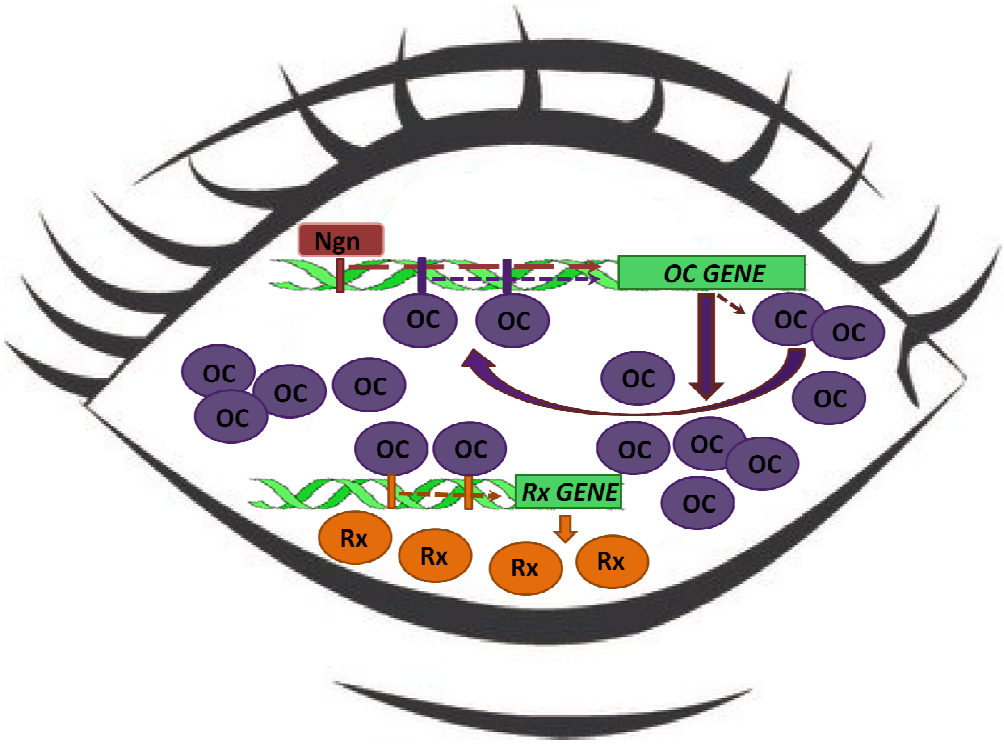
UNIVERSITA' DI NAPOLI FEDERICO II

DOTTORATO DI RICERCA IN
BIOCHIMICA E BIOLOGIA CELLULARE E MOLECOLARE
XXV CICLO

Maria Rosa Pezzotti

Ci-Onecut genetic pathway: functional studies



Academic Year 2011/2012

UNIVERSITA' DI NAPOLI FEDERICO II

DOTTORATO DI RICERCA IN
BIOCHIMICA E BIOLOGIA CELLULARE E MOLECOLARE
XXV CICLO

Ci-Onecut genetic pathway: functional studies

Candidate
Maria Rosa Pezzotti

Tutor
Prof. Francesco Aniello

Co-Tutor
Dr. Margherita Branno

Coordinator
Prof. Paolo Arcari

Academic Year 2011/2012

A me stessa

RIASSUNTO

I geni *Rx (Retinal Homeobox)*, identificati in molte specie di vertebrati ed invertebrati, posseggono un pattern di espressione ed una funzione molto conservate nello sviluppo del Sistema Nervoso Centrale (SNC) anteriore. Inoltre nei vertebrati è stato dimostrato il loro ruolo anche nella formazione dell'occhio ed in particolare nel processo di specificazione delle cellule della retina e/o nella loro proliferazione. Il gruppo di ricerca in cui ho svolto il mio lavoro di Dottorato aveva precedentemente dimostrato che nell'invertebrato cordato basale *Ciona intestinalis* questo gene ha un ruolo fondamentale nella formazione dell'organo di senso della visione (l' ocello) e nel differenziamento dei fotorecettori (D'Aniello *et al*., 2006). La caratterizzazione della regione regolativa del gene *Ci-Rx* aveva portato all'isolamento di due frammenti minimi in grado di dirigere l'espressione del gene reporter negli stessi territori in cui è espresso il trascritto endogeno. Dall'analisi bioinformatica condotta su tali sequenze si erano evidenziati i siti di legame per il fattore di trascrizione Ci-Onecut (OC), espresso, a partire dallo stadio di neurula fino allo stadio larvale, nella vescicola sensoria, nel ganglio viscerale e in alcune cellule del cordone nervoso caudale. L'analisi era dunque proseguita con la dimostrazione della co-localizzazione di tali geni, e la diretta interazione tra il fattore di trascrizione Ci-Onecut e le sequenze enhancer di *Ci-Rx.*

