of the most basal bilaterian clades (*sensu* Hyman, 1951). Other phylogenetic schemes, however, consider them as derived spiralian protostomes; that is, morphologically simplified from a more complex (and likely coelomate) ancestor. In the first case, acoels should be considered one of the earliest branching extant bilaterians; in the later, acoels (and acoelomorphs) should be taken as a simplified, albeit early branching, spiralian plathelminthes.

3.1. MORPHOLOGICAL AND DEVELOPMENTAL FEATURES THAT SEPARATE ACOELOMORPHA FROM THE REST OF THE PLATYHELMINTHES

Besides the lack of morphological synapomorphies between acoelomorphs and plathelminthes, acoelomorphs have several autapomorphies that help us to distinguish them from the rest of plathelminthes. First, acoelomorphs lack protonephridia. Second, the rootlets of cilia in the epidermis are interconnected, forming a root system peculiar only to them. Third, in the locomotory cilia four of the nine peripheral double tubules end at a considerable distance below the tip of the cilium (bent cilia; Ax, 1995), a peculiarity of the Acoelomorpha. And finally, the general structure of the nervous system in acoelomorphs is much simpler than in most plathelminthes (Reuter *et al.*, 2001).

With regard to the embryonic development, two main features separate acoelomorphs from the rest of plathelminthes. First, and waiting for additional data from nemertodermatids, acoels have a duet spiral embryonic cleavage (Henry *et al.*, 2000) instead of the quartet spiral pattern that characterizes embryos of the phyla belonging to the so-called Spiralia clade (e.g. annelids, molluscs, and nemertines, among several others, including plathelminthes). In the duet cleavage, pairs of micromeres are generated starting at two-cell stage from two macromeres, whereas in quartet spiral cleavage the quartets of micromeres arise from the four macromeres produced after the first and second cleavages (Boyer *et al.*, 1996). Second, acoelomorphs produce only endomesoderm from both third duet macromeres whereas in quartet spiral cleavers mesoderm forms from both endomesoderm derived from the fourth quartet macromeres and ectomesoderm which derives from second quartet micromeres (Henry *et al.*, 2000).

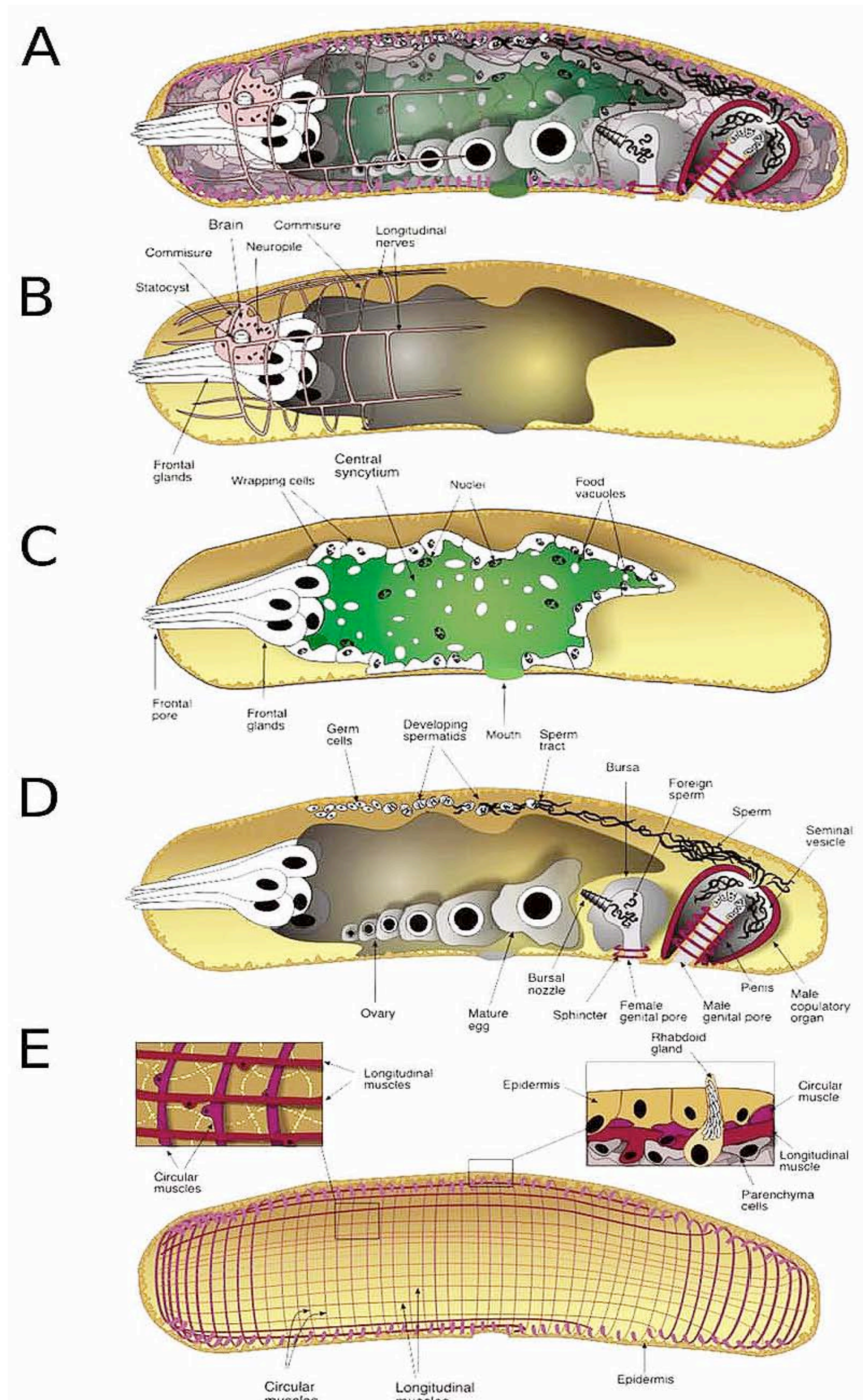


Figure 11. Morphology of the Acoela **A.** General morphology. **B.** Nervous system. **C.** Digestive system. **D.** Male and female reproductive systems. **E.** A closer view of the epidermis and the musculature. In all diagrams the anterior part of the body is placed to the left and the posterior part to the right; the dorsal side is up and the ventral side is down.
Adapted from the web site http://devbio.umesci.maine.edu/styler/globalworming/lateralview.html

3. 2. THE ADVENT OF MOLECULAR PHYLOGENIES. ARE THE PLATYHELMINTHES BASAL BILATERIANS AND MONOPHYLETIC?

The introduction of molecular phylogenies based on characters drawn from sequences of the ribosomal 18S subunit (Field *et al*., 1988) represented a turning point in the analysis and understanding of animal phylogenetic relationships. In a mere 10 years, the traditional scheme of increasing complexity from non-bilaterians to acoelomate, pseudocoelomate and finally coelomate bilaterians, changed substantially (for a review, see Adoutte *et al*., 1999). Most acoelomate (Platyhelminthes among them) and pseudocoelomate clades were found displaced from the base into more internal positions within two new superclades, the Lophotrochozoa and the Ecdysozoa. Together with a third superclade, the former Deuterostomata, these three superclades comprise all bilaterians.

An additional analysis using the first broad taxon sampling of Platyhelminthes corroborated its more internal position within the lophotrochozoan Bilateria, but also strongly suggested they were not monophyletic but polyphyletic (Carranza *et al*., 1997). Indeed, the bulk of the Platyhelminthes (the Catenulida and the so-called Rhabditophora) were found to belong to the Lophotrochozoa while the order Acoela was found to represent the earliest branching bilaterian group, and clearly separated from the rest of Platyhelminthes. Because acoels were represented in that study by a single species and because they had a high or very high rate of nucleotide substitution, which might result in an artifactual placement due to the so-called Long Branch Attraction (LBA) effect, a much denser sampling of acoels was analyzed over the following years. The results (Ruiz-Trillo *et al*., 1999, 2002) clearly showed that Platyhelminthes were indeed polyphyletic and that acoels represent the earliest offshoot of the Bilateria, well separated from the plathelminthes.

As a consequence, Acoelomorpha was erected as a new phylum (Baguñà and Riutort, 2004), made by two former orders of the Platyhelminthes: the Acoela and the Nemertodermatida, which now are ranked as classes. In turn, the rest of the

Platyhelminthes retain its former name and its rank of phylum, but it is now formed by two out of its three main former clades: the Catenulida and the Rhabditophora, holding the bulk of the former phylum. Therefore, and from now on, the name Platyhelminthes should be considered equivalent to Catenulida+Rhabditophora, and its phylogenetic position to be within the superclade Lophotrochozoa.

3. 3. PHYLOGENETIC, GENOMIC AND PHYLOGENOMIC DATA SUPPORT THE POLYPHYLY OF THE PLATYHELMINTHES AND THE BASAL POSITION OF ACOELOMORPHA WITHIN THE BILATERIA

In the last 10 years, the position of Acoelomorpha as earliest branching extant bilaterians separate from the Platyhelminthes has been further corroborated using different molecular datasets, namely sequences of 18S and 28S ribosomal genes (Jondelius *et al*., 2002; Telford *et al*., 2003; Paps *et al*., 2009a), single nuclear gene (Myosin Heavy Chain type II; Ruiz-Trillo *et al*., 2002) and larger sets of nuclear genes (Paps *et al*., 2009b), mitochondrial genes (Ruiz-Trillo *et al*., 2004), microRNAs sets (Pasquinelli *et al*., 2003; Sempere *et al*., 2007; Wheeler *et al*., 2009).

However, the introduction of phylogenomic approaches based on the use of high-throughput sequencing strategies that collect data from many genes (more than 100) at random (from expressed sequence tags or "ESTs" or whole-genome annotations) has provided the best evidence available so far that acoels are not members of the former phylum Platyhelminthes (Philippe *et al*., 2007, Hejnol *et al*., 2009). However, it is important to point out that these phylogenomic analyses still differ in a very important result: the precise position of the acoelomorphs within the Bilateria. While in Hejnol *et al*. (2009) Acoelomorpha appear with high support, together with the Xenoturbellida, as sister-group to the rest of the Bilateria, the analysis by Philippe *et al*. (2007) found more support (albeit rather low) for a placing of Acoelomorpha close to the deuterostomates, or even within them. Upcoming massive sequencing of whole genomes and a broader taxon sampling will bring new insights on that issue and, likely, unearth still a few surprises.

If acoelomorphs branched earlier than the protostome-deuterostome split it could be anticipated that they will bear an intermediate number of Hox genes between those of cnidarians and those of the P-DLCA. In fact, previous analyses of the Hox and ParaHox gene complement in acoelomorphs have shown the presence of three Hox genes in acoels, one corresponding to each of the anterior, central and posterior classes; plus one Cdx ParaHox (Cook *et al*., 2004). In nemertodermatids, instead, two central and one posterior Hox gene were found but not anterior relatives. Nemertodermatids also have an Xlox and a Cdx gene (Jiménez-Guri *et al*., 2006). Altogether, these results indicate that the set of Hox genes in Acoelomorpha might be considered representatives of an intermediate stage in the evolution from the small set of Hox genes present in the C-BLCA (2 genes) to the larger set (7-9 genes) present in the P-DLCA.

However, neither the arrangement of these genes in the genome, nor their functions in Acoelomorpha have been analyzed to date. To fill this gap, we searched in the acoel *Symsagittifera roscoffensis* whether their Hox genes are linked in a genomic cluster and whether they show a collinear expression along the (major) AP body axis. The results obtained, which address two of the main aims of this Thesis have already been published (Moreno *et al*., 2009) and will be discussed in the General Discussion section.

Besides, we have analyzed by RNA interference (Fire *et al*., 1998) the developmental role of a posterior Hox gene, called *IpHoxPost,* during the postembryonic development, and in regenerating and intact adults of the species *Isodiametra pulchra*. Results from this analysis have also been published (Moreno *et al*., 2010) and will be discussed in the General Discussion section. The use of a different species here is justified by the fact that, so far, is not possible to apply the RNAi methodology in *S. roscoffensis* due to difficulties with its culture at the laboratory. However this technology works perfectly in *I. pulchra*.

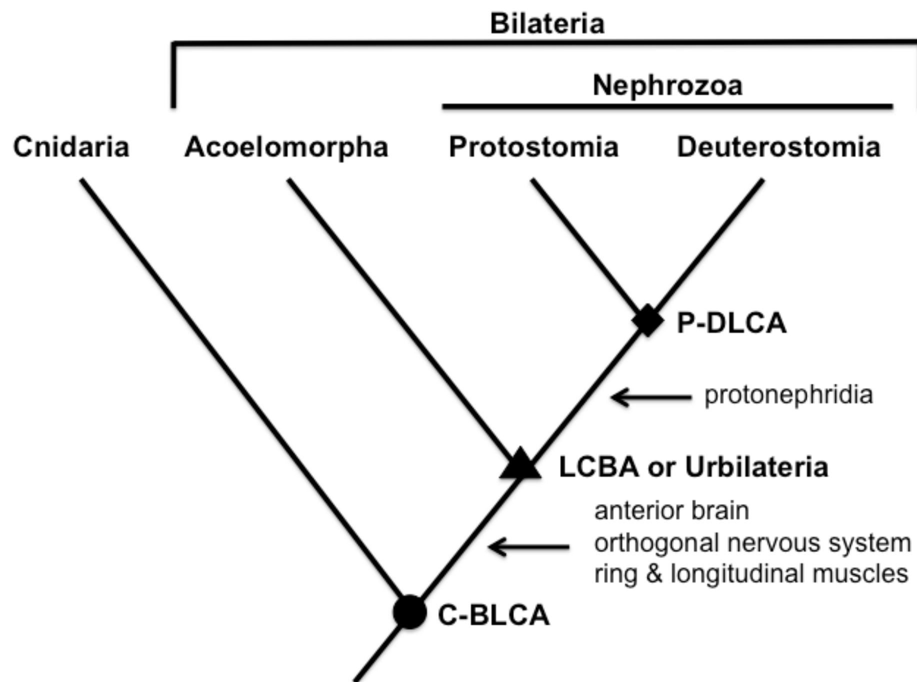


Figure 12. Phylogenetic position of the phylum Acoelomorpha as sister-group to the Nephrozoa (protostomes + deuterostomes) bilaterians. Main synapomorphies are indicated by arrows. LCBA originated when an anterior brain, an orthogonal nervous tract system, and circular plus longitudinal muscles developed. After Acoelomorpha branched out, protonephridia appeared giving rise to the Nephrozoa. Modified from Hejnol and Martindale, 2008a.

3.4. A CAUTIONARY NOTE FROM PHYLOGENOMICS AND OTHER MOLECULAR DATA

Although the phylogenetic position of acoelomorphs as the earliest branching extant bilaterian appears to rest on solid grounds (ribosomal, nuclear, and mitochondrial genes, miRNA sets, Hox data sets, etc.) the controversial results from recent phylogenomic studies (Philippe *et al*., 2007; Hejnol *et al*., 2009) ask from us to adopt a cautionary attitude. Indeed, due to its rather high or very high nucleotide/aminoacid rate changes, the branch lengths of most acoel species in phylogenetic trees are not short but moderate to long, in some cases very long. Under specific evolutionary models and when small taxon samples and long distant outgroups are used, long-branch taxa could be artifactually attracted either externally to the outgroup or internally to other long branch taxa. This is the well know LBA effect (Felsenstein, 1978). Ways to overcome it go from devising best evolutionary models, using broad taxon sampling, and using few, short branched, close outgroup taxa.

Waiting for upcoming wider taxon and gene sampling phylogenomic analyses and from whole genome sequencing data (under way), and taken for granted that acoels are definitely not plathelminthes, its final phylogenetic position within the Bilateria will be, most probably, and taking into account all published (and unpublished) data, one the following two: 1) earliest branching extant bilaterians as argued at length along this Introduction; or 2) sister-group to the deuterostomates or, a bit further in, sister-group to the deuterostomate Ambulacraria (Hemichordatea+Echinodermata). If the first scenario is true, the acoelomorphs would represent extant relatives of the LCBA and, hence, a good model system to analyze the putative Radial-Bilateral Transition (RBT). If the second outcome is the correct, acoelomorphs should be taken as simplified organisms originated, likely by progenesis, from the common ancestor of deuterostomates or from the common ancestor to the Ambulacraria+Acoelomorpha clade. And if this turns out to be the case, the reduced number of Hox genes and the reduced number of miRNAs in acoelomorphs should be better interpreted as the result of specific and sustained losses from an ancestor with a larger complement of both Hox and miRNAs. In this particular case, acoelomorphs would become a good model system, for instance, to analyze and understand how simplified morphologies are paralleled (or not) by reduced or modified genomes.

4. THE ACOEL SPECIES *SYMSAGITTIFERA ROSCOFFENSIS* AND *ISODIAMETRA PULCHRA*, NEW ANIMAL MODELS IN EVOLUTION AND DEVELOPMENT

As already mentioned, acoels are small, flat, soft-bodied, unsegmented worm-like organisms of increasing interest, due to its new phylogenetic position at the base of the bilaterian tree, as models to analyze evolutionary developmental biology questions. Acoels live in marine habitats around the world, from tropical to polar regions (Friedrich and Hendelberg, 2001; Haapkylä *et al*., 2009). Most live in between sand grains on beaches, although there are also some benthic and pelagic species. Acoels are hermaphroditic and nearly all species reproduce sexually by mating and internal fertilization. Some species are even able to reproduce asexually by fission or budding. In addition, some families like the Convolutidae and Sagittiferidae show a very

interesting feature: the establishment of symbiotic relationships with unicellular algae which live inside the animal's body (Bourlat and Hejnol, 2009).

The class Acoela is composed by 22 families and approximately 370 species (Tyler *et al*., 2006). They have been fairly well described at the morphological and ultrastructural levels. However, knowledge of its developmental and morphological features is, so far, restricted to a handful of species. In this thesis, two acoel species have been studied in depth: *Symsagittifera roscoffensis* and *Isodiametra pulchra*. A brief introduction on the biology of both species and its significance as models for Evo-Devo studies follows.

4.1. *SYMSAGITTIFERA ROSCOFFENSIS*. KOSTENKO AND MAMKAEV, 1990 (GRAFF, 1891)

Symsagittifera roscoffensis belongs to the family Sagittiferidae. Its main feature is the presence of an especial male copulatory organ and needle-shaped secretory products called sagittocysts, which range from 5 to 50 µm in length and from 1 to 5 µm in diameter (Mamkaev and Kostenko, 1991). Sagittocysts are distributed at the ventral side of the body, often close to the genital openings in sexually mature animals, indicative of a putative relationship between sagittocysts and sexual behaviour.

The *S. roscoffensis* adults measure around 2.5-3 mm long. The body is stretched and flat (about 450 microns) during gliding but the lateral sides of the body are completely bent towards the central ventral side when swimming. The anterior part of the body contains a set of several large mucus-secreting glands called the frontal organ, two ocelli, the statocyst, and a bilobed brain, from which spring six longitudinal nerve tracts towards the posterior end of the body (Bery *et al*., in press; Rieger *et al*., 1991). The epidermis is single layered encasing an innermost gut (also called central parenchyma) bearing a single mouth (no anus) and an intermediate layer of tissue called peripheral parenchyma. The oval-shaped mouth opens ventrally at the end of the first eighth part of the body length and is connected to a pharynx, which is a simple deep invagination of the tegument forming a short tube in the parenchyma. The peripheral parenchyma is formed by the insunk bodies of epithelial and gland cells, by muscle cells

and neoblasts, and by true parenchyma cells of two general types: fixed parenchyma cells, highly branched and electrolucent, and chordoid cells, also highly branched but vacuolated (Rieger *et al.*, 1991). A third type, the wrapping cells, form a layer that separates peripheral and central parenchyma. The posterior part of the body contains the male and female gonopores and the penis. The ovaries and testis extend from just posterior to the mouth opening down to the female and male gonopore, respectively.

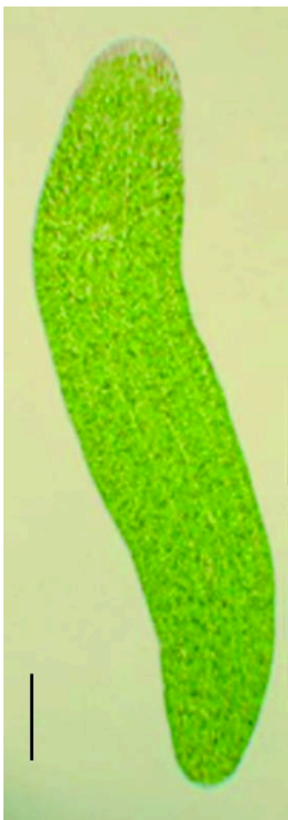


Figure 13. Dorsal view of an adult *specimen of S. roscoffensis*.
Light microscopy.
Scale bar: 220µm.
Anterior is up.

S. roscoffensis adults are green in colour (Figure 13) due to the presence of green extracellular endosymbiotic algae of the species *Tetraselmis (Platymonas) convolutae*, which are distributed along the whole body length (Keeble and Gamble, 1907; Selosse, 2000). The presumed benefits of this symbiosis for the algae are shelter against predation, a constant milieu, optimal photosynthetic conditions, and interchange of nutrients as N-compounds (Selosse, 2000). While the juveniles could feed from the medium they must ingest the algae during growth; otherwise they eventually die. For the adults, the symbiosis implies stopping active feeding and developing a strict dependence on the algae for life and survival (Keeble and Gamble, 1905).

S. roscoffensis lives in sheltered sand beaches, distributed all long the European Atlantic coast from the English Channel to Portugal. The endosymbiosis determines the ecology and behaviour of *S. roscoffensis*. Since the algae must be in contact with the sunlight in order to perform the photosynthesis, the acoels inhabit in the small layer of water between the sand grains and the sea surface. At low tide they form green pools between the tide-marks, resembling layers of algae (Keeble, 1910; Selosse, 2000). The great abundance and the broad distribution along the European Atlantic coast have contributed to the use of this species as a model organism in different biological studies. The beaches of Roscoff, a small village in the north of Bretagne

(France), are one of the best places to collect specimens, due to their high abundance and easy collection.

Like other acoel species, *S. roscoffensis* are hermaphrodites. They reproduce through internal fertilization after the copula, often carried out by sperm injection through the body wall (Apelt, 1969). The eggs are laid in clusters embedded in jelly, known as cocoons. The cocoon contains from 2 to 12 eggs (Bresslau, 1909; Georgévitch, 1899). From each egg hatches a colourless worm. They acquire the free-living algae from the surrounding seawater three days after hatching. After several weeks, the juveniles develop the copulatory organs and become fertile adults.

Interestingly, *S. roscoffensis* is able to regenerate missing parts of the body. This feature is closely related with the regeneration capability of all acoel species, on account of the presence of neoblasts. In the case of *S. roscoffensis*, amputated animals are able to regenerate their missing parts after one week. The regeneration ability is another character that makes *S. roscoffensis* a very useful model organism for developmental and genetic studies.

Last but not least, it is worth to highlight two powerful genetic tools available for this species recently developed in our laboratory (University of Barcelona, Department of Genetics). Firstly, a collection of 30.000 ESTs clones from juvenile specimens, which were chosen instead of adults in order to avoid contamination from the DNA of the algae. This ESTs collection is an array library stored in microtiter plates, already sequenced by Genoscope (Centre National de Séquençage, France), which have been very useful for phylogenomic analyses (Hejnol *et al.*, 2009) as well as to spot genes of interest in developmental and evolutionary terms. Secondly, a bacterial artificial chromosomes (BACs) library containing 100.000 independent clones has been produced using the pBACe3.6 vector, which is the only one available in acoels at present. The 100.000 clones correspond, approximately, to 10 copies of the whole genome, enough coverage to recover most genes of interest (Frengen *et al.*, 1999). The inserts in the BACs have an average size of 100 kb pairs. The availability of BAC libraries is a key aspect for genetic linkage analyses and future genome sequencing.

4.2. *ISODIAMETRA PULCHRA*. HOOGE AND TYLER, 2005 (SMITH AND BUSH, 1991)

Isodiametra pulchra (Figure 14) is an acoel belonging to the Isodiametridae family, characterized by a male copulatory organ with a muscular, isodiametric, tubular penis (Hooge and Tyler, 2005). *I. pulchra* is distributed widely along the East Coast of North America.

Its general morphology is similar to other acoels species. Two main features help to distinguish it from *S. roscoffensis*. First of all, the body length of *I. pulchra* adults is about only 360 μm; that is, about 9-10 times smaller than *S. roscoffensis*. On the other hand, *I. pulchra* is unpigmented because they do not hold any symbiotic algae. However, well-fed individuals have often a green digestive parenchyma, likely due to the photopigment of ingested diatoms (Smith and Bush, 1991). *I. pulchra* usually feeds on these diatoms; and they can be easily kept in laboratory cultures feeding on the diatom *Nitzschia curvilineata*. In culture, every fertile adult release an average of 1-2 eggs per day. Eggs are released individually and not in cocoons as in *S. roscoffensis*.

Figure 14. Dorsal view of an adult specimen of *I. pulchra*. Light microscopy. Scale bar: 100mm. Anterior is to the left.

Their easy culture (Andersen *et al*., 2005), extensive knowledge of its morphology (Smith and Bush, 1991), and their ability to regenerate, makes of this species a very good laboratory model for biological studies. In addition, knockdown technologies based on RNA interference have recently been developed for *I. pulchra* (De Mulder *et al*., 2009a). In particular, it has enabled us to analyze the function of the posterior Hox

gene, *IpPostHox*, during the post-embryonic development, regeneration, and adulthood (Moreno *et al*., 2010).

Finally, it is important to highlight that we have generated in the laboratory a relatively small ESTs collection that contains 2.300 clones belongs to adult specimens, some of which encode for interesting regulators of development. This collection has been sequenced at the Max Planck Institute in Berlin.

Both ESTs collections, from *S. roscoffensis* and *I. pulchra*, have been used to infer phylogenetic relationships of metazoan groups, being specially important for the phylogenomic analyses in which acoels result to branch as the earliest bilaterians (see Hejnol *et al*., 2009).

2

AIMS OF THIS STUDY

Recently phylogenetic and phylogenomic analyses strongly suggest that acoel flatworms are the earliest offshoots of the Bilateria, holding an intermediate phylogenetic position between radial cnidarians and the rest of the bilaterians. As such, the study of developmental regulatory genes in acoels should provide us with key insights on the putative radial-bilateral transition. Among developmental regulatory genes, Hox and ParaHox genes, involved in antero-posterior axial patterning, are particularly important to tackle this evolutionary issue.

Accordingly, the main aims of this thesis are:

1. To determine the Hox and ParaHox gene complement in the acoel flatworm *Symsagittifera roscoffensis*.

2. To analyze whether *S. roscoffensis* Hox genes are arranged in a genomic cluster or not.

3. To sequence the genomic regions containing Hox genes in *S. roscoffensis* to determine the distribution of exons and introns, to find out conserved regions and regulatory domains, and to analyze the possible existence of syntenic relationships with the HOX clusters in other Bilateria.

4. To analyze the Hox and ParaHox gene expression patterns in juvenile and adult *S. roscoffensis* specimens by whole mount *in situ* hybridization.

5. To study the role of the Hox genes in the acoel species *Isodiametra pulchra* during its postembryonic development, regeneration, and adulthood by RNA interference methods (techniques that only work in this species), with the objective of inferring the ancestral role of the Hox genes within the Bilateria.

3

GENERAL DISCUSSION

1. THE HOX-PARAHOX COMPLEMENT IN ACOEL FLATWORMS, AND ITS GENOMIC ARRANGEMENT AND STRUCTURE IN THE ACOEL *SYMSAGITTIFERA ROSCOFFENSIS*

Hox genes control the specification of tissues along the major (AP) body axis in all, extant, bilaterian groups. It is unclear, though, when this role was established for the first time in evolutionary history. Because acoel flatworms represent the first offshoot within the Bilateria (Baguñà and Riutort, 2004; Hejnol and Martindale, 2008a; Hejnol *et al*., 2009; Ruiz-Trillo *et al*., 1999), we decided to analyze its Hox gene complement and to explore whether the different Hox genes are arranged in a genomic cluster. Our aim is to explore if the small set of Hox genes in acoels might be instrumental in explaining how the expanded set of 7-8 Hox genes postulated for the last common ancestor of protostomes and deuterostomes (P-DLCA) (Balavoine *et al*., 2002; de Rosa *et al*., 1999) arose from the very small set of Hox genes in Cnidaria (Chourrout *et al*., 2006; Kamm *et al*., 2006; Ryan *et al*., 2007).

1. 1. THE HOX-PARAHOX COMPLEMENT IN ACOELS

Using the PCR methodology to screen genomic DNA and cDNA samples from *S. roscoffensis* (Sr) and *I. pulchra* (Ip) we recovered an identical set of three Hox genes in both species, with clear affinities to the bilaterian paralogues groups 1 (*SrHox1*, *IpHox1*), 5 (*SrHox5*, *IpHox5*), and 9–10 (*SrHoxPost*, *IpHoxPost*). A Cdx ParaHox gene was also recovered from *S. roscoffensis* (see Cameron *et al*., 2006, for the study of gene relationships). Cook *et al*. (2004) had already published partial sequences for these three genes in *S. roscoffensis* but genome mapping was missing. We extended the length of our sequences using both rapid amplification of cDNA ends (RACE) methodologies and genomic sequencing.

Furthermore, two recent and independent studies have recovered the same set of Hox genes from the acoels *Convolutriloba longifissura* (Hejnol and Martindale, 2009)

and *Convolutriloba retrogemma* (Sikes and Bely, 2008). Although the precise number of genes belonging to any particular gene group within any genome can not be determined before sequencing the whole genome, the independent finding of the same Hox gene complement in four acoel species likely means that no more Hox genes are present in the genome of this group. The presence of two posterior Hox genes in the acoel species *Paratomella rubra* (Cook *et al*., 2004) is most probably the result of a specific gene duplication in Paratomellidae.

As regards ParaHox genes, only Cdx orthologues have been found in *S. roscoffensis* and *C. longifissura*. However, until sequences of whole genomes are available, the exact number of ParaHox genes also will remain uncertain. It is important to note that nemertodermatids, the putative sister-group of acoels have, besides a Cdx gene, one Xlox gene (central ParaHox) (Jiménez-Guri *et al*., 2006). This would suggest that acoels might have also Xlox genes.

The set of three Hox genes (from the anterior, central and posterior classes of paralogues groups) found in the four acoel species might therefore represent the ancestral condition in Bilateria and the starting point for the use of a ‘‘Hox code’’ as a source of positional information along the AP axis. Indeed, we would suggest that one gene from each of the three major groups of Hox genes would represent the minimal set compatible with a bilaterian grade of structural organization. An analysis supporting this view come from a recent work by Ogishima and Tanaka (2007), who used motif-based reconstruction methodologies to infer with confidence the sequences of the proto-central Hox (and other) proteins. Strikingly, the sequence from the central Hox of acoels perfectly match that predicted for the central gene in a hypothetical ancestral ProtoHox cluster. From these results the authors suggest that the Hox gene set in acoels represents, in fact, an intermediate stage in the evolution of the Hox group between the cnidarian-bilaterian last common ancestor (C-BLCA) and the extant bilaterian groups.

Alternatively, the three gene set might also be the result of a secondary reduction (the acoel Hox complement being not ancestral but derived) from a larger Hox gene set in the bilaterian ancestor (Deutsch, 2008). There are several arguments against this

scenario. First, it seems unlikely that an ancestor with a large complement of Hox genes would lose all but one copy of each class. We would expect, instead, a salt-and-pepper pattern of Hox retention, as it actually occurs in other taxa where gene loss has left unequal numbers of Hox genes belonging to these three classes (e.g., nematodes (Aboobaker and Blaxter, 2003), ascidians (Ikuta *et al*., 2004), and Platyhelminthes (Koziol *et al*., 2009)). Second, the Hox gene reduction in acoels would be more plausible, and palatable, if linked to a, yet to be reported, general reduction in gene content during the evolutionary history of this group. At first, the free-living (non-parasitic) lifestyle of acoels and the lack of embryological and morphological evidences for secondary reductions argue against such big, and specific, losses. Finally, and as a consequence, if the small set of Hox genes in acoels resulted from secondary gene losses, this will leave unabridged the large gap between the set of 7-8 genes Hox genes postulated for the P-DLCA and the very small set of Hox genes in Cnidaria (Chourrout *et al*., 2006; Kamm *et al*., 2006; Ryan *et al*., 2007). Hence, unless acoels and cnidarians share a massive Hox reduction from a hypothetical large HOX cluster at the C-BLCA, such a large gap is left unexplained. On the contrary, if the phylogenetic position of acoels as extant members of the earliest divergent Bilateria holds, their simple Hox gene set and their simple morphology might represent the core from which more elaborate morphological patterns gradually arose in "higher" bilaterians in parallel to the expansion of central and posterior Hox classes. This, in essence, will support a more parsimonious scenario for Hox gene (and morphological) complexification.

1. 2. THE ARRANGEMENT OF THE GENOMIC HOX GENE CONTAINING REGIONS IN *S. ROSCOFFENSIS*

To analyze whether the three Hox genes found in acoels are linked and organized in a single genomic cluster, we screened a BAC library, developed from adult specimens of *S. roscoffensis* in our laboratory, with the sequences previously isolated using the RACE experiments. None of the isolated BAC clones bore more than one Hox gene. This became very clear after sequencing completely the three positive BACs, each containing a different Hox gene. In order to understand whether these Hox genomic regions were in close proximity, within the whole genome, the three BAC clones were

used as probes in fluorescence *in situ* hybridization (FISH) assays, performed both on metaphase chromosomes and in interphase nuclei. These experiments clearly showed that the three Hox genes in *S. roscoffensis* are located in different chromosomes. Hence, in this species, these Hox genes are not linked in a single cluster.

In *S. roscoffensis*, and maybe in the whole acoel clade, Hox genes seem to have been dispersed across the genome, likely from a hypothetical entire cluster. Reasons for such cluster dispersion in acoels are not clear. As pointed out by Duboule (2007), compact and organized cluster only occur in vertebrates and cephalochordates. In other bilaterians, Hox gene clusters are disorganized (e.g. the sea urchin cluster (Cameron, 2006)), split (in *Drosophila,* the original cluster has been divided in two (Kaufman *et al*., 1990; Lewis, 1978); in the tunicate *Ciona intestinalis* the HOX cluster split apart into five pieces (Ikuta *et al*., 2004)) or atomized (Hox genes scattered all over the genome; e.g. the urochordate *Oikopleura dioica* (Seo *et al*., 2004)). Because acoels seems to be to a very ancient linage, which appeared, more than 550 Myr ago during or even before the Cambrian age (Ruiz-Trillo *et al*., 1999), such a long time span would allow for the cluster HOX to disintegrate, provided that there have been no constraints to maintain it intact (e.g. global regulators, (Duboule, 2007)).

Finally, is important to note that to assign each gene to a specific chromosome could not be done with confidence. The main reason is the almost identical size and form of the metaphasic set of 20 chromosomes of *S. roscoffensis* (see Fig. 3 in Moreno *et al*., 2009).

1.3. THE STRUCTURE OF THE GENOMIC REGIONS CONTAINING HOX GENES IN *S. ROSCOFFENSIS*

The genomic sequences around the three Hox genes cover a total area of 300 kb, which should allow us the analysis of syntenic relationships, if any, with the corresponding genomic regions of other organisms. Unfortunately, neither Evx nor Mox genes are present in the three BAC clones sequenced, suggesting a lack of syntenic relationships. Moreover, the microRNAs associated with HOX clusters in other

bilaterians have not been found either in our annotation using the miRBase tool (Sanger Institute). Looking at these genomic regions with more detail, other genes adjacent to the Hox genes were annotated using the prediction software GENESCAN and, when BLASTX similarity searches gave hits with E values <10 E-10, selected for further analyses (see Moreno *et al*., 2009, Fig. 2 for more information about these genes). These predicted transcriptional units we found to be functional since they were amplified by PCR using specific primers for each gene on cDNA samples of different developmental stages. These genes, however, do not have orthologues in the equivalent containing-Hox genomic regions of other bilaterians. Therefore, the genomic areas where *S. roscoffensis* Hox genes are placed differ substantially from those of other bilaterian Hox regions. This has been demonstrated to be a usual feature in dispersed or broken clusters (Ikuta *et al*., 2004).

Interestingly, two transposon-related proteins were found at both sides of the *SrHox5* gene: the Pol polyprotein, which contains two retrotransposon protein domains (RT-LTR and rve or integrase), and the HASI, which also contains two retrotransposon protein domains (rve and Pox_A32). At one side of *SrHoxPost,* two transposon-related proteins were also found: the barrier-to-auto integration factor and the Pol polyprotein, which in this case only contains the rve protein domain (Moreno *et al*., 2009, Fig. 2). These sequences would indicate that the split and further dispersal of a hypothetical ancestral HOX cluster might have resulted from unequal recombinations between these and other similar sequences in different chromosomes (Lim and Simmons, 1994). As a result, and irrespective whether it took place by this mechanism or by other shuffling mechanisms, the Hox gene complement in the acoel *S. roscoffensis* appears dispersed in the genome. Deeper analysis of flanking sequences (for instance repetitive sequences) should provide us with more insights on the process that lead to this gene dispersion.

The introns and exons numbers, intron positions and exon distances within the *SrHox5* and *SrHoxPost* have also been determined from genomic sequence predictions and comparison with cDNA sequences. However, for the anterior Hox, this analysis was not possible because the sequences of the BAC clone containing this gene begin, actually, at the end of the Hox gene, although the final exon of *SrHox1* was included.

We have cloned new BAC clones containing the Hox1 gene and these are being now sequenced. It is assumed that they will contain the rest of the transcriptional unit for these genes. The predictions from the genomic sequences and the cDNA sequences can be checked in the Appendix C of this thesis. A schematic diagram of the structure of both *SrHox5* and *SrHoxPost* genes is shown next.

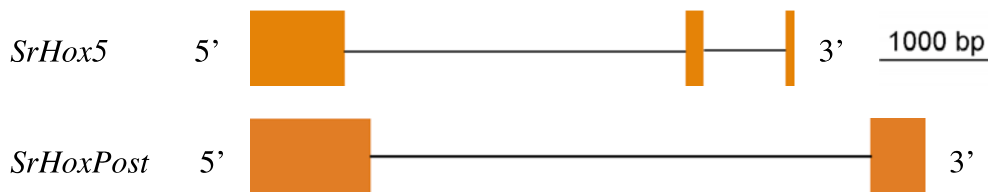


Figure 1. Structure of the coding regions of *SrHox5* and *SrHoxPost* genes. Orange squares represent the exons and thin black lines the intronic regions linking the exons. Scale bar indicates base pairs length.