Nella mia attività di ricerca, mi sono occupata dello studio della cascata genica del fattore trascrizionale Ci-Onecut nella regolazione di *Ci-Rx.*
Lo studio si è suddiviso in due fasi parallele:
1. studio della relazione funzionale tra *Ci-Onecut* e *Ci-Rx*.
2. caratterizzazione del promotore di *Ci-Onecut*, allo scopo di trovare i fattori di trascrizione a monte del gene, in grado di regolarne l'espressione tessuto specifica.

1. Per lo studio funzionale ho messo a punto esperimenti di perturbazione promotore guidata, utilizzando come *enhancer* la regione regolativa del gene *Ci-Etr*, come strumento per esprimere in tutto il sistema nervoso, compresi i territori che includono i preursori delle cellule pigmentate e fotorecettrici, e già a partire dallo stadio di 110-cellule (ovvero prima ancora che abbia inizio la trascrizione di *Ci-Onecut*) una forma costitutivamente attiva (ovvero fusa "in frame", al *C*-terminale, al dominio di attivazione VP16) e una forma costitutivamente repressiva (fusa "in frame", al *C*-terminale, al dominio di inattivazione WRPW) della sequenza codificante di *Ci-Onecut*. Da tali esperimenti sono emerse una serie di anomalie fenotipiche riguardo la morfologia, la pigmentazione o il numero delle cellule pigmentate (otolite ed ocello). Dal momento che la/le cellule alterate somigliavano talvolta all'otolite, talvolta all'ocello, allo scopo di chiarire la natura degli organi pigmentati ho quindi eseguito saggi

immunoistochimici per la βγ-cristallina, una proteina strutturale espressa specificamente nei palpi e nell'otolite. Dagli esperimenti è emerso che, come il gene *Ci-Rx*, anche *Ci-Onecut* potrebbe essere coinvolto nella regolazione dei geni coinvolti nella specificazione delle cellule pigmentate, e più in particolare della cellula che compone l'ocello. Per poter dimostrare una diretta connessione funzionale tra questi due geni ho quindi effettuato esperimenti di ibridazione *in situ* usando come sonda ad RNA una sequenza complementare a quella del gene *Ci-Rx,* per verificare, nelle larve in cui la proteina Ci-Onecut era stata perturbata, l'abilità di Ci-Onecut di influenzare l'espressione del gene *Ci-Rx* endogeno. Dagli esperimenti si è evinto che l'espressione genica di *Ci-Rx* dipende da quella della proteina Ci-Onecut e che Ci-Onecut agisce come attivatore trascrizionale, essendo esso necessario per l'espressione di *Ci-Rx* nella vescicola sensoria (D'Aniello *et al.,* 2011).