The second intron of *SrHox5* is a phase-2 intron in the codon 53 within the homeodomain. Examples of introns within the homeobox have been also found in some bilaterian anterior and posterior Hox genes (Bürglin, 1996), as well as in their putative sister ParaHox genes (Arnone *et al*., 2006), but not in central Hox genes. However, in the plathelminthes species *Echinococcus multilocularis*, *Schistosoma mansoni* and *Schmidtea mediterranea*, there are two introns within the homeodomain of the Lox5 gene (Koziol *et al*., 2009). Interestingly, the second intron is a phase-0 intron between the codons 51 and 52, being this position very close to the position of the *SrHox5* intron. However, Lox5 genes from plathelminthes possess the Lox5 or "spiralian parapeptide", a diagnostic amino-acid motif which is N-terminal to the homeodomain used previously to assign lophotrochozoan affinities (Balavoine *et al*., 2002). However, the Hox5 genes of acoels lack this parapeptide, which is another evidence they are not plathelminthes (see Figure 3). Besides, the planarian *PlHox5* has a glutamine in the position 6 of the homeodomain, which is characteristic of Lox2, Lox4, Lox5, and Hox7 genes, instead of threonine 6, as found in all other Hox5 genes (Koziol *et al*., 2009). Acoels bear the characteristic T6 of Hox5 genes.

In addition, sequencing of genomic regions containing Hox genes has allowed spot other conserved domains in the protein products (Figures 2, 3 and 4). The main

region of conservation between these and other Hox genes is, obviously, the homeobox (Moreno *et al*., 2009. Fig.1). Hejnol and Martindale (2009) have analyzed the distribution in acoels of conserved domains at both sides of the homeodomain; the so-called N- and C- domains (Ogishima and Tanaka, 2007). The N-domain of the anterior Hox gene is a Na-motif, similar to that found in other bilaterian Hox1 orthologues (both share the Na-signature; motif description appears in Figure 2). However the C-domain of the anterior Hox orthologue in acoels is shorter than in all other Hox genes and is not similar to any of the previous C-motifs described (Figure 2). The central class orthologue of acoels contains the Nb N-motif (motif description appears in Figure 3), which characterizes the bilaterian central Hox orthologues. Besides, it contains a C-motif called Ca motif, which has greatest similarity with the motifs of bilaterian Hox5 orthologues (Figure 3). Finally, acoel posterior Hox genes possess the Nbx N-motif (motif description appears in Figure 4), specific of posterior Hox proteins in bilaterians, and a non-conserved C-terminal domain.

Using motif-based reconstruction methodologies, Ogishima and Tanaka (2007) have proposed that three genes, a protoHox1, a protoHox2 and a protopost formed the ancestral cluster, in the C-BLCA. Later on, in the bilaterian lineage, the protocentral gene arose. Interestingly, motifs present in the Hox genes of acoels are similar to those proposed for the ProtoHox genes by these authors. Altogether, these data strongly indicate that the Hox genes of acoels could correspond to the actual Hox genes of the LCBA.

The next three figures that I have introduced show the sequence alignments of conserved regions in the acoel Hox proteins, with those published sequences from other bilaterians. The homeodomain and the HP, N- and C- domains are shown schematically and shaded in different colours. Acoels possess the specific motives or signatures Na, Nb, Nbx and Ca (Ogishima and Tanaka, 2007), indicated by black lines. Asterisks indicate the end of the protein. Dots indicate the existence of no conserved sequences between adjacent domains. Hyphens indicate conserved positions in relation with the sequences from *S. roscoffensis*. All sequences are continuous; white gaps between amino acids have been inserted to facilitate seeing the domains in the alignment. The

species used for the alignment are: Bl, *Branchiostoma lanceolatum*; Cl, *Convolutriloba longifissura*; Cr, *Convolutriloba retrogemma*; Dm, *Drosophila melanogaster*; Ip, *Isodiametra pulchra*; Mm, *Mus musculus*; Nv, *Neanthes virens*; Sk, *Saccoglossus kowalevskii*; Sp, *Strongylocentrotus purpuratus*; Sr, *Symsagittifera roscoffensis*.

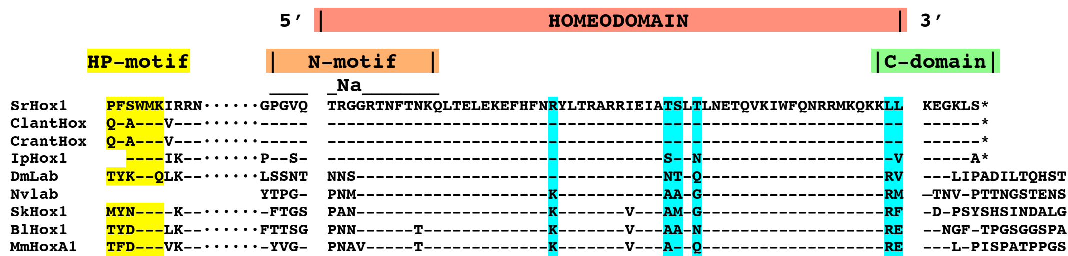


Figure 2. Sequence alignment of anterior Hox orthologues. Positions of phylogenetic interest are highlighted in blue: 24, 36, 37, 39, 59 and 60. In position 24, all acoels share an R, also featuring in *Drosophila*, whereas the rest of bilaterians have a K. The position 36-37 shows the amino acid pair TS in three acoel species and SS in the fourth, *I. pulchra*. The position 39 is occupied by a T in three acoel species, whereas there is a G, an N or a Q in the rest of bilaterians, though *I. pulchra* also has an N. The last two amino acids of the homeodomain (positions 59 and 60) also differ between acoels and other bilaterians. The HP motif (yellow) is well conserved and placed 39 positions before the homeodomain in *S. roscoffensis*.

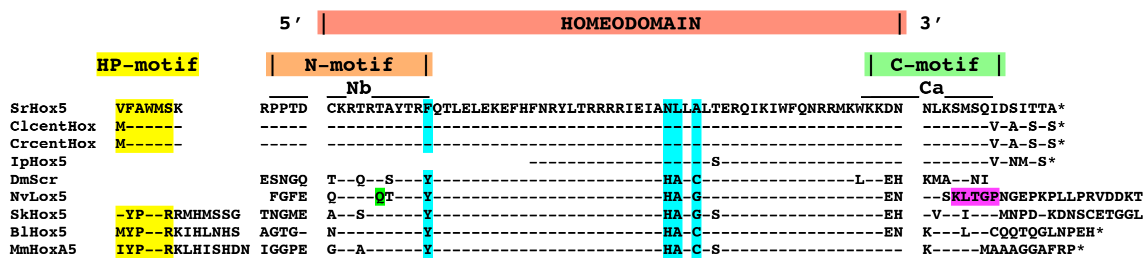


Figure 3. Sequence alignment of central Hox orthologues. Positions of phylogenetic interest are highlighted in blue: 11, 36, 37 and 39. In position 11 all acoels share an F (this region of the gene is unknown in *I. pulchra*), whereas the rest of animals there have a Y. The position 36-37 shows the amino acid pair NL in all acoel species, whereas is an HA pair in the rest of bilaterians. An A in all acoel species occupies the position 39, being a G or a C in the rest of bilaterians. The lophotrochozoan ‘‘lox5 parapeptide’’ found in all Lox5 genes is here shown for the Lox5 gene of *Neanthes virens* (shaded in purple). Lack of this signature in acoels is a further proof they not belong to the phylum Platyhelminthes. Moreover lophotrochozoan lox5 genes have a Q in position 6 (green), whereas Hox5 genes have a T. The HP motif (yellow) is well conserved and placed in acoels only 6 positions before the homeodomain.

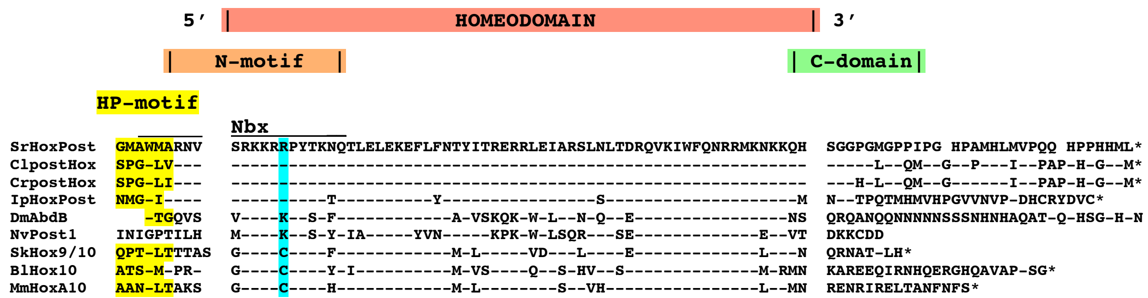


Figure 4. Sequence alignment of posterior Hox orthologues. The homeodomain of this gene is less conserved than the former two. The only position exclusive of acoels is the R6 (blue); other bilaterians having either a K or a C. The HP motif (yellow) is poorly conserved and placed only 3 positions before the homeodomain.

It is important to point out that after these analysis no significant similarities were found outside the homeodomain besides the N- and C- motives, and a few amino acids in the N-terminal region of all Hox proteins similar to the hexapeptide motif (HP) used in the Hox-PBX interactions. The hexapeptide is a six-amino acid-long motif lying upstream the homeodomain, necessary for PBX binding and, apart from the homeodomain itself, is the most characteristic feature of Hox genes. PBX is a Hox protein partner that acts by increasing the binding specificity of Hox proteins (Morgan *et al*., 2000). Interestingly, these motifs are located in acoels at the predicted distance from the homeodomain (In der Rieden *et al*., 2004), suggesting a very ancient Hox-PBX interaction. The appearance of a PBX-motif in the posterior Hox genes *Che9-14A* and *Che9-14C* of *Clytia* and in the *Nematostella Anthox1a*, which is located directly or close (a single amino acid) to the homeodomain, suggests that this interaction could be already present in the common ancestor of cnidarians and bilaterians.

The hexapeptide motif contains a conserved core sequence of four residues, YPWM, in all Hox protein classes except in the posterior class, which keeps only the tryptophan (Merabet *et al*., 2009). In acoels, this motif is also well conserved in the anterior and central Hox genes. However, instead of a D immediately before the WMK, as happens in vertebrates, the Hox1 proteins of acoels have an S in *S. roscoffensis* and *I. pulchra*, and an A in *C. longifissura* and *C. retrogemma* (Figure 2). As for Hox5, the amino acids just before the WMK are FA instead of YP, as it happens in vertebrates (Figure 3).

In addition, finding a PBX clone in the *S. roscoffensis* EST collection was of great importance to analyze the interaction Hox-PBX in acoels (data not published). ISH experiments revealed a match between the PBX expression pattern in juvenile specimens of *S. roscoffensis* and the neural system. This pattern overlaps with the *SrHox1* expression area, suggesting that both proteins may interact and that these genes could share a role in the specification of the nervous system (Figure 5), although functional analyses are needed to confirm this hypothesis.

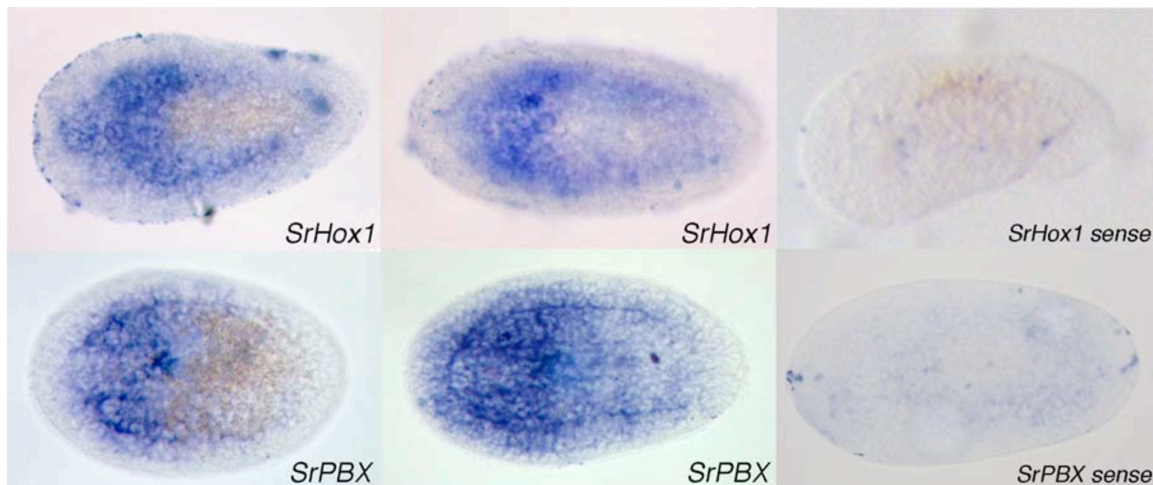


Figure 5. Comparative pictures of *SrHox1* expression patterns (up) and *SrPBX* (down). Both expression patterns overlap though *SrPBX* is more specific of the nervous system (see the labelled thin nerve tracts along the AP axis). All pictures show a dorsal view, anterior is to the left and posterior to the right.

Moreover it has been already described in zebra fish that PBX proteins interact primarily with PG1 and PG2 Hox genes in order to specify regional identities (rhombomere identities) throughout all but the most posterior hindbrain (Waskiewicz *et al*., 2002). In fact, it is both in mouse and zebra fish, that the interactions between PBX and Hox1 proteins specify directly the identity of rhombomere 4 via Hox1-PBX binding sites in the regulatory regions of target genes (Pöpperl *et al*., 1995; reviewed in Nonchev *et al*., 1997). If the role of the complex *Hox1-PBX* in the specification of the nervous system in acoels was confirmed, it might suggest a conserved role for this heterodimer associated to neural patterning in the whole Bilateria clade.

Finally, one homeodomain response HB-1 element has also been found around 10 kb before the *SrHoxPost* transcriptional start site, which would suggest auto or cross-

regulatory interactions with other homeobox genes. This evolutionary conserved element, containing a cluster of three homeodomain binding sites (named each one "HB1"), is present in the introns of Hox genes ranging from fish to humans as well as in the Ultrabithorax (Ubx) and Decapentaplegic (Dpp) genes of *Drosophila*. The HB1 element is also bound by *Drosophila* CAD homeoprotein and its homologue in the mouse, CDX-1; therefore, it is supposed to be a target for various homeodomain proteins of the Hox and ParaHox classes, in both vertebrates and invertebrates (Haerry and Gehring, 1997; Santini *et al*., 2003). The HB1 motif of *S. roscoffensis* is more similar to the vertebrate's HB1 than to the *Drosophila*'s ones (Figure 6). However, it is not placed within the introns of Hox genes but before the *SrHoxPost* transcriptional start site (predicted by the GENSCAN software). Unfortunately, the proteins that bind (and regulate) this gene through the HB1 site are still unknown. Oligo-binding chromatography with *S. roscoffensis* protein extracts would help in the future identification.

```
SrHoxPost     GTGTAATTACTAAC CCATAAACA TTTTATCG
HsHoxA4       CC-C-------CT-C--------TT -------A-C
MmHoxB4       ACA--------TCGC--------T  -------G-C
MmHoxA4       CA--------TAGC-C--------   ---A--T-CC
MmHoxA7       AACCC----A-TGGG--------A-G-------GAG
DfUbx         TCCAC----TC-G-GG---T-TT G--A-TGGTA
DmUbx         TCCAC----TC-G-GG---T-TT G--A-TATTGG
```

Figure 6. Alignment of HB1 elements from mouse (Mm), human (Hs), two *Drosophila* species (*D. melanogaster* and *D. funebris*), and *S. roscoffensis* gene *SrHoxPost*. HB1 of *S. roscoffensis* is more similar to the vertebrate's HB1 than to *Drosophila*'s ones. The HB1 element is placed in the 5' region before the *SrHoxPost* transcriptional start site according to the GENSCAN prediction. Conserved positions are highlighted in yellow.

2. HOX GENE EXPRESSION PATTERNS IN JUVENILE SPECIMENS OF *SYMSAGITTIFERA ROSCOFFENSIS*

The results presented so far give support to the hypothesis of a bilaterian ancestral HOX cluster made by three genes, orthologues to anterior, central and posterior classes. This small set would work as a first "Hox code" within the Bilateria, specifying regional identities along the main body axis (AP). To throw some light on the origin of the Hox patterning system in Bilateria and to explore how this patterning role could have been established over evolutionary time, the spatial domains of Hox gene expression in juvenile specimens of *S. roscoffensis* were analyzed by whole mount *in situ* hybridization.

SrHox1 is expressed in a broad anterior domain that includes the brain and areas of peripheral parenchyma surrounding the statocyst. Further down, *SrHox1* is also expressed, in two broad bands of peripheral parenchyma that end at about the beginning of the last third of the body length. *SrHox5* is expressed in two thin lateral stripes of peripheral parenchyma (perhaps including some neural structures) beginning at the level of the statocyst and going down to the last third of the body length. Finally, *SrHoxPost* is expressed in peripheral parenchyma areas of the posterior part of the body. No sharp borders were detected between the different Hox expression domains. The lack of specific markers labelling peripheral parenchyma precluded the definition of cell types that express the different Hox genes. However we can state that while these genes are all clearly active at the peripheral parenchyma layer, they are not expressed in the epidermis and in central (digestive) parenchyma (Moreno *et al*., 2009).

These results show the three Hox genes are expressed in nested domains along the AP axis in the juvenile worm, which is indicative of the presence of a Hox based vectorial patterning system, or "Hox code", in the acoel *S. roscoffensis*. This vectorial system may provide, as in other bilaterians, positional information along the major body axis of the animal (Figure 7). Interestingly, a recent study on *C. longifissura* has shown similar nested domains of expression patterns (Hejnol and Martindale, 2009).

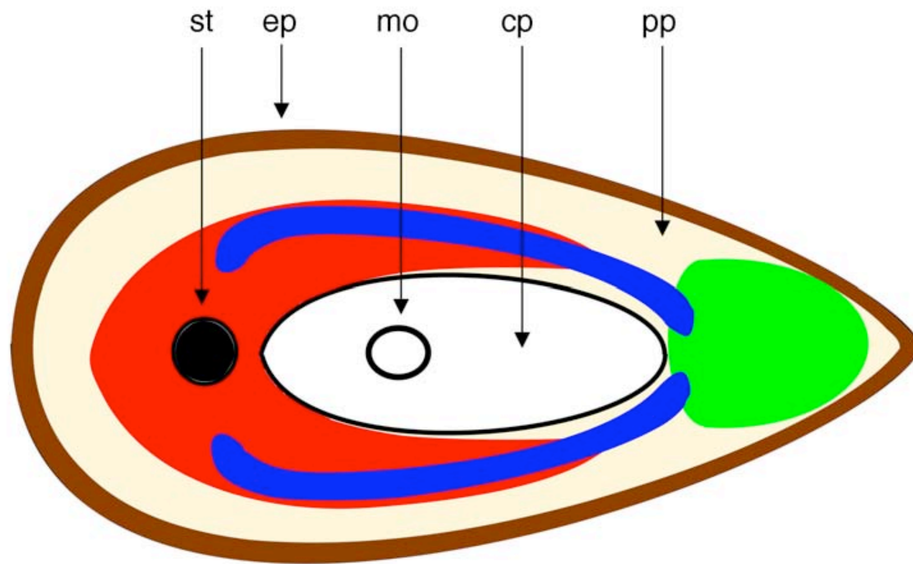


Figure 7. Diagrammatic dorsal view of the general morphology of one juvenile specimen of *S. roscoffensis* showing the nested domains of Hox gene expression. Colour code: *SrHox1*, red; *SrHox5*, blue; *SrHoxPost*, green. Head is on the left and tail on the right. cp, central parenchyma or gut; ep, epidermis; mo, mouth opening; pp, peripheral parenchyma; st, statocyst. For further information, see Moreno *et al*., 2009, Fig. 4.

Given the phylogenetic position of the acoels, this would represent the first example of collinear expression of Hox genes along the AP body axis for any metazoan; indicating that the spatial collinearity of Hox gene expression might have been present already in the LCBA. The set of three Hox genes found in the four acoel species so far analyzed and its nested expression along the AP axis in *S. roscoffensis and C. longifissura* could be considered a good proxy to the ancestral number and mode of expression in the LCBA. During bilaterian evolution, and despite the expansion of the HOX cluster, the system that regulates its spatial collinearity was maintained.

Its important to point out that what we have found in *S. roscoffensis* is another example of the so-called transcollinearity, i.e.; keeping the correct sequence of nested expression domains along the axis, correlated to their paralogues groups, but in the absence of a proper genomic clustering (Duboule, 2007). This represents yet another proof that the presence of a nested set of expression domains does not require the presence of an intact cluster (Aboobaker and Blaxter, 2003; Cameron *et al*., 2006; Ikuta *et al*., 2004; Seo *et al*., 2004).

As regards temporal Hox expression in the acoels, Hejnol and Martindale (2009) have shown in *Convolutriloba longifissura* that Hox expression starts simultaneously after gastrulation showing no temporal collinearity. As stated by some authors, temporal collinearity seems associated with functionally intact HOX clusters (Monteiro and Ferrier, 2006). Since recent evidence shows that temporal and spatial patterning of Hox expression is controlled by distinct mechanisms (Soshnikova and Duboule, 2009; Tarchini and Duboule, 2006; Tschopp *et al*., 2009), this could explain why spatial but not temporal collinearity is maintained in acoels despite cluster disintegration.

3. NUCLEOTIDE SEQUENCE AND EXPRESSION PATTERN OF THE CDX PARAHOX GENE IN *SYMSAGITTIFERA ROSCOFFENSIS*

The Caudal gene homologue in *S. roscoffensis*, *SrCdx*, was cloned by PCR using degenerated primers against conserved peptides within the homeobox (Hox) region. Subsequently, it was amplified by means of the RACE methodology. The sequence of *SrCdx* is very similar to the Cdx gene identified in *Convolutriloba longifissura* (Hejnol and Martindale, 2008b), the only Cdx gene recovered from acoels so far. Sequence alignment of Caudal/Cdx homologues from acoels and other bilaterians shows that the main conserved region of the protein is again the homeodomain. However, when Cdx from the cnidarians *Clytia hemisphaerica* (*CheCdx*) and *Eleutheria dichotoma* (*Cnox-4 Ed*) are included, it is fairly evident that cnidarian and bilaterian Cdx genes are quite divergent (Figure 8).

Expression of Cdx in both protostomes (*Drosophila*, Wu and Lengyel, 1998) and deuterostomes (vertebrates, James *et al*., 1994; and amphioxus, Brooke *et al*., 1998), suggests a conserved role in the specification of hindgut and anal structures. This functional conservation indicates that Cdx was likely involved in patterning the posterior gut and anal regions in the Urbilateria (de Rosa *et al*., 2005).

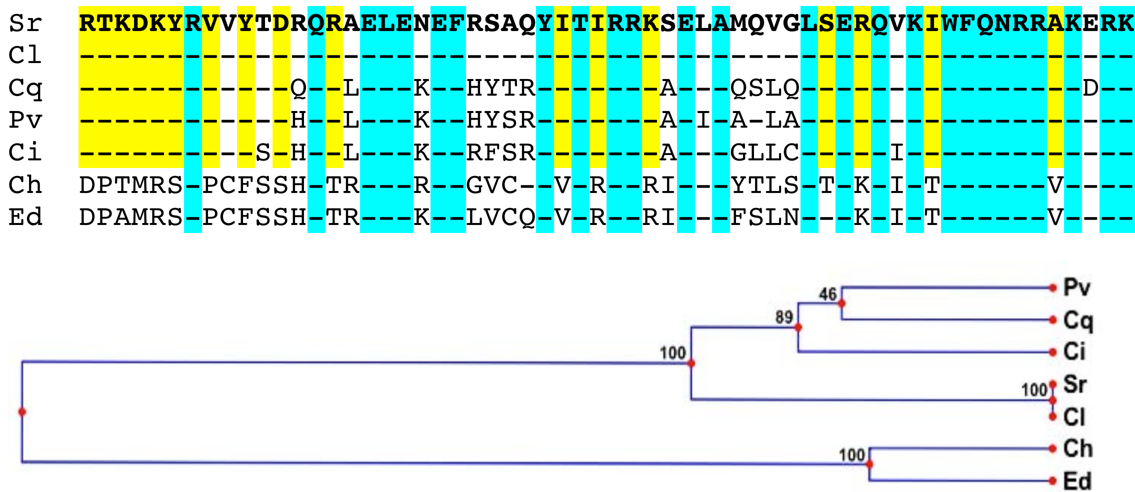


Figure 8. Cdx homeobox alignment (top) and the corresponding phylogenetic tree built with the software CLC Sequence Viewer 6.3 by UPGMA with 100 replicates for the bootstrap analysis (bottom). Conserved amino acids between all species are shaded in blue. Conserved amino acids between bilaterian species but not in cnidarians are shaded in yellow. Notice that Cdx proteins among cnidarian are very divergent whereas among bilaterians are well conserved. Ci, *Ciona intestinales*; Ch, *Clytia hemisphaerica*; Cl, *Convolutriloba longifissura*; Cq, *Culex quinquefasciatus*; Pv, *Patella vulgata*; Sr, *Symsagittifera roscoffensis*.

In the acoel *C. longifissura*, *ClCdx* is expressed in the most posterior ectodermal area of adult animals, a region that will form the male gonopore. However, acoels lack an anal opening. Therefore, if a through gut was already present in the LCBA, the anus might have been lost in the Acoelomorpha lineage. According to Hejnol and Martindale (2008b) because the metazoan mouth evolved first, it is more parsimonious to consider that the anal opening arose independently in different groups by co-opting hindgut genes in posterior domains at the ectodermal–endodermal boundary.

Our (limited) data from the *SrCdx* expression patterns does not help to choose between both alternatives. In *S. roscoffensis*, expression around the male gonopore is evident only in adults but not in juveniles (Figure 9, arrow). Besides this posterior expression domain, *SrCdx* is also expressed in two rows of cells that run along the AP axis in the peripheral parenchyma of the adult. Interestingly, a third domain of expression seems to correspond to the central (digestive) parenchyma, though this assertion needs further proof (Figure 9, C). In juveniles, the pattern is clearly different from the adults, and similar to patterns described in juveniles of *C. longifissura* (Hejnol

and Martindale, 2008b, 2009). Strikingly here *SrCdx* is expressed within the nervous system. The domain of expression contains some commissures located around the statocyst and it extends down the nerve cords along the whole AP body axis (Figure 9, A & B, arrowheads). These results indicate that while Cdx plays a role in the specification of the nerve cords during the post-embryonic development, it plays a different role during adulthood, regulating the formation of the male gonopore region and probably the gut/central parenchyma region.

Expression patterns of the Cdx homologue in cnidarians do not clarify how the ancestral expression pattern of this gene would actually look like. This is because the pattern is highly variable among different species. In the anthozoan *Nematostella vectensis*, the Xlox/Cdx paralogue is expressed in the early planula stages in two thin endodermal stripes along the ventral midline (Ryan *et al*., 2007), but in the hydrozoans *Eleutheria dichotoma* (Kamm *et al*., 2006) and *Clytia hemisphaerica* (Chiori *et al*., 2009) the expression is restricted only to the ectodermal layer. More, in the planula larvae of *C. hemisphaerica* it is expressed at both the oral and aboral poles, and in the polyp of *E. dichotoma* expression is solely restricted to the aboral pole. This plasticity calls for further analysis before suggesting commonalities in the Cdx activities.

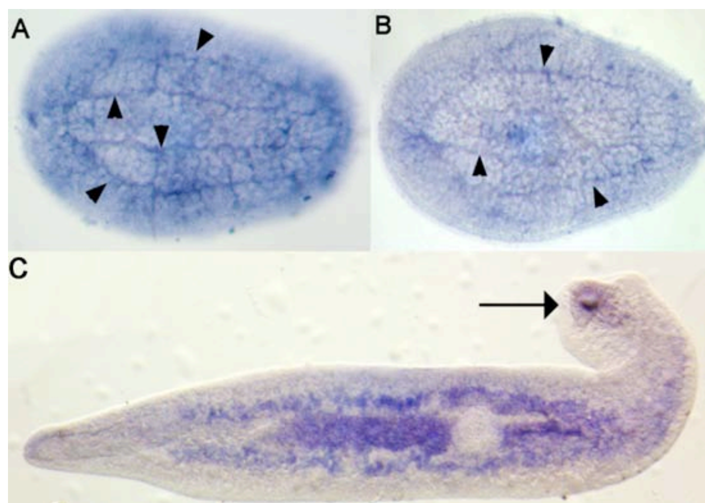


Figure 9. *SrCdx* expression patterns in juveniles (**A** and **B**) and adult specimens (**C**) of *S. roscoffensis*. In juveniles, *SrCdx* is mainly expressed in the nervous system. From the commissures around the statocyst it labels thin nerve tracts running in parallel along the AP body axis (arrowheads). In adults, expression is found around the male gonopore (arrows), in two rows of cells of peripheral parenchyma running along the AP axis, and in the central parenchyma. All pictures show a dorsal view, with anterior placed to the left.

4. NUMBER, TYPES, ARRANGEMENT, AND EXPRESSION OF HOX AND PARAHOX GENE COMPLEMENTS IN CNIDARIA

The next step to track the origin of the bilaterian Hox-ParaHox patterning system and to infer the Hox-ParaHox complement in the C-BLCA and its evolutionary history from this ancestor to the LCBA, is to analyze the number, types and functions of these genes in the sister group of the Bilateria, i.e.; the Cnidaria. Despite several studies have already been published (Chiori *et al*., 2009; Chourrout *et al*., 2006; Kamm *et al*., 2006; Quiquand *et al*., 2009; Ryan *et al*., 2007) there is no agreement on the main issues. In other words, the number and type of genes supposed to make the so-called ProtoHox cluster (from which HOX and ParaHox clusters arose in the C-BLCA) remain nowadays unknown. In fact, a major difficulty in analyzing Hox genes in the Cnidaria is that it remains very difficult to trace the affinities of them with any other bilaterian relatives (see below).

Three features of the Hox-ParaHox genes in cnidarians must be considered to determine whether these organisms bear a true HOX-ParaHox cluster acting as an axial vectorial patterning system:

1. The number of Hox-ParaHox genes and its orthologies to their bilaterian counterparts.

2. The genomic arrangement of these genes, looking for conserved clusters.

3. The expression patterns of these genes, looking for a Hox code.

1. The number of Hox-ParaHox genes in cnidarians and its assigned orthologies to bilaterians vary among different studies, mainly due to the different methodologies used for phylogenetic reconstruction. Moreover, alignable Hox-ParaHox sequences are too short and have very little phylogenetic signal to establish clearly these orthology assignments. Hox-ParaHox related genes found in cnidarians and its orthologies according to different authors are shown in the Table 1.

Most reports concur that true Hox genes do occur in cnidarians; hence, the Hox-ParaHox family was already present and diversified in the C-BLCA (Chiori *et al*., 2009; Chourrout *et al*., 2006; Kamm *et al*., 2006; Quiquand *et al*., 2009; Ryan *et al*., 2007). The best agreement is on the presence of cnidarian genes orthologues to the anterior Hox and ParaHox classes (e.g. Gsx) (Chiori *et al*., 2009; Chourrout *et al*., 2006; Finnerty *et al*., 2004; Gauchat *et al*., 2000; Kamm *et al*., 2006; Quiquand *et al*., 2009; Ryan *et al*., 2007). In other words, the C-BLCA already did have ''anterior'' Hox (PG1 and PG2) and ParaHox (Gsx) genes. But in contrast, neither group 3 genes nor central Hox genes have already been found, despite the recent report of two Xlox (central ParaHox) related genes in hydrozoans and the new proposal (Quiquand *et al*., 2009) that the ParaHox gene *NVHD065* from *N. vectensis,* previously classified as an intermediate gene between Xlox and Cdx, is a bona fide actual Xlox gene.

Finally, the existence of ''posterior'' genes in cnidarians is even more controversial issue. Some authors are convinced of their presence (Finnerty *et al*., 2004; Gauchat *et al*., 2000; Ryan *et al*., 2007), but others do not (Chourrout *et al*., 2006; Kamm *et al*., 2006). In a recent analysis, Chiori *et al*. (2007) have suggested that the Hox9-14 group of Hox genes in metazoans can be organised in four sub-groups: one bilaterian and three cnidarian sub-groups (A, B and C), the last two branching paraphyletically with respect the bilaterian ''posterior'' Hox genes, and forming what the authors call a cnidarian ''posterior'' Hox group. In sheer contrast, Ryan *et al*., (2007) consider the posterior Hox genes in Cnidaria fully homologous to its bilaterian counterparts whereas, in a very recent reanalysis of this problematic issue done by Quiquand *et al*. (2009) leads them to state that the relationships between these genes in cnidarians and bilaterians are still unclear. Hejnol and Martindale (2009) give special importance to the HP motif present in the posterior Hox genes *Che9-14A* and *Che9-14C* of *Clytia* and *Anthox1a* of *Nematostella*. Such motif is directly adjacent or very close (a single amino acid) to the 3' end of the homeodomain and is similar to the HP motif present in the posterior Hox genes of bilaterians. On this base, these authors favour the homology between cnidarian and bilaterian posterior Hox genes.

	PG1	PG2	PG3	PG9/13	Gsx	Xlox	Cdx
1	*Anthox6* *Anthox6a*	*Anthox7* *Anthox8aA* *Anthox8b*		*Anthox1* *Anthox1a*	*Anthox2*		*NVHD065*
2	*Anthox6* *Anthox6a*		*Anthox7* *Anthox8a* *Anthox8b*	*Anthox1* *Anthox1a*	*Anthox2*	*NVHD065*	
3	*Anthox6* *Anthox6a*	*Anthox7* *Anthox8a* *Anthox8b*		*Anthox1* *Anthox1a*	*Anthox2*	*NVHD065*	
4	*CheHox1* *Anthox6* **among others**	*Anthox6a* *Anthox7* *Anthox8a* *Anthox8b*		**Cnidarian posterior subgroups A, B and C**	*CheGsx* *Anthox2* **among others**		*CheCdx* **among others**
5	*Anthox6* *Anthox6a* *Anthox6Ms* *Cnox4 Hm* *Cnox1 Cv* *Cnox5 Ed* *Cnox1 Pc*	*Anthox7* *Anthox8a* *Anthox8b*		**"PG9/Cdx"** *Anthox1* *Anthox1a* *Cnox4 Pc* *HoxB Hm* *HoxC2 Hm* *HoxC3 Hm* *CnoxC Ch* *Cnox1 Ed* *Cnox3 Ed* *CnoxD Hm* *Scox3 Cx* *Scox1 Cx* *Scox4 Cx*	*Anthox2* *Cnox2 Am* *Cnox2 Cv* *Cnox2 Hv* *Cnox2 Ed* *Cnox2 Ssp* *Gsx Pc* *Cnox2 Hys* *Scox2 Cx*	*Pdx Td* *Pdx Ch* *NVHD065*	*Cnox4 Ed*

Table 1. Hox and ParaHox gene orthologies between cnidarians and bilaterians according to different authors: 1. Kamm *et al*., 2006; 2. Chourrout *et al*., 2006; 3. Ryan *et al*., 2007; 4. Chiori *et al*., 2009; 5. Quiquand *et al*., 2009. *Nematostella vectensis* Hox and Hox-related genes are here named according to the consensus nomenclature taken from the first column of table 1 in Ryan *et al*., 2007. Genes from other cnidarian species have been named using the nomenclature from the table 1 in Quiquand *et al*., 2009; and from Chiori *et al*., 2009. Colour code: green, clear orthologies with bilaterian genes; yellow, uncertain orthologies with bilaterian genes; orange, no orthologies with bilaterian genes (genes originated in Cnidaria by independent internal duplications); white, no orthologues found. Genes placed in fused cells are those which could not be distinguished between the two corresponding paralogues groups in phylogenetic analyses.

In 1998, Brooke and colleagues made the proposal that a four-gene ProtoHox cluster (formed by one anterior, one group 3, one central, and one posterior ProtoHox genes, the result of a series of ancient tandem duplications of a first Hox gene) underwent a duplication giving rise to the HOX and ParaHox sister clusters. Later on, the resulting clusters had independent gene tandem duplications which resulted into the great variety of extant HOX and ParaHox clusters (Chourrout *et al*., 2006; García-Fernández, 2005a).

However, most phylogenetic analyses of the Hox-ParaHox genes in cnidarians (Chiori *et al*., 2009; Hejnol and Martindale, 2009; Ryan *et al*., 2007) question this hypothesis. Brooke's proposal was founded on molecular phylogenetic analysis using sequences from amphioxus, which suggested that the ParaHox gene Gsx was most closely related to Hox1-2, Xlox was related to Hox3, and Cdx to the posterior Hox. Although this scenario was assumed as very plausible, new phylogenetic analysis by Chiori *et al*. (2009) indicated that ParaHox genes in cnidarians classified as members of the ''anterior'' (Gsx) or ''posterior'' (Cdx) classes were not phylogenetically related to the so-called ''anterior'' or ''posterior'' Hox groups of bilaterians. Instead, Gsx and Cdx were found to be sister groups. Following a different set of analysis, and according to Ryan *et al*. (2007) Gsx forms an independent lineage from the anterior Hox genes, the posterior Hox genes and Cdx do not form a monophyletic group, and the mean statistical support for a sister group relationship between Xlox and Hox3 is not very high.

In contrast to the Brooke and collaborator's model, Ryan and collaborators suggest a different scenario in which a single 'Proto-Hox' gene gives rise first to one Hox and one ParaHox gene, the ancestral Hox1 and Gsx orthologues. Later on, tandem duplications independently extended the Hox and ParaHox gene complements in the C-BLCA, which was followed by a translocation event that set apart, in the genome, the ParaHox and the HOX clusters. Under this scenario, the ParaHox and the HOX cluster are not sister clusters. Finally, in cnidarians and bilaterians there were lineage specific gene duplications, deletions and rearrangements, giving rise to the current HOX and ParaHox clusters in the extant groups of animals.

Thus, since many authors with contradictory results have used different phylogenetic approaches, data sets, and evolutionary methods, the number and classes of Hox-ParaHox genes present in the C-BLCA are still clearly undefined. More data, likely coming from current whole genome sequencing from cnidarian groups and more bilaterians will be needed to clarify this rather confusing situation.

2. As today, the only example of genomic linkage between Hox genes in Cnidaria is found in the *Nematostella* genome, where anterior genes are linked on a single genomic scaffold and a Cdx/Xlox and Gsx genes are linked in another (Ryan *et al*., 2007). In *Hydra*, molecular analyses has not uncovered any clustering between PG1 and PG9-like genes over a 250 kb scale (Gauchat *et al*., 2000) while in the hydromedusae *Eleutheria dichotoma* Hox genes are physically dispersed in the genome (Kamm *et al*., 2006). No further information is currently available for any other cnidarian species.

3. If cnidarians were using a "Hox code" as a vectorial positional system along any body axis, they would be expressing different Hox genes in distinct domains along this particular axis. Ryan *et al*. (2007) postulated the presence of such a Hox code from the expression patterns of two Hox genes acting along the OA axis in larvae and polyps of *N. vectensis*. *Anthox6* (an orthologue of the bilaterian Hox1) was expressed at the oral pole, whereas *Anthox1* (a putative orthologue of bilaterian posterior Hox genes) labelled the opposite, aboral pole. Other five *Anthox* genes were found, however, they were widely expressed in the endodermic wall, along the whole OA axis, a fact that was contrary to the presence of a trans-collinearity mechanism. Besides, in a surprising result it was shown that the *Anthox6* orthologues of the hydrozoans *Podocoryne* (*PcaCnox1*) and *Eleutheria* (*EdiCnox5*) were expressed in opposite sides to those of *N. vectensis* in their planula larvae (Kamm *et al*., 2006; Yanze *et al*., 2001). Furthermore, and contrary to Ryan *et al*., (2007), recent analyses of Hox gene expression patterns in the planula larvae of *Clytia* (Chiori *et al*., 2009) clearly indicates that expression of these Hox genes is not trans-collinear.

Three main sets of arguments go against the existence of a Hox code in cnidarians. First, orthologue genes are often not expressed at the same stage of the life cycle. For example, the *Clytia* gene *CheHox9-14A* is expressed in the oral pole of the planula larvae whereas *EdiCnox3*, its counterpart in *Eleutheria*, is only expressed in the medusa stage (Chiori *et al*., 2009; Kamm *et al*., 2006). Second, even when orthologue genes are expressed at the same stage, their transcripts are often localised at the opposite ends of the OA axis. For instance, *Clytia CheHox9-14B* and its *Podocoryne* orthologue *PcaCnox4* are expressed in the oral pole (Chiori *et al*., 2009; Yanze *et al*., 2001), but

their *Nematostella* relative, *NveAnthox1*, has an aboral expression domain (Finnerty *et al*., 2004; Ryan *et al*., 2007). Finally, orthologue genes having similar expression domains often label different tissues: *PcaCnox1* is expressed in both the endoderm and the ectoderm of the planula larva (Yanze *et al*., 2001), while *EdiCnox5* is restricted to the ectoderm (Kamm *et al*., 2006).

Summing up, collinearity of Hox gene expression does not seem to be a property of Hox genes in cnidarians and their expression domains along the OA axis are not conserved. Altogether, this strongly argues against a conserved role for HOX cluster genes to pattern the OA axis in Cnidaria as it does in the AP patterning Hox system in Bilateria (Ball *et al*., 2007). Therefore, the role of Hox genes in patterning the main body axis was very likely an innovation of that came with the origin of bilateral animals.

The differences above discussed between the use of the Hox group of genes, as such, in cnidarians and bilaterians (Ball *et al*., 2007; Chiori *et al*., 2009) suggests a very important difference in the construction of animals such as the LCBA (using a Hox code) and the cnidarians (bearing Hox genes but not an integrated Hox code). In consequence, if the RBT hypothesis is correct and the C-BLCA was a radial animal, the appearance of the bilateral symmetry in Bilateria could be causally related to the emergence of a very simple (three to four genes) Hox based vectorial patterning system in the lineage leading to the LCBA (Figure 10.1). The set of three Hox genes found in four acoel species and its nested domains of expression along the AP axis in *S. roscoffensis* (and *C. longifissura*) could be considered a good proxy to the ancestral number and mode of expression of Hox genes in the LCBA.

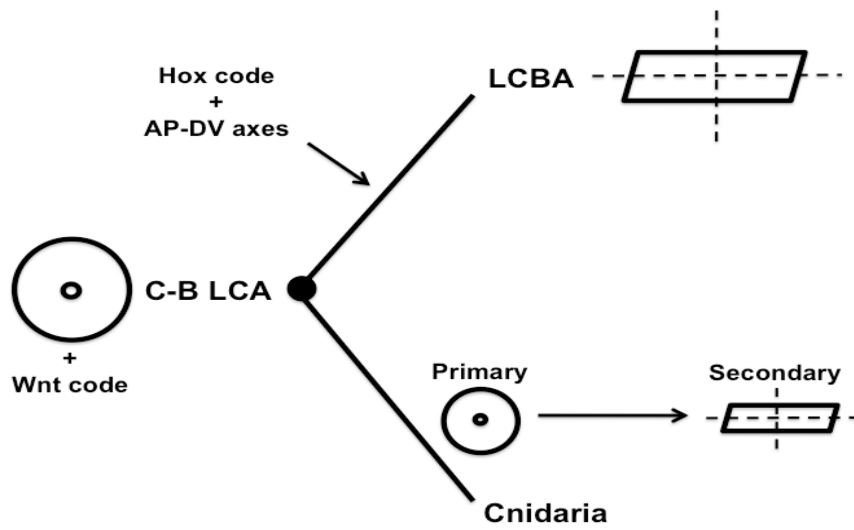
Two other alternatives could also be contemplated, though. If we take the so-called Turbellarian theory, the C-BLCA would be a bilateral animal, which could bear a simple Hox system similar to that proposed here for the LCBA. Cnidarians would then have lost bilateral symmetry secondarily due to its newly acquired sessile life style. In parallel, their Hox system would have disintegrated due to the release of selective pressures to maintain a collinear expression along the AP axis in the new radial animal

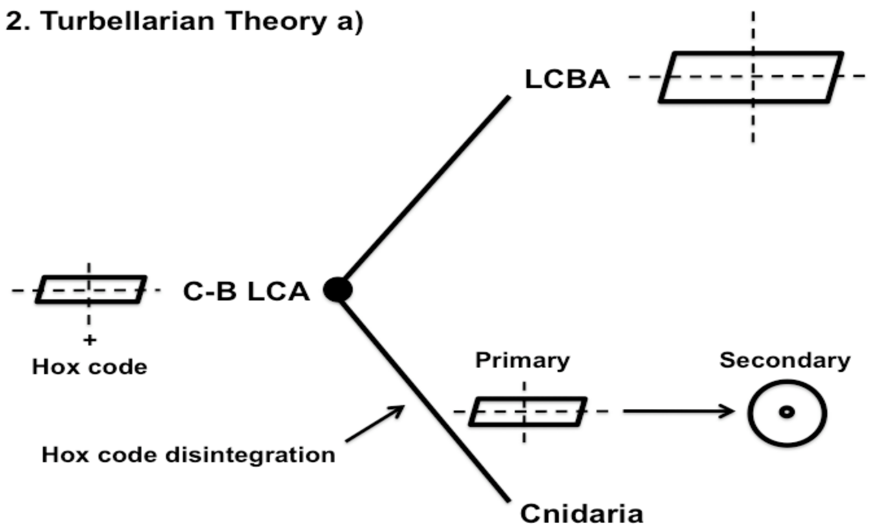
(Figure 10.2). However, we think it is more parsimonious to assume that cnidarians never have possessed a Hox system but used a different patterning mechanism, perhaps a Wnt system, for organizing the OA axis (see General Discussion section 7). Hox genes were co-opted for the AP axis patterning in the lineage leading to the LCBA, and after the split from cnidarians (Figure 10.3).

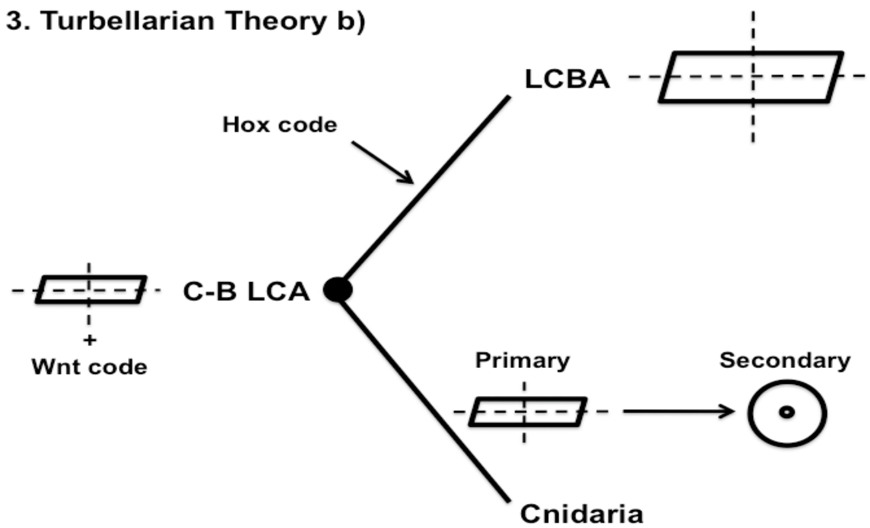
As a final addendum, it is pertinent to mention that the expression domains of Hox-ParaHox along the OA axis in cnidarian planula larvae and polyps are extremely plastic, despite its very similar body plan. This situation is very different to that present in bilaterians where Hox gene expression along the AP axis is conserved despite a high disparity in body plan structure.

Next page

Figure 10. Three possible scenarios for the evolution of the Hox gene axial patterning system and bilateral symmetry in metazoans. **1.** Radial-Bilateral Transition scenario. The C-BLCA was a radially symmetrical organism with a Wnt code patterning the OA body axis. From this ancestor, bilateral symmetry evolved in the lineage leading the LCBA, which implied co-option of Hox genes to pattern the AP axis and development of a Hox code. In turn, cnidarians remained symmetrically radial though some bilateral structures secondarily evolved. **2.** Turbellarian theory, scenario a) The C-BLCA already was a bilateral organism bearing a Hox code which patterned the AP body axis. In the cnidarian lineage, the Hox code disintegrated due to the adoption of a sessile life style and radial symmetry was subsequently developed in the last common cnidarian ancestor. In this case, bilateral structures in cnidarians would be remnant structures inherited from its bilateral ancestor. **3.** Turbellarian theory, scenario b) The C-BLCA already was a bilateral organism bearing a Wnt code patterning the AP body axis. Hox genes were then co-opted for the AP axial patterning system in the Bilateria lineage and a Hox code subsequently evolved. Change from a free-living to a sessile life style in the cnidarian lineage produced a shift from bilateral to radial symmetry in the lineage leading to the last common cnidarian ancestor. Concentric circles represent radial body symmetry. A rectangle with two orthogonal lines represents bilateral body symmetry.


1. Radial-Bilateral Transition
Hox code
+
AP-DV axes
LCBA
C-B LCA
+
Wnt code
Primary
Secondary
Cnidaria


2. Turbellarian Theory a)
LCBA
C-B LCA
+
Hox code
Primary
Secondary
Hox code disintegration
Cnidaria


3. Turbellarian Theory b)
LCBA
Hox code
C-B LCA
+
Wnt code
Primary
Secondary
Cnidaria

5. SCENARIOS FOR HOX AND PARAHOX CLUSTER ORIGIN AND EVOLUTION WITHIN THE METAZOA

Different models have been proposed for the origin and evolution of Hox-ParaHox genes and clusters. According to Larroux *et al*. (2007) the origin of the whole Hox-ParaHox system was a unique ANTP gene in the lineage leading to the C-BLCA, after they split from sponges (Figure 11). This ProtoHox gene duplicated in tandem several times giving rise to the ancestral gene cluster present in the C-BLCA. However, it is important to point out that a new phylogenetic analysis have suggested the presence in the ancestor of the sponge class Demospongiae of at least one Hox-ParaHox or EgHBox gene (an ANTP-class homeobox gene proposed to be ancestrally related to the Hox genes). This gene (or perhaps several genes), from which Hox and ParaHox genes in Eumetazoa originated would have been lost in the Demospongiae lineage (Peterson and Sperling, 2007).

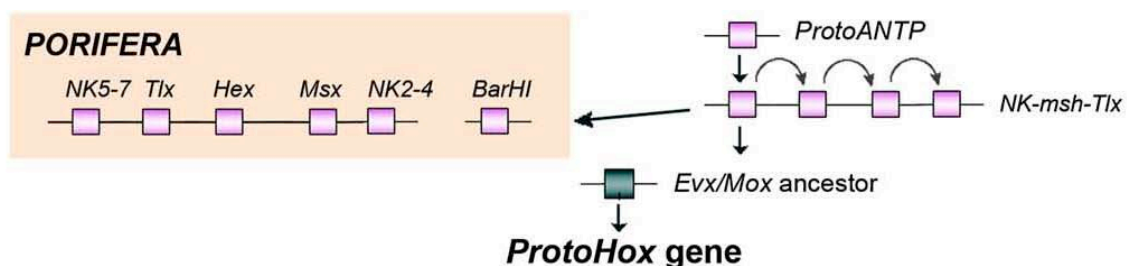


Figure 11. Origin of the ProtoHox gene within the metazoans. Given the absence of Hox and ParaHox genes in Porifera, ProtoHox genes likely appeared after Porifera diverged as a result of a cis-duplication event of a non-Hox ANTP-class gene, possibly an Evx/Mox ancestor gene. Extracted from Quiquand *et al*., 2009.

Assuming a ProtoHox gene cluster arising from a single ANTP Hox gene, two are the main theories on the evolutionary path from the ANTP ProtoHox gene to the HOX and ParaHox clusters: the Multigene Duplication model (Brooke *et al*., 1998; Chourrout *et al*., 2006; García-Fernández, 2005a) and the Tandem Duplication model (Chiori *et al*., 2009; Ryan *et al*., 2007), which have been explained above (see Figure 12).

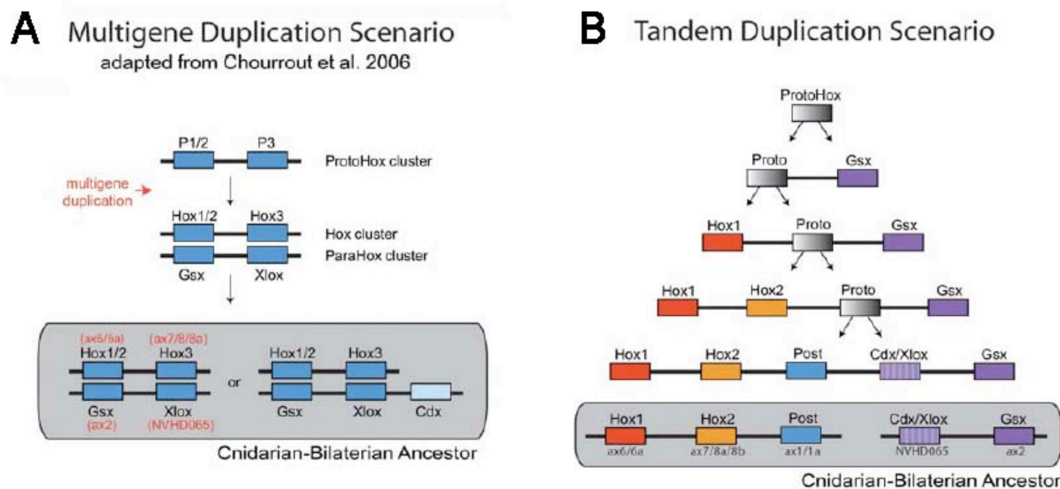


Figure 12. Multigene (or Segmental) Duplication (**A**) and Tandem Duplication (**B**) scenarios to explain the origin and evolution of HOX and ParaHox gene clusters at the C-BLCA. Modified from Ryan *et al*., 2007. **A.** Series of tandem duplications from an ancestral gene originated an ancestral ProtoHox cluster (Lewis, 1978; Gehring *et al*., 1994b; Schubert *et al*., 1993; Zhang and Nei, 1996; Brooke *et al*., 1998), that later on had a segmental duplication, and further up in evolution split to originate the HOX and ParaHox clusters (García-Fernández, 2005a; García-Fernández, 2005b). Under this scenario, HOX and ParaHox are sister clusters, being Gsx most closely related to Hox1-2, Xlox to Hox3, and Cdx to Hox9-14. **B.** HOX and ParaHox clusters originated by tandem duplications and finally split and shifted to different genome regions by a translocation event. The ParaHox genes would represent detached Hox genes rather than sister cluster genes (Chiori *et al*., 2009; Hejnol and Martindale, 2009; Ryan *et al*., 2007).

Irrespective of which mechanism produced these genes, the most likely scenario suggests the presence in the C-BLCA of anterior Hox genes, Gsx, Xlox, Cdx and proto-posterior Hox genes (Figure 12). Because orthologies between posterior genes in cnidarian and bilaterians are still not clear, it is fair to suppose that cnidarian and bilaterian posterior Hox genes were produced independently in both lineages from the proto-posterior Hox gene present in the C-BLCA.

According to current phylogenetic studies, we would suggest a model, based on the Tandem Duplication model that would explain the origin of the Hox-ParaHox complement in the C-BLCA:

ProtoHox gene

↓ *tandem duplications*

PG1-PG2-Gsx-Xlox-Cdx-ProtoPostHox

↓ *detachment of the ParaHox genes*

C-BLCA ⇒ **PG1-PG2-ProtoPostHox** + **Gsx-Xlox-Cdx**

The next step is model the change from the Hox-ParaHox complement in the C-BLCA to the complement in the LCBA. To that end, acoelomorph flatworms may hold the key.

Assuming the phylogenetic position of Acoelomorpha at the base of the Bilateria and including the data reported on the complement of Hox genes in the nemertodermatid flatworms (Jiménez-Guri *et al*., 2006), the sister group of the acoels, the LCBA had a ParaHox cluster with the 3 genes (Gsx, Xlox and Cdx) and a simple HOX cluster made by four (PG1-PG2-PG5-PostHox) or five genes (PG1-PG2-PG3-PG5-PostHox). If we assume that Brooke *et al*. (1998) model is true, meaning that Xlox and PG3 Hox are paralogues; the number of Hox genes in the LCBA should be five (PG1-PG2-PG3-PG5-PostHox). It is important to note, though, that PG3 has yet to be found in acoelomorphs, though it is assumed to be present in the LCBA because Xlox has been recovered in nemertodermatids and tentatively also in cnidarians. On the contrary, if the gene tandem duplication model holds true the number of Hox genes in the LCBA would be then four (PG1-PG2-PG5-PostHox). This is because under this scenario Xlox and PG3 are unrelated and therefore it is not necessary to assume the presence of a PG3 gene based on the existence of an Xlox relative. Since PG3 has not been found either in cnidarians or in acoelomorphs, a single origin for this gene in the P-DLCA seems more parsimonious than two independent losses in cnidarians and acoelomorphs. In summary, in our view, the most likely number of Hox genes in the LCBA (the Urbilateria) is four (Figure 13).

Independently of whether the ancestral HOX cluster did have four or five genes, the expression patterns of Hox genes in *S. roscoffensis* suggests the presence of a Hox based vectorial patterning system for the AP axis at the LCBA. In turn, the lack of a

Hox based vectorial patterning system in cnidarian suggests that collinearity along the main body axis was a bilaterian innovation (Kamm *et al*., 2006), possibly linked to novel requirements placed upon by directional locomotion or by the need of the antero-posterior patterning of the central nervous system (Deutsch and Le Guyader, 1998). In addition, the central Hox class represents another innovation of bilaterian animals; since genes homologous to central Hox genes have been found only in Bilateria and not in cnidarians.

The Cdx ParaHox, cloned from *S. roscoffensis* and *C. longifissura*, is the only ParaHox gene reported from the acoels. Because nemertodermatid flatworms and, tentatively, the cnidarians, bear a Xlox ParaHox, this suggest the loss of this gene in acoels or the presence of highly divergent homeodomain sequences which could not be recovered so far by PCR screens. Xlox genes are involved in midgut patterning in bilaterians as well as in the epithelial digestive system in general. Transformation of a true epithelial gut into a syncitial digestive system in acoels (Smith and Tyler, 1985) might explain the loss or modification of the Xlox from acoel genomes, though this could be still a simplistic scenario. Finally, the absence of a Gsx ParaHox and anterior Hox genes in nemertodermatids is most likely due to limited sampling (so far, a single screening having been carried out and in a single species), although it could also be due to gene loss in this lineage. Identical arguments apply to the PG2 and Gsx absences in acoels.

All in all, the phylogenetic position of acoelomorphs at the base of the bilaterian tree has allowed us to have a glimpse on how the Hox-ParaHox set could have been at the LCBA, the intermediate ancestor between the C-BLCA and the more complex P-DLCA (which is considered identical to the LCBA by the supporters of the Archicoelomate theory). Further up in the evolution of the Bilateria, the basic complement of Hox-ParaHox genes present in the LCBA duplicated in tandem giving rise first to the PG3, and later on to the five central paralogues groups (PG4-8) and, in the deuterostomate lineage to five posterior paralogues groups (PG9-13) (Figure 13). This path is summarized in the following diagram:

LCBA ⇒ **PG1-PG2-PG5-PostHox + Gsx-Xlox-Cdx**

↓ *Tandem duplications of Hox genes*

P-DLCA ⇒ **PG1-PG2-PG3-PG4-PG5-PG6/8-PostHox + Gsx-Xlox-Cdx**

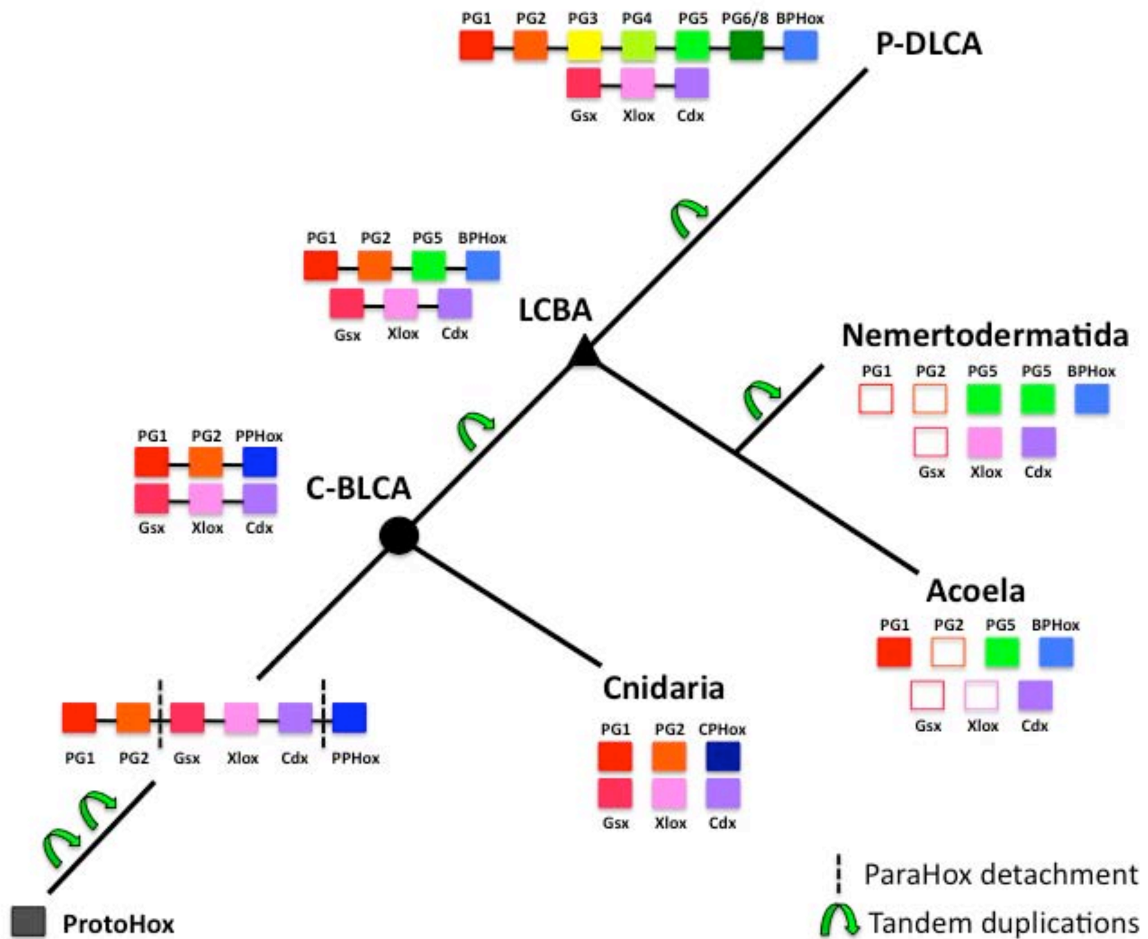


Figure 13. Evolution of HOX and ParaHox clusters in metazoans from a single ProtoHox cluster after the divergence of Porifera. The model integrates data from recent phylogenetic studies in Cnidaria, and tries to combine the most parsimonious hypothesis for Hox-ParaHox evolution. A unique ANTP gen in the lineage leading to the cnidarian-bilaterian last common ancestor (C-BLCA) duplicated several times in tandem giving rise to an ancestral HOX-ParaHox gene cluster, which later split (broken line) and moved to different regions of the genome. Since orthologies between posterior Hox genes in cnidarians (CPHox) and bilaterians (BPHox) are still not clear, it is fair to suppose that cnidarian and bilaterian posterior Hox genes were produced independently in both lineages from the proto-posterior Hox gene (PPHox) present in the C-BLCA. In the lineage leading to the last common bilaterian ancestor (LCBA), another tandem duplication originated the PG5 gene. From the LCBA, two sister-groups formed: that leading to present day acoelomorphs, and another giving rise to the protostome-deuterostome last common ancestor (P-DLCA). In the acoel lineage, the original cluster might have disintegrated (at least in *S. roscoffensis*) and some genes lost (tentatively but not proven, the genes PG2, Gsx and Xlox), here represented by empty boxes. A similar process (yet to be proven) could have occurred in nemertodermatids, though absence of anterior genes (empty boxes) is more probably the result of limited sampling. A further tandem duplication in the Nemertodermatida lineage originated a second PG5 gene. Finally, a series of tandem duplications, namely involving the central Hox class originated the extended HOX cluster present in the P-DLCA.

6. FUNCTIONAL ANALYSIS OF THE *IPHOXPOST* GENE IN THE ACOEL *ISODIAMETRA PULCHRA* AND THE ANCESTRAL FUNCTION OF THE POSTERIOR HOX GENE IN THE LCBA

The analyses of sequences, genomic arrangements, and the expression patterns of the three Hox genes found in different species of acoels suggest the presence of a Hox based vectorial patterning system or Hox code that could be the most ancient example of Hox code used by the Bilateria. Functional studies are, however, necessary to understand how the functions of Hox genes evolved over time in different taxa. This was the main reason to start this analysis in acoels with the final objective of exploring the ancestral function of these genes in the LCBA, and how its presumed role has changed in the transitions from radial to bilateral animals and from "basal" to "higher" bilaterians.

To carry out this analysis we used the RNA interference technique (De Mulder *et al*., 2009b; Fire *et al*., 1998) on post-embryonic, adult, and regenerating organisms of the acoel *Isodiametra pulchra*. Because some methodological problems hindered cloning the central paralogue, analyses could only be performed in two of the three Hox genes: the anterior and the posterior Hox paralogues. However, only the posterior one, the *IpHoxPost*, could properly be knocked-down by RNAi. This study is the first analysis of a Hox knockdown phenotype in a basal bilaterian. Changing species was necessary since the RNAi methodology doesn't still work in *S. roscoffensis*.

The ISH assay in control organisms (non treated by RNAi) reveals *IpHoxPost* to be expressed in the posterior third of the body in juveniles, a pattern reminiscent of those observed in two other acoels: *Symsagittifera roscoffensis* and *Convolutriloba longifissura* (Hejnol and Martindale, 2009; Moreno *et al*., 2009). In adult organisms, this gene is expressed in two well-defined rows of cells corresponding to longitudinal muscle bundles and/or peripheral parenchyma cells, that run parallel to the male and female copulatory apparatus (see Moreno *et al*., 2010, Fig. 1). *IpHoxPost* expression

pattern becomes progressively restricted around the gonopore area during growth and sexual maturation. In addition, we also detected expression in the nucleolus and in a small patch of cytoplasm surrounding the nuclei of mature egg cells (Moreno *et al*., 2010, Fig. 2, panel C). Nucleolar localization of Hox proteins has previously been observed in vertebrate proliferating cells (Corsetti *et al*., 1995). In *I. pulchra*, though, localization of the message seems restricted to mature, post-mitotic, eggs, ruling out an equivalent role in oocyte proliferation control. It is important to point out that this is one of the few reports in which Hox genes have been shown to be expressed as maternal messages (Dearden *et al*., 2000; Peterson *et al*., 2000a; Rangini *et al*., 1989; Yanze *et al*., 2001).

The phenotypic consequences of knocking down the *IpHoxPost* expression are evidenced by deep changes in body morphology and tissue organization, primarily in the posterior half of the animal. As expected, the anterior part of the body is unaffected by the treatment. Moreover, the epidermis, including posterior epidermis, is also unaffected. Furthermore, in regenerating animals we observe that the wound closes properly, and the epidermis seals the area correctly, suggesting that the epidermis is not under the direct control of the posterior Hox gene.

Light microscopy analyses of RNAi-treated animals reveal several phenotypic changes caused by the *IpHoxPost* knockdown: a slight increase in average body size, malformations in the chordoid vacuoles, defects in the structure of the male and female copulatory organs, and the presence of higher amounts of sperm cells accumulating always in terminal areas (animals unable to spawn given the anatomical defects in the copulatory organs). Interestingly, more than 50% of treated animals contain an aggregate of diatom's frustules localized in the central parenchyma area, filling an area that runs from the mouth to the posterior tip of the body.

Although the body wall musculature seems not generally affected by double stranded (ds)-RNA treatments, phalloidin immunostaining assay reveals disorganization, reduction or even loss of mouth specific muscles as well as muscles located in the area of copulatory organs (see Moreno *et al*., 2010, Fig. 5, 1). This

suggests a major role of *IpHoxPost* in the specification of posterior musculature. The role of the posterior Hox regulating the development and diversification of postembryonic mesoderm and musculature is well known in deuterostome (Wellik, 2007) and protostome (Capovilla *et al.*, 2001; Liu and Fire, 2000) animals. Our results suggest that the role of posterior Hox genes in the specification of mesodermal derivatives within the Bilateria might represent its ancestral role.

Accumulation of processed food (e.g. diatom's frustules) in posterior body cavities could result from disorganization of the musculature around the mouth. While epidermal cilia (unaffected by *IpHoxPost* knockdown) help organisms to feed on diatoms, the activities of mouth specific muscles are necessary to expel the diatom's frustules, once diatom cores have been digested. Disorganization of these muscles would prevent the correct expelling of processed food. Another specific case for Hox genes regulating the patterning of the pharyngeal area has been reported from amphioxus, though the mechanistic details remain obscure (Holland and Holland, 1996). Additional, serotonin immunostaining assays reveals that the animals have some neural fibres also affected, all in the posterior half of the body, and specially the dorsal nerve cords (see Moreno *et al.*, 2010, Fig. 5, 2). These effects of posterior Hox knockdowns in the mesoderm and surrounding neural tissue are, somehow, reminiscent of phenotypes generated in the absence of posterior Hox genes in mice, where limbs skeleton and their innervation patterns are both affected simultaneously (Wahba *et al.*, 2001).

Using BrdU and Phospho-Histone H3 techniques we detected changes of proliferation in different body regions. Interestingly, the number of proliferating cells in the posterior part of the body increases dramatically in ds-RNA treated animals (see Moreno *et al.*, 2010, Fig. 6, 1). Thus, *IpHoxPost* seems tightly involved in the control of cell proliferation. Such a role for Hox genes has been already shown in different animal systems and in several tissues (Lei *et al.*, 2006; Rogulja-Ortmann and Technau, 2008). However, the underlying mechanistic link between Hox gene activities and the regulation of proliferation remains an open issue (see, for instance Chu *et al.*, 2004; Raman *et al.*, 2000; Salsi *et al.*, 2009). In this context we are performing now BrdU

tests with in situ hybridization, at the same time. This will allow us to follow the activities of Hox genes in mitotic and post-mitotic cells.

A further phenotypic consequence of knocking down *IpHoxPost* is the presence of degenerating eggs. Production of mature eggs is completely halted at 10-15 days of ds-RNA treatment. Because *IpHoxPost* transcripts accumulate in mature eggs during egg development this may indicate that this gene is needed for a correct maturation of egg cells. Whereas normal developing eggs are found at the anterior half of the body, in the midbody region more mature stages appear malformed.

In *I. pulchra*, the antibody against the PIWI protein labels specifically a subpopulation of somatic stem cells, spermatogonia and spermatocytes, as well as the developing eggs (De Mulder *et al*., 2009a). The use of this antibody has convincingly shown that it is the maturing eggs and not the earlier stages that are mostly affected (see Moreno *et al*., 2009, Fig. 6, 2). Similarly, it has been observed that in mice the progression from precursors to mature eggs correlates with the transcriptional repression of several genes, among which are the Hox (Saitou *et al*., 2003).

As mentioned earlier, expression of *IpHoxPost* labels the nucleolus of mature egg cells and a small patch of the cytoplasm surrounding the nucleus. These findings suggest that expression of *IpHoxPost* and maturation of the eggs might be directly linked (see Moreno *et al*., 2010, Fig. 2, A). This is by no means surprising because other transcription factors that directly control proliferative activities seem localized in nucleolar compartments (Su *et al*., 1993). This link, however, remains to be tested in the germline context. Alternatively, *IpHoxPost* might have only an indirect effect on the maturation of eggs. For instance, if eggs need positional information cues (or nurture) from neighbouring cells, any alteration at in the posterior part of the body, like those driven by inhibition of *IpHoxPost*, could indirectly alter their maturation process and, as a consequence, eggs will degenerate.

To summarize; when *IpHoxPost* functions are knockdown by RNAi, body form, tissue organization and cellular proliferation are clearly affected. The most dramatic

changes occur at the posterior end of the animal where this gene is expressed. Interestingly, no obvious effects are detected in the anterior half of the body. Altogether, these results allow us to conclude that the posterior Hox gene specifies some mesodermal (peripheral parenchyma) and neural structures in the posterior half of the body of the acoel *I. pulchra*. Another key developmental event controlled by this gene is the control of the egg maturation process. It is important to emphasize that within the posterior half of the animal other structures seem unaffected (some muscles or the epidermis), showing that within the large domains of the embryo or the adult Hox genes still work as "micro-managers". In *I. pulchra*, *IpHoxPost* carry out its functions through a tight control of posterior cell proliferation and cellular identity, a role shared with many other bilaterians.

Based on expression patterns obtained with ISH analyses, Hejnol and Martindale (2009) suggest that the in the acoel *C. longifissura* Hox genes might have a major role in the axial patterning of the nervous system. Further, they suggest this role as ancestral within bilaterians, being subsequently co-opted for patterning roles in other tissues along the AP axis, such as paraxial mesoderm, paired appendages, and genitalia (Deutsch and Le Guyader, 1998).

In contrast to Hejnol and Martindale (2009) proposal, the results obtained in this work, and those obtained in other bilaterians, indicate that posterior Hox genes are mainly involved in the regulation of the postembryonic mesoderm and musculature. This would indicate that such functions arose early on in bilaterian evolution. However, Hox genes in acoels could also be involved primarily in patterning the nervous system, as attested by reduced nerve fibers found in *IpHoxPost* knockdowns and from the co-expression patterns of *SrHox1* and *SrPBX* genes in *S. roscoffensis* juveniles. The decision, at this point relies on obtaining more data from other acoels and a better, cellular, resolution of the *in situ* patterns. We have to indicate, though, that our data has been obtained in juveniles and adults and miss an important, very informative, piece of information, the expression in the embryo. This would obviously help in understanding the primary roles of Hox genes.

Likewise, the control of cell proliferation and the egg maturation process could be other example of ancestral roles of the posterior Hox genes within the Bilateria.

In a phylogenetic context, however, the lack of knowledge on the function of posterior Hox genes in cnidarians makes difficult to trace any parallelisms between the uses of Hox genes in cnidarians and acoels (or bilaterians) and to produce an educated guess as regards its functional role and evolution of Hox activities from the C-BLCA onwards. The only knockdown of posterior-like Hox genes in Cnidaria was done by inhibiting the activity of *Cnox-1* and *Cnox-3* of *Eleutheria dichotoma*. This experiment showed some phenotypic effects concentrated in the oral pole structures of the medusa (Jakob and Schierwater, 2007). This suggests again that the oral pole of cnidarians and the posterior pole of bilaterians might be homologous, and suggests a conserved role for posterior Hox genes in patterning the structures of the oral pole of cnidarians and the posterior pole of bilaterians. However, lack of conservation between expression patterns among cnidarians makes difficult to extrapolate this homology model from the results of a single species.

7. AXIAL BODY PATTERN MECHANISMS BEFORE THE ADVENT OF HOX GENES: THE WNT GENES AND THE CO-OPTION OF HOX GENES

The results obtained in different laboratories and our results in acoels indicate that the Hox axial patterning system arose, most probably, at the very beginning of the bilaterian lineage. This proposal leads to a new question: before the advent of the Hox system, which genetic system was used to pattern the main body axis?

As of today, a fair amount of data points to the Wnt signalling pathway as the main signalling system involved in the axial patterning of ancestral metazoans (Lee *et al*., 2006). Some arguments support this assertion. First, the Wnt pathway is involved in axial (AP) patterning in many bilaterians. In vertebrates, hemichordates, and echinoderms, Wnt signalling is essential for posterior patterning (Holland, 2002), while

in vertebrates the Dickkopf (Dkk) family of Wnt antagonists are required to specify anterior structures (Niehrs, 2006). Also in several protostomes (insects and polychaetae annelids), there is ample evidence for a general role of Wnt expression in the posterior end of the body (Seaver and Kaneshige, 2006).

At the same time, Wnt signalling also plays a significant role in patterning the cnidarian primary body axis (Hobmayer *et al*., 2000; Wikramanayake *et al*., 2003; Broun *et al*., 2005; Kusserow *et al*., 2005; Augustin *et al*., 2006; Guder *et al*., 2006; Lee *et al*., 2006; Plickert *et al*., 2006; Ryan and Baxevanis, 2007). Interestingly, Wnt signalling pathway seems also to operate in sponge larvae, as shown in two species of demosponges (Nichols *et al*., 2006; Adell *et al*., 2003). In fact, in a recent study of Wnt expression in the demosponge *Amphimedon queenslandica* it has been shown that the domain of activity is consistent with a role in patterning the primary body axis during development (Adamska *et al*., 2007).

All this data, taken together would suggest that the first (ancestral) mechanism responsible to pattern the primary body axis of metazoans involved the Wnt signalling pathway. Later on, Hox genes were somehow co-opted into this role, although it is difficult to say whether they were co-opted prior to the C-BLCA or subsequently, in the lineage leading to the LCBA (Ryan and Baxevanis, 2007). The transition to a system governed by Hox genes from one dominated by the members of the Wnt family would still be observable in the many examples of Wnt genes controlling Hox function, a role clearly demonstrated in the axial specification of vertebrates (Lohnes, 2003). Also, recent data from protostomes and deuterostomes indicates that Wnt signalling partitions Hox and ParaHox domains to specify unique cell fates during development (Bilder *et al*., 1998; Merabet *et al*., 2005; Arata *et al*., 2006; Bondos, 2006; Bondos *et al*., 2006).

8. ON THE LCBA AND THE "CAMBRIAN EXPLOSION". A SPECULATIVE SCENARIO

I believe that the results presented in this work challenge two of the fundamental beliefs of the "new animal phylogeny" (Adoutte *et al*., 2000; Adoutte *et al*., 1999): the placement of all acoelomates among the descendant of the P-DLCA and the ensuing representation of the LCBA as a complex organism (the complex Urbilateria) bearing most of the features of the extant bilaterians (Arendt *et al*., 2001; Carroll *et al*., 2005; De Robertis and Sasai, 1996). According to the "new animal phylogeny" (Adoutte *et al*., 2000; Adoutte *et al*., 1999), the LCBA would have to be a relatively large coelomate and segmented worm-like animal, with a blood circulatory system, segmental nephridia, gonads, simple eyes, appendages, a heart tube and a centralized nervous system with a regionalized brain and ventral nerve chord. Besides, it would have to develop indirectly from a microscopic ciliated larva (Balavoine and Adoutte, 2003).