2. Per lo studio della regione promotrice del gene *Ci-Onecut* ho effettuato un'analisi per delezione di una regione intergenica di circa 3 kb, che separa *Ci-Onecut* dal gene che lo precede al 5'. I vari frammenti parzialmente sovrapposti, ottenuti per PCR, sono stati clonati a monte del gene reporter, ed elettroporati in uova fecondate di *Ciona intestinalis*. L'analisi mi ha portato all'isolamento di un frammento minimo di 262 bp in grado di ricapitolare l'espressione del gene reporter nei territori endogeni di *Ci-Onecut*. Tale frammento è stato poi sottoposto ad un'analisi *in silico*, dalla quale sono emersi Ci-Onecut e Ci-Neurogenina (Ngn) come putativi fattori di trascrizione in grado di riconoscere il frammento minimo del promotore e quindi di regolare l'espressione genica di *Ci-Onecut*. Dopo averne dimostrato la parziale co-localizzazione genica, per dimostrare la diretta interazione *in vivo* tra i fattori OC e Ngn coi loro specifici siti di legame, ho messo a punto esperimenti di iper-espressione ectopica delle proteine OC e Ngn nella notocorda. A tale scopo mi sono servita di una strategia promotore guidata, utilizzando il promotore del gene *Brachyury* (Corbo *et al*., 1997; Yamada *et al*., 2003) come strumento per esprimere le proteine OC e Ngn nelle cellule della notocorda. Esperimenti di co-lettroporazione nelle uova fecondate di *Ciona* delle proteine iper-espresse, insieme al costrutto contenente l'enhancer del promotore di *Ci-Onecut* a monte del gene reporter hanno dimostrato che le proteine OC e Ngn, espresse nella notocorda, sono in grado di riconoscere i propri siti di legame sul frammento minimo di promotore e sono in grado di attivarli per avere l'espressione del gene reporter in un territorio dove normalmente né *Ci-Onecut* né *Ci-Neurogenina* si esprimono. L'analisi del meccanismo regolativo è stata effettuata mediante la mutagenesi sito-diretta. Varie combinazioni di frammenti mutati nei siti di legame per Ngn e OC e clonati a monte del gene reporter sono stati saggiati *via* elettroporazione in uova fecondate di *Ciona* e, dall'analisi fenotipica e statistica degli embrioni osservati, si è evinto che esiste una cooperazione tra i fattori Ngn e OC. Dai

dati riportati si è dimostrato un meccanismo regolativo secondo cui la Ngn agisce da fattore necessario e sufficiente all'attivazione di *Ci-Onecut* nei suoi territori endogeni e OC come fattore necessario al suo mantenimento; una volta tradotta, infatti, la proteina OC è in grado di autoregolarsi, innescando un meccanismo di *feedback* positivo attraverso il quale riconosce i suoi siti di legame contenuti sul suo promotore minimo e mantiene la sua espressione tessuto-specifica.

SUMMARY

Functional similarity and conservation of basic gene expression patterns led some authors to infer that the ascidian ocellus sensory organ, the vertebrate eye and the pineal organ could have been derived from a common archetypal "visual organ". Genes known to be involved in vertebrate eye and pineal gland development and function appear analogously expressed in tunicate pigmented cell territories. Transcription factors, such as Rx (Retinal Homeobox), play a fundamental role in vertebrate regionalization and eye development. This gene is highly conserved between vertebrates and invertebrates. It specializes in the development of eyes starting from the pre-chordates to the vertebrates, and it is involved in the formation of the central nervous system in invertebrates.

Although several studies have been done on the function of the nervous system patterning, little is known about the factors that bind *Rx* promoters and activate their expression.

Ciona intestinalis represents a simple model system to investigate on the network in which this important factor is involved and only one *Rx* gene (*Ci-Rx*) has been identified in this ascidian (D'Aniello et al., 2006).

In the group in which I worked, it was already demonstrated that the *Ci-Rx*, expressed in photoreceptor cells, has a fundamental role in the formation of the photosensitive ocellus sensory organ and in the differentiation of photoreceptor cells. Morpholino against *Ci-Rx* in fact showed that the ocellus does not develop (D'Aniello *et al.,* 2006). The characterization of the *Ci-Rx* regulatory sequence led to the isolation of two minimal enhancers, able to direct the expression of the reporter gene in the same territories of the endogenous transcript. Through bioinformatics analysis, it was identified, on both these sequences, a possible binding site for the Ci-Onecut (OC) transcription factor. In order to demonstrate a possible co-localization of *Ci-Rx* and *Ci-Onecut* expressions in the same cells, a preliminary study was carried out demonstrating that these two genes partially co-localize in the same cells in the anterior part of the brain. The effective interaction between OC and the *Ci-Rx* minimal regulatory regions was demonstrated by an Electrophoretic Mobility Shift Assay (D'Aniello *et al.,* 2011).

The first goal of my project was focused on the analysis of *Ci-Rx* as a downstream gene of the Ci-Onecut cascade and to study the functional relationship between these two genes. In order to assess the ability of Ci-Onecut protein to influence the *Ci-Rx* endogenous expression, I used over-expression experiments of perturbed Ci-Onecut proteins, composed by the *Ci-Onecut* coding sequence fused in frame in a case with an active domain (VP16), and in another case with a repressor domain (WRPW). I used a promoter-driven strategy in order to express these perturbed proteins in the

whole nervous system, including photoreceptor and pigmented cells precursors. *In situ* hybridization experiments on the transgenic larvae, electroporated with the perturbed Ci-Onecut proteins, using an RNA probe complementary to the *Ci-Rx* gene, were performed. The results demonstrate a functional connection between *Ci-Onecut* and *Ci-Rx* expression and that Ci-Onecut acts as a transcriptional activator, being it necessary for the *Ci-Rx* expression in the anterior part of the nervous system (D'Aniello *et al.,* 2011). Furthermore the phenotypes obtained, analyzed by microscopy, revealed strong alteration in morphology, pigmentation and number of the pigmented cells. In particular, by using of immunohistochemical assay experiments my data seem to indicate that, like Ci-Rx*,* also Ci-Onecut could be a good candidate involved in the regulation of pigmented cells genes and in particular it could have a functional role in the ocellus specification.

The second goal of my project focused mainly on the characterization of the regulatory elements underlying the precise spatio-temporal activation of *Ci-Onecut* gene during embryogenesis. This study revealed new informations on the genetic pathway that controls central nervous system patterning during the early stages of development. *Ci-Onecut* regulatory elements were analyzed by electroporation of embryos with a series of constructs containing various *Ci-Onecut* 5' promoter fragments fused to a reporter gene. These studies allowed the identification of a 262 bp region able to induce the expression of the reporter gene, from neurula to larva stage in the same territories in which the endogenous transcript is expressed. This 262 bp sequence contains different motifs recognized by various transcription factors. Detailed *in silico* and *in vivo* studies, based on ectopic and promoter-driven strategy expression experiments, allowed me to demonstrate that the transcription factors responsible for the *Ci-Onecut* promoter activation in *C. intestinalis* are Ci-Neurogenin (Ngn) and Ci-Onecut (OC).

All together my results suggest that once activated by Ngn, an autoregulatory mechanism maintains the *Ci-Onecut* expression in its endogenous territories. My study identified, for the first time among the Ci-Onecut transcription factors, a positive *feedback* regulation, responsible for the spatial-temporal gene expression.

SUMMARY

LIST OF TABLES AND FIGURES

1. INTRODUCTION

1.1 Ciona intestinalis as model system

Tunicates or urochordates (larvacean, salps and ascidian), cephalochordates (lancelets) and vertebrates (including lamprey and hagfish) belong to the phylum of chordates. The embryologist Alexander Kowalewski noted, in the 1866, the striking similarities between ascidians and vertebrates embryos. His observation of a notochord and a dorsal neural tube in ascidian larva suggested that ascidians are member of the phylum *Chordata*. Despite the evolutionary distance between extant ascidians and modern vertebrates such as human, mouse, chick, frog and zebrafish, the basic features of the vertebrate body plan are recognizable in the ascidian larvae. It is vertebrate-like in having a dorsal hollow nerve cord, notochord and segmental muscle. Until recently cephalochordates have been considered to be the closest group to vertebrates while tunicates are thought to be representative of the basal chordate lineage. New phylogenetic analyses, supported the idea that cephalochordates are the most basal extant chordate while tunicates are the closest living invertebrate relatives to vertebrates (Delsuc *et al*., 2006; Dunn *et al*., 2008; Vienne and Pontarotti, 2006). Nevertheless, this hypothesis is still controversial (Bourlat *et al*., 2006) and needs to be verified by further evidence.

The ascidian *Ciona intestinalis* has emerged in the last years as a model system to understand the evolution of chordate and to explain the complex molecular processes underlying chordate embryogenesis.

1.2 Ciona intestinalis development

Thanks to the peculiar evolutionary position, ascidian represents a simple experimental model to investigate the molecular mechanisms responsible for cell-fate specification during embryonic development. In contrast to the blastomeres of vertebrates, those of most ascidians are fate-restricted, and give rise to a single tissue from the beginning of gastrulation, shortly before the 110-cell stage (Munro *et al*., 2006) (Fig. 1.1). Ascidians have been considered for more than 100 years the classical organism to study what is known as “mosaic development” (Conklin, 1905a). The mosaic behaviour of isolated blastomeres has been taken as evidence that prelocalized ooplasmic factors (cytoplasmatic determinants) specify tissue precursor cells during embryogenesis (Conklin, 1905b).