The phylogenetic position of Acoelomorpha at the base of the bilaterian phylogenetic tree (showed by phylogenomic analysis and microRNAs complementary studies), plus its simple morphology and direct mode of development, challenge the theory of the complex Urbilateria. The concept of the complex Urbilateria was born from the comparison between protostomes and deuterostomes, when Acoelomorpha were still consider plathelminthes Therefore, this complex ancestor would represent most probably the last common ancestor of protostomes and deuterostomes, but not the last common ancestor of all remaining bilaterians. Once that the Acoelomorpha have been showed to branch before the protostome-deuterostome divergence, the hypothesis about their simplification from a complex ancestor is neither parsimonious nor necessary.

On the other hand, regressive evolution (secondary simplifications) is accompanied by narrowing the ecological niche in a monotonous and extremely simplified environment and ceasing the mobile life. This is the main evolutionary pathway of parasitic and sedentary Metazoa. Conversely, acoelomorphs are active, free-

living hunters, and there is no reason for supposing a regressive evolution in its lineage (Kaufman, 2008).

Finally, there are not embryological evidences for secondary reductions in Acoelomorpha, as happens in groups of parasitic flatworms with a simplified morphology. For example, the cleavage pattern is irregular and anarchic in the parasitic flatworms belonging to Neodermata (tapeworms, flukes, etc.), which is determined by the general disintegration of the body caused by parasitism; indeed their free-living ancestors had a spiral cleavage (Kaufman, 2008). However, Acoelomorpha develop through a regular duet pattern of spiral cleavage (Boyer *et al*., 1996). In conclusion, there is no obvious reason for the regressive evolution of Acoelomorpha, and their simple morphology should be considered to be primary.

It is worth to highlight that very recently, another free-living simple group of worms has found a place at the base of the bilaterian tree: the genus *Xenoturbella* (Hejnol *et al*., 2009). *Xenoturbella* is a completely ciliated worm with only a ventral mouth opening to the digestive system and a basiepidermal nervous system. Both the general anatomy and several ultrastructural features (as the epidermal ciliary rootlets including the unique ciliary tips (Franzén and Afzelius, 1987; Lundin, 1998)), suggest the relation between *Xenoturbella* and Acoelomorpha (Haszprunar, 1996). If a group of complex animals were the sister group of Acoelomorpha, this would be a strong evidence for a secondary reduction in the Acoelomorpha lineage; however this does not seem to be the case.

In conclusion, as substantiated by recent molecular data, the stem species of the Bilateria was likely a small, acoelomate, unsegmented acoeloid organism with only one digestive opening in ventral position (Hejnol and Martindale, 2008a). This acoel-like ancestor might have evolved through paedogenesis from a cnidarian planula larva, as the Planuloid–Acoeloid hypothesis of bilaterian evolution of von Graff (1891) predicted. Although it is also possible that both cnidarians and bilaterians may have evolved from a planulomorph ancestor with traits of bilateral symmetry, as the Planulozoa' hypothesis suggests (Wallberg *et al*., 2004). If this hypothesis is confirmed,

the Urbilateria would not correspond to the LCBA but to the C-BLCA, and the LCBA should be named instead as LCTA, being the T for Triploblastica (Baguñà *et al*., 2008).

In order to clarify the issue of bilaterian origins a reasonable question to ask ourselves would be: How informative is the fossil record with respect to the origin of Bilateria? Although there are probes of the existence of complex bilaterians in the fossil record from 40 to 55 Myr before the Lower Cambrian boundary (*Vernanimalcula guizhouena*, Chen *et al*., 2004), fossils of organisms similar to the planuloid-acoeloid ancestor (or with a similar grade of complexity) have not been reported. However, the trace fossils of the earliest bilaterians belonging to the Neoproterozoic age (about 570 Myr ago) are indicating the small body size of these ancestors (Valentine and Collins, 2000; Peterson *et al*., 2005). A likely explanation for the shortage of fossils belonging to the planuloid-acoeloid ancestor is that these small soft-bodied and benthic stem-group bilaterians fossilize and left trace fossils with difficulty due to the lack of skeletonization (Conway Morris, 1993). This would suggest the existence of a large cryptic Precambrian evolutionary history of these stem-group bilaterians during which the ancestors of all groups whose fossils appeared during the Cambrian have evolved (Erwin, 2006).

Indeed, as our results about the Hox complement in acoels suggest, the likely gradual expansion of genes in the HOX cluster in simple early bilaterians challenges old and new attempts to link the apparently sudden appearance of most bilaterian phyla during the so-called Cambrian explosion with a quick expansion of the HOX cluster (Erwin *et al*., 1997; García-Fernández, 2005a; Valentine *et al*., 1996). The piecemeal expansion of the HOX cluster strongly contradicts the hypothesis that the sudden appearance of most bilaterian phyla during the Cambrian (the so-called Cambrian 'explosion') was causally related to the sudden expansion of the HOX cluster (Valentine *et al*., 1996). In fact, suggests an alternative scenario with a progressive incorporation of new characters during the early evolution of bilateral animals, contributing to recover the old idea of "intermediate taxa"; since the Hox-ParaHox complement in the LCBA would represent an intermediate step between the complement in the C-BLCA and the P-DLCA. Whereas the P-DLCA would possess the large cluster HOX with 7-9 genes

(similar to the proposed cluster for the complex Urbilateria for the defenders of the Archicoelomate theory), the LCBA would possess a simple set of 3-4 Hox genes similar to the Hox complement found in acoels. The results of this thesis, plus other significant work done recently in acoels (for instance: Hejnol and Martindale, 2008a), challenge the Archicoelomate hypothesis as well as the concept of the complex Urbilateria.

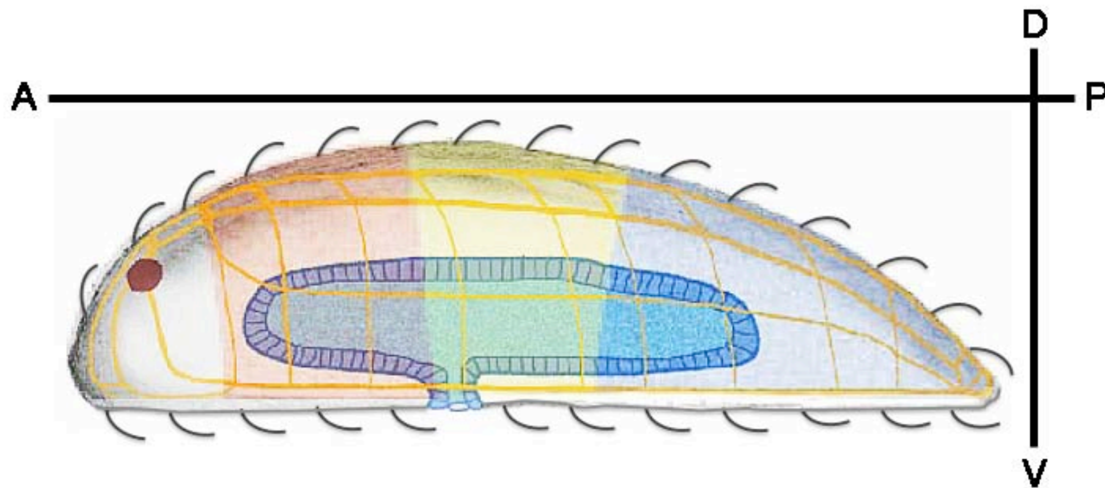


Figure 14. Main features of the acoeloid-like LCBA. The most likely LCBA was a simple and small (less than 1 cm, Erwin and Davidson, 2002) acoeloid-like ancestor. It was a triploblastic bilateral animal with AP and DV defined body axes, the former established by a simple Hox code made by 3-4 genes (indicated by red, yellow, and blue colours along the AP axis). This ancestor also had a blind gut (blue cavity) with single (mouth) opening (Hejnol and Martindale, 2008a & b), and a simple nervous system (Haszprunar, 1996) formed by an anterior concentration of neurons, and a diffuse set of A-P directed (and transversal) nerve tracts (yellow lines), ocelli (red spot) and other sensory organs. It lacked coelom, nephridia, segments, and appendages. The epidermis was probably multiciliated. It was a benthic (Peterson *et al*., 2005) direct-developer (Raff, 2008), probably with internal fertilization (Buckland-Nicks and Scheltema, 1995), and hermaphroditic. Diagram modified from Hejnol and Martindale, 2008a. A, anterior; D, dorsal; P, posterior; V, ventral.

Another interesting evolutionary issue revolves around the mode of development of the LCBA. Here there are two opposite hypotheses being contemplated by different authors. While some authors have described the LCBA as an indirect developer (Nielsen, 1995; Peterson *et al*., 2000b) others describe it as direct developer (Baguñà and Riutort, 2004; Cook *et al*., 2004; Jondelius *et al*., 2002; Ruiz-Trillo *et al*., 1999; Wallberg *et al*., 2007). The indirect developer hypothesis suggests that the Urbilateria was a ciliated planktotrophic larval-like organism with and a benthic adult form developed indirectly. On the contrary, the direct developer hypothesis suggests that adult bilaterian body plans evolved first and that larval body plans arose subsequently

by stepwise intercalation of genes already used in the adult to generate features of the larva (Valentine and Collins, 2000; Sly *et al*., 2003). Taken into account that phylogenetic analyses do support a metazoan phylogeny in which basal clades are direct developers (Acoelomorpha), most probably the direct development is primitive in bilaterians, and planktonic larvae were independently intercalated into an existing direct-developing strategy multiple times (Hadfield *et al*., 2001).

Moreover, the fossil record seems to be consistent with the primitive direct development of bilaterian adults. Whereas recent estimates suggest that the origin of bilaterians lies in the Late Precambrian, between 580 and 600 Myr ago (Peterson *et al*., 2005), an estimate of the timing for evolution of planktonic larvae of approximately 500 Myr ago is emerging, which if correct puts the origin of these second body plans 100 Myr later than the divergence of the basal bilaterian benthic adult (Signor and Vermeij, 1994; Raff, 2008). Therefore the first bilaterians were probably small benthic animals with a degree of complexity similar to living acoel flatworms.

9. PERSPECTIVES

To validate acoels (and nemertodermatids) as good Evo-Devo model-systems at the origin of the Bilateria, it is tantamount to know the actual number of their Hox and ParaHox gene sets. Otherwise, whether the small number of these gene sets reflects an ancestral situation or whether is the result of a simplification process from a more complex ancestor bearing full sets of these genes, will remain an open question. To that end, sequencing a whole genome will validate, or not, the conclusions here attained. If proved true, they will provide additional information on the genomic regions adjacent to Hox and ParaHox genes, on the presence of pseudogenes, and on the presence of other ANTP genes usually related with Hox genes in HOX clusters of other organisms. Likewise, sequencing whole genomes of one or more nemertodermatid flatworms and one or both xenoturbellid species would also be extremely informative.

At the functional level, the analysis carried out here should be made extensive to the central Hox gene, and to Cdx and PBX genes. Moreover, and because they did not produce phenotypes in *I. pulchra*, functional analysis of anterior Hox genes by RNA interference needs to be performed in other acoels species. Likewise, studying the roles of these genes during the embryonic development of acoels will bring fresh data to compare it to that of cnidarians and other bilaterians. To that end, developing microinjection techniques for acoels eggs and embryos are of prime importance.

Finally, detection of target genes regulated by Hox genes in acoels will also be very important. Using chromatin immunoprecipitation assays (ChIP) it is possible nowadays to determine the genes directly downstream from Hox genes in genetic regulatory networks. These analyses would be essential to understand how Hox gene regulatory networks evolved along metazoan history, and how changes in these networks and changes in morphological structures are linked. Last but not least, they could also throw some light into the most important morphological innovation in the history of the Metazoa: the origin of bilateral symmetry and the axial (AP and DV) coordinating system.

4

CONCLUSIONS

I. Acoel flatworms have three Hox genes, orthologous to the bilaterian PG1, PG5 and PG9-14. This set represents the minimal complement of Hox genes reported so far in the Bilateria. The phylogenetic position of acoels, as sister-group to the rest of bilaterians makes this Hox gene set a good proxy for the complement used in the LCBA.

II. The three Hox genes in the acoel *Symsagittifera roscoffensis* are located in separate chromosomes. Hence, Hox genes in this species form a so-called disintegrated cluster, *sensu* Duboule (2007).

III. The Hexapeptide motif is well conserved in SrHox1 and SrHox5. On the other hand, expression patterns of *SrPBX* and *SrHox1* overlap in the anterior part of the body, suggesting a role for a complex between PBX-HOX1 in patterning the nervous system of acoels.

IV. The three Hox genes in *S. roscoffensis* are expressed in nested domains along the AP axis. Taking into account the phylogenetic position of acoels, this would represent the earliest example of trans-collinear expression in Bilateria and the first example of a "Hox code" used by the metazoans.

V. The posterior Hox gene of the acoel *Isodiametra pulchra*, *IpHoxPost*, is involved in the development of the posterior part of the body. Specific structures under its control are the mouth specific muscles and also the muscles located around the copulatory organs. Therefore, the control of the development and diversification of postembryonic mesoderm and musculature would represent an ancient role for the posterior Hox. In addition, *IpHoxPost*, is also involved in egg maturation processes and in the regulation of cell proliferation in the posterior part of the body. This gene plays no specific roles in the anterior half of the body.

VI. A Cdx ParaHox gene is also present in the genome of acoels. It is expressed in the nervous system of juveniles, and in the gonopore area in adult organisms.

5

REFERENCES

Abate-Shen, C. 2002. Deregulated homeobox gene expression in cancer: cause or consequence? *Nature Reviews Cancer* 2: 777-785.

Aboobaker, A. and Blaxter, M. L. 2003. Hox gene loss during dynamic evolution of the nematode cluster. *Current Biology* 13: 37-40.

Adamska, M., Degnan, S. M., Green, K. M., Adamski, M., Craigie, A., Larroux, C., Degnan, B. M. 2007. Wnt and TGF-β expression in the sponge *Amphimedon queenslandica* and the origin of metazoan embryonic patterning. *PLoS ONE* 2: e1031.

Adell, T., Nefkens, I., and Muller, W. E. G. 2003. Polarity factor 'Frizzled' in the demosponge *Suberites domuncula*: identification, expression and localization of the receptor in the epithelium/pinacoderm. *FEBS Letters* 554: 363-368.

Adoutte, A., Balavoine, G., Lartillot, N., and Rosa, R. d. 1999. Animal evolution: the end of the intermediate taxa? *Trends in Genetics* 15 (3): 104-108.

Adoutte, A., Balavoine, G., Lartillot, N., Lespinet, O., Prud'homme, B., and de Rosa, R. d. 2000. The new animal phylogeny: Reliability and implications. *Proceedings of the National Academy of Sciences of the United States of America* 97 (9): 4453-4456.

Amemiya, C. and Wagner, G. P. 2006. Animal evolution: when did the 'Hox system' arise? *Current Biology* 16: R546-548.

Amores, A., Force, A., Yan, Y.-L., Joly, L., Amemiya, C., Fritz, A., Ho, R. K., Langeland, J., Prince, V., Wang, Y.-L., Westerfield, M., Ekker, M., and Postlethwait, J. H. 1998. Zebra fish *hox* clusters and vertebrate genome evolution. *Science* 282: 1711-1714.

Andersen, R. A., Berges, J. A., Harrison, P. J., and Watanabe, M. M. 2005. Recipes for freshwater and seawater media. In *Andersen RA (ed) Algal culturing techniques*. Amsterdam: Elsevier. pp. 429-538.

Apelt, G. 1969. Fortpflanzungsbiologie, Entwicklungszyklen und vergleichende Frühentwicklung acoeler Turbellarien. *Marine Biology* 4: 267-325.

Arata, Y., Kouike, H., Zhang, Y., Herman, M. A., Okano, H., and Sawa, H. 2006. Wnt signaling and a Hox protein cooperatively regulate psa-3/Meis to determine daughter cell fate after asymmetric cell division in *C. elegans*. *Developmental cell* 11: 105-115.

Arenas-Mena, C., Martinez, P., Cameron, R. A., and Davidson, E. H. 1998. Expression of the *Hox* gene complex in the indirect development of a sea urchin. *Proceedings of the National Academy of Sciences of the United States of America* 95: 13062-13067.

Arendt, D., Technau, U., and Wittbrodt, J. 2001. Evolution of the bilaterian larval foregut. *Nature* 409: 81-85.

Arnone, M. I., Rizzo, F., Annunciata, R., Cameron, R. A., Peterson, K. J., and Martínez, P. 2006. Genetic organization and embryonic expression of the ParaHox genes in the sea urchin *S. purpuratus*: Insights into the relationship between clustering and colinearity. *Developmental Biology* 300: 63-73.

Augustin, R., Franke, A., Khalturin, K., Kiko, R., Siebert, S., Hemmrich, G., and Bosch, T. C. G. 2006. Dickkopf related genes are components of the positional value gradient in Hydra. *Developmental Biology* 296: 62-70.

Ax, P. 1995. Platyhelminthes aus dem Eulittoral von Godhavn (Disko, Grönland). *Microfauna Marina* 10: 249-294.

Baguñà, J. and Riutort, M. 2004. The dawn of bilaterian animals. The case of acoelomorph flatworms. *BioEssays* 26: 1046-1057.

Baguñà, J., Martinez, P., Paps, J., and Riutort, M. 2008. Back in time: a new systematic proposal for the Bilateria. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 363: 1481-1491.

Balavoine, G. and Adoutte, A. 2003. The segmented Urbilateria: a testable scenario. *Integrative and Comparative Biology* 43 (1): 137-147.

Balavoine, G., Rosa, R. d., and Adoutte, A. 2002. Hox clusters and bilaterian phylogeny. *Molecular Phylogenetics and Evolution* 24: 366-373.

Ball, E. E., de Jong, D. M., Schierwater, B., Shinzato, C., Hayward, D. C., and Miller, D. J. 2007. Implications of cnidarian gene expression patterns for the origins of bilaterality - is the glass half full or half empty? *Integrative and Comparative Biology* 45 (5): 701-711.

Beklemishev, V. N. 1969. *Principles of comparative anatomy of invertebrates*. Edinburgh: Oliver & Boyd.

Bery, A., Cardona, A., Martinez, P. and Hartenstein, V. Fine structure of the central nervous system of a juvenile acoel, *Symsagittifera roscoffensis*. *Development, Genes and Evolution*. In press.

Bilder, D., Graba, Y., and Scott, M. P. 1998. Wnt and TGFbeta signals subdivide the AbdA Hox domain during *Drosophila* mesoderm patterning. *Development* 125 (9): 1781-1790.

Bode, A., Salvenmoser, W., Nimeth, K., Mahlknecht, M., Adamski, Z., Rieger, R. M., Peter, R., and Ladurner, P. 2006. Immunogold-labeled S-phase neoblasts, total neoblast number, their distribution, and evidence for arrested neoblasts in *Macrostomum lignano* (Platyhelminthes, Rhabditophora). *Cell and Tissue Research* 325: 577-587.

Bondos, S. E. 2006. Variations on a theme: Hox and Wnt combinatorial regulation during animal development. *Science's signal transduction knowledge environment* 2006: pe38.

Bondos, S. E., Tan, X.-X., and Matthews, K. S. 2006. Physical and genetic interactions link Hox function with diverse transcription factors and cell signaling proteins. *Molecular Cellular Proteomics* 5: 824-834.

Bourlat, S. J. and Hejnol, A. 2009. Acoels. *Current Biology* 19 (7): R1-R3.

Boyer, B. C., Henry, J. Q., and Martindale, M. Q. 1996. Modified spiral cleavage: the duet cleavage pattern and early blastomere fates in the acoel turbellarian *Neochildia fusca*. *Biological Bulletin* 191: 285-286.

Brend, T., Gilthorpe, J., Summerbell, D., Rigby, P. W. 2003. Multiple levels of transcriptional and post-transcriptional regulation are required to define the domain of *Hoxb4* expression. *Development* 130: 2717-2728.

Bresslau, E. 1909. Die Entwicklung der Acoelen. *Verh. Dtsch. Zool. Gesamte* 19: 314-324.

Brooke, N. M., Garcia-Fernández, J., and Holland, P. W. H. 1998. The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. *Nature* 392: 920-922.

Broun, M., Gee, L., Reinhardt, B., and Bode, H. R. 2005. Formation of the head organizer in hydra involves the canonical Wnt pathway. *Development* 132: 2907-2916.

Budd, G. E. 2008. The earliest fossil record of the animals and its significance. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 363: 1425-1434.

Buckland-Nicks, J., and A. H. Scheltema. 1995. Was internal fertilization an innovation of early Bilateria? Evidence from sperm structure of a mollusc. *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain)* 261: 11-18.

Bürglin, T. R. 1996. Homeodomain Proteins. In *Encyclopedia of Molecular Biology and Molecular Medicine*. Weinheim: R. A. Meyers. pp. 55-76.

Cameron, R. A., Rowen, L., Nesbitt, R., Bloom, S., Rast, J. P., Berney, K., Arenas-Mena, C., Martinez, P., Lucas, S., Richardson, P. M., Davidson, E. H., Peterson, K. J., and Hood, L. 2006. Unusual gene order and organization of the sea urchin Hox cluster. *Journal of Experimental Zoology (Mol Dev Evol)* 306B: 45-58.

Capovilla, M., Cambris, Z., and Botas, J. 2001. Direct regulation of the muscle-identity gene *apterous* by a Hox protein in the somatic mesoderm. *Development* 128: 1221-1230.

Carranza, S., Baguñà, J., and Riutort, M. 1997. Are the platyhelminthes a monophyletic primitive group? An assessment using 18s rDNA sequences. *Molecular Biology and Evolution* 14 (5): 485-497.

Carroll, S. B., Grenier, J. K., and Weatherbee, S. D. 2005. *From DNA to diversity: Molecular Genetics and the Evolution of Animal Design, 2nd Edition*. Malden, MA: Blackwell Publishing.

Carroll, S. B., DiNardo, S., O'Farrell, P. H., White, R. A., and Scott, M. P. 1988. Temporal and spatial relationships between segmentation and homeotic gene expression in *Drosophila* embryos: distributions of the fushi tarazu, engrailed, Sex combs reduced, Antennapedia, and Ultrabithorax proteins. *Genes & Development* 2: 350-360.

Chen, J.-Y., Bottjer, D. J., Oliveri, P., Dornbos, S. Q., Gao, F., Ruffins, S., Chi, H., Li, C.-W., and Davidson, E. H. 2004. Small bilaterian fossils from 40 to 55 million years before the Cambrian. *Science* 305, 218-222.

Chiori, R., Jager, M., Denker, E., Wincker, P., Da Silva, C., Le Guyader, H., Manuel, M., and Quéinnec, E. 2009. Are Hox genes ancestrally involved in axial patterning? Evidence from the hydrozoan *Clytia hemisphaerica* (Cnidaria). *PloS one* 4 (1): e4231.

Chourrout, D., Delsuc, F., Chourrout, P., Edvardsen, R. B., Rentzsch, F., Renfer, E., Jensen, M. F., Zhu, B., Jong, P. d., Steele, R. E., and Technau, U. 2006. Minimal ProtoHox cluster inferred from bilaterian and cnidarian Hox complements. *Nature* 442: 684-687.

Chu, M. C., Selam, F. B., and Taylor, H. S. 2004. HOXA10 regulates p53 expression and matrigel invasion in human breast cancer cells. *Cancer Biology & Therapy* 3 (6): 568-572.

Conway Morris, S. 1993. The fossil record and the early evolution of the Metazoa. *Nature* 361: 219-225.

Cook, C. E., Jiménez, E., Akam, M., and Saló, E. 2004. The Hox gene complement of acoel flatworms, a basal bilaterian clade. *Evolution & Development* 6 (3): 154-163.

Corsetti, M. T., Levi, G., Lancia, F., Sanseverino, L., Ferrini, S., Boncinelli, E., and Corte, G. 1995. Nucleolar localisation of three Hox homeoproteins. *Journal of Cell Science* 108: 187-193.

Cuvier, G. 1817. *Le règne animal distribué d'après son organisation, pour servir de base à l'histoire naturelle des animaux et d'introduction à l'anatomie comparée*. Vol. 2. Paris: Deterville.

De Mulder, K., Pfister, D., Kuales, G., Willems, M., Salvenmoser, W., Thaler, M., Hrouda, M., Gorny, A., Borgonie, G., and Ladurner, P. 2009a. Characterization of the stem cell system of the acoel *Isodiametra pulchra*. *BMC Developmental Biology* 9: 69.

De Mulder, K., Pfister, D., Kuales, G., Egger, B., Salvenmoser, W., Willems, M., Steger, J., Fauster, K., Micura, R., Borgonie, G., and Ladurner, P. 2009b. Stem cells are differentially regulated during development, regeneration and homeostasis in flatworms. *Developmental Biology* 334: 198-212.

De Robertis, E. M. 2008. Evo-devo: variations on ancestral themes. *Cell* 132 (2): 185-195.

De Robertis, E. M. and Sasai, Y. 1996. A common plan for dorsoventral patterning in Bilateria. *Nature* 380: 37-40.

de Rosa, R., Prud'homme, B., and Balavoine, G. 2005. *caudal* and *even-skipped* in the annelid *Platynereis dumerilii* and the ancestry of posterior growth. *Evolution & Development* 7 (6): 574-587.

de Rosa, R., Grenier, J. K., Andreeva, T., Cook, C. E., Adoutte, A., Akamk, M., Carroll, S. B., and Balavoine, G. 1999. Hox genes in brachiopods and priapulids and protostome evolution. *Nature* 399: 772-776.

Dearden, P., Grbic, M., Falciani, F., and Akam, M. 2000. Maternal expression and early zygotic regulation of the *Hox3/zen* gene in the grasshopper *Schistocerca gregaria*. *Evolution & Development* 2 (5): 261-270.

Denes, A. S., Jékely, G., Steinmetz, P. R., Raible, F., Snyman, H., Prud'homme, B., Ferrier, D. E., Balavoine, G., Arendt, D. 2007. Molecular architecture of annelid nerve cord supports common origin of nervous system centralization in bilateria. *Cell* 129 (2): 277-288.

Deutsch, J. S. 2008. Do acoels climb up the ''Scale of Beings''? *Evolution & Development* 10 (2): 135-140.

Deutsch, J. S. and Le Guyader, H. 1998. The neuronal zootype. An hypothesis. *Comptes Rendus de l'Academie des Sciences - Series III* 321: 713-719.

Dixit, R., VijayRaghavan, K., and Bate, M. 2008. Hox genes and the regulation of movement in *Drosophila*. *Developmental Neurobiology* 68 (3): 309-316.

Duboule, D. 1994. Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Development, Supplement*: 135-142.

Duboule, D. 2007. The rise and fall of Hox gene clusters. *Development* 134: 2549-2560.

Egger, B., Steinke, D., Tarui, H., Mulder, K. D., Arendt, D., Borgonie, G., Funayama, N., Gschwentner, R., Hartenstein, V., Hobmayer, B., Hooge, M., Hrouda, M., Ishida, S., Kobayashi, C., Kuales, G., Nishimura, O., Pfister, D., Rieger, R., Salvenmoser, W., Smith, J., Technau, U., Tyler, S., Agata, K., Salzburger, W., and Ladurner, P. 2009. To be or not to be a flatworm: the acoel controversy. *PloS one* 4 (5): e5502.

Ehlers, U. 1985. *Das Phylogenetische System der Plathelminthes*. Stuttgart: Gustav Fischer.

Ehlers, U. 1991. Comparative morphology of statocysts in the Plathelminthes and the Xenoturbellida. *Hydrobiologia* 227: 263-271.

Erwin, D. H. 2006. The developmental origins of animal body plans. In *S. H. Xiao & A. J. Kaufman (ed) Neoproterozoic geobiology and paleobiology*. New York, NY: Plenum. pp. 159-197.

Erwin, D. H. and Davidson, E. H. 2002. The last common bilaterian ancestor. *Development* 129: 3021-3032.

Erwin, D. H., Valentine, J. W., and Jablonski, D. 1997. The origin of animal body plans. *American Scientist* 85: 126-137.

Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology* 27: 401-410.

Ferrier, D. E. K. and Holland, P. W. H. 2002. *Ciona intestinalis* ParaHox genes: evolution of Hox/ParaHox cluster integrity, developmental mode, and temporal colinearity. *Molecular Phylogenetics and Evolution* 24: 412-417.

Ferrier, D. E. K., Dewar, K., Cook, A., Chang, J. L., Hill-Force, A., and Amemiya, C. 2005. The chordate ParaHox cluster. *Current Biology* 15: R820-R822.

Field, K. G., Olsen, G. J., Lane, D. J., Giovannoni, S. J., Ghiselin, M. T., Raff, E. C., Pace, N. R., and Raff, R. A. 1988. Molecular phylogeny of the animal kingdom. *Science* 239 (4841): 748-753.

Finnerty, J. R. 2005. Did internal transport, rather than directed locomotion, favor the evolution of bilateral symmetry in animals? *Bioessays* 27 (11): 1174-1180.

Finnerty, J. R. 2003. The origins of axial patterning in the metazoa: how old is bilateral symmetry? *International Journal of Developmental Biology* 47: 523-529.

Finnerty, J. R., Pang, K., Burton, P., Paulson, D., and Martindale, M. Q. 2004. Origins of bilateral symmetry: Hox and Dpp expression in a sea anemone. *Science* 304: 1335-1337.

Finnerty, J. R., Paulson, D., Burton, P., Pang, K., and Martindale, M. Q. 2003. Early evolution of a homeobox gene: the parahox gene *Gsx* in the Cnidaria and the Bilateria. *Evolution & Development* 5 (4): 331-345.

Finnerty, J. R., Master, V. A., Irvine, S., Kourakis, M. J., Warriner, S., and Martindale, M. Q. 1996. Homeobox genes in the Ctenophora: identification of paired-type and Hox homologues in the atentaculate ctenophore, *Beroë ovata*. *Molecular Marine Biology and Biotechnology* 5: 249-258.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391: 806-811.

Franzén, A. and Afzelius, B. 1987. The ciliated epidermis of *Xenoturbella bocki* (Platyhelminthes, Xenoturbellida) with some phylogenetic considerations. *Zoologica Scripta* 16: 9-17.

Frengen, E., Weichenhan, D., Zhao, B., Osoegawa, K., Geel, M. v., and Jong, P. J. d. 1999. A modular, positive selection bacterial artificial chromosome vector with multiple cloning sites. *Genomics* 58 (3): 250-253.

Friedrich, C. and Hendelberg, J. 2001. On the ecology of Acoela living in the Arctic Sea ice. *Belgian journal of zoology* 131 (Supplement 1): 213-216.

Fröbius, A. C. and Seaver, E. C. 2006. *ParaHox* gene expression in the polychaete annelid *Capitella sp. I*. *Development Genes and Evolution* 216: 81-88.

Furlong, R. F., Younger, R., Kasahara, M., Reinhardt, R., Thorndyke, M., and Holland, P. W. H. 2007. A degenerate ParaHox gene cluster in a degenerate vertebrate. *Molecular Biology & Evolution* 24: 2681-2686.

García-Fernández, J. 2005a. Hox, ParaHox, ProtoHox: facts and guesses. *Heredity* 94: 145-152.

García-Fernández, J. 2005b. The genesis and evolution of homeobox gene clusters. *Nature Review Genetics* 6: 881-892.

Gauchat, D., Mazet, F., Berney, C., Schummer, M., Kreger, S., Pawlowski, J., and Galliot, B. 2000. Evolution of Antp class genes and differential expression of *Hydra Hox-ParaHox* genes in anterior patterning. *Proceedings of the National Academy of Sciences of the United States of America* 97: 4493-4498.

Gegenbaur, C. 1859. *Grundzüge der Vergleichenden Anatomie*. Leipzig: W. Engelmann.

Gehring, W. J. 1985. The homeobox: A key to the understanding of development? *Cell* 40: 3-5.

Gehring, W. J. 1998. *Master control genes in development and evolution: The homeobox story*. New Haven: Yale University Press.

Gehring, W. J., Affolter, M., and Bürglin, T. 1994b. Homeodomain proteins. *Annual review of biochemistry* 63: 487-526.

Gehring, W. J., Qian, Y. G., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G., and Wiithricht, K. 1994a. Homeodomain-DNA recognition. *Cell* 78: 211-223.

Georgévitch, J. 1899. Etude sur le développement de la *Convoluta roscoffensis* Graff. *Archives De Zoologie Experimentale Et Generale* 7: 343-361.

Gilbert, S. F. 2006. *Developmental Biology (8th edition)*. Sunderland, MA: Sinauer Associates.

Gilbert, S. F., Opitz, J. M., and Raff, R. A. 1996. Resynthesizing evolutionary and developmental biology. *Developmental Biology* 173 (32): 357-372.

Guder, C., Pinho, S., Nacak, T. G., Schmidt, H. A., Hobmayer, B., Niehrs, C., and Holstein, T. W. 2006. An ancient Wnt-Dickkopf antagonism in *Hydra*. *Development* 133: 901-911.

Haapkylä, J., Seymour, A. S., Barneah, O., Brickner, I., Hennige, S., Suggett, D., and Smith, D. 2009. Association of *Waminoa sp*. (Acoela) with corals in the Wakatobi Marine Park, South-East Sulawesi, Indonesia. *Marine Biology* 156: 1021-1027.

Hadfield, M. G., Carpizo-Ituarte, E. J., Del Carmen, K., and Nedved, B. T. 2001. Metamorphic competence, a major adaptive convergence in marine invertebrate larvae. *American Zoologist* 41: 1123-1131.

Hadzi, H. 1944. *Turbellarijska teorija knidarijev (Turbellarien-Theorie der Knidarien)*. Ljubljana, Slovenia.

Haeckel, E. 1866. *Generelle Morphologie der Organismen: allgemeine Grundzüge der organischen Formen-Wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte Descendenz-Theoriee*. Berlin: George Reimer.

Haerry, T. E. and Gehring, W. J. 1997. A conserved cluster of homeodomain binding sites in the mouse *Hoxa-4* intron functions in *Drosophila* embryos as an enhancer that is directly regulated by *Ultrabithorax*. *Developmental Biology* 186: 1-15.

Hand, C. 1959. On the origin and phylogeny of the coelenterates. *Systematic Zoology* 8: 191-202.

Haszprunar, G. 1996. Plathelminthes and Plathelminthomorpha—paraphyletic taxa. *Journal of Zoological Systematics & Evolutionary Research* 34: 41-48.

Hatschek, B. 1888. *Lehrbuch der Zoologie*. Jena: G. Fischer.

Hejnol, A. and Martindale, M. Q. 2008a. Acoel development supports a simple planula-like urbilaterian. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 363: 1493-1501.

Hejnol, A. and Martindale, M. Q. 2008b. Acoel development indicates the independent evolution of the bilaterian mouth and anus. *Nature* 456: 382-386.

Hejnol, A. and Martindale, M. Q. 2009. Coordinated spatial and temporal expression of *Hox* genes during embryogenesis in the acoel *Convolutriloba longifissura*. *BMC Biology* 7: 65.

Henry, J. Q., Martindale, M. Q., and Boyer, B. C. 2000. The unique developmental program of the acoel flatworm, *Neochildia fusca*. *Developmental Biology* 220: 285-295.

Hejnol, A., Obst, M., Stamatakis, A., Ott, M., Rouse, G. W., Edgecombe, G. D., Martinez, P., Baguñà, J., Bailly, X., Jondelius, U., Wiens, M., Müller, W. E. G., Seaver, E., Wheeler, W. C., Martindale, M. Q., Giribet, G., and Dunn, C. W. 2009. Assessing the root of bilaterian animals with scalable phylogenomic methods. *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain)* 276 (1677): 4261-4270.

Hobmayer, B., Rentzsch, F., Kuhn, K., Happel, C. M., von Laue, C. C., Snyder, P., Rothbacher, U., and Holstein, T. W. 2000. WNT signalling molecules act in axis formation in the diploblastic metazoan *Hydra*. *Nature* 407: 186-189.

Holland, L. Z. 2000. Body-plan evolution in the Bilateria: early antero-posterior patterning and the deuterostome–protostome dichotomy. *Current Opinion in Genetics & Development* 10: 434-442.

Holland, L. Z. 2002. Heads or tails? Amphioxus and the evolution of anterior-posterior patterning in deuterostomes. *Developmental Biology* 241: 209-228.

Holland, L. Z. and Holland, N. D. 1996. Expression of *AmphiHox-1* and *AmphiPax-1* in amphioxus embryos treated with retinoic acid: insights into evolution and patterning of the chordate nerve cord and pharynx. *Development* 122: 1829-1838.

Holland, L. Z., Kene, M., Williams, N. A., and Holland, N. D. 1997. Sequence and embryonic expression of the amphioxus engrailed gene (*AmphiEn*): the metameric pattern of transcription resembles that of its segment polarity homolog in *Drosophila*. *Development* 124: 1723-1732.

Holland, P. W. H. 1998. Major transitions in animal evolution: a developmental genetic perspective. *American Zoologist* 38: 829-842.

Hooge, M. D. and Tyler, S. 2005. New tools for resolving phylogenies: a systematic revision of the Convolutidae (Acoelomorpha, Acoela). *Journal of Zoological Systematics & Evolutionary Research* 43 (2): 100–113.

Hyman, L. H. 1940. *The invertebrates (1st edition)*. New York: McGraw-Hill.

Hyman, L. H. 1951. *The Invertebrates. Vol II. Platyhelminthes and Rhynchocoela. The Acoelomate Bilateria*. New York; Toronto: McGraw-Hill.

Ikuta, T., Yoshida, N., Satoh, N., and Saiga, H. 2004. *Ciona intestinalis* Hox gene cluster: Its dispersed structure and residual colinear expression in development. *Proceedings of the National Academy of Sciences of the United States of America* 101 (42): 15118-15123.

In der Rieden, P. M. J., Mainguya, G., Wolteringa, J. M., and Durston, A. J. 2004. Homeodomain to hexapeptide or PBC-interaction-domain distance: size apparently matters. *Trends in Genetics* 20: 76-79.

Jakob, W. and Schierwater, B. 2007. Changing hydrozoan bauplans by silencing Hox-like genes. *PLoS ONE* 2 (8): e694.

Jakob, W., Sagasser, S., Dellaporta, S., Holland, P., Kuhn, K., and Schierwater, B. 2004. The *Trox-2* Hox/ParaHox gene of *Trichoplax* (Placozoa) marks an epithelial boundary. *Development Genes and Evolution* 214: 170-175.

James, R., Erler, T., and Kazenwadel, J. 1994. Structure of the murine homeobox gene *cdx-2*. Expression in embryonic and adult intestinal epithelium. *Journal of biological chemistry* 269: 15229-15237.

Jiménez-Guri, E., Paps, J., García-Fernández, J., and Saló, E. 2006. *Hox* and *ParaHox* genes in Nemertodermatida, a basal bilaterian clade. *International Journal of Developmental Biology* 50: 675-679.

Jondelius, U., Ruiz-Trillo, I., Baguñà, J., and Riutort, M. 2002. The Nemertodermatida are basal bilaterians and not members of the Platyhelminthes. *Zoologica Scripta* 31: 201-215.