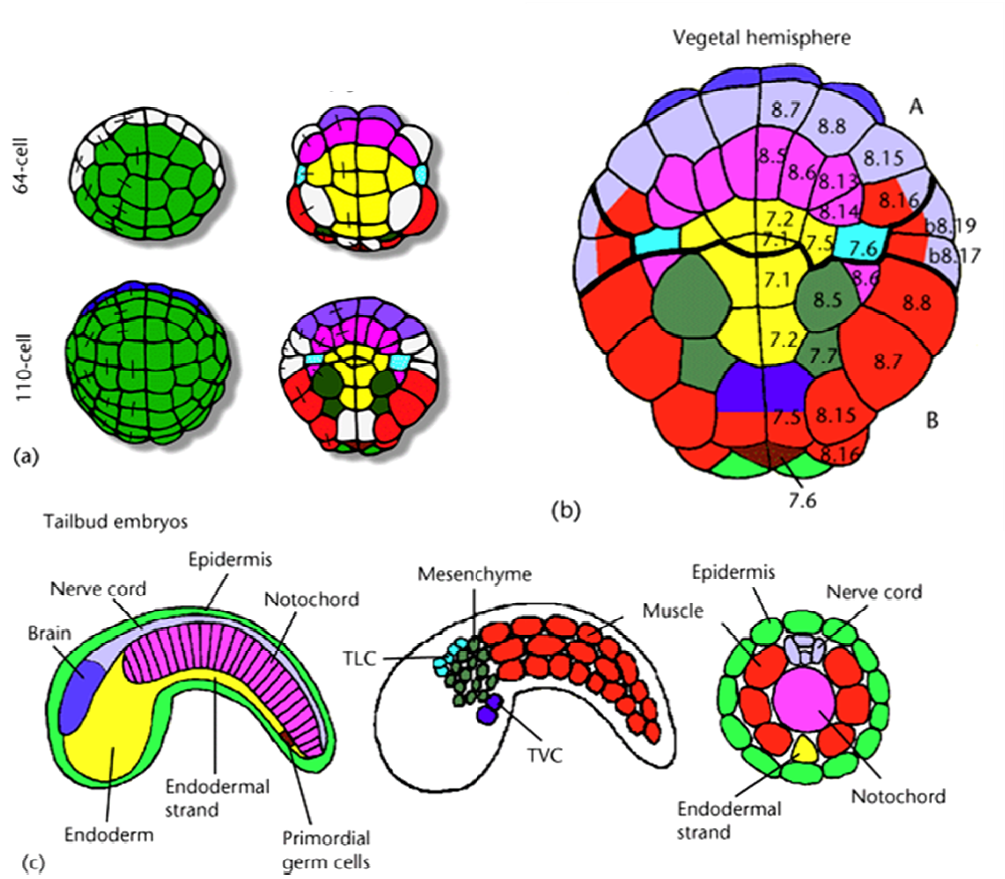


Figure 1.2: ***Fate maps of ascidian embryos.*** (A) (B) Fate map of the 64-cell stage ascidian embryo. Animal and vegetal hemispheres are shown. Circum-notochord side is to the left. Bars connect sister blastomeres generated during the previous cleavage. Colouring with a single colour indicates that their fate is restricted to a single tissue type. (C) Organization of ascidian tailbud embryos (Schematic overview of the major tissue types. From Munro *et al*., 2006. (The colour code is the same as in (A). Mid-sagittal plane, sagittal plane, and transverse section of the tail, showing the phenotypic features of chordates, including a dorsal nervous system and a tail with a notochord flanked by muscle. Adapted from "Lemaire *et al*., 2008". TLC, trunk lateral cells (precursors of blood and adult body wall muscle); TVC, trunk ventral cells (precursors of heart and adult body wall muscle).

When blastomeres are dissociated from an early cleavage stage until the 110-cell stage and are allowed to develop in partial embryos, they show that the differentiation of epidermis, muscle and endoderm is autonomous without the involvement of any cell-cell signalling. In contrast, the notochord, brain and sensory pigment cells failed to differentiate when embryonic cells were dissociated, suggesting that the development of those tissues is non-autonomous and requires cell interactions (Nishida, 1997). At tailbud stage the main tissues of *C. intestinalis* are already specified (Fig. 1.1B).

The *C. intestinalis* larva is composed of only 2500 cells and, when fully developed, its major components are essentially: the notochord, muscle, epidermis, endoderm, mesenchyme and the nervous system (Katz, 1983). The notochord is made of only 40 cells, the muscle of 36 cells and the central nervous system of approximately 350 cells, of which 100 are neurons (Satoh, 1994) (Fig. 1.1C). Ascidians develop in a relatively short time, in fact a swimming tadpole hatches within 24 hours after fertilization at 18°C (Fig. 1.2B-L).

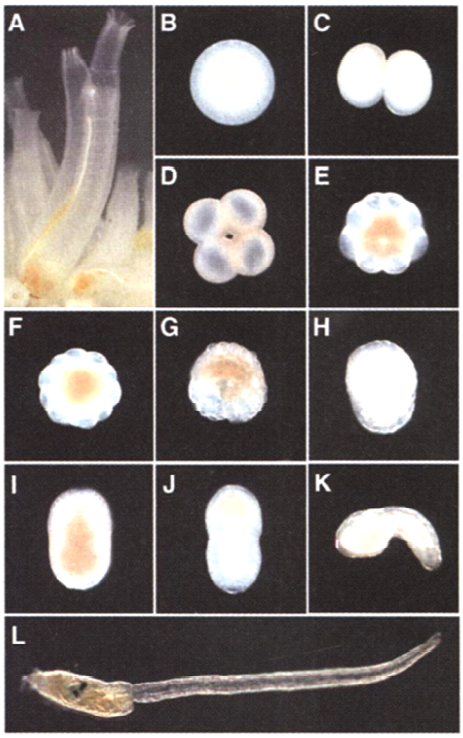


Figure 1.2: ***The sea squirt Ciona intestinalis.*** (A) Adults with incurrent and outcurrent siphons. The white duct is the sperm duct, while the orange duct paralleling it is the egg duct. (B to L) Embryogenesis. (B) Fertilized egg, (C) 2-cell embryo, (D) 4-cell embryo, (E) 16-cell embryo, (F) 32-cell embryo, (G) gastrula (about 150 cells), (H) neurula, (I to K) tailbud embryos, and (L) tadpole larva. Embryos were dechorionated to show their outer morphology clearly. From Dehal *et al*., 2002.

The nonfeeding larva swims for a few hours and then metamorphoses into a juvenile. After 1 or two months depending on the temperature of the water, the juvenile develops in an adult with reproductive capacity (Satoh, 1994) (Fig. 1.2A).

1.3 Ciona intestinalis genome organization

The *C. intestinalis* genome is ~155 Mb in size, one-twentieth the size of mouse genome and contains ~16000 genes distributed over 14 chromosomes. The urochordate lineage diverged before the extensive gene duplications occurred in vertebrates and in most cases *C. intestinalis* genome contains a single homologue of the multiple paralogous genes present in vertebrates. Only six fibroblast growth factor (FGF) genes, for example, are present in the *C. intestinalis* genome while 22 FGF genes are present in the vertebrate genomes (Satou *et al*., 2002). This peculiarity, together with the small size of

the genome, is a unique advantage to understand genome organization, gene function and regulation avoiding the functional redundancy typical of vertebrates (Satoh *et al*., 2003). The genome of *C. intestinalis* has undergone a considerable DNA loss not only of intergenic regions, but also of genes, indicative of a rapid genome evolution. For example, *Ciona* has only nine *Hox* genes, lacking *Hox*7, *Hox*8, *Hox*9 and *Hox*11 (Spagnuolo *et al*., 2003). Moreover, *Ciona Hox* genes are not linked together in a single cluster and are not expressed temporally in the expected colinear pattern (Ferrier and Holland, 2002; Holland and Gibson-Brown, 2003). In addition, also the appendicularian *Oikopleura dioca* lacks these four *Hox* genes suggesting that this loss occurred at the base of tunicate lineage (Seo *et al*., 2001).