Kamm, K., Schierwater, B., Jakob, W., Dellaporta, S. L., and Miller, D. J. 2006. Axial patterning and diversification in the Cnidaria predate the Hox system. *Current Biology* 16: 920-926.

Kaufman, Z. S. 2008. Some problems of regressive evolution. *Biology Bulletin* 35: 318-326.

Kaufman, T. C., Seeger, M. A., and Olsen, G. 1990. Molecular and genetic organization of the antennapedia gene complex of *Drosophila melanogaster*. *Advances in genetics* 27: 309-362.

Keeble, F. 1910. In *Plant-animals a study in symbiosis*. Cambridge: University Press.

Keeble, F. and Gamble, F. W. 1905. On the isolation of the infecting organism ("Zoochlorella") of *Convoluta roscoffensis*. *Proceedings of the Royal Society of London*. *Series B, Containing papers of a Biological character*. *Royal Society (Great Britain)* 77 (514): 66-68.

Keeble, F. and Gamble, F. W. 1907. The origin and nature of the green cells of *Convoluta roscoffensis*. *Quarterly Journal of Microscopical Science* 51: 167-219.

Kostenko, A. G. and Mamkaev, Y. V. 1990. The position of "green convoluts" in the system of acoel turbellarians (Turbellaria, Acoela). 1. *Simsagittifera* gen. n. 2. Sagittiferidae fam. n. *Zoologicheskii Zhurnal* 69 (6): 11-21.

Koziol, U., Lalanne, A. I., and Castillo, E. 2009. *Hox* genes in the parasitic platyhelminthes *Mesocestoides corti*, *Echinococcus multilocularis*, and *Schistosoma mansoni*: evidence for a reduced *Hox* complement. *Biochemical Genetics* 47: 100-116.

Kusserow, A., Pang, K., Sturm, C., Hrouda, M., Lentfer, J., Schmidt, H. A., Technau, U., von Haeseler, A., Hobmayer, B., Martindale, M. Q., and Holstein, T. W. 2005. Unexpected complexity of the *Wnt* gene family in a sea anemone. *Nature* 433: 156-160.

Ladurner, P. and Rieger, R. 2000. Embryonic muscle development of *Convoluta pulchra* (Turbellaria-Acoelomorpha, Platyhelminthes). *Developmental Biology* 222: 359-375.

Ladurner, P., Rieger, R., and Baguñà, J. 2000. Spatial distribution and differentiation potential of stem cells in hatchlings and adults in the marine platyhelminth *Macrostomum sp.*: A Bromodeoxyuridine analysis. *Developmental Biology* 226: 231-241.

Larroux, C., Fahey, B., Degnan, S. M., Adamski, M., Rokhsar, D. S., and Degnan, B. M. 2007. The NK Homeobox gene cluster predates the origin of Hox genes. *Current Biology* 17: 706-710.

Lee, P. N., Pang, K., Matus, D. Q., and Martindale, M. Q. 2006. A WNT of things to come: Evolution of Wnt signaling and polarity in cnidarians. *Seminars in Cell & Developmental Biology* 17: 157-167.

Lei, H., Juan, A. H., Kim, M. S., and Ruddle, F. H. 2006. Identification of a *Hoxc8*-regulated transcriptional network in mouse embryo fibroblast cells. *Proceedings of the National Academy of Sciences of the United States of America* 103 (27): 10305-10309.

Lemons, D. and McGinnis, W. 2006. Genomic evolution of Hox gene clusters. *Science* 313: 1918-1922.

Lewis, E. B. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* 276: 565-570.

Lim, J. K. and Simmons, M. J. 1994. Gross chromosome rearrangements mediated by transposable elements in *Drosophila melanogaster*. *BioEssays* 16: 269-275.

Liu, J. and Fire, A. 2000. Overlapping roles of two Hox genes and the *exd* ortholog *ceh-20* in diversification of the *C. elegans* postembryonic mesoderm. *Development* 127: 5179-5190.

Lohnes, D. 2003. The Cdx1 homeodomain protein: an integrator of posterior signaling in the mouse. *BioEssays* 25: 971-980.

Lundin, K. 1998. The epidermal ciliary rootlets of *Xenoturbella bocki* (Xenoturbellida) revisited: new support for a possible kinship with the Acoelomorpha (Platyhelminthes). *Zoologica Scripta* 27: 263-270.

Mamkaev, Y. V. and Kostenko, A. G. 1991. On the phylogenetic significance of sagittocysts and copulatory organs in acoel turbellarians. *Hydrobiologia* 227: 307-314.

Manuel, M. 2009. Early evolution of symmetry and polarity in metazoan body plans. *Comptes rendus biologies* 332: 184-209.

Marques, A. C. and Collins, A. G. 2004. Cladistic analysis of Medusozoa and cnidarian evolution. *Invertebrate Biology* 123 (1): 23-42.

Martindale, M. Q., Finnerty, J. R., and Henry, J. Q. 2002. The Radiata and the evolutionary origins of the bilaterian body plan. *Molecular Phylogenetics and Evolution* 24: 358-365.

Martindale, M. Q., Pang, K., and Finnerty, J. R. 2004. Investigating the origins of triploblasty: 'mesodermal' gene expression in a diploblastic animal, the sea anemone *Nematostella vectensis* (phylum, Cnidaria; class, Anthozoa). *Development* 131: 2463-2474.

Martinez, P., Lee, J. C., and Davidson, E. H. 1997. Complete sequence of *SpHox8* and its linkage in the single Hox gene cluster of *Strongylocentrotus purpuratus*. *Journal of Molecular Evolution* 44: 371-377.

McGinnis, W. and Krumlauf, R. 1992. Homeobox genes and axial patterning. *Cell* 68: 283-302.

Merabet, S., Hudry, B., Saadaoui, M., and Graba, Y. 2009. Classification of sequence signatures: a guide to Hox protein function. *BioEssays* 31 (5): 500-511.

Merabet, S., Hombria, J. C., Hu, N., Pradel, J., and Graba, Y. 2005. Hox-controlled reorganisation of intrasegmental patterning cues underlies *Drosophila* posterior spiracle organogenesis. *Development* 132 (13): 3093-3102.

Meyer, A. and Malaga-Trillo, E. 1999. Vertebrate genomics: more fishy tales about *Hox* genes. *Current Biology* 9: R210-R213.

Monteiro, A. S. and Ferrier, D. E. K. 2006. Hox genes are not always colinear. *International Journal of Developmental Biology* 2 (3): 95-103.

Monteiro, A. S., Schierwater, B., Dellaporta, S. L., and Holland, P. W. H. 2006. A low diversity of ANTP class homeobox genes in Placozoa. *Evolution & Development* 8 (2): 174-182.

Moreno, E., Nadal, M., Baguñà, J., and Martínez, P. 2009. Tracking the origins of the bilaterian *Hox* patterning system: insights from the acoel flatworm *Symsagittifera roscoffensis*. *Evolution & Development* 11 (5): 572-579.

Moreno, E., De Mulder, K., Salvenmoser, W., Ladurner, P., and Martínez, P. 2010. Inferring the ancestral function of the posterior Hox gene within the Bilateria: controlling the development of reproductive structures and the musculature in the acoel flatworm *Isodiametra pulchra*. *Evolution & Development* In press.

Morgan, R., Rieden, P. I. d., Hooiveld, M. H. W., and Durston, A. J. 2000. Identifying HOX paralog groups by the PBX-binding region. *Trends in Genetics* 16 (2): 66-67.

Mulley, J. F., Chiu, C.-h., and Holland, P. W. H. 2006. Breakup of a homeobox cluster after genome duplication in teleosts. *Proceedings of the National Academy of Sciences of the United States of America* 103 (27): 10369-10372.

Nichols, S. A., Dirks, W., Pearse, J. S., and King, N. 2006. Early evolution of animal cell signaling and adhesion genes. *Proceedings of the National Academy of Sciences of the United States of America* 103 (33): 12451-12456.

Niehrs, C. 2006. Function and biological roles of the Dickkopf family of Wnt modulators. *Oncogene* 25 (57): 7469-7481.

Nielsen, C. 1995. *Animal evolution: interrelationships of the living phyla*. Oxford: Oxford Univ. Press.

Nielsen, C. and Martinez, P 2003. Patterns of gene expression: homology or homocracy? *Developmental Genes Evolution* 213 (3): 149-154.

Nonchev, S., Maconochie, M., Gould, A., Morrison, A., and Krumlauf, R. 1997. Cross-regulatory interactions between Hox genes and the control of segmental expression in the vertebrate central nervous system. *Cold Spring Harbor Symposia on Quantitative Biology* 62: 313-323.

Nüsslein-Volhard, C., Frohnhöfer, H. G., and Lehmann, R. 1987. Determination of anteroposterior polarity in *Drosophila*. *Science* 238 (4834): 1675-1681.

Ogishima, S. and Tanaka, H. 2007. Missing link in the evolution of *Hox* clusters. *Gene* 387: 21-30.

Pang, K. and Martindale, M. Q. 2008. Developmental expression of homeobox genes in the ctenophore *Mnemiopsis leidyi*. *Development Genes and Evolution* 218: 307-319.

Paps, J., Baguñà, J., and Riutort, M. 2009a. Lophotrochozoa internal phylogeny: new insights from an up-to-date analysis of nuclear ribosomal genes. *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain)* 276: 1245-1254.

Paps, J., Baguñà, J., and Riutort, M. 2009b. Bilaterian phylogeny: A broad sampling of 13 nuclear genes provides a new lophotrochozoa phylogeny and supports a paraphyletic basal Acoelomorpha. *Molecular Biology & Evolution* 26: 2397-2406.

Paré, A., Lemons, D., Kosman, D., Beaver, W., Freund, Y., and McGinnis, W. 2009. Visualization of individual *Scr* mRNAs during *Drosophila* embryogenesis yields evidence for transcriptional bursting. *Current Biology* 19: 2037-2042.

Pasquinelli, A. E., McCoy, A., Jimenez, E., Saló, E., Ruvkun, G., Martindale, M. Q., and Baguñà, J. 2003. Expression of the 22 nucleotide *let-7* heterochronic RNA throughout the Metazoa: a role in life history evolution? *Evolution & Development* 5 (4): 372-378.

Pearson, J. C., Lemons, D., and McGinnis, W. 2005. Modulating Hox gene functions during animal body patterning. *Nature Reviews Genetics* 6: 893-904.

Pedersen, K. J. 1964. The cellular organization of *Convoluta convoluta*, an acoel turbellarian: a cytological, histochemical and fine structural study. *Zellforschung Und Mikroskopische Anatomie* 64: 655-687.

Peterson, K. J. and Sperling, E. A. 2007. Poriferan ANTP genes: primitively simple or secondarily reduced? *Evolution & Development* 9 (5): 405-408.

Peterson, K. J., Cameron, R. A., and Davidson, E. H. 2000b. Bilaterian origins: significance of new experimental observations. *Developmental Biology* 219: 1-17.

Peterson, K. J., McPeek, M. A., and Evans, D. A. 2005. Tempo and mode of early animal evolution: inferences from rocks, Hox, and molecular clocks. *Paleobiology* 31 (Suppl.): 36-55.

Peterson, K. J., Irvine, S. Q., Cameron, R. A., and Davidson, E. H. 2000a. Quantitative assessment of *Hox* complex expression in the indirect development of the polychaete annelid *Chaetopterus sp*. *Proceedings of the National Academy of Sciences of the United States of America* 97 (9): 4487-4492.

Pfister, D., Mulder, K. D., Philipp, I., Kuales, G., Hrouda, M., Eichberger, P., Borgonie, G., Hartenstein, V., and Ladurner, P. 2007. The exceptional stem cell system of *Macrostomum lignano*: Screening for gene expression and studying cell proliferation by hydroxyurea treatment and irradiation. *Frontiers in Zoology* 4, 9.

Philippe, H., Brinkmann, H., Martinez, P., Riutort, M., and Baguñà, J. 2007. Acoel flatworms are not platyhelminthes: evidence from phylogenomics. *PloS one* 2 (8): e717.

Plickert, G., Jacoby, V., Frank, U., Muller, W. A., and Mokady, O. 2006. Wnt signalling in hydroid development: Formation of the primary body axis in embryogenesis and its subsequent patterning. *Developmental Biology* 298: 368-378.

Pöpperl, H., Bienz, M., Studer, M., Chan, S.K., Aparicio, S., Brenner, S., Mann, R.S., and Krumlauf, R. 1995. Segmental expression of *Hoxb-1* is controlled by a highly conserved autoregulatory loop dependent upon *exd/pbx*. *Cell* 81 (7): 1031-1042.

Quiquand, M., Yanze, N., Schmich, J., Schmid, V., Galliot, B., and Piraino, S. 2009. More constraint on *ParaHox* than *Hox* gene families in early metazoan evolution. *Developmental Biology* 328: 173-187.

Raff, R. A. 2000. Evo–devo: the evolution of a new discipline. *Nature Review Genetics* 1: 74-79.

Raff, R. A. 2008. Origins of the other metazoan body plans: the evolution of larval forms. *Philosophical transactions of the Royal Society of London*. *Series B, Biological sciences* 363: 1473-1479.

Raman, V., Martensen, S. A., Reisman, D., Evron, E., Odenwald, W. F., Jaffee, E., Marks, J., and Sukumar, S. 2000. Compromised HOXA5 function can limit p53 expression in human breast tumours. *Nature* 405: 974-978.

Rangini, Z., Frumkin, A., Shani, G., Guttmann, M., Eyal-Giladi, H., Gruenbaum, Y., and Fainsod, A. 1989. The chicken homeo box genes *CHox1* and *CHox3*: cloning, sequencing and expression during embryogenesis. *Gene* 76 (1): 61-74.

Remane, A. 1950 (1949). Die Entstehung der Metamerie der Wirbellosen, *Verh. Deutsch. Zool. Ges.*, S. 16–23.

Reuter, M., Raikova, O. I., Jondelius, U., Gustafsson, M. K. S., Maule, A. G., and Halton, D. W. 2001. Organisation of the nervous system in the Acoela: an immunocytochemical study. *Tissue & Cell* 33 (2): 119-128.

Rieger, R. M., Tyler, S., Smith III, J. P. S., and Rieger, G. E. 1991. Platyhelminthes: Turbellaria. In: F.W. Harrison and B. J. Bogitsh, Editors, *Microscopic anatomy of invertebrates. Platyhelminthes and nemertinea* Vol. 3. Manchester, New York: Wiley-Liss, Inc. pp. 7-140.

Rogulja-Ortmann, A. and Technau, G. M. 2008. Multiple roles for *Hox* genes in segment-specific shaping of CNS lineages. *Fly* 2 (6): 316-319.

Ruiz-Trillo, I., Riutort, M., Littlewood, D. T. J., Herniou, E. A., and Baguñà, J. 1999. Acoel flatworms: earliest extant bilaterian metazoans, not members of platyhelminthes. *Science* 283: 1919-1923.

Ruiz-Trillo, I., Riutort, M., Fourcade, H. M., Baguñà, J., and Boore, J. L. 2004. Mitochondrial genome data support the basal position of Acoelomorpha and the polyphyly of the Platyhelminthes. *Molecular Phylogenetics and Evolution* 33: 321-332.

Ruiz-Trillo, I., Paps, J., Loukota, M., Ribera, C., Jondelius, U., Baguñà, J., and Riutort, M. 2002. A phylogenetic analysis of myosin heavy chain type II sequences corroborates that Acoela and Nemertodermatida are basal bilaterians. *Proceedings of the National Academy of Sciences of the United States of America* 99 (17): 11246-11251.

Ryan, J. F. and Baxevanis, A. D. 2007. Hox, Wnt, and the evolution of the primary body axis: insights from the early-divergent phyla. *Biology Direct* 2:37.

Ryan, J. F., Mazza, M. E., Pang, K., Matus, D. Q., Baxevanis, A. D., Martindale, M. Q., and Finnerty, J. R. 2007. Pre-bilaterian origins of the Hox cluster and the Hox code: Evidence from the sea anemone, *Nematostella vectensis*. *PloS one* 1 (1): e153.

Saitou, M., Payer, B., Lange, U. C., Erhardt, S., Barton, S. C., and Surani, M. A. 2003. Specification of germ cell fate in mice. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 358: 1363–1370.

Salsi, V., Ferrari, S., Ferraresi, R., Cossarizza, A., Grande, A., and Zappavigna, V. 2009. HOXD13 binds DNA replication origins to promote origin licensing, and is inhibited by Geminin. *Molecular and cellular biology* 29 (21): 5775–5788.

Salvini-Plawen, L. V. 1978. On the origin and evolution of the lower Metazoa. *Journal of Zoological Systematics and Evolutionary Research* 16 (1): 40-87.

Santini, S., Boore, J. L., and Meyer, A. 2003. Evolutionary conservation of regulatory elements in vertebrate *Hox* gene clusters. *Genome Research* 13: 1111-1122.

Schubert, F. R., Nieselt-Struwet, K., and Gruss, P. 1993. The Antennapedia-type homeobox genes have evolved from three precursors separated early in metazoan evolution. *Proceedings of the National Academy of Sciences of the United States of America* 90 (1): 143-147.

Seaver, E. C. and Kaneshige, L. M. 2006. Expression of 'segmentation' genes during larval and juvenile development in the polychaetes *Capitella sp. I* and *H. elegans*. *Developmental Biology* 289: 179-194.

Sedgwick, A. 1884. On the origin of metameric segmentation and some other morphological questions. *The Quarterly journal of microscopical science* 21: 43-82.

Seipel, L. and Schmid, V. 2006. Mesodermal anatomies in cnidarian polyps and medusae. *The International Journal of Developmental Biology* 50: 589-599.

Selosse, M. A. 2000. Un exemple de symbiose algue-invertébré à Belle-Isle-en-Mer: La planaire *Convoluta roscoffensis* et la prasinophycée *Tetraselmis convolutae*. *Acta Botanica Gallica* 147 (4): 323-331.

Sempere, L. F., Martinez, P., Cole, C., Baguñà, J., and Peterson, K. J. 2007. Phylogenetic distribution of microRNAs supports the basal position of acoel flatworms and the polyphyly of Platyhelminthes. *Evolution & Development* 9 (5): 409-415.

Seo, H.-C., Edvardsen, R. B., Maeland, A. D., Bjordal, M., Jensen, M. F., Hansen, A., Flaat, M., Weissenbach, J., Lehrach, H., Wincker, P., Reinhardt, R., and Chourrout, D. 2004. *Hox* cluster disintegration with anteroposterior order of expression in *Oikopleura dioica*. *Nature* 431: 67-71.

Seto, A. G., Kingston, R. E., and Lau, N. C. 2007. The coming of age for Piwi proteins. *Molecular Cell* 26 (5): 603-609.

Signor, P. W. and Vermeij, G. J. 1994. The plankton and the benthos: origins and early history of an evolving relationship. *Paleobiology* 20: 297-319.

Sikes, J. M. and Bely, A. E. 2008. Radical modification of the A–P axis and the evolution of asexual reproduction in *Convolutriloba* acoels. *Evolution & Development* 10 (5): 619-631.

Sikes, J. M. and Bely, A. E. 2009. Making heads from tails: Development of a reversed anterior–posterior axis during 3 budding in an acoel. *Developmental Biology* 338 (1): 86-97.

Sly, B. J., Snoke, M. S., and Raff, R. A. 2003. Who came first? Origins of bilaterian metazoan larvae. *International Journal of Developmental Biology* 47: 623-632.

Smith, J. P. S. and Bush, L. 1991. *Convoluta pulchra* n. sp. (Turbellaria: Acoela) from the East Coast of North America. *Transactions of the American Microscopical Society* 110 (1): 12-26.

Smith, J. P. S. and Tyler, S. 1985. The acoel turbellarians: Kingpins of metazoan evolution or a specialized offshoot? In *Conway Morris, S.; George, C.D.; Gibson, R.; Platt, R.H. (ed), The Origins and Relationships of Lower Invertebrates*. Oxford: Clarendon Press. pp. 123-142.

Smith, J. P. S., Tyler, S., Rieger, R. M. 1986. Is the Turbellaria polyphyletic? *Hydrobiologia* 132: 13-21.

Soshnikova, N. and Duboule, D. 2009. Epigenetic temporal control of mouse *Hox* genes in vivo. *Science* 324 (5932): 1320-1323.

Su, L., Hershberger, R. J., and Weissman, I. L. 1993. LYAR, a novel nucleolar protein with zinc finger DNA-binding motifs, is involved in cell growth regulation. *Genes & Development* 7 (5): 735-748.

Swalla, B. J. 2006. Building divergent body plans with similar genetic pathways. *Heredity* 97: 235-243.

Tarchini, B. and Duboule, D. 2006. Control of *Hoxd* genes collinearity during early limb development. *Developmental Cell* 10: 93-103.

Tautz, T. 2004. Segmentation. *Developmental Cell* 7: 301–312.

Telford, M. J. 2008. Xenoturbellida: the fourth Deuterostome phylum and the diet of worms. *Genesis* 46 (11): 580-586.

Telford, M. J., Lockyer, A. E., Cartwright-Finch, C., and Littlewood, D. T. J. 2003. Combined large and small subunit ribosomal RNA phylogenies support a basal position of the acoelomorph flatworms. *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain)* 270: 1077-1083.

Tschopp, P., Tarchini, B., Spitz, F., Zakany, J., and Duboule, D. 2009. Uncoupling time and space in the collinear regulation of *Hox* genes. *PloS Genetics* 5 (3): e1000398.

Tyler, S., Schilling, S., Hooge, M., and Bush, L. F. 2006. Turbellarian taxonomic database. Version 1.5. Available at http://turbellaria.umaine.edu.

Valentine, J. W. and Collins, A. G. 2000. The significance of moulting in ecdysozoan evolution. *Evolution & Development* 2 (3): 152-156.

Valentine, J. W., Erwin, D. H., and Jablonski, D. 1996. Developmental evolution of metazoan bodyplans: the fossil evidence. *Developmental Biology* 173: 373-381.

von Graff, L. 1891. *Die Organisation der Turbellaria Acoela*. Leipzig: W. Engelmann.

von Graff, L. 1882. *Monographie der Turbellarien. I. Rhabdocoela*. Leipzig: W. Engelmann.

Wada, H., Garcia-Fernandez, J., and Holland, P. W. H. 1999. Colinear and segmental expression of amphioxus Hox genes. *Development* 213: 131-141.

Wahba, G. M., Hostikka, S. L., and Carpenter, E. M. 2001. The paralogous *Hox* genes *Hoxa10* and *Hoxd10* interact to pattern the mouse hindlimb peripheral nervous system and skeleton. *Developmental Biology* 231 (1): 87-102.

Wallberg, A., Thollesson, M., Farris, J. S., and Jondelius, U. 2004. The phylogenetic position of the comb jellies (Ctenophora) and the importance of taxonomic sampling. *Cladistics* 20: 558-578.

Wallberg, A., Curini-Galletti, M., Ahmadzadeh, A., and Jondelius, U. 2007. Dismissal of Acoelomorpha: Acoela and Nemertodermatida are separate early bilaterian clades. *Zoologica Scripta* 36: 509-523.

Waskiewicz, A. J., Rikhof, H. A., and Moens, C. B. 2002. Eliminating zebrafish Pbx proteins reveals a hindbrain ground state. *Developmental Cell* 3: 723-733.

Wellik, D. M. 2007. *Hox* patterning of the vertebrate axial skeleton. *Developmental Dynamics* 236 (9): 2454-2463.

Wheeler, B. M., Heimberg, A. M., Moy, V. N., Sperling, E. A., Holstein, T. W., Heber, S., and Peterson, K. J. 2009. The deep evolution of Metazoan microRNA. *Evolution & Development* 11 (1): 50-68.

Wikramanayake, A. H., Hong, M., Lee, P. N., Pang, K., Byrum, C. A., Bince, J. M., Xu, R., and Martindale, M. Q. 2003 An ancient role for nuclear β-catenin in the evolution of axial polarity and germ layer segregation. *Nature* 426: 446-450.

Willmer, P. 1990. *Invertebrate relationships: patterns in animal evolution*. Cambridge; New York: Cambridge University Press.

Wu, L. H. and Lengyel, J. A. 1998. Role of *caudal* in hindgut specification and gastrulation suggests homology between *Drosophila* amnioproctodeal invagination and vertebrate blastopore. *Development* 125: 2433-2442.

Yanze, N., Spring, J., Schmidli, C., and Schmid, V. 2001. Conservation of *Hox/ParaHox*-related genes in the early development of a cnidarian. *Developmental Biology* 236 (1): 89-98.

Zhang, J. and Nei, M. 1996. Evolution of Antennapedia-class homeobox genes. *Genetics* 142: 295-303.

6

APPENDICES

Appendix A

Publications derived from this thesis

1. **Eduardo Moreno, Marga Nadal, Jaume Baguñà and Pedro Martínez.** 2009. Tracking the origins of the bilaterian Hox patterning system: insights from the acoel flatworm *Symsagittifera roscoffensis*. *Evolution & Development* **11** (5): 572-579

2. **Eduardo Moreno, Katrien De Mulder, Willi Salvenmoser, Peter Ladurner and Pedro Martínez.** 2010. Inferring the ancestral function of the posterior Hox gene within the Bilateria: controlling the development of reproductive structures and the musculature in the acoel flatworm *Isodiametra pulchra*. *Evolution & Development* **12** (3): 258-266

Appendix B

1. Table with the list of primers used to extend the homeodomain sequences towards the 3' and 5' ends using the Smart RACE cDNA Amplification technology (Clontech) from *Symsagittifera roscoffensis* cDNA.

RACE	Primer name	Sequence 5'→ 3'
3'	SrHox1F1	TCTTACACGAGCTAGGAGGATC
3'	SrHox1F2	TAGCTACTTCGCTGACCTTGAA
3'	SrHox5F1	AGTGGCGGCGGAATTAGGGCAACAC
3'	SrHox1F5	TGGCACTTCGGTTTTCTACCCACTTCGC
3'	SrHoxPostF1	GCAGCATTTCACCCAGCTCTCTCTACGC
3'	SrHoxPostF2	GCAGCATTTCACCCAGCTCTCTCTACGC
5'	SrHox1R1	CACCTTCGACTTCTAGTTAAGCATCTTTTGC
5'	SrHox1R2	CCTGAGTCTCATTCAAGGTCAGCGAAGTAG
5'	SrHox5R1	GCCTCTGTTTGTTCACGCAGTTGTGATG
5'	SrHox5R2	CCATATTTTGATCTGTCGCTCCGTCAGC
5'	SrHoxPostR1	TGATGAGGCGGATGTTGTTGAGGCAC
5'	SrHoxPostR2	CGGAGTGCTGCTTTTTGTTCTTCATCCTC

2. Table with the list of primers used to extend the homeodomain sequences towards the 3' and 5' ends using the Smart RACE cDNA Amplification technology (Clontech) from *Isodiametra pulchra* cDNA.

RACE	Primer name	Sequence 5'→ 3'
3'	IpHox1F1	TGATTCGGATCCCTGGAACTGGAG
3'	IpHox1F2	AATTCCACTTCAACCGCTACCTCACC
3'	IpHox1F3	CAACCTCAACGAGACCCAGGTCAAGA
3'	IpHox5F1	ACCTAACCCGCCGGAGAGCATC
3'	IpHox5F2	GCATCGAGATCGCCAACCTCCTC
3'	IpHox5F3	CGCCCTCTCCGAGCGACAGATA
3'	IpHoxPostF1	AGACTGGAAATCGCCCGCAGTC
3'	IpHoxPostF2	TGGAAATCGCCCGCAGTCTCAG
3'	IpHoxPostF3	GCCCGCAGTCTCAGCCTCAC
5'	IpHox1R1	TCCACTTCAACCGCTACCTCACCCG
5'	IpHox1R2	GAACCGCCGCATGAAGCAGAAGA
5'	IpHox1R3	TTTAGTCTTCGGTCGTTGCCGCG
5'	IpHoxPostR1	AAATCGCCCGCAGTCTCAGCCT
5'	IpHoxPostR2	TGGAAATCGCCCGCAGTCTCAG
5'	IpHoxPostR3	AGAAGACTGGAAATCGCCCGCA
5'	IpHoxPostR4	TCAAACCAGTCGCGCTCTCCTGAA
5'	IpHoxPostR5	GCACATGGTTCACCCAGGAGTGGT

Appendix C

Sequences

1. EMBL/GenBank accession numbers for Hox genes of *Symsagittifera roscoffensis* and *Isodiametra pulchra*.

Gene	gb accession number	gi accession number
SrHox1	FJ619530.1	222876505
SrHox5	FJ619532.1	222876509
SrHoxPost	FJ619531.1	222876507
IpHox1	FJ619527.1	222876499
IpHox5	FJ619528.1	222876501
IpHoxPost	FJ619529.1	222876503

2. GenBank accession numbers for the 3 genomic regions adjacent to Hox genes in the acoel *Symsagittifera roscoffensis*.

Gene	gb accession number
SrHox1	HM177432
SrHox5	HM177433
SrHoxPost	HM177434

3. *Symsagittifera roscoffensis* Hox cDNA sequences (assembled from RACE fragments) plus ORFs predicted by means of the GENSCAN software from BAC clone sequences. The corresponding amino acid sequence for each protein is shown after the DNA sequence.

SrHox1

cDNA sequence containing the ORF predicted by GENSCAN. Colour code: yellow, final exon predicted from the genomic sequence; grey, cDNA sequence not found in the genome because the BAC clone ends just in the intron before the final exon; red, stop codon. The number of exons in this gene is thus unknown and the *SrHox1* sequence is incomplete at the N-terminus.

```
AACTTCAAACTTTGGATCTCACGGGTACCACAACCTAACTCCGACTCCGTCATTTTACTCAAAC
TATCACAATTCTAACCAATCTGGCTCCAATTATAACATTTACACCAATGGGCTCAACTCCAACT
CAACATACAATCTCGGTTCGGCGCCCTCCAATTACTTGCCCGGAAACGAGACTAATCCTCAAAG
TTACTCGAAATCGGCAAGTTATGACTTCGGATCCGCGGGAGCCTACGTCTCTTCGGACCTGCTC
TACCCTTCAAACACAATTCTGCACGAGTATTTTTTCCCCAATAACAACCCTCAGTACTACAGCT
ACCCTTCCACGGCTAACAATGGGCACGCACAACAAGCATATCCCACTAACGTCTTGTCCAACTA
CATGGCCAGAGGTGGTACTAACCCTGGTTCTCAGACACACGCGCCAACTGCGGCCGTAAACATA
AATGTCAGCTGTAATATAGCGGGAAACGGACCCACTTCCATGAGTCCCGAAGAACACTCGAGTT
CGCAGACCACTCCACATGGCGGGAACCACTCGGAACATACGACATCACCAGTAGCCAAGTACAA
CACACATACAAGCCCTGTATACGACTACATAGATGATATGAAACAATCCAACTTTGCCAATATT
TGCAATAACAATCCAATGTATCCAGGAACTATTCCCCACGAGACGTCGCCGGAAGCTGTAGGAT
ACCCTTCTGGCTTGCACATGGTTGAATCGGACTACATGAAGACGGTCCCAGGGTCACAAGAAAC
TAGAAAACTCGGTATTCAGGGGGGTCACTTGTTGGGATCCGAGGGTAAACTAGGCTCGGAAATG
GTGGTGACCCATCACCCTCACGTGGCAAATGCGACCATATCAAATGCTGCTGCTAATCATGTGG
CACATGCTGCAAATGGAACTGCCGCTGTCCCTCCCTTCTCTTGGATGAAGATCAGAAGGAACCA
GTCACACAACTTATGGGGCAAAAGCGCATCTCTGACACAATACCACGCGGGTGCAGCTGGTACG
CACTTAGCCAGCGGATGTCACCTGGGACCAGGAGTGCAAACCAGGGGAGGGAGGACGAACTTCA
CAAACAAACAGCTGACAGAACTGGAAAAGGAGTTTCACTTCAATAGATATCTTACACGAGCTAG
GAGGATCGAGATAGCTACTTCGCTGACCTTGAATGAGACTCAGGTCAAGATATGGTTTCAAAAC
AGAAGAATGAAGCAGAAGAAACTGCTTAAAGAAGGAAAATTGAGCTAA
```

```
TSNFGSHGYHNLTPTPSFYSNYHNSNQSGSNYNIYTNGLNSNSTYNLGSAPSNYLPGNETNPQS
YSKSASYDFGSAGAYVSSDLLYPSNTILHEYFFPNNNPQYYSYPSTANNGHAQQAYPTNVLSNY
MARGGTNPGSQTHAPTAAVNINVSCNIAGNGPTSMSPEEHSSSQTTPHGGNHSEHTTSPVAKYN
THTSPVYDYIDDMKQSNFANICNNNPMYPGTIPHETSPEAVGYPSGLHMVESDYMKTVPGSQET
RKLGIQGGHLLGSEGKLGSEMVVTHHPHVANATISNAAANHVAHAANGTAAVPPFSWMKIRRNQ
SHNLWGKSASLTQYHAGAAGTHLASGCHLGPGVQTRGGRTNFTNKQLTELEKEFHFNRYLTRAR
RIEIATSLTLNETQVKIWFQNRRMKQKKLLKEGKLS
```

SrHox5

ORFs predicted by GENSCAN (all) plus cDNA sequence (in bold). Colour code: grey, exon1; yellow, exon2; blue, exon 3; green, start codon; red, stop codon.

ATGTACGAAAACAATCCGTCGCCAACTGCCTGCTCACCCAGTCCGAACTCAACCCAATC
ACAATTAGTTCTCAGCAACCAAAACCATCTCCAAAATGCAACCAGTAATCATCAATCTT
TGGGCGGAAATTTCTACTCCCAAAGCAATCAAAACGGACCAGCAGCTGTTGATTATAAT
CTTCACGTGACCTCATTAACTGTTAACTACCCCGCGATAGCACCCCCACCACCACCACC
CGCCGGATCcACCTCAGTCGTCCCGGAAGCGCAtCAgCACcACGGGAACGCGCC**ATCGA**
aCCCACTGGCGGCGGCATATCAGAGCACTCCTTACTACTCGCCcGCCAGCGCAGTGGCG
GCGGAATTAGGGCAACACTATCCACACTCACTTTTACCCGCAGCTGCAGCTACCCTGTC
TTCATACAACCCTTATTACCCCTCAATGGCCGCGACCCAAATGTACCAGTCATCCGGCG
CTCTTCAATCCAACCCAGCTTTACAACAGATTACGAACTCAGCCACCGCTGGATCAAAC
CCGATGAACTGGCACTTCGGTTTTCTGCCCACTTCGCTTGATGCCAGATCAATGGGCGC
TAATATTTCACACCACGCAAGTTCGGGATCTTTGAAACACTCCCAGATCAACGAAGCAC
CCCACCATTTTGGACACCAGTTCAAGTTCCAGGGTTCCAATCCCGGTCTGGACAGCAAT
GGGTTGATGATGATTGCGTCAGGAGGAGAGGGTAATGGATCCGAATTAGATGGGGGAGA
TCACATGAATACTTGCCACGGTCAGATTTCGCCCTCGCATAACCAGAACGGATCGAAAG
TGTTTGCGTGGATGAGCAAAAGACCACCGACAGACTGCAAACGCACGCGCACTGCCTAC
ACGCGGTTCCAGACACTTGAGCTGGAGAAAGAGTTCCACTTCAACCGGTACCTCACTCG
CAGACGGAGGATAGAGATCGCCAATCTACTCGCACTGACTGAGCGACAGATCAAAATAT
GGTTCCAGAACAGACGAATGAAGTGGAAAAAAGACAACAACCTGAAGAGCATGTCTCAA
ATAGACTCCATCACAACTGCGTGA

MYENNPSPTACSPSPNSTQSQLVLSNQNHLQNATSNHQSLGGNFYSQSNQNGPAAVDYN
LHVTSLTVNYPAIAPPPPPPAGSTSVVPEAHQHHGNAP**SNPLAAAYQSTPYYSPASAVA**
AELGQHYPHSLLPAAAATLSSYNPYYPSMAATQMYQSSGALQSNPALQQITNSATAGSN
PMNWHFGFLPTSLDARSMGANISHHASSGSLKHSQINEAPHHFGHQFKFQGSNPGLDSN
GLMMIASGGEGNGSELDGGDHMNTCHGQISPSHNQNGSKVFAWMSKRPPTDCKRTRTAY
TRFQTLELEKEFHFNRYLTRRRRIEIANLLALTERQIKIWFQNRRMKWKKDNNLKSMSQ
IDSITTA

SrHoxPost

ORFs predicted by GENSCAN (all) plus cDNA sequence (in bold). Colour code: grey, exon1; yellow, exon2; green, start codon; red, stop codon.

ATGAGTTgGGGGGCATCTGCGCACtACGCAACTCACGACAACAACACGGCAACACAACA
TCACCTAAACCACTTTGCCACCGCTCCTCAGCTTCCGGTCGAGTTTGACCCCTTGAATA
GGTCAAATAACAAATCTTCAGGGGTCAATGACCTCAATTTGAACTCCTCAGGTCAACAA
ATTACACCTGTAGATACAATGTTAGCGTCACCTGGCTACGTGAACACCCAAAACTTCAT
GGCAAATcAATCGGTCGATTCGAAGCTTtCGCCTTtCGAAGACGCCGGCCAGAGCCAAG
CGCAGCAGTTGTACTGGACCCACGCGAACGCCTATTCGAATTACTACCACTCAAACATG
GCTGCGATCGATGGGTGCAATAGGGTCGGGTTCTACGGGTCACGTGATACTTCAAAATT
CATGCCCAATCCAGCAGCCGCCGCTAATTTAAACTTATACGAGCAGCACATTAATAACT
ACCGAAACATGGTCACAGGGGCAAATAACCCAAACGGGTTGGTAATTAACTCTTCATTT
TTGCCCCTTTCCAACCACCCGCATACTAATAACAACAACAACAACAACAACAATATTGT
ACAGCATCCACATAACAATCACACCACGTACGATCATACCAACCAAACGGCGTTCGAGT
TCGGAACGAGTTTCAACTCCCAGATTCAAAACAACGCCATCATTCATACGAACAATGTG
AATTTCGTGTCGCCTTCGCTGCTCGGCGCCACTACCAACACCAGCATCAGCCAATCAAA
TCAGAGTACATTAGTCACAGCTGTACCACCGGTGAACCAATCACAAGCGGGCACCCCAA
TTACGCCAACCAACCGGAACGCTTCCTGGTCAACAACCAATCACAACACAGCTGCAGCG
GACCAACGCAACTCGCTGGCTCCAAACGTCAAAACTTGCATCAAAAAGGAGGCGAGCCT
AAGTCAGGGATCCGGAAGCCCCCCGCGAGAGGTTtccGGGGGGTCTCTTGgAGTGGgAG
GCTgCTTC**AACTCGgCAGAGAGCTCgCCGgCCGaTTCGtCGACTGAATCTGTCGGcAaT**
ATTATGAACTCGCCAGAaCACGCcGGCaGCATTTcACCCAGCTCTCTCTACGCTCAGAT
GACCTCTGACCcTTCGGgTGTTCTGCAGGGAAATCAACAGGTCATGTCATCCAGCggTC
ACTCACAGGGTTCCCACAATCACCACCTGCTCATGACACCCGGAAACACTTCAGTGCCG
CCCCAGACACCCTCCAATGGCGGGACCAACCCCTCGTCCGGGGGAGGCATGGCATGGAT
GGCCAGAAACGTCAGTCGGAAAAAGCGAAGACCATACACCAAGAACCAAACTCTGGAAT
TGGAGAAAGAGTTtCTCTTCAaCACGTACaTAACcAGAGAACGTAGAcTGGAAaTAGCG
AGATCACTCAACCTGACAGATCGCCAGGTCAAGATCTGGTtCCAAAACCGGAGGAtGAA
GAACAAAAaGCAGCACTCCGGAGGACCTGGTATGGACCCccGATTCCTGGTCATCCGG
CGATGCACTtGAtGGTGCCTCAACAACATCCGCCTCATCACATGTTGTGA

MSWGASAHYATHDNNTATQHHLNHFATAPQLPVEFDPLNRSNNKSSGVNDLNLNSSGQQ
ITPVDTMLASPGYVNTQNFMANQSVDSKLSPFEDAGQSQAQQLYWTHANAYSNYYHSNM
AAIDGCNRVGFYGSRDTSKFMPNPAAAANLNLYEQHINNYRNMVTGANNPNGLVINSSF
LPLSNHPHTNNNNNNNNNIVQHPHNNHTTYDHTNQTAFEFGTSFNSQIQNNAIIHTNNV
NFVSPSLLGATTNTSISQSNQSTLVTAVPPVNQSQAGTPITPTNRNASWSTTNHNTAAA
DQRNSLAPNVKTCIKKEASLSQGSGSPPREVSGGSLGVGGCF**NSAESSPADSSTESVGN**
IMNSPEHAGSISPSSLYAQMTSDPSGVLQGNQQVMSSSGHSQGSHNHHLLMTPGNTSVP
PQTPSNGGTNPSSGGGMAWMARNVSRKKRRPYTKNQTLELEKEFLFNTYITRERRLEIA
RSLNLTDRQVKIWFQNRRMKNKKQHSGGPGMGPPIPGHPAMHLMVPQQHPPHHML

4. ***Symsagittifera roscoffensis*** **Hox cDNA sequences assembled from RACE fragments only.**

ORF in bold, 3’ UTR in italic, stop codon in red.