1.4 Comparison of ascidian and vertebrate nervous systems

Many studies revealed that ascidian larva and vertebrate embryo show anatomical and molecular homologies. The ascidian larva has a miniature CNS with many chordate credentials. It contains less than 100 neurons and 250 glial cells whilst the peripheral nervous system is composed of 20-30 epidermal sensory neurons (Meinertzhagen and Okamura, 2001) and structurally it is composed of four main regions along the anteroposterior axis: the sensory vesicle, the neck, the visceral ganglion, and the caudal nerve cord. The CNS develops from a neural plate, but unlike vertebrates, its anterior most region does not roll up and internalize but gives rise to the dorso anterior epidermis which includes the adhesive organ, head sensory neurons and pharynx. The posterior part of the neural plate develops into the four regions of the dorsally located CNS (Meinertzhagen *et al.*, 2004; Nishida, 1987). The sensory vesicle is the most anterior part of the CNS and contains two pigmented cells: the anterior spherical otolith and the ocellus. The otolith is considered the gravity organ, is a spherical mass of pigmented granules connected to the floor of the sensory vesicle by a narrow stalk. In the right posterior wall of the sensory vesicle is situated the ocellus that is considered the photo sensing organ of the *C. intestinalis* larvae. It is formed of three parts: a cup-shaped cell containing many small pigmented granules, three lens cells and about 30 photoreceptor cells (Horie *et al*., 2005).

The molecular mechanisms regulating ascidian pigment cell development are not yet clear. The two pigment cells arise from the paired a8.25 blastomeres that are positioned bilaterally in the gastrula embryo and will give rise to the a9.49 pair at neural plate stage (Fig. 2.3A, B). It is known that determination of the a8.25 cells as pigment cell precursors requires direct inductive influence from the nerve cord precursor cells at gastrula stage. At this stage, the a8.25 blastomeres constitute an equivalence group in the sense that both have the potential to form either an ocellus or otolith pigmented cell (Darras and Nishida, 2001).

The choice to adopt either otolith or ocellus cell fate is made after neural tube closure at early tailbud stage and requires cell-cell interactions. As the neural tube closes, the four cells derived from the division of the two a8.25 cells, the a10.97 and a10.98 pairs, converge and intercalate along the anterior-posterior axis in order to align along the midline (Fig. 2.3C). The a10.98 pair gives rise to part of the brain vesicle, the anterior a10.97 forms the otolith while the posterior one differentiates into the ocellus pigment cell. In the 2004 Cole and Meinertzhagen examined the mitotic history of the embryonic cells forming the CNS in the *Ciona* larva. They described that 17 dorsal cells, out of about 32, progeny of the right a9.33 and a 9.37 blastomeres become photoreceptor cells (Fig. 1.3B).

Just posteriorly to the sensory vesicle there is the neck containing only six cells whose function are still unclear. The visceral ganglion contains 45 cells with 18 neurons, of which five pairs are motoneurons (Nicol and Meinertzhagen, 1991). Finally, the tail nerve cord extends along the entire larval tail and is formed by four cells.

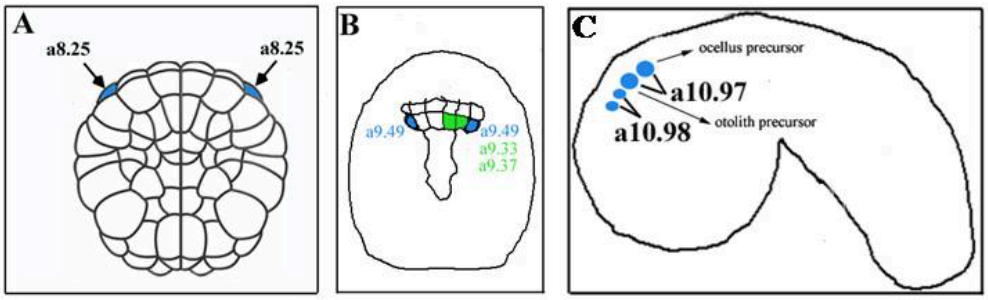


Figure 1.3: ***Cell lineage of pigmented (light blue) and photoreceptor cells (green) precursors in ascidian embryos at 110 cells.*** (A), neural plate (B) and tailbud stages (C). Adapted from Locascio *et al*., 1999.