SrHox1

AACTTCAAACTTTGGATCTCACGGGTACCACAACCTAACTCCGACTCCGTCATTTTACT
CAAACTATCACAATTCTAACCAATCTGGCTCCAATTATAACATTTACACCAATGGGCTC
AACTCCAACTCAACATACAATCTCGGTTCGGCGCCCTCCAATTACTTGCCCGGAAACGA
GACTAATCCTCAAAGTTACTCGAAATCGGCAAGTTATGACTTCGGATCCGCGGGAGCCT
ACGTCTCTTCGGACCTGCTCTACCCTTCAAACACAATTCTGCACGAGTATTTTTTCCCC
AATAACAACCCTCAGTACTACAGCTACCCTTCCACGGCTAACAATGGGCACGCACAACA
AGCATATCCCACTAACGTCTTGTCCAACTACATGGCCAGAGGTGGTACTAACCCTGGTT
CTCAGACACACGCGCCAACTGCGGCCGTAAACATAAATGTCAGCTGTAATATAGCGGGA
AACGGACCCACTTCCATGAGTCCCGAAGAACACTCGAGTTCGCAGACCACTCCACATGG
CGGGAACCACTCGGAACATACGACATCACCAGTAGCCAAGTACAACACACATACAAGCC
CTGTATACGACTACATAGATGATATGAAACAATCCAACTTTGCCAATATTTGCAATAAC
AATCCAATGTATCCAGGAACTATTCCCCACGAGACGTCGCCGGAAGCTGTAGGATACCC
TTCTGGCTTGCACATGGTTGAATCGGACTACATGAAGACGGTCCCAGGGTCACAAGAAA
CTAGAAAACTCGGTATTCAGGGGGGTCACTTGTTGGGATCCGAGGGTAAACTAGGCTCG
GAAATGGTGGTGACCCATCACCCTCACGTGGCAAATGCGACCATATCAAATGCTGCTGC
TAATCATGTGGCACATGCTGCAAATGGAACTGCCGCTGTCCCTCCCTTCTCTTGGATGA
AGATCAGAAGGAACCAGTCACACAACTTATGGGGCAAAAGCGCATCTCTGACACAATAC
CACGCGGGTGCAGCTGGTACGCACTTAGCCAGCGGATGTCACCTGGGACCAGGAGTGCA
AACCAGGGGAGGGAGGACGAACTTCACAAACAAACAGCTGACAGAACTGGAAAAGGAGT
TTCACTTCAATAGATATCTTACACGAGCTAGGAGGATCGAGATAGCTACTTCGCTGACC
TTGAATGAGACTCAGGTCAAGATATGGTTTCAAAACAGAAGAATGAAGCAGAAGAAACT
GCTTAAAGAAGGAAAATTGAGCTAA*AGTGGAAAATGCTCAGTTCGGATCAGCTACGGCT*
CGAAAATGGCGCAAAAGATGCTTAACTAGAAGTCGAAGGTGTAACAGAATTTTTATTTG
ACGTGGAGTCATTTTGTATTAATTTGTAATTACTACAATTGCTGCAGCTTTTGATCAAA
CAACTGAGTAGAATCCAAAGTTTTAGGTTTGTATTCTGTCATCTTAGTAATTCTCACTG
ATGTTCACTGATAAACCTAACTGTTTATGAATGATAAATTATAATATATTTTTGTTCGT
TAA

SrHox5

ATCGAaCCCACTGGCGGCCGCATATCAGAGCACTCCTTACTACTCGCCcGCCAGCGCAG
TGGCGGCGGAATTAGGGCAACACTATCCACACTCACTTTTACCCGCAGCTGCAGCTACC
CTGTCTTCATACAACCCTTATTACCCCTCAATGGCCGCGACCCAAATGTACCAGTCATC
CGGCGCTCTTCAATCCAACCCAGCTTTACAACAGATTACGAACTCAGCCACCGCTGGAT
CAAACCCGATGAACTGGCACTTCGGTTTTCTGCCCACTTCGCTTGATGCCAGATCAATG
GGCGCTAATATTTCACACCACGCAAGTTCGGGATCTTTGAAACACTCCCAGATCAACGA
AGCACCCCACCATTTTGGACACCAGTTCAAGTTCCAGGGTTCCAATCCCGGTCTGGACA
GCAATGGGTTGATGATGATTGCGTCAGGAGGAGAGGGTAATGGATCCGAATTAGATGGG
GGAGATCACATGAATACTTGCCACGGTCAGATTTCGCCCTCGCATAACCAGAACGGATC
GAAAGTGTTTGCGTGGATGAGCAAAAGACCACCGACAGACTGCAAACGCACGCGCACTG
CCTACACGCGGTTCCAGACACTTGAGCTGGAGAAAGAGTTCCACTTCAACCGGTACCTC
ACTCGCAGACGGAGGATAGAGATCGCCAATCTACTCGCACTGACTGAGCGACAGATCAA
AATATGGTTCCAGAACAGACGAATGAAGTGGAAAAAAGACAACAACCTGAAGAGCATGT
CTCAAATAGACTCCATCACAACTGCGTGA*ACAAACAGAGGCTGTCGAAATGAAACAAAA*
ACAGAAACAGAAAAAAAACTTCATTAGAAAGAAGGGCTTTAAATTCGCTCATTCTTCGG
CAATTCAAAACCCAATACATGCAGGGTACTCTTAAATCTATAAAACATTGCTCGAGATA
GTCGAGTCCCAATAACTTCTTAATTTACTCCGTCAAATCGGAGATAATGTTTTGGAAAC
TTGGTGTTTAATTTAAAACAGACAGCGCGAGGAAATTTTTTATTAATGTTTGTAAATTT
TTAATATATTTTGCTAGAGCATCAAAAAATATGAATTGAAATTCTTTGAAGCAG

SrHoxPost

AACTCGgCAGAGAGCTCgCCGgCCGaTTCGtCGACTGAATCTGTCGGcAaTATTATGAA
CTCGCCAGAaCACGCcGGCaGCATTTcACCCAGCTCTCTCTACGCTCAGATGACCTCTG
ACCcTTCGGgTGTTCTGCAGGGAAATCAACAGGTCATGTCATCCAGCggTCACTCACAG
GGTTCCCACAATCACCACCTGCTCATGACACCCGGAAACACTTCAGTGCCGCCCCAGAC
ACCCTCCAATGGCGGGACCAACCCCTCGTCCGGGGGAGGCATGGCATGGATGGCCAGAA
ACGTCAGTCGGAAAAAGCGAAGACCATACACCAAGAACCAAACTCTGGAATTGGAGAAA
GAGTTtCTCTTCAaCACGTACaTAACcAGAGAACGTAGAcTGGAAaTAGCGAGATCACT
CAACCTGACAGATCGCCAGGTCAAGATCTGGTtCCAAAACCGGAGGAtGAAGAACAAAA
aGCAGCACTCCGGAGGACCTGGTATGGACCCccGATTCCTGGTCATCCGGCGATGCAC
TtGAtGGTGCCTCAACAACATCCGCCTCATCACATGTTGTGA*CTATGAGTCGGACCAAA*
TTCGAAAAAAAaCTCaTtCGAGCGCAGGAGAaGTtCCGGGAaTCGAaGTTTtCCGAGAa
TCACTTGCTTTAAGTCTCCATTCCCAGTTATGCCCGGACTTTTGCTGGGCTCATTTCCG
GTTTAATAGCTTtATTAGTAATACCAATTCGCAGTACCAAACTTTCGAAAAGCTTTATT
TTCTCTCTTCCTATTTtATTAATAGTTtCTTTCTTTTGTAAAaCTACAACTTGCTTGAT
GCAAaTGTTTTCTAAGTGCTGATTGTCTTCCATACATAATCCCTCTATTCTAaTtATAa
CAGTAAaCACTAaTTAAAaTTGAAAaTGAGCAAAaTATAATTTTGATCTATTTCAATAT
AAGTGTCTAAATTATTTTGAGATAGTTTTGTACGCATTATATCGATGAAATATGCTTGC
TTGTCTGTTGAATATCTCTTCAAAATTACTGTTATTTGAGTTCCATGCAAGGTGGGGAT
GAAAAAAAATTGAAAATTGACTTTCAAATGGTGCAGTTGCTTAAATATTGTTCTATGTA
AAAGCTAAGTAGCATATTAAATTCGTAAAAAAa

5. *Symsagittifera roscoffensis* Cdx cDNA sequence assembled from RACE fragments.

ORF in bold, 5' and 3' UTRs in italic, star codon in green, stop codon in red.

The corresponding amino acid sequence is shown after the DNA sequence.

TGGGGGATAGATCTAACTGCTCGCCGCCCGAAGGGAGAAAAAAGCAGAAGAACAGAGTATCACACAAAGG
CTTCATACAATCCTGCACAGAACCATAAACCGGATCCAGACTTTTTCAAAAATTATATTGCCAGCTCTGA
TCAGCTCTAAAACTCGTTTACAACCACATTTATAGCCGCAATTCTGGTCTGTAACGGCCATCACGTTGTA
AATTTGATATAATATTCGCAATAAAATGCATCTCGACGGTGGTAATTCCGAAGTCATCGACAGTATGAGC
CGGGAATCCATCCAGACCGGCTGGATCCATCACCTAGCACACAACTTCTCCAATGTCGCGTCCAGTGCAC
TCGTCAACAACCCGGCCGGCAACTACCCGGGATCCGGATCCACCATACCCGGACTGTCCGGCAACACTCA
TCACTCGTTGAAGTTCGCAGCCGACCATCACAGTCCGACAGGGTTGAAACTCCACTCGGATCATAGCCCA
GTCGACCCGAACAACCCTCACTCGCAATTCACCGAGCTCCGAACAGCGGCTCAGAATGCAGCCTCGGTCG
*ACGCTGCCGGAGCCTGGCCACCGACC***ATGCCCAACTCCTACTTGGCTGTTGCGGCCGCGAGTCATAACCA**
GGCGACTGGCATGATGCACCCGGCCTCTCGCAACTCGTTCGCTGCGTCGGCAGTTCACCCGAATGCATCG
TGGGCACACGCCAGGAATGCGTACGGCACGCAATTCACATCCGCTGAACAGGTTGCGGCCGCGTACAGAG
GATGCAACCTGAACATAATGAATGCCGTGGTTGCGGGAGGAGACGCCGCAGCCGCACAACACCACCACAA
CTCCACCGCCGCAGCAGCCGCAGCCGCCATGCACGCCATGCCATTCAACATGACACACCCGGCAGCCGGA
AACCCCATGGCACCAGGCGGACATCCAGGGGTACATCCGGGTCAAAACCCATTCGTCAGTCCACACCACC
CCTATGCTGCTTCAGTCGCGGCAGCTGAGATGTACGGGGATCACTCCAGTTCAACTTTCAGATATAACCC
AACAACCGGAAAGACGAGAACAAAAGACAAATATCGAGTTGTCTACACCGACCGTCAACGAGCCGAACTC
GAGAACGAATTCCGAAGTGCGCAATACATCACCATCCGACGCAAGTCCGAACTGGCCATGCAAGTCGGCT
TAAGCGAGCGACAAGTTAAGATCTGGTTCCAGAACCGACGAGCTAAGGAACGAAAGGTGTCCCGAAAGGT
TCCTGGAGGCGGGAATCACAGTTCTTCAGACATCGAAGACTCCGATAATGAAATAGACGATGAAGATGAA
GAAAACCARTTAAGGTCACCTGCTAAACACCAATCATTATCACCTGGTTCCCAAAACACATGTCATAACG
GTGTTCAGTTACCTCTAATCAAAAGCGAAAACAATAGCTTACTACGGCCGCCTTCGAATGAGTTATACGG
GGCGAAGATGTCATGA*TAATTACGTAATCATCAACGCAATAGTTACGTAATGGTCGGTACTTCATAAACC*
TTTCTTCGTTGAAATGAGGACAACTTTTAGCTCATAAACCTCAATCAGAGAACGTTTTACAGAGAACTAT
TTTTGTACGATATCACTGCAACCCTATATTGTAATCAATTTACCTGCTGGCGGAAAAAGTTGAACAAAGC
CACAAWGCCTGTTTTTTTTTCAARCCACTTCAGCGATTGCTCAAGAAACTGAACTTTTTTCAGAGCACCG
CTTMTGACGAACATAGGCCAGAAGCCAAGTAATAATTTTCGGCTGACCGCAAGATTAGAAATTTCAACCT
CCTYAAAACTTGTACACAGTTTCAAGTTTAATGAAATTCTTGAAATTAGACATACAGACTTTAAAGGGTG
GATTTCACAAATTCGATTCTGAAWTTGGAATGAAAGGAAGATCTTGATTTGTCAGATWAGAGTCGGTTAT
TTGAAGTTCGAGGAACCTGAGAACCCTGTTTTCCTTCACTCAACGTGGGGAGATCTAGTCTGGAGTTTTC
AATAGGGTTAACTTGACTCAATTTTGTCGCAACCAATAGTAAAGAAGAAACGGGAGCGAGTTATGCGAGA
TACYTGGAAATCTGATTTGACGAATTTARAAGAAAGAGACTTGACTTGGGGTTTGAGAGCTGATCTGGGG
YTYCAATTTGGGTGCTGAGATTTGAGCGGATTCTTTCTTTTGACTGGGTTTAAGTTGGTCGTTTGATATG
GATNCCCAAAAGTAGATCAAACGGGATTGTTTTGTCGCAGCTGGGGTTCTTGGCGGATCATGGCGAAAGG
TGATTAGGATCTGGGTTGTTTTAGATGAAGTTTAATTCGATTAGAGTCTGAAACGCCAGTCGTAAATTGG
ATTCGAGAACACTTCACAGCTCAGTCGCCGCAGCCACGCAGCTGGAAATCTTAGAGTGAATAAATGTCCA
TTCAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

MPNSYLAVAAASHNQATGMMHPASRNSFAASAVHPNASWAHARNAYGTQFTSAEQVAAA
YRGCNLNIMNAVVAGGDAAAAQHHHNSTAAAAAAAMHAMPFNMTHPAAGNPMAPGGHPG
VHPGQNPFVSPHHPYAASVAAAEMYGDHSSSTFRYNPTTGKTRTKDKYRVVYTDRQRAE
LENEFRSAQYITIRRKSELAMQVGLSERQVKIWFQNRRAKERKVSRKVPGGGNHSSSDI
EDSDNEIDDEDEENXLRSPAKHQSLSPGSQNTCHNGVQLPLIKSENNSLLRPPSNELYG
AKMS

6. ***Symsagittifera roscoffensis*** **PBX (extradenticle) cDNA sequence from the EST collection. The corresponding amino acid sequence is shown after the DNA sequence.**

EST clone: WH0AAA59YG18RM1

```
GCAACGCGTACGATAATCCTCAAAATATGGATGATCAACACAGATTCGCACTTGCCTTT
CTTCAACAGCAGCAGCAACAACAGATGAATGCTGCTGCAATGCACTCCGGTGTGGACCC
CTCGGGAGTCAACATGCACGGGGGTGTTTCTGTCTCCGGAGCGTCCGGAGGGTGGGGGC
ATCTTGCACACCAGGGGGTTCCCACTTCAGTAGGGCTTGGCGACGGATCCGGTCAAGAT
CTAGGCCTAGGTTCTCTGGGAAATAACTCAGGAATGTTGACGGAAGAAACTAGAAAGCG
ACAGTTGCAAGAAATCTTACAGCAGATTATGACGATAACCGAACAGTCTCTTGACGCTG
CACAGGCAAGGAAGCAGACGTTGAATATCCATAGGATGCGCCCTGCCCTCTTCAGTGTC
CTCTGTGAGATCAAAGAAAAAACGGGAACCTTTCTGAACATGCGTAATCAGAATGATGA
TGACGCACCGGACCCGCAAATAGTGCGTCTGGATAACATGTTGGCGGCCGAGGGAGTTT
CAGGGGATGGTAAGTCACCCACTGGCAGCTCTACAACCGGGGGAGCAGGACAACCAGAT
AACACTATAGAACACTCCGACTACAAGGCGAAGCTGGGGCC
```

```
NAYDNPQNMDDQHRFALAFLQQQQQQMNAAAMHSGVDPSGVNMHGGVSVSGASGGWGH
LAHQGVPTSVGLGDGSGQDLGLGSLGNNSGMLTEETRKRQLQEILQQIMTITEQSLDAA
QARKQTLNIHRMRPALFSVLCEIKEKTGTFLNMRNQNDDDAPDPQIVRLDNMLAAEGVS
GDGKSPTGSSTTGGAGQPDNTIEHSDYKAKLG
```

7. *Isodiametra pulchra* Hox cDNA sequences assembled from RACE fragments only.

ORF in bold, 3’ UTR in italic, stop codon in red.

The corresponding amino acid sequence is shown after the DNA sequence.

IpHox1

TCGTGGATGAAGATCAAGAGGAATCAACCTCATCTACAATGGTCCAAGTCGGCGACAGC
AGGGCTGCCTGGCTCTACGCACGCTTACCCGCCTGGGTCCCAGACCCGCGGAGGACGCA
CCAACTTCACCAACAAGCAGCTGACGGAGCTGGAGAAGGAGTTCCACTTCAACCGCTAC
CTCACCCGCGCCCGCCGCATCGAGATCGCCTCCAGTCTCAACCTCAACGAGACCCAGGT
CAAGATCTGGTTCCAGAACCGCCGCATGAAGCAGAAGAAGCTCGTCAAGGAGGGCAAAC
TCGCCTAG*CCAATCAAATCCAACCATTCCATTCATCAGCCGACACCTGATTCCAACCAA*
TCAGCTTTCGATAGGACTCGTTCTGTCAGTCAGATTTCAGAAGTTTAGAAGGTTCTGAC
TGAACTTTCAGTTTAGTCTTCGGTCGTTGCCGCGACCAATTATATTGAAATGGGGGAAT
GTTCGAAAGTTTTGCAAACTTGGATAGTTCAATTGATGTCTGTAAATATGTTCTGAATT
CTTGATGTATTTTTGTGTAAAACAAATTTAAAAATCTTTAGTAAAAAAAAAAAAAAAAA
AAAA

SWMKIKRNQPHLQWSKSATAGLPGSTHAYPPGSQTRGGRTNFTNKQLTELEKEFHFNRY
LTRARRIEIASSLNLNETQVKIWFQNRRMKQKKLVKEGKLA

IpHox5

ACTTCAACCGGTACCTAACCCGCCGGAGACGCATCGAGATCGCCAACCTCCTCGCCCTC
TCCGAGCGACAGATAAAGATCTGGTTCCAGAACCGACGGATGAAGTGGAAGAAGGACAA
CAACCTGAAGAGCATGTCGCAGGTCGACAACATGACCTCATAG*TCGACCCCGAGGCTCC*
CATCAAACNCAAATTTGTAATATTTGTTGCATTAGCATTAATTCACTGTNTGAAAAAAA
AAANAAAAAAAAAAAAAAAAAAAAAAA

FNRYLTRRRRIEIANLLALSERQIKIWFQNRRMKWKKDNNLKSMSQVDNMTS

IpHoxPost

AACATGGGATGGATAGCCCGCAATGTGAGCCGCAAGAAGCGGAGACCGTACACCAAGAC
CCAAACTCTGGAACTGGAGAAAGAATTCCTCTACAACACATATATCACTCGAGAGAGAA
GACTGGAAATCGCCCGCAGTCTCAGCCTCACCGACCGGCAGGTCAAGATCTGGTTCCAG
AACCGGAGGATGAAGAACAAGAAGCAGATGAACGGAGGAACTCCTCAAACTATGCACAT
GGTTCACCCAGGAGTGGTCAATGTTCCCATGGATCATTGCAGATATGACGTGTGCTAA*C*
CGAGCTTTCTCCTGATTGGTCCAAATCTGGTCACGTGATTGAATCGCGCAGGCCAATCA
TGTGTCAAACCAGTCGCGCTCTCCTGAAGATTGGAGATTCGCTTTATCCGGTACAATAG
CTTTAACCAATGCAGTACCAAATTATCAGAGCTTTATTTGGACTATGCTAATTTTAATT
TTATAATGTCCCTCTGTTTGTAAATGTTGGCATTTAGCTCTTCTTCTAGGTGCAGTTCT
TTCTCGGTTGTTGTTTTGTTGTCTTTTTTCCGGGAGGAAAGACACATAAATAAACTGTT
CTACTCCGAAAAAAAAAAAAAAAAAAAAAAA

NMGWIARNVSRKKRRPYTKTQTLELEKEFLYNTYITRERRLEIARSLSLTDRQVKIWFQ
NRRMKNKKQMNGGTPQTMHMVHPGVVNVPMDHCRYDVC

Appendix D

Nucleostemin is essential for neoblast regulation in the acoel *Isodiametra pulchra*

Katrien De Mulder[1,2], Eduardo Moreno[3], Pedro Martinez[3], Gaetan Borgonie[2], Peter Ladurner[1]

[1]University of Innsbruck, Institute of Zoology, Technikerstrasse 25, A-6020 Innsbruck, Austria
[2]University of Ghent, Department of Biology, Ledeganckstraat 35, B-9000 Ghent, Belgium
[3]University of Barcelona, Department of Genetics, Av. Diagonal 645, 08028 Barcelona, Spain

Abstract

Flatworms are astonishing regarding their regeneration capacity, based upon their unique pluripotent stem cell system, which is maintained throughout adulthood. Recently, a detailed morphological and molecular characterization of the acoel stem cell system was given, showing that acoels might share this pluripotent stem cell system with rhabditophoran flatworms. How this pluripotent stem cell system is maintained throughout adulthood however, still remains unsolved. Since acoels are supposed to be the earliest branching Bilaterians, developmental research on these animals will be of great evolutionary importance. *Nucleostemin* (*nclst*) was recently defined as a conserved stem cell regulator gene being originally found in mammalian nervous and cancer stem cells and described to play a crucial role in the regulation of cell cycle progression. Here, we describe the expression dynamics and function of *ipnclst* in the stem cell system of the acoel flatworm *Isodiametra pulchra* (Acoela, Acoelomorpha). *Ipnclst* was expressed in testes, ovaries and a subpopulation of somatic stem cells, and drastically upregulated during regeneration. Functional down regulation of *Ipnclst* resulted in a complete failure of tissue homeostasis. Stem cells gradually lost their proliferation capacity, which resulted in a dramatic reduction in tissue turn over, loss of regeneration and reproduction ability. Furthermore, prolonged RNAi treatment finally led to the death of the animals. A similar phenotype was observed using the conserved proliferation marker *PCNA*, suggesting a crucial role of *ipnclst* in stem cells. We propose that these findings could provide a base for further addressing the molecular events concerning pluripotent stem cell regulation in acoel flatworms, platyhelminths and higher organisms.

Introduction

Stem cells are unique in the ability to differentiate into various mature cell types while retaining their potential to self renew, and by this are crucial for tissue maintenance throughout life (Morrison and Spradling, 2008). Since in flatworms, stem cells are the only diving cells within the organism (Brondsted, 1969; Reddien and Sanchez, 2004; Sanchez and Kang, 2005), the number of stem cells, their proliferation activity and decision to differentiate must be tightly controlled to maintain tissue turn over and to avoid tumour formation or premature ageing.

Although many studies have addressed the topic of stem cell differentiation, the knowledge about their capacity to self renew and how to maintain pluripotency is still relatively young. Recently, the highly conserved nucleolar protein Nucleostemin (Nclst) was identified as a novel candidate stem cell regulator (Tsai and McKay, 2002). *Nclst* was originally identified in both embryonic- and adult neuronal stem cells, and later described in primitive bone marrow and cancer cell lines. In these cells, *nclst* plays a multiplex regulator of cell cycle progression (Cada *et al*., 2007; Jafarnejad *et al*., 2008; Ma and Pederson, 2008; Tsai and McKay, 2002; Ye *et al*., 2008). When stem cells undergo differentiation, *nclst* expression drastically drops, before cells exit the cell cycle (Tsai and McKay, 2002). Although *nclst* was supposed to be stem cell specific, it is drastically upregulated in differentiated cells during regeneration in newt (Maki *et al*., 2007). After limb amputation, transdifferentiating muscle fibers start to re-express *nclst*, even before dedifferentiation was morphologically visible. These observations suggest that *nclst* is not only necessary to maintain stem cells, but is also associated with the dedifferentiation process of cells, required to produce the multipotent stem cells during newt limb regeneration (Maki *et al*., 2007).

How *nclst* exactly fulfils its function is still unclear, but it does act on different pathways. *Nclst* regulates cell cycle progression both in a p53-dependent and -independent manner (Dai *et al*., 2008; Ma and Pederson, 2008; Ma and Pederson, 2007). In addition, *nclst* might positively regulate telomere synthesis, by inhibiting TRF1 (telomeric repeat binding factor, an inhibitor of telomerase), and so avoiding senescence of stem cells (Zhu *et al*., 2006). Although knowledge on Nclst function is increasing, most work to date has been conducted on mammalian stem cells, *Drosophila* and *C. elegans* (Chen *et al*., 2008; Kaplan *et al*., 2008; Ma and Pederson,

2008). A comprehensive study of *nclst* function in more basal animals however, is completely missing.

Our research is focussing on the stem cell system of the acoel *Isodiametra pulchra* (Acoela, Acoelomorpha). Like rhabditophoran flatworms, acoels possess an exceptional stem cell system (so called “neoblast” system), which lay at the base of their astonishing regeneration capacity (Gaerber *et al*., 2007; Steinbock, 1963; Steinbock, 1967; Gschwentner *et al*., 2001; De Mulder *et al*., in revision). The acoel *I. pulchra* was recently put forward as a novel model organism to study evolutionary developmental processes (Ladurner and Rieger, 2000; De Mulder *et al*., in revision). The simplicity of culturing and the unlimited access to eggs the whole year through is a great advantage when different stages of development have to be analysed. The morphological knowledge, absence of symbionts, and transparency of the animal allows us to analyse the effect of different biological conditions on morphology. Furthermore, as recently described, stem cell proliferation analysis using the thymidine homolog BrdU (Bromodeoxyuridine) as well as RNA interference can be easily performed by soaking (De Mulder *et al*., in revision). With the availability of 14.110 EST’s, *I. pulchra* is growing to become a potential new model organism in evolutionary developmental biology research.

In this study, we report on the expression dynamics and function of the evolutionary conserved stem cell marker *ipnclst* in the acoel *I. pulchra* during homeostasis, regeneration, postembryonic development, upon radiation and prolonged starvation. Since in acoels and rhabditophoran flatworms, stem cells (so called neoblasts) are the only proiliferating cells within the organism, proliferation markers can be used to determine stem cell localisation. Therefore, during this study, the expression pattern of the proliferation marker *PCNA* was examined as a comparison. Our results show that *ipnclst* is essential for stem cell maintenance and crucial during the process of regeneration. We suppose that our findings might contribute to understand how pluripotency is regulated.

Material and Methods

Animal culture

Isodiametra pulchra (Acoela, Acoelomorpha), originally described as *Convoluta pulchra* (Smith and Bush 1991) is kept in petri dishes, filled with nutrient-enriched f/2 artificial sea water (ASW). Culture conditions were originally taken over from *Macrostomum lignano* (Rieger *et al*., 1988), except that *I. pulchra* is kept continuously in the dark. During the whole experiments, *I. pulchra* was fed ad libitum on the diatom *Nitzschia curvilineata*, unless stated otherwise.

Ipnclst, IpPCNA **and** ***IpMCM2*** **cloning and sequence analysis**

From an earlier *I. pulchra* EST project (Ladurner and Agata, unpublished), four clones were identified, showing high similarity (e-value ≤ $3e^{-23}$) to the stem cell marker *nucleostemin* (Cpu_aW_0030_p17; Cpu_aW_028_D15; Cpu_aW_007_B06; Cpu_aW_007_J04). Resequencing and alignment of the available clones resulted in the full length sequence of a single *nucleostemin like*-gene, which we named *ipnclst* (GenBank accession number FM992878).

Within the same EST collection, seven clones were identified, showing significant similarity (e-value ≤ e^{-68}) to *Xenopus tropicalis PCNA* (Cpu_aW_021_H07; Cpu_aW_009_B12; Ipu_eW_004_N18; Cpu_aW_013_I08; Cpu_aW_018_A18; Ipu_eW_006_M23; Ipu_eW_002_L14). Subsequent concatenation resulted in the full length sequence of a single *PCNA*-like gene, which we named *ipPCNA* (GenBank accession number FM992877). PSORT* analysis (http://cubic.bioc.columbia.edu/cgi/var/nair/loctree/query) and NLS prediction were performed on both sequences, in order to predict subcellular localization of the corresponding proteins.

In addition, three clones (Cpu_aW_028_E07, Cpu_aW_005_D14, Cpu_aW_005_I18) were identified showing high similarity with the *MCM2 like*-gene of *Xenopus laevis* (e-value ≤ e^{-79}). Concatenation of the corresponding EST sequences resulted in the complete 3'end of a single *MCM2*-like gene, which we named *ipMCM2*. Full length sequence of the gene was obtained by a 5' RACE PCR approach (generacer kit, Invitrogen), using the successive primers: R4: 5'-

AGCACCTCGTTGGTGTCCTT-3’, R3: 5’-GGGTCTCGATGAATGACTCC-3’ and R2: 5’-TTGTCCTGGTCGGTGTTGGT-3’. The obtained fragment resulted in an upstream additional sequence of 547 bp. Within this newly obtained sequence, 3 new primers were designed and used for an additional nested PCR round, R8: 5’-GGTCCGCGTCGTTCATCTTGTCG-3’, R7: 5’-TGACCGGGCTCTTCTGCACGTAGG and R6: 5’-CGGGTCCCCACAGATCAGGATGTT-3’, which resulted finally in the complete ORF of *ipMCM2* (GenBank accession number FM 993911).

Whole mount in situ hybridization

Whole mount in situ hybridization (ISH) on acoels was carried out as described previously (Pfister *et al*., 2007), except for proteinase K treatment, which was performed for only 7 min. Sense and antisense riboprobes were generated using the DIG RNA labeling KIT SP6/T7 (Roche), following the manufactories protocol. The following primer couples were used for generating *in situ* riboprobe template: 5’-GCACGAGGAGAAATGGGTAAG-3’ and 5’-AGTTCCCAAGAAGACGCATCA-3’ for *ipnclst* (762bp); 5’-TGGTTGATGCCCATGTCTTC-3’ and 5’-ATGTTCTGGGTCTGCGACTG-3’ for *ipPCNA* (795bp); 5’-ACCCCAAGCTCACCAACAC-3’ and 5’-GACATGACACGGGTAGTGC-3’ for *ipMCM2* (678bp). During hybridization, riboprobes were used at a final concentration of 0,05 ng/µl for both *ipPCNA, ipMCM2* and 0,0375ng/µl for *ipnclst* respectively. Pictures were made using a Leica DM5000 microscope and a Pixera Penguin 600CL digital camera.

Double labelling of S-phase (BrdU) and *ipnclst/ipPCNA* expressing cells

Preceding fixation, animals were labelled with Bromodeoxyuridine (BrdU) by 30 min incubation in ASW, containing 5 mM BrdU (Ladurner *et al*., 2000). After washing animals and subsequent fixation, *in situ* hybridization was performed as described earlier (De Mulder *et al*., in revision). Following colour development, a standard BrdU immunohistochemical protocol was performed (Ladurner *et al*., 2000), except for pronase XIV treatment, which was done at a final concentration of 0,1mg/ml for 10 min at 37°C.

Immuno histochemistry

In order to localize Nucleostemin protein, we produced a polyclonal antibody, (Genscript, USA), recognizing the N-terminal peptide KKNPKKKNRKDPGVC of the

protein, starting at amino acid 33. Since this epitope was not accessible following standard immunohistochemistry, an adapted “antigen retrieval” protocol was needed. Specimens were gradually relaxed in 7,14% $MgCl_2$ and 30 min fixed with 4% PFA (in PBS, pH 7,4 at RT). Multiple PBS-T (0,1%) washes (3 x 5 min, 1h at RT) were followed by an initial antigen retrieval treatment for 10 min at 95°C in antigen retrieval buffer (Dako, K5336). Specimen were washed shortly in 2 x SSC buffer (3 x 5 min), followed by two extended washing steps (2 x SSC-T (0,1%) for 2 x 15 min). Animals were rinsed briefly 3 x 5 min in PBS-T preceding blocking in PBS-BSA (1%)-T (30 min RT). Primary antibody was incubated overnight at a final concentration of 1/100 in PBS-BSA-T (4°C). After washing with PBS-T (0,1%) (3 x 5 min), specimen were incubated in secondary antibody (1/200 FITC-swine-α-rabbit, 1 h RT, DAKO) and excessive antibody was removed by 3 x 5 min washing steps in PBS-T. Specimens were mounted in Vectashield (Vector Laboratories) and analyzed using a Leica DM5000. Details were taken with a Zeiss LSM 510.

Animal staging during post embryonic development, regeneration and starvation

Staging for postembryonic development, regeneration and starvation was performed as described earlier (De Mulder *et al.*, in revision).

Hard X-ray irradiation

Intact worms (1-2 months old) were exposed to 80 Gray Hard X ray, using a linear Accelerator (8MeV, 400 cGy/min; Radio-Oncology, Medical Hospital, Innsbruck). Animals were fixed 1 hour, 1 day and 1 week postirradiation and examined for *ipnclst* and *ipPCNA* expression respectively.

Gene specific RNA interference

RNA interference by soaking was performed as described earlier (De Mulder *et al.*, in revision), with the following modifications; dsRNA probes were generated using an *in vitro* transcription system (T7 Ribomax large scale RNA synthesis, Promega #1300), subsequently annealed and precipitated using a standard phenol/chloroform protocol. The dsRNA probe completely overlapped in sequence with the ISH probe (bp55-817 for *ipnclst*; bp56–851 for *ipPCNA*). Since earlier experiments confirmed that no mock effect was obtained using the exogenous *luciferase* gene (De Mulder *et al.*, in

revision), either *luciferase* or ddH_2O were used as a control. dsRNA was diluted in f/2 culture medium at a final concentration of 15 ng/µl and supernatant was changed every 24 hours.

A first batch of worms (n= 75) were 4-5 weeks old adults. Throughout the whole experiment, animals were fed ad libitum, in order to minimize influence of metabolism fluctuation on gene expression. A second batch of worms (n=75) were regenerates, amputated at the midregion of the body at the beginning of the experiment. Animals were left to regenerate under continuous RNAi treatment. After one week, amputation was repeated.

A third and forth batch of treated animals were embryos, collected from one day and one week pretreated animals respectively. Hatchlings were kept under continuous RNAi treatment during postembryonic development. Specimens were examined for cell proliferation, *ipnclst* and *PCNA* expression as well as the influence on the expression of the third stem cell marker *MCM2*. In addition, survival, reproducibility, postembryonic development and regeneration capacity were followed during the whole experiment (d=24).

Semithin sectioning and Transmission electron microscopy

Both techniques were performed as described earlier (Bode *et al*., 2006).

Results

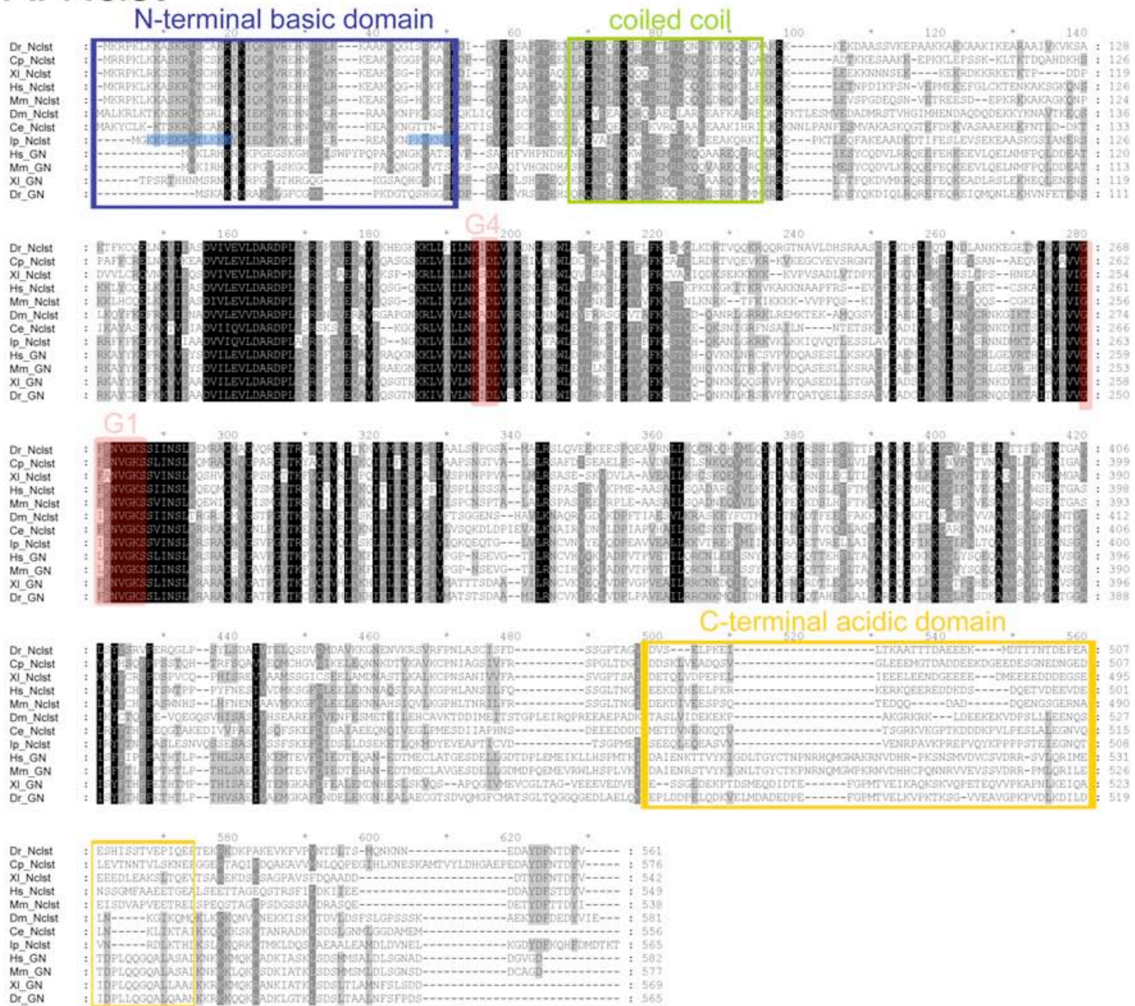
Gene cloning and sequence analysis

The full length sequence of *ipnclst* (1875bp) was obtained from an EST project by concatenation of four overlapping clones (accession number FM992878). *Ipnclst* showed highest similarity with *Tribolium castaneum nclst*-like (42,1% identity versus 57% similarity) and honeybee *Apis mellifera nclst*-like genes (42,1% versus 54,3%) respectively. *Ipnclst* mRNA codes for an open reading frame of 565 amino acids, with an estimated molecular weight of 64,17 kDa. Alignment of *ipnclst* with *nclst*-like genes of other species demonstrated the presence of the conserved N-terminal basic domain essential for nucleolar localization of the protein and p53 interaction (Fig. 1A) (Tsai and McKay, 2002). *Ipnclst* further possessed the preserved coiled coil domain, two GTP-binding motifs (G4 and G1) crucial for dynamic shuttling of the protein between nucleolus and nucleoplasm as well as a C-terminal acid domain (Fig. 1A) (Tsai and McKay, 2002). Within the N-terminal basic domain, two potential nuclear localisation signals (NLS) were found (Fig. 1A, highlighted in blue). Both PSORT* and LOCtree* analysis predicted the nuclear localization of Ipnclst, which was further confirmed by immunocytochemistry (see below).

In addition, we isolated a full length putative *PCNA*-like gene. Sequence analysis revealed a cDNA sequence of 1004bp with an open reading frame (ORF) encoding a 270-amino-acid putative PCNA protein (accession number FM992877). Both PSORT and LOCtree analysis predicted a nuclear localization of the protein. Alignment of *ipPCNA* with other *PCNA*-like genes confirmed a high degree of similarity within the *PCNA* members (Fig. 1B). *IpPCNA* showed highest similarity with the *Xenopus tropicalis PCNA*-like (49,8% identity versus 68,2% similarity) and the crustaean *Fenneropenaeus chinensis PCNA*-like gene (50,2% versus and 67,9%), respectively. *IpPCNA* possesses 24/24 conserved amino acids for trimerization and 9/9 residues forming the conserved PCNA-DNA binding site. All conserved residues are highlighted within the alignment (Fig. 1B).

The third gene *ipMCM2*, was only partially available within the EST collection. 5' RACE PCR resulted in a full length cDNA sequence of 2967 bp, coding for a putative IpMCM2 protein of 886 amino acids (FM993911). *IpMCM2* showed highest similarity with a *Branchiostoma floridae MCM2*-like gene (64,7% identity versus 76,8% similarity)

and a *Xenopus laevis MCM2*-like gene (64,4% identity versus 75,6% similarity) respectively. Within the *IpMCM2* sequence, the highly conserved Zinc finger motif and characteristic cdc46/MCM domain were identified (Fig. S1).



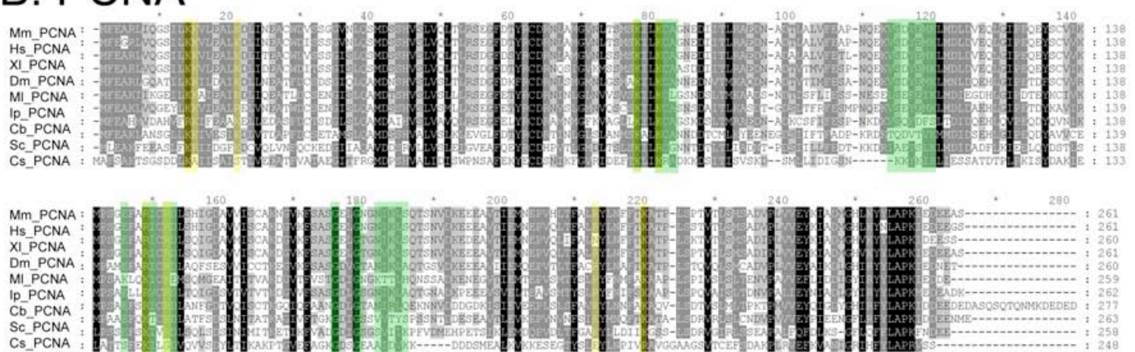


Fig. 1: Alignment of *ipnclst* with other *nclst*-like genes; functional domains of the protein are highlighted. The N-terminal basic domain (indicated in blue) is essential for nucleolar localization; two predicted NLS signals within the N-terminal domain of Ipnclst are marked in blue. The evolutionary conserved coiled coil domain is indicated in green; two GTP binding motifs, crucial for dynamic protein shuttling between nucleolus and nucleoplasm are indicated in red. Note Nclst is a permutated GTPase, with the G4 domain lying upstream of G1. In addition, the conserved C- terminal acidic domain is indicated in yellow. Accession numbers *Nclst*-like genes: *Isodiametra pulchra* Ipnclst (FM992878); *Homo sapiens* HsNclst (BC001024); *Mus musculus* MmNclst (AY181025); *Danio rerio* DrNclst (AY648717); *Xenopus laevis* XlNclst (BC045248); *Caenorhabditis elegans* CeNclst (NP_495749); *Homo sapiens* Guanine nucleotide binding protein-like 3 HsGN (BC011720); *Mus musculus* MmGN (BC057033); *Xenopus laevis* XlGN (BC087521); *Danio rerio* DrGN (AY648737). (B) Alignment of *ipPCNA* with other *PCNA-like* genes; functional domains are marked. In *ipPCNA*, 24/24 of the conserved residues forming the trimerization interface are present and highlighted in green. 9/9 conserved amino acids, outlining putative DNA binding sites, are marked in yellow. Accession number *PCNA*-like genes: *Isodiametra pulchra* IpPCNA (FM992877); *Mus musculus* MmPCNA (NP035175); *Homo sapiens* (HsPCNA CAG46598); *Xenopus laevis* (XlPCNA NP001081011); *Drosophila melanogaster* DmPCNA (NM057557); *Macrostomum lignano* MlPCNA (Angu7882); *Sacharomyces cerevisae* CsPCNA (EDN64705); *Caenorhabditis briggsae* (XP_001678472); *Sponge Cenarchaeum symbiosum* (YP_876348).

Ipnclst **positive cells define a neoblast subpopulation**

In order to analyze *ipnclst* expression and function in *Isodiametra pulchra*, we initially performed whole mount *in situ* hybridization on intact adults. *Ipnclst* was expressed in testes, ovaries, developing eggs as well as in a subpopulation of mesodermal located cells (Figs. 2A-F). To determine if those *ipnclst* positive cells were neoblasts, we performed double labelling, in which proliferating cells in S-phase were stained by BrdU incorporation. This method was shown earlier to be useful, since neoblasts are the only dividing cells within the organism (De Mulder *et al*., in revision). Analysis of this double labelling revealed the presence of *ipnclst* single labelled, *ipnclst*/BrdU double labelled as well as BrdU single labelled cells (Figs. 2D-F). These observations confirmed *ipnclst* is expressed in a subpopulation of somatic stem cells.

In adult animals, *ipPCNA* expression significantly resembles the expression of *ipnclst* (Figs. 2 G-L). *IpPCNA* is expressed in both testes and ovaries, as well as in a subpopulation of mesodermally located cells (Fig. 2I). Double labelling of *ipPCNA* with a 30 min BrdU pulse, revealed also for *ipPCNA* the existence of both *ipPCNA* single labelled, BrdU single labelled as well as *PCNA*/ BrdU double labelled cells, although double labelled cells were significantly more abundant (Figs. 2J-L).

On a morphological level, *ipnclst* and *ipPCNA* positive cells are small in size, possessing only a thin rim of cytoplasm (Figs. 2C-D, I-J). Gene expression of both markers was completely absent within the epidermal cell layer, as well as in the midbody (central syncytium) region and anterior to the statocyst. These properties meet the morphological characteristics and distribution pattern of the stem cell population of *I. pulchra* (De Mulder *et al*., in revision). Whole mount ISH with sense probes did not show any significant staining, confirming the specificity of the probes (Figs. S2 A, B).

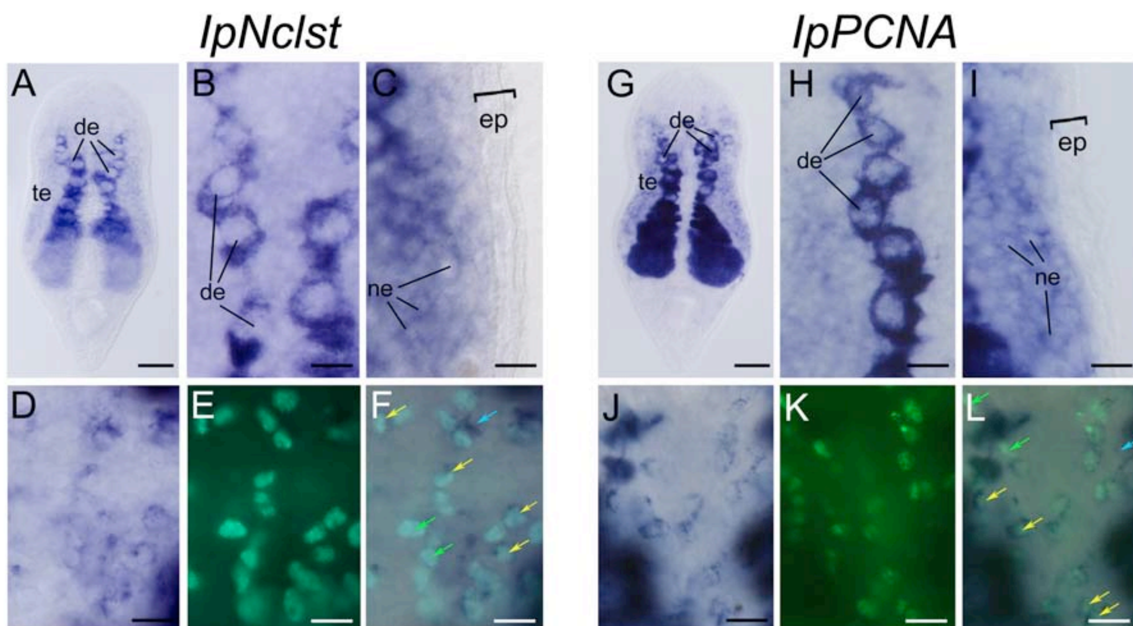


Fig. 2: *Ipnclst* (A-F) and *ipPCNA* (G-L) expression during homeostasis in adults. (A-C) Whole mount *ipnclst* ISH in intact adults. *Ipnclst* is expressed in developing eggs, testes and in a subpopulation of somatic stem cells. (B) Detail of *ipnclst* expression in two chains of developing eggs. (C) Detail of *ipnclst* expression in somatic stem cells. Note the complete lack of specific signal within the epidermis. (D-F) *ipnclst*/BrdU double labelling: (D) *ipnclst* positive mesodermal cells. (E) Distribution of S-phase cells (focal plane corresponding to picture D). *Ipnclst*/BrdU overlay (F). Note the presence of mainly double labelled cells (yellow arrows), as well as BrdU single labelled (green arrows) respectively *ipnclst* single labelled cells (blue arrows). (G-L) Whole mount *ipPCNA* mRNA localization in intact adults. *IpPCNA* mRNA is expressed in developing eggs, testes and in a subpopulation of mesenchymal located cells (G). Detail of *ipPCNA* expression in a chain of developing eggs (H). Detail of *ipPCNA* mRNA localization in a subpopulation of mesodermal located cells. As for *Ipnclst*, *ipPCNA* expression is completely absent within the epidermal cell layer. (J-L) Detail of *ipPCNA* positive cells, counterstained with BrdU. (J) *IpPCNA* positive cells. (K) BrdU positive cells, corresponding to picture J. (L) Overlay *IpPCNA* positive cells and cells in S-phase. In all pictures, anterior is to the top. (te) testes, (de) developing eggs, (ne) neoblast, (ep) epidermis. Scale bar 100µm in A and G, 20µm in all other pictures.

To further determine the subcellular localisation of Ipnclst, a polyclonal Ipnclst antibody was generated, showing specific immunoreactivity in a subpopulation of laterally located mesodermal cells, as well as in developing eggs (Figs. S2 C, D). Further investigation in the subcellular localization of Ipnclst protein, revealed the nuclear subcellular localization. Within the nucleus, Ipnclst was recognized within the large nucleoli of developing eggs as well as in the small nucleoli of a subpopulation of parenchymal located cells (Fig. S2 D).

***Ipnclst* is accumulated within blastemal cells during regeneration**

In earlier studies, it was shown that acoels possess, like other flatworms, a remarkable regeneration capacity (Gäerber *et al*., 2007; Steinböck, 1969; De Mulder *et al*., in revision). Here, we present an extended molecular analysis of the regeneration process using the stem cell marker *ipnclst* (Fig. 3). During the first three hours after initial cutting, the wound surface closes and no significant upregulation of *ipnclst* could be detected yet at that time. After 10-17 hours, *ipnclst* positive cells were present within the blastema, gradually expanding within the next 24 hours (Figs. 3B, C). *Ipnclst* upregulation reached a peak at around two days postamputation (Fig. 3E). During the following days, cell differentiation within the regeneration blastema was paralleled by a simultaneous downregulation of *ipnclst* expression (Figs. 3F-H). After 50 hours, *ipnclst* expression became below detection level in differentiating and differentiated cells, but remained high within the genital blastema, as these structures differentiate later than the surrounding tissue (Figs. 3F, G). After four days, *ipnclst* expression resembled the default expression pattern (Fig. 3H).

Comparison of *ipnclst* with *ipPCNA* expression during the regeneration process revealed the paralleled expression dynamics in time and location (compare upper and lower panel in Fig. 3). During the entire regeneration process, neither *ipnclst* nor *ipPCNA* expression could be observed within the epidermal layer. Moreover, amputation of the posterior end resulted only in a local proliferative response of somatic stem cells. No significant effect of tailplate amputation was observed in stem cells in the anterior part of the animal. These observations confirm a local response of amputation on the stem cell population, a hypothesis that was supposed earlier when *ipiwi1* expression dynamics during regeneration were studied (De Mulder *et al*, in revision).

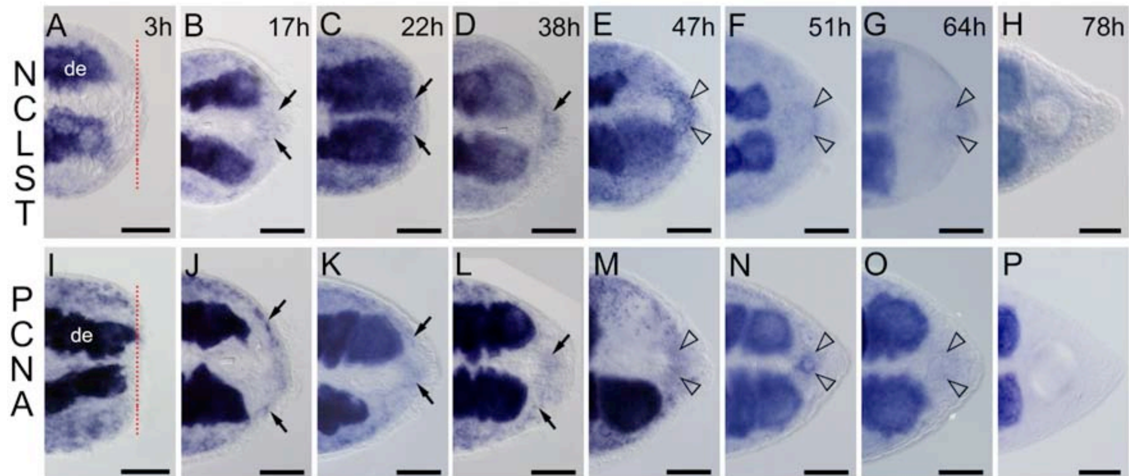


Fig. 3: *Ipnclst* (A-H) and *ipPCNA* (I-P) expression dynamics during tail regeneration in *Isodiametra pulchra*. (A, I) Regenerating tail, 3 hours after cutting. Red line indicates the initial amputation side. Wound surface is closed, but no local upregulation of both genes could be detected yet. (B, J) Forming blastema, 17 hours after cutting. (B) *Nclst* expression upregulation, scattered within the posterior tissue. Arrows indicate local expression upregulation. (J) At this time point, only a small rim of *PCNA* positive cells could be detected, just beneath the epidermal layer. (C-E, K-M) 2 days old regenerating blastema. During the second day of regeneration, a gradual increase in *ipnclst*/*ipPCNA* positive cells could be detected. Note the complete absence of signal within the epidermis. (F, N) Specific expression of both *ipnclst* and *ipPCNA* within the genital blastema after 51 hours regeneration (arrowheads). At this time point, most expression within the tailplate is reduced, paralleled by gradual differentiation of those cells. After 4 days (H, P), all structures are fully regenerated and gene expression resembles the steady state expression. In all pictures, anterior is to the left. (de) developing eggs .Scale bar 100 µm.

Ipnclst expression dynamics during development and germline formation

In order to follow the expression of the stem cell population during postembryonic development, *ipnclst* expression dynamics were followed during successive developmental stages until adulthood (Fig. 4). During the first days posthatching, *ipnclst* was expressed in a subpopulation of small mesodermal located cells, characterized by a small rim of cytoplasm. *Ipnclst* positive cells were distributed in a bilateral pattern, but were completely absent anterior to the statocyst, as well as in the midregion of the animals and the epidermis. Distribution, as well as morphology of labelled cells resembled the neoblast population as shown by 30 min BrdU incorporation and *PCNA* or *piwi* expression at corresponding developmental stages (Figs. 4A-B, I-J and De Mulder *et al*, in revision). The gradual growth of the juvenile was paralleled by *nclst* expression expansion (Figs. 4C-F). After 4-7 days post hatching, *ipnclst* was additionally expressed in several larger cells at both lateral sides which give rise to germline cells (Fig. 4E). This chain of developing eggs could be

followed during the whole postembryonic development through adulthood (Figs. 4D-H). After 10 days, a condensation of *nclst* expression could be observed in the developing genital blastema. This suggests the presence of active stem cells, which was confirmed by a similar *PCNA* and *ipiwi1* expression upregulation (Figs. 4F, N and De Mulder *et al*., in revision). Development was completed at 19 days, and gene expression resembled the adult expression pattern (Figs. 4H, P).

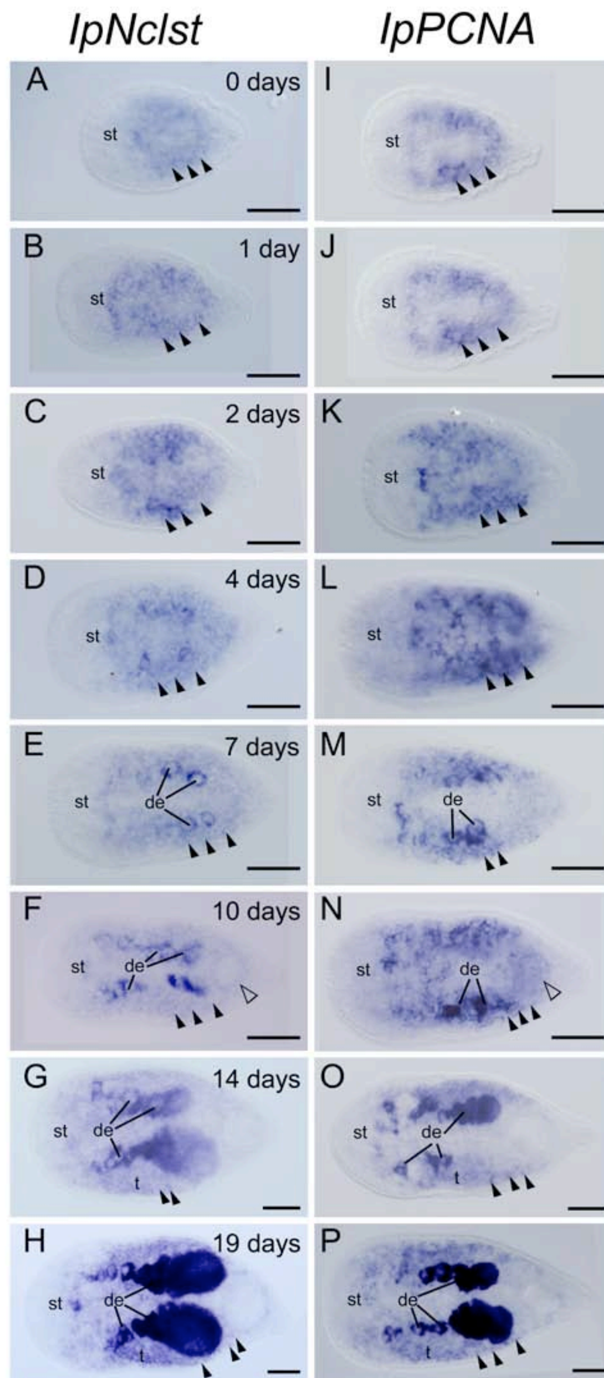


__Fig. 4__: *Ipnclst* (A-H) and *ipPCNA* (I-P) expression dynamics during postembryonic development. During the first days posthatching, both *ipnclst* and *ipPCNA* are expressed in a subpopulation of small mesodermal located cells (arrowheads) (A-B, I-J). Note the bilateral distribution, as well as the complete absence of positive cells anterior to the statocyst as well as in the epidermis, resembling neoblast distribution (A-B, I-J). The gradual growth of the juvenile was paralleled by *ipnclst* (C-F) and *ipPCNA* (K-N) expression expansion. After 4-7 days post hatching, *ipnclst/ipPCNA* were additional expressed in several larger cells at both lateral sites, which later gave rise to developing eggs (E, M). This chain of developing eggs could be followed during the whole postembryonic development (D-H, L-P). After 10 days, expression accumulation of both genes could be weakly observed within the developing genital blastema (open arrowheads) (F, N). At the time gene expression resembled the adult expression pattern, postembryonic development was completed (H, P).

High X ray sensitivity of *nclst* expressing cells

In those animals where stem cells are the only proliferating cell population such as in acoels and rhabditophoran flatworms, they can be physiologically distinguished from other cells by their high sensitivity to radiation (Eisenhoffer *et al*., 2008; Orii *et al*., 2005; Rossi *et al*., 2007; Salvetti *et al*., 2009; De Mulder *et al*, in revision). In this work, we used Hard X ray radiation to confirm the stem cell specific gene expression of *ipnclst* (Fig. 5A). Already three hours postradiation, *ipnclst* gene expression was drastically reduced, remaining at low expression level within the testes up to 1 day post radiation (Figs. 5B, C). Within one week (Fig. 5D), *Ipnclst* expression was completely abolished and remained below detection level. Finally, animals died within one month. In contrast, expression levels of the housekeeping gene *ipefalpha* (Elongation factor alpha), did not show clear signs of expression downregulation up to one week pos radiation (Figs. 5E, M).

To confirm the radiation specific elimination of stem cells, we analysed in parallel the effect of radiation on the proliferation marker *ipPCNA* (Fig. 5I). Comparable to *ipnclst, ipPCNA* expression was largely affected upon radiation (Figs. 5J-L). Already one hour after initial radiation exposure, *ipPCNA* expression was drastically reduced (Fig. 5J) and dropped below detection level within 24 hours post radiation (Figs. 5K, L). In general, gonadal stem cells seemed to be more resistant in which remnant expression could be detected during the first days.

***Nclst* expression dynamics confirms acoel stem cell plasticity upon starvation**

Complementary to their famous regeneration capacity, flatworms are well known for the plasticity of their stem cell system upon starvation (Baguñà, 1981; Nimeth *et al*., 2004; Oviedo *et al*., 2003; Pellettieri and Sanchez, 2007; Pfister *et al*., 2008). Recently, also acoel flatworms were shown to be able to adapt astonishingly to prolonged food deprivation (De Mulder *et al*., in revision). In this study, we further examined the distribution and activity of neoblasts in de-growing acoels, based on *ipnclst* and *ipPCNA* expression dynamics (Figs. 5F-H, N-P). After one week of food deprivation, animals were slightly reduced in size, although both *ipPCNA* and *nclst* expression still resembled control adult expression (compare Figs. 5A, I with F, N). Both genes were still detectable in testes, ovaries, developing eggs as well as in a subpopulation of mesodermal located cells (Figs. 5F, N). After two weeks of starvation, animals were significantly reduced in size (≤ 500µm) and chains of eggs were shortened in length

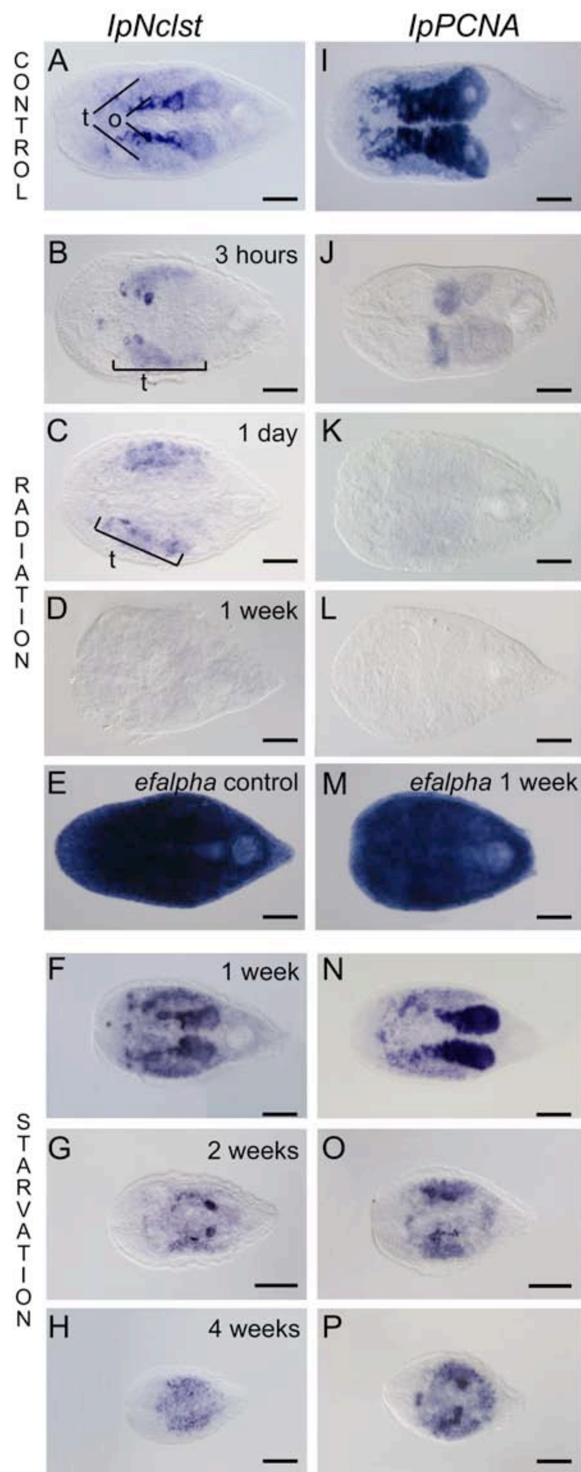


Fig. 5: Expression dynamics of *ipnclst* and *ipPCNA* following X ray-radiation exposure (B-D, J-L) and prolonged starvation (F-H, N-P). Upper panel: control expression pattern of *ipnclst* (A) and *ipPCNA* (H), showing expression of both genes in testes, ovaries as well as in a subpopulation of somatic stem cells. Middle panel: X-ray sensitivity of *ipnclst* (B-D) and *ipPCNA* (J-L) expressing cells. Already one hour after Hard X-ray exposure, expression of both genes was drastically reduced (B, J). Remnant expression could be detected within the gonads during the first days postradiation (B-C, J-K). After one week, both *ipPCNA* and *ipnclst* expression reached detection level (D, L). In contrast, *ipefalpha* expression remained unchanged up to 1 week post radiation (E, M). Lower panel: Expression dynamics of *ipnclst* (F-H) respectively *ipPCNA* (N-P) during prolonged starvation. After one week of food deprivation (F, N), gene expression resembled controls, although animals reduced their length. Prolonged starvation resulted further in significant reduction of body size (G-H, O-P). However, significant decrease in gene expression could not be observed. Interestingly, gene expression of one month starved animals resembled the expression pattern found in hatchlings, confirming the enormous plasticity of *I. pulchra*. In all figures, anterior is to the left. (t) testes, (o) ovaries. Scale bar 100 µm.

(Figs. 5G, O). Like in control animals, expression of both *ipnclst* and *ipPCNA* during prolonged starvation could not be detected anterior to the statocyst, neither in the epidermal layer. Prolonged starvation for one month resulted in further drastic degrowth of the animals, shrinking to a final size of ≤ 300µm in length, thereby regaining their juvenile shape (compare Figs. 4 and 5). Neoblasts were predominantly

found at the lateral sites of the animal. Small clusters, resembling the original gonads, remained detectable. Refeeding of the animals resulted in complete reversion of the starvation effect and worms produced viable offspring within the next month. These data confirm the earlier proposed hypothesis that the stem cell system of acoels is characterized by an enormous plasticity to adapt on environmental conditions (De Mulder *et al.*, in revision and data not shown).

Loss of *ipnclst* leads to cell cycle arrest and consequent failure of tissue homeostasis

In order to elucidate the biological function of *ipnclst* in *I. pulchra,* an established RNA-interference method by soaking was used to specifically downregulate *ipnclst* expression during homeostasis (Figs. 6-7), regeneration (Fig. 8) and post embryonic development. *Luciferase* dsRNA treated controls were confirmed not to show any mock effect, neither on gene expression, nor on cell proliferation or morphological level (De Mulder *et al.*, in revision). Hence, control animals were treated either with *luciferase* or ddH_2O (Figs. 6A-D).

In contrast, whole mount in situ hybridisation of specimens, treated for 21 days with *ipnclst* confirmed the complete mRNA downregulation of the respective gene (Fig. 6F). In order to further analyse the effect of *ipnclst* treatment on the stem cell system, proliferation activity was analysed. Proliferation in *ipnclst* treated worms was abolished, as shown by the complete absence of BrdU incorporating cells after three weeks of treatment (compare Figs. 6A, D). In order to follow the effect of *ipnclst* treatment on the stem cell system, expression dynamics of two additional proliferation markers (*ipPCNA*, *ipMCM2*) was examined (Figs. 6G, H). mRNA levels of both proliferation markers dropped below detection level within three weeks of treatment. Only in some animals, little expression in degenerating eggs could still be observed (Fig. 6H). At this timepoint, none of the treated adults laid viable eggs anymore. To summarize, loss of *ipnclst* expression resulted in a complete down regulation of the proliferation markers *ipPCNA* and *ipMCM2*, suggesting a crucial role of *ipnclst* in cell proliferation control.

As a comparison to analyse the effect of *ipnclst* RNAi on the stem cell system, *ipPCNA* RNAi was performed. A completely similar effect on the stem cell population was obtained during a prolonged *ipPCNA* treatment (Figs. 6I-L). No dividing cells were observed after 3 weeks of *ipPCNA* RNAi treatment (Fig. 6I). Additionally, also in these animals none of the stem cell gene expression (*ipnclst, ipPCNA, ipMCM2*) was

maintained (Figs. 6J-L). These observations confirm the role of both *ipnclst* and *ipPCNA* in stem cell regulation.

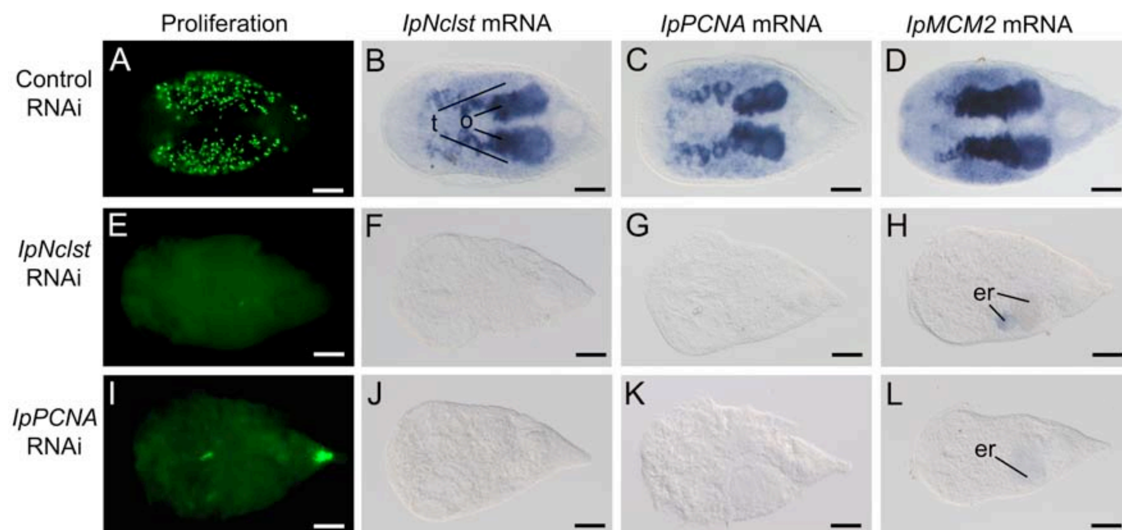


Fig. 6: Effect of 21 days *ipnclst* respectively *ipPCNA* RNA interference on stem cell gene expression. (A-D) Control animals. (E-H) 21days *ipnclst* treated animals. (I-L) 21 days *ipPCNA* treated animals. Controls show a normal expression pattern of *ipnclst*, *ipPCNA* as well as *ipMCM2*, confirming dsRNA treatment results in no mock effect on gene expression level. (E-H) Effect of *ipnclst* treatment on the stem cell population. Cell proliferation was totally abolished, as shown by the complete absence of BrdU incorporation (E). Lack of proliferation was further demonstrated by the absence of two fundamental proliferation markers. (G, H) After 21 days of *ipnclst* treatment, neither *ipPCNA*, nor *ipMCM2* expression could be observed. (I-L) Effect of *ipPCNA* treatment on the stem cell population. After 21 days of *ipPCNA* RNA treatment, no S-phase cells could be determined, as shown by the lack of BrdU labelled cells (I). Like for *ipnclst* RNAi, a prolonged treatment of *ipPCNA* resulted in the complete lack of stem cell gene expression (J-L). In all pictures, anterior is to the left. (t) testes, (o) ovaries, (er) egg remnant (degenerating egg). Scale bar 100 µm.

The effect of stem cell failure on tissue homeostasis was in parallel followed on a morphological level (Fig. 7, Fig. S3). After two weeks of RNA interference, no developing eggs could be detected in *ipnclst* treated adults. Only in some cases, few eggs were observed in *PCNA* treated eggs, suggesting the crucial function of both genes during oogenesis (Figs. 2 C, F, I). Remarkably, *ipnclst* treatment showed a significant earlier effect on tissue homeostasis, compared to *ipPCNA* treated worms (Fig. 7). In addition to the absence of developing eggs (Figs. 7 D, F), no mature sperm could be observed within the seminal vesicle after two weeks of treatment (Fig. 7E). This is in contrast with *ipPCNA* treated worms, which possessed at that time point still some remnant sperm (Figs. 7G-I). The obtained phenotype became more extreme on morphological level after three weeks of *Ipnclst* or *IpPCNA* depletion (Fig. S3). Treated animals completely lacked any gonadal structures (Figs. S3D-F). Seminal vesicles

were completely empty and no developing sperm, neither egg remnants could be detected (Fig. S3E). Instead, large vacuoles, filled with liquid, were present within the whole body (Figs. S3E, F). After three weeks, also *ipPCNA* treated worms, in which a visible phenotype was slightly delayed, exhibited a severe effect on morphological level (Figs. S3G-I). Due to failure of normal tissue turnover, animals died within one month.

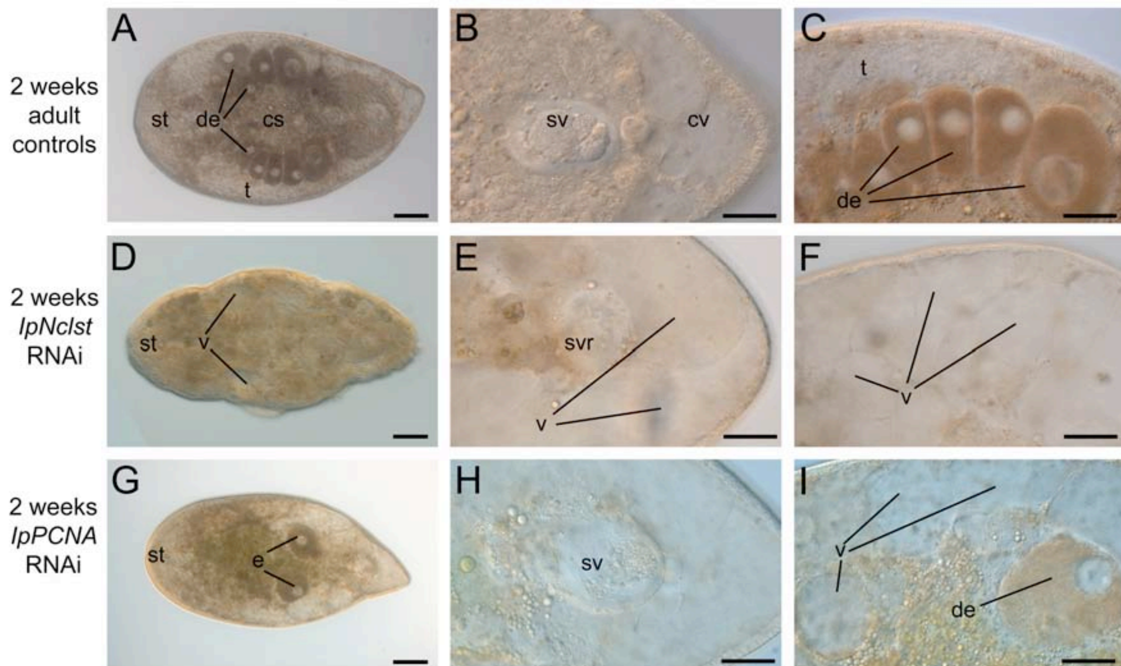


Fig. 7: Gradual loss of tissue homeostasis during prolonged *ipnclst* or *ipPCNA* RNA interference. (A-C) Control animals. (D-F) Two weeks *ipnclst* treated adults. (G-I) Two weeks *ipPCNA* treated animals. Overview of the morphology of a control adult, showing clear chains of developing eggs (A). Detail of the seminal vesicle, filled with mature sperm (B). Higher magnification of developing eggs. Note the clear visualisation of the nucleus (C). Effect of prolonged Ipnclst downregulation (D-F). After twee weeks of treatment, no developing eggs could be observed (D), neither mature sperm could be detected (E). Note the absence of developing eggs as well as the increasing presence of vacuoles gradually within the body (F). Two weeks *ipPCNA* RNAi treated animals (G-I). Egg production is drastically reduced and only single eggs can be observed (G). Reduced sperm can still be observed within the seminal vesicle (H). Similarly, mature sperm was detected within the seminal vesicle, although significantly reduced (H). Although less drastic, vacuoles, filled with liquid gradually situated within the body (I). (st) statocyst. (de) developing eggs, (cs) central syncytium, (t) testes, (sv) seminal vesicle, (cv) chordoid vacuole, (v) vacuole, (svr) seminal vesicle remnant. In all figures, anterior is to the left. Scale bar 50 μm, except in A, D and G (100μm).

Loss of regeneration capacity in *Nclst* treated worms

Since neoblasts play a crucial role during regeneration, we analyzed the effect of *nclst* RNAi mediated gene silencing on the regeneration capacity of *I. pulchra* (Fig. 8). In contrast to controls, which regenerated the tail plate completely within one week (Figs. 8A-C, S4A-C), no blastema formation could be observed in treated worms (Figs. 8D-F, S4D-F), although specimens were able to close the wound surface, a process independent of stem cells. Lost tissue however was not regenerated, but instead vacuoles, filled with liquid were abundant at the posterior part of the animal (Figs. S4D-F). After a second round of transection, an even more severe effect was observed (Figs. 8D-F). Again, regenerates were able to close the wound, but completely failed in rebuilding the lost body part. Genital organs gradually degradated and egg formation was completely abolished (Figs. 8D-F). Formed vacuoles disappeared, animals progressively disintegrated and died within four weeks.

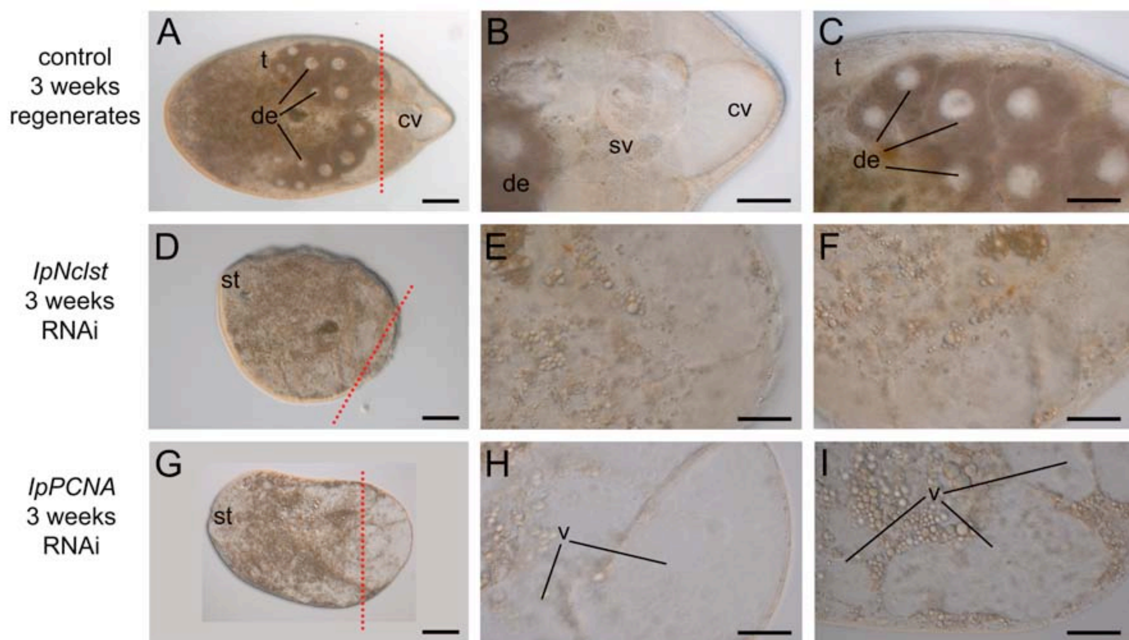


__Fig. 8__: Complete loss of regeneration capacity of *ipPCNA*/*ipnclst* treated worms, following 21 days of RNAi. (A-C) control animals. Red line marks the initial amputation level. (D-F) 21d *ipPCNA* treated regenerates. (G-I) 21 days *ipnclst* treated regenerates. Left panel: overview. Note the complete absence of developing eggs and tailplate (D, G). Middle panel: detail of the posterior part. In *ipnclst* or *ipPCNA* treated worms, genital organs are complete missing and no sperm could be detected (E, H). Right panel: detail on the female gonads. A chain of developing eggs in controls is visible (C), in strong contrast with treated worms, were only loose tissue, filled with vacuoles could be observed (F, I). In all pictures, anterior is to the left. Scale bar 50 µm, except in A, D, G (100 µm).

A similar situation was observed during the regeneration process of *PCNA* depleted animals, although the penetration of the phenotype was slightly delayed, compared to *ipnclst* (Figs. 8G-I, S4G-I). After the first round of amputation, also *ipPCNA* treated worms were able to close the wound, although only large vacuoles filled the posterior tissue and no regeneration of the lost body parts took place (Figs. S4D, E). After two weeks, few developing eggs were still visible (Fig. S4F), although none of these animals produced viable offspring anymore at this time point (n=56). After a second round of amputation, the effects became even more drastic (Figs. 8G-I). No developing eggs were present and posterior organs such as seminal vesicle, glands and bursa were not rebuilt (Fig. 8H). Also in these animals, the parenchyma was filled with vacuoles instead (Fig. 8I). Due to the loss of tissue homeostasis, regenerates died within one month.

Nucleostemin plays a crucial role during postembryonic development

Besides their role in tissue homeostasis and regeneration, neoblasts are believed to be the sole source for growth during postembryonic development in acoels. A first batch of embryos (n=40), collected from adults which were pre-treated with dsRNA for 24 hours, did still hatch. However, none of the hatchlings did grow or further develop, but instead died within the first seven days of development (data not shown). A second batch of embryos (n=43), collected from one week pre-treated adults, disintegrated before hatching. Within two weeks of continuous *ipnclst* RNAi treatment versus three weeks of continuous *ipPCNA* RNAi treatment, egg production dropped to zero, demonstrating the crucial function of these genes in development.

Discussion

Ipnclst, IpPCNA and *IpMCM2* are conserved stem cell genes

In this study, we have isolated and characterized three evolutionary conserved stem cell genes *ipnclst*, *ipPCNA* and *ipMCM2* in the acoel *Isodiametra pulchra*. In each of the three genes, we could identify all conserved residues within the sequences. *Ipnclst* possesses all conserved domains which characterize *nclst*-like genes (Tsai and McKay, 2002): the N-terminal basic domain, including two nuclear localization signals; a coiled coil domain; 2 GTP binding motifs and the C-terminal acidic domain. Similarly, alignment of *PCNA*-like genes confirmed the presence of all conserved residues for DNA binding and protein trimerization within the *ipPCNA* sequence (Jonsson *et al*., 1995). Finally, also *ipMCM2* showed extreme high similarity with other *MCM2*-like genes, possessing both the Zinc finger motif and the cdc46/MCM characteristic domain (Yan *et al*., 1991). To our knowledge, this is the first time that *nclst*-like genes were studied in more basal organisms. We conclude that *nclst* is highly conserved during evolution.

Ipnclst is expressed in a neoblast subpopulation

In this study, we examined the expression dynamics and function of the conserved stem cell marker *ipnclst* during different biological processes in *I. pulchra*. *Ipnclst* positive cells show clear stem cell morphology (small cells, with high nucleus/cytoplasm ratio). The distribution of *ipnclst* positive cells within the body was further in accordance to the stem cell system in *I. pulchra*, described earlier by S-phase distribution and *ipiwi1* expression (De Mulder *et al*., in revision), as well as the expression pattern of the stem cell markers *ipPCNA* and *ipMCM2* (data shown here). The fact that *ipnclst* expression was completely abolished after radiation, a classical tool used in flatworm stem cell research, further confirms the stem cell specific character of *ipnclst* positive cells (Orii *et al*., 2005; Reddien *et al*., 2005; Rossi *et al*., 2007; Salvetti *et al*., 2005; Salvetti *et al*., 2002). The presence of *ipnclst* single, BrdU single as well as *ipnclst*/BrdU double-labelled cells in *I. pulchra* is comparable to the situation in mice, where nucleostemin was not expressed in a cell cycle dependent manner (Ohmura *et al*., 2008). Subcellular Ipnclst localization using a polyclonal antibody against Ipnclst showed the nucleolar accumulation of the protein, confirming the LocTREE nuclear prediction of Ipnclst and further suggesting a conserved function

of the nucleolus in cell cycle progression (Ma and Pederson, 2008; Politz *et al*., 2005; Tsai and McKay, 2002).

Unexpectedly, double labelling of *ipPCNA* positive cells with BrdU incorporation, also revealed the existence of *ipPCNA* single, *ipPCNA*/BrdU double as well as BrdU single labelled stem cells, although double labelled cells were much more abundant. This is surprising, since *ipPCNA* is a member of the DNA replication fork, supposed to be crucial in cells, going through S-phase (Moldovan *et al*., 2007). However, one could explain the presence of BrdU single labelled cells by the fact that *ipPCNA* was detected only on mRNA level, and cells, *ipPCNA* negative on mRNA level, still could posses PCNA protein. The presence of *ipPCNA* single labelled cells on the other hand, can be clarified by a possible posttranscriptional/translational regulation as well as regulation of PCNA activity on protein level.

***Ipnclst* plays a conserved role during cell cycle progression and regeneration**

In this study, we have shown for the first time the local accumulation of *ipnclst* expression in stem cells during acoel flatworm regeneration. This is not surprising, since *nclst* is supposed to be a multiplex regulator of cell cycle progression (Ma and Pederson, 2008). A crucial function of *nclst* during regeneration was shown earlier in newts, where *nclst* was accumulated in dedifferentiating cells within the blastema (Maki *et al*., 2007). However, the regeneration process in both organisms significantly differs in the origin of blastemal cells. Regeneration in flatworms and acoels is based upon the proliferation of neoblasts (Egger *et al*., 2007), whereas in newts, blastemal cells arise "*de novo"* by dedifferentiation of muscle fibers (Straube *et al*., 2004). This fundamental difference in the origin of blastemal cells in acoels and newts could explain the time difference in proliferation and *nclst* accumulation. In flatworms, the presence of adult stem cells could explain the simultaneous *ipnclst* and *ipPCNA* expression upregulation during regeneration. In contrast to acoels, in which cell proliferation and *nclst* expression upregulation almost happens simultaneously, *nclst* accumulation in newts occurs significantly before S-phase re-entry (Maki *et al*., 2007). A comparable situation was also observed during adult heart regeneration. *Nclst* was significantly upregulated in the heart tissue after acute myocardial infarction (Tjwa and Dimmeler, 2008), where it might be crucial for the proliferative response of CSCs and cardiomyocytes after injury.

Finally, the fact that *nclst* expression during regeneration in *I. pulchra* is already down regulated before final differentiation has taken place, confirms the hypothesis that *nclst* expression drops before cell cycle exit and differentiation (Tjwa and Dimmeler, 2008). Also this observation is in agreement with studies of rodent neuronal and human bone marrow stem cells, where *nclst* is drastically down regulated in maturing and differentiated cells (Kafienah *et al*., 2006; Tsai and McKay, 2002).

Effect of radiation and prolonged food deprivation on *ipnclst* expression

The fact that *nclst* expression in some organisms is extended to a subpopulation of differentiated cells (Kudron and Reinke, 2008) as well as in cells with dedifferentiation capacity (Maki *et al*., 2007), raised the question if *ipnclst* was stem cell specifically expressed in *I. pulchra*. Hard X ray exposure, which specifically eliminates neoblasts, confirmed the stem cell specific expression of *ipnclst*. *Ipnclst* expression drastically reduced during the first day postradiation and remained below detection level afterwards. Due to the lack of normal tissue turnover, as a result of the absence of functional stem cells, animals died within one month, comparable with our own earlier observations (De Mulder *et al*., in revision).

On the other hand, the drastical shrinkage and regrowth of the animals upon starvation and refeeding, paralleled by the expression dynamics of *ipnclst,* further demonstrates the enormous plasticity of the acoel stem cell system. The ability to adapt to environmental changes was demonstrated already in the same organism using the stem cell marker *piwi* (De Mulder *et al*., in revision), a character that acoels share with rhabditophoran flatworms (Nimeth *et al*., 2004; Oviedo *et al*., 2003; Pellettieri and Sanchez, 2007; Pfister *et al*., 2008; De Mulder *et al*., in revision).

***Ipnclst* and *ipPCNA* are essential for stem cell maintenance, development and regeneration capacity**

In the current study, we showed the drastic reduction of cell proliferation upon loss of *ipnclst* or *ipPCNA.* These findings further support earlier observations that loss of *nclst* induces cell cycle arrest (Zhu *et al*., 2006; Jafarnejad *et al*., 2008). This hypothesis is further supported by our data showing that elimination of *ipnclst* expression was followed by a drastic reduction in the expression of *ipPCNA* and *ipMCM2*, two well conserved regulators of cell proliferation. The lethal phenotype, obtained by prolonged *ipnclst* RNAi is comparable to the effect of Hard-X ray radiation,

which also led to loss of a functional proliferating stem cell population (De Mulder *et al*., in revision) and is comparable with the effects, earlier described in mouse (Tsai and McKay, 2002).

Conclusion

In this study, we have shown that *nclst*-like genes are highly evolutionary conserved in the most basal bilaterians, playing a crucial role in stem cell maintenance of acoel flatworms. *Nclst* expression was significantly upregulated upon regeneration and drastically reduced before cell differentiation within the blastema took place. In addition, *nclst* expression dynamics could be used to further demonstrate the fundamental difference in the regeneration process between flatworms and newts. Instead, our data further showed the astonishing similarity between the acoel and rhabditophoran stem cell system.

However, the molecular program involved in regulating the transition of stem cells between the actively dividing and differentiated state still remains unclear. Further investigation will be needed to unravel how *nclst* is exactly regulated in acoels, and if this might be linked to the adult pluripotency of neoblasts in acoels and flatworms in general.

Acknowledgments

The authors want to thank Prof. Dr. K. Agata (Kyoto University, Japan) for the *I. pulchra* EST-library collaboration. This work was supported by a predoctoral FWO grant to K.D.M (Belgium) and a FWF grant 18099 to P.L (Austria). E.M was supported by the Spanish Ministerio de Ciencia e Innovación, the Generalitat de Catalunya, and the Marine Genomics Europe (Network of Excellence).

References

Baguñà, J., 1981. Planarian neoblast. *Nature* 290, 14-15.

Bode, A., Salvenmoser, W., Nimeth, K., Mahlknecht, M., Adamski, Z., Rieger, R.M., Peter, R., and Ladurner, P., 2006. Immunogold-labeled S-phase neoblasts, total neoblast number, their distribution, and evidence for arrested neoblasts in *Macrostomum lignano* (Platyhelminthes, Rhabditophora). *Cell Tissue Res*. 325, 577-587.

Brondsted, H.V., 1969. Planarian regeneration. Pergamon Press, Oxford.

Cada, Z., Boucek, J., Dvorankova, B., Chovanec, M., Plzak, J., Kodets, R., Betka, J., Pinot, G.L., Gabius, H.J., and Smetana, K., Jr., 2007. Nucleostemin expression in squamous cell carcinoma of the head and neck. *Anticancer Res*. 27, 3279-3284.

Chen, L., McCloskey, T., Joshi, P.M., and Rothman, J.H., 2008. *ced-4* and proto-oncogene *tfg-1* antagonistically regulate cell size and apoptosis in *C. elegans. Curr. Biol*. 18, 1025-1033.

Dai, M.S., Sun, X.X., and Lu, H., 2008. Aberrant expression of Nucleostemin activates p53 and induces cell cycle arrest via inhibition of MDM2. *Mol. Cell. Biol*. 28, 4365-4376.

Eisenhoffer, G.T., Kang, H., and Sanchez, A.A., 2008. Molecular analysis of stem cells and their descendants during cell turnover and regeneration in the planarian *Schmidtea mediterranea. Cell Stem Cell* 3, 327-339.

Gschwentner, R., Ladurner, P., Nimeth, K., and Rieger, R., 2001. Stem cells in a basal bilaterian. S-phase and mitotic cells in *Convolutriloba longifissura* (Acoela, Platyhelminthes). *Cell Tissue Res*. 304, 401-408.

Jafarnejad, S.M., Mowla, S.J., and Matin, M. M., 2008. Knocking-down the expression of nucleostemin significantly decreases rate of proliferation of rat bone marrow stromal stem cells in an apparently p53-independent manner. *Cell. Prolif*. 41, 28-35.

Jonsson, Z.O., Podust, V.N., Podust, L.M., and Hubscher, U., 1995. Tyrosine 114 is essential for the trimeric structure and the functional activities of human proliferating cell nuclear antigen. *EMBO J*. 14, 5745-5751.

Kafienah, W., Mistry, S., Williams, C., and Hollander, A.P., 2006. Nucleostemin is a marker of proliferating stromal stem cells in adult human bone marrow. *Stem Cells* 24, 1113-1120.

Kaplan, D.D., Zimmermann, G., Suyama, K., Meyer, T., and Scott, M.P., 2008. A nucleostemin family GTPase, NS3, acts in serotonergic neurons to regulate insulin signaling and control body size. *Genes Dev*. 22, 1877-1893.

Kudron, M.M. and Reinke, V., 2008. *C. elegans* nucleostemin is required for larval growth and germline stem cell division. *PLoS. Genet*. 4, e1000181.

Ladurner, P. and Rieger R, 2000. Embryonic muscle development of *Convoluta pulchra* (Turbellaria-acoelomorpha, Platyhelminthes). *Dev. Biol*. 222, 359-375.

Ladurner, P., Rieger, R., and Bagunà, J., 2000. Spatial distribution and differentiation potential of stem cells in hatchlings and adults in the marine platyhelminth *Macrostomum* sp.: a bromodeoxyuridine analysis. *Dev. Biol*. 226, 231-241.

Ma, H. and Pederson, T., 2007. Depletion of the Nucleolar Protein Nucleostemin causes G1 cell cycle arrest via the p53 pathway. *Molecular Biology of the Cell* 18, 2630-2635.

Ma, H. and Pederson, T., 2008. Nucleostemin: a multiplex regulator of cell-cycle progression. *Trends in Cell Biology* In Press, Corrected Proof.

Maki, N., Takechi, K., Sano, S., Tarui, H., Sasai, Y., and Agata, K., 2007. Rapid accumulation of nucleostemin in nucleolus during newt regeneration. *Dev. Dyn*. 236, 941-950.

Moldovan, G.L., Pfander, B., and Jentsch, S., 2007. *PCNA*, the maestro of the replication fork. Cell 129, 665-679.

Morrison, S.J. and Spradling, A.C., 2008. Stem cells and niches: mechanisms that promote stem cell maintenance throughout fife. *Cell* 132, 598-611.

Nimeth, K.T., Mahlknecht, M., Mezzanato, A., Peter, R., Rieger, R., and Ladurner, P., 2004. Stem cell dynamics during growth, feeding, and starvation in the basal flatworm *Macrostomum* sp. (Platyhelminthes). *Dev. Dyn*. 230, 91-99.

Ohmura, M., Naka, K., Hoshii, T., Muraguchi, T., Shugo, H., Tamase, A., Uema, N., Ooshio, T., Arai, F., Takubo, K., Nagamatsu, G., Hamaguchi, I., Takagi, M., Ishihara, M., Sakurada, K., Miyaji, H., Suda, T., and Hirao, A., 2008. Identification of stem cells during Prepubertal spermatogenesis via monitoring of Nucleostemin promoter activity. *Stem Cells* 26, 3237-3246.

Orii, H., Sakurai, T., and Watanabe, K., 2005. Distribution of the stem cells (neoblasts) in the planarian *Dugesia japonica. Dev. Genes Evol*. 215, 143-157.

Oviedo, N.J., Newmark, P.A., and Sanchez, A.A., 2003. Allometric scaling and proportion regulation in the freshwater planarian *Schmidtea mediterranea. Dev. Dyn*. 226, 326-333.

Pellettieri, J. and Sanchez, A.A., 2007. Cell turnover and adult tissue homeostasis: from humans to planarians. *Annu. Rev. Genet*. 41, 83-105.

Pfister, D., De Mulder, K., Hartenstein, V., Kuales, G., Borgonie, G., Marx, F., Morris, J., and Ladurner, P., 2008. Flatworm stem cells and the germ line: developmental and evolutionary implications of macvasa expression in *Macrostomum lignano. Dev. Biol*. 319, 146-159.

Politz, J.C., Polena, I., Trask, I., Bazett-Jones, D.P., and Pederson, T., 2005. A nonribosomal landscape in the nucleolus revealed by the stem cell protein nucleostemin. *Mol. Biol. Cell*. 16, 3401-3410.

Reddien, P.W., Oviedo, N.J., Jennings, J.R., Jenkin, J.C., and Sanchez, A.A., 2005. SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. *Science* 310, 1327-1330.

Reddien, P.W. and Sanchez, A.A., 2004. Fundamentals of planarian regeneration. *Annu. Rev. Cell Dev. Biol*. 20, 725-757.

Rieger, R., Gehlen, M., Hazpruhnar, G., Homlund, M., Legnitit, A., Salvenmoser, W., and Tyler, S., 1988. Laboratory cultures of marine Macrostomida (Turbellaria). *Forts. Zool*. 36, 525.

Rossi, L., Salvetti, A., Marincola, F.M., Lena, A., Deri, P., Mannini, L., Batistoni, R., Wang, E., and Gremigni, V., 2007. Deciphering the molecular machinery of stem cells: a look at the neoblast gene expression profile. *Genome Biol*. 8, R62.

Salvetti, A., Rossi, L., Lena, A., Batistoni, R., Deri, P., Rainaldi, G., Locci, M.T., Evangelista, M., and Gremigni, V., 2005. *DjPum*, a homologue of *Drosophila Pumilio*, is essential to planarian stem cell maintenance. *Development* 132, 1863-1874.

Salvetti, A., Lena, A., Rossi, L., Deri, P., Cecchettini, A., Batistoni, R., and Gremigni, V., 2002. Characterization of *DeY1*, a novel Y-box gene specifically expressed in differentiating male germ cells of planarians. *Gene Expr. Patterns* 2, 195-200.

Salvetti, A., Rossi, L., Bonuccelli, L., Lena, A., Pugliesi, C., Rainaldi, G., Evangelista, M., and Gremigni, V., 2009. Adult stem cell plasticity: Neoblast repopulation in non-lethally irradiated planarians. *Developmental Biology* In Press, Accepted Manuscript.

Sanchez, A.A. and Kang, H., 2005. Multicellularity, stem cells, and the neoblasts of the planarian *Schmidtea mediterranea. Exp. Cell Res*. 306, 299-308.

Sanchez Alvarado, A., 2006. Planarian Regeneration: Its end Is its beginning. *Cell* 124, 241-245.

Siddiqi, S., Gude, N., Hosoda, T., Muraski, J., Rubio, M., Emmanuel, G., Fransioli, J., Vitale, S., Parolin, C., D'Amario, D., Schaefer, E., Kajstura, J., Leri, A., Anversa, P., and Sussman, M.A., 2008. Myocardial induction of nucleostemin in response to postnatal growth and pathological challenge. *Circ. Res.* 103, 89-97.

Straube, W.L., Brockes, J.P., Drechsel, D.N., and Tanaka, E.M., 2004. Plasticity and reprogramming of differentiated cells in amphibian regeneration: partial purification of a serum factor that triggers cell cycle re-entry in differentiated muscle cells. *Cloning Stem Cells* 6, 333-344.

Tjwa, M. and Dimmeler, S., 2008. A nucleolar weapon in our fight for regenerating adult hearts: nucleostemin and cardiac stem cells. *Circ. Res*. 103, 4-6.

Tsai, R.Y. and McKay, R.D., 2002. A nucleolar mechanism controlling cell proliferation in stem cells and cancer cells. *Genes Dev*. 16, 2991-3003.

Yan, H., Gibson, S., and Tye, B.K., 1991. Mcm2 and Mcm3, two proteins important for ARS activity, are related in structure and function. *Genes & Development* 5, 944-957.

Ye, F., Zhou, C., Cheng, Q., Shen, J., and Chen, H., 2008. Stem-cell-abundant proteins Nanog, Nucleostemin and Musashi1 are highly expressed in malignant cervical epithelial cells. *BMC Cancer* 8, 108.

Zhu, Q., Yasumoto, H., and Tsai, R.Y., 2006. Nucleostemin delays cellular senescence and negatively regulates TRF1 protein stability. *Mol. Cell Biol*. 26, 9279-9290.

Supplementary Figures

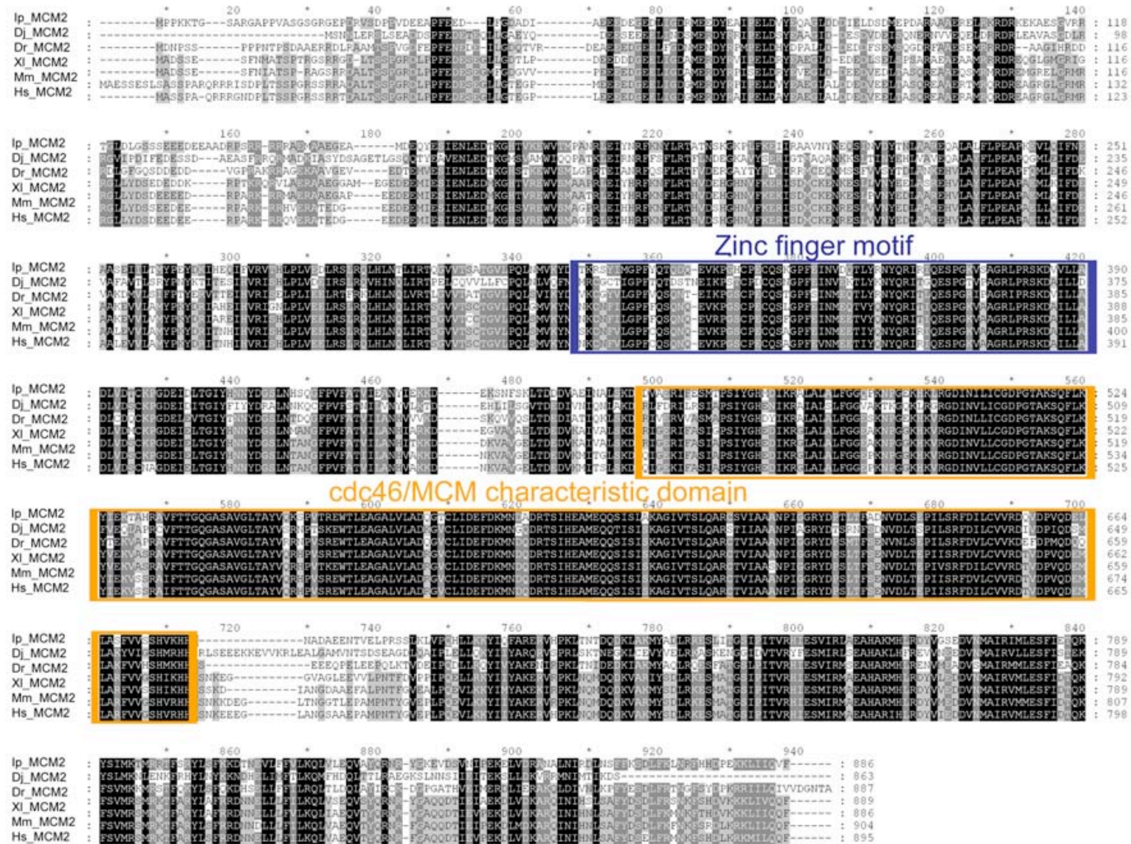


Fig. S1: Alignment of *MCM2*- like genes; evolutionary conserved domains of the protein are highlighted. Blue, Zinc finger like motif; Orange, cdc46/MCM characteristic domain. Accession numbers *MCM2*-like genes: *Isodiametra pulchra* IpMCM2 (FM993911); *Dugesia japonica* DjMCM2 (CAC36296); *Danio rerio* DrMCM2 (AAH48026); *Xenopus laevis* XlMCM2 (P55861); *Mus musculus* MmMCM2 (P97310); *Homo sapiens* HsMCM2 (BAA12177).

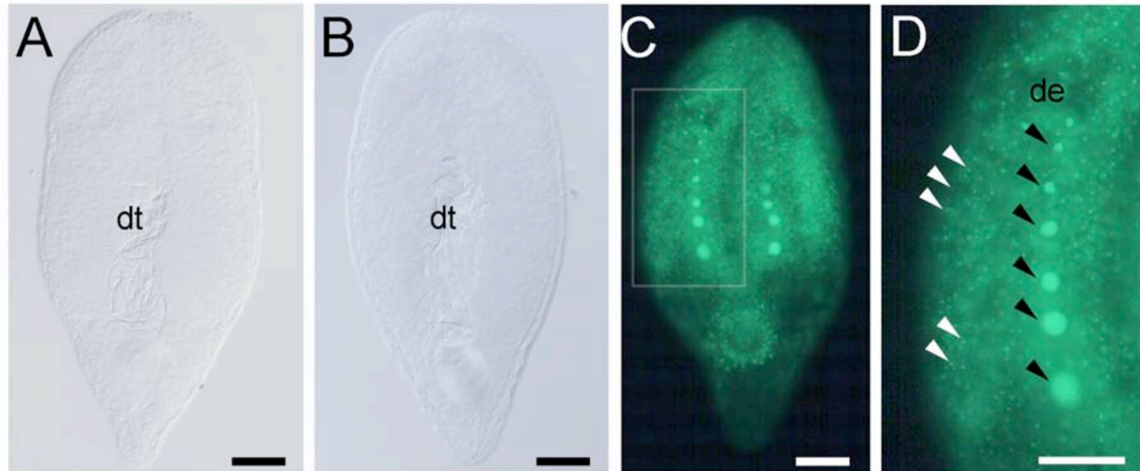


Fig. S2: A) *Ipnclst* sense ISH. B) *IpPCNA* sense ISH. Note the absence of any specific signal in A and B, confirming the specificity of the DIG-labelled probes. C-D) Ipnclst subcellular localization. C) Whole mount immunohistological localization of Ipnclst protein. Ipnclst is clearly visible in large nucleoli of developing eggs, as well as in a subpopulation of somatic stem cells. D) Magnification of (C): black arrowheads point towards the nucleoli of a chain of developing eggs. White arrowheads indicate the small nucleoli of somatic stem cells. (de) developing eggs; (dt) diatoms. Scale bars 100 µm, except in picture D (50 µm).

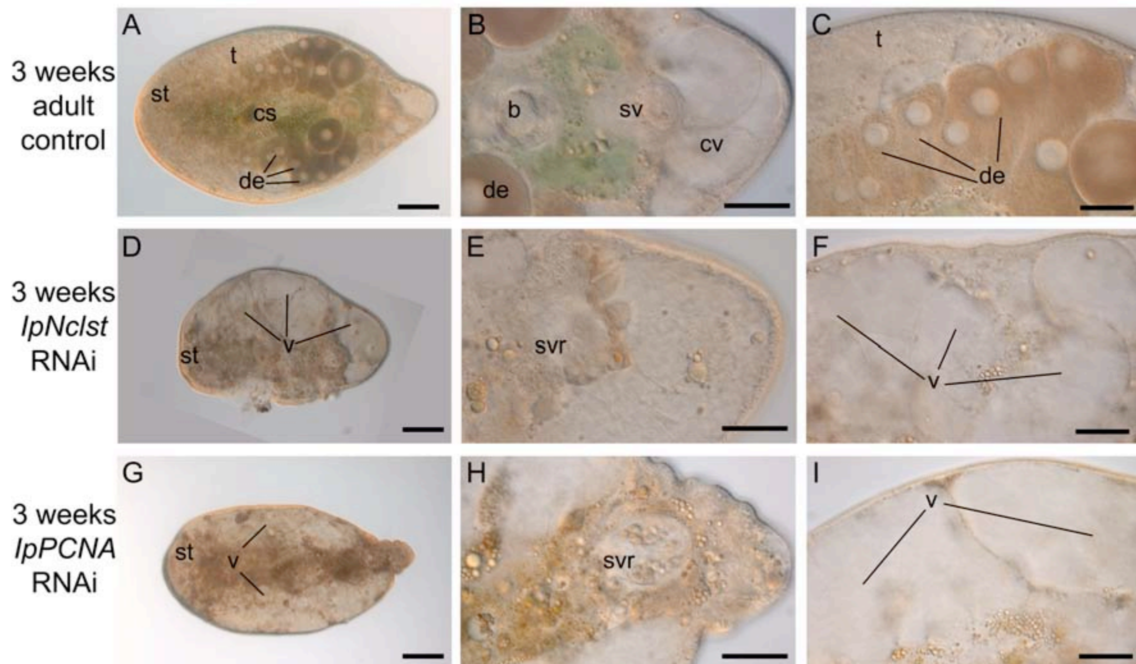


Fig S3: Loss of tissue homeostasis, as a result of prolonged RNAi treatment. (A-C) Control treated specimen, showing normal morphology. (A) Overview of control adult, with clear developing eggs and intact central syncytium. (B) Detail of the posterior end. Note the presence of distinct chordoid vacuoles, seminal vesicle, filled with mature sperm and intact bursa seminalis. (C) Detail of chain of developing eggs and sperm, located laterally within the testes. (D-F) Drastic effect on morphological level after 21 days *ipnclst* RNAi treatment. (D) Overview of the morphological effect upon *ipnclst* RNAi treatment. Note the degenerated status, expansion of vacuoles, filled with liquid as well as the significant shrinkage of the animal. (E) Detail of the posterior end. Note the absence of mature sperm within the degenerated seminal vesicle. (F) Morphology of the lateral side, lacking any developing eggs. Instead, several vacuoles are present. (G-I) Outcome of *ipPCNA* RNAi on morphological level. (G) Overview, 21 days treated *ipPCNA* adult. Note the significantly similarity of lateral distributed vacuoles, comparable with the *ipnclst* phenotype. (H) Magnification of the posterior end. No mature sperm could be observed within the seminal vesicle. (I) Enlargement of the laterally formed vacuoles. No developing eggs, neither sperm could be detected within the parenchym. In all figures, anterior is to the left. (st) statocyst, (cs) central syncytium, (de) developing eggs, (b) bursa seminalis, (cv) chordoid vacuole*, (t) testes, (v) vacuoles, (svr) seminal vesicle remnant. Scale bar 50 µm, except in A, D, G (100 µm).

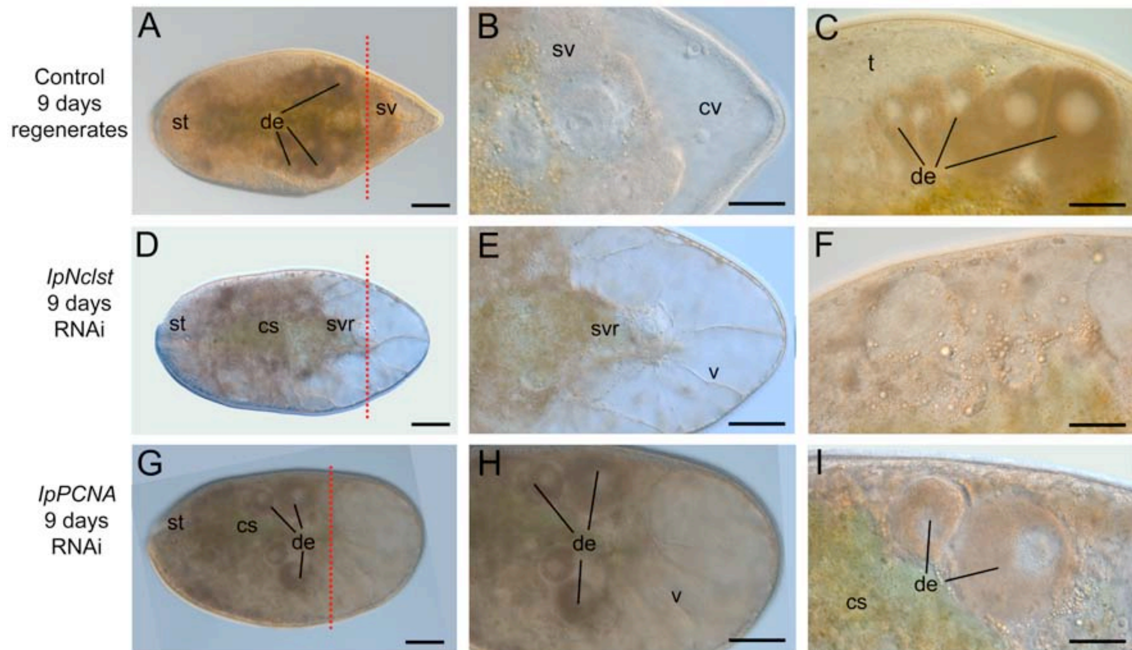


Fig. S4: Severe effect on regeneration capacity upon *ipnclst* or *ipPNCA* RNAi, nine days of initial amputation of the posterior end. Red line marks the initial amputation level. (A-C) control animals. (D-F) 9d *ipnclst* regenerates. (G-I) 9 days *ipPCNA* treated regenerates. Left panel: overview. Note the complete absence of developing eggs and tailplate in *nclst* treated worms (D) and only few eggs in *PCNA* treated specimens. Middle panel: detail of the posterior part. In *ipnclst* or *ipPCNA* treated worms, large vacuoles are present (E, H). Right panel: detail on the female gonads. A chain of developing eggs in controls is visible (C), in strong contrast with *ipnlcst* treated worms, were only loose tissue, filled with vacuoles could be observed (F). In *ipPCNA* regenerates, single eggs could still be observed (I). In all pictures, anterior is to the left. (st) statocyst, (de) developing eggs, (sv) seminal vesicle, (cv) chordoid vacuole, (cs) central syncytium, (svr) seminal vesicle remnant, (v) vacuoles. Scale bar 50 μm, except in A, D, G (100 μm).