The structure of ascidian CNS shows considerable similarities to those of vertebrates. For example, the territories of expression of different genes, responsible for vertebrate CNS patterning, have been investigated in the ascidian nervous system. In *C. intestinalis* the *Otx* homologue is expressed in the anteriormost region of the larval CNS, the sensory vesicle, and this expression corresponds to those of *Otx1* and *Otx2* in the anterior vertebrate CNS suggesting that the ascidian sensory vesicle can be considered homologous to the forebrain of Vertebrates (Lemaire *et al*., 2002).

Nevertheless, the *Emx* gene, a marker of vertebrate telencephalon, is expressed in the dorsoanterior epidermis but not in the sensory vesicle (Oda and Saiga, 2001). One interpretation of this finding may be that ascidians

lack the telencephalon as it has been supposed also for amphioxus (Holland and Holland, 2001).

Interestingly, the expression pattern of *Pax-2/5/8,* and of its vertebrate homologues are also conserved. In the ascidian *Halocynthia roretzi,* this gene is expressed in the neck region of the larval CNS, while in vertebrates their expression is in the midbrain and anterior hindbrain suggesting that the neck region could be considered the ascidian homolog of the vertebrate midhindbrain boundary (Gruss and Walther, 1992; Wada *et al*., 1998). Proceeding along the antero-posterior axis of the ascidian tadpole, the visceral ganglion expresses *Hox1* anteriorly (Katsuyama *et al*., 1995) and *Hox3* posteriorly (Locascio *et al*., 1999). Even thought, the visceral ganglion does not show a visible segmentation, this structure may therefore be homologous to the vertebrate caudal rhomboencephalon (rhombomeres 4-6) (Lemaire *et al*., 2002). The ascidian anterior neural tube can be compared with the vertebrate anterior spinal cord since these two regions are both marked by *Hox5* expression (Gionti *et al*., 1998).

Taken together, the morphological and molecular similarities between the vertebrate and ascidian CNS suggest a common evolutionary origin of these structures. Although the vertebrate CNS has undergone complex elaboration it seems that many of the basic mechanisms involved in establishing and patterning the CNS have been set early in the evolution of chordates (Passamaneck and Di Gregorio, 2005).

The conservation of the structures at morphological level does not correspond to all these molecular data of homology, indicating that developmental pathways for common structures involve the same molecules but used in a vastly different way (Lemaire *et al*., 2008).

1.5 Ciona intestinalis as a model system to analyze expression and function of developmental genes

The ascidian *C. intestinalis* has been studied for over a century as a model system for embryological studies. As model, ascidians offer several advantages. First, a single animal can produce thousand of embryos and their transparency makes them ideal for live-imaging studies.

Gene expression analysis by whole mount *in situ* hybridization is very simple and fast. At early stages of development, the signals for zygotic expression are first seen in the nuclei and, as development proceeds, are distributed over the entire cytoplasm.

Transgenic experiments in *C. intestinalis* are possible by the introduction of plasmid DNA into the fertilized eggs by an electroporation technique (Corbo *et al*., 1997). This method allows rapid screening of genomic DNA fragments linked to a reporter gene with the aim to identify regulatory elements controlling gene expression. Taking advantage of this technique it is

also possible to misexpress a gene in an ectopic tissue or to disrupt a specific gene activity through the over expression of a corresponding dominant negative protein (Di Gregorio and Levine, 2002). Moreover, electroporation has also been used to assess if orthologues genes, isolated from other organisms, are able to mimic the activities of ascidian genes (Satoh *et al*., 2000). It is also possible to transform individual embryos with two different transgenes. This co-electroporation method is very useful to label specific tissues and at the same time analyze the consequences of the altered expression or function of a desired gene. The only disadvantage of the electroporation technique is the mosaic incorporation of the construct, after a variable number of cell divisions. The transgene can segregate into only one of the two daughter cells and, as consequence, at late stages of development a variable number of cells within a tissue will lack the transgene (Di Gregorio and Levine, 2002).

Recently it has been demonstrated that also in *C. intestinalis* it is possible the disruption of gene function taking advantage of the powerful method of the RNA interference (Nishiyama and Fujiwara, 2008).

1.6 Homeobox genes

Homeobox genes play a key role in the tissue-specific, spatial and temporal regulation of genes required for cell differentiation. The genes belonging to the family of homeobox encode transcription factors highly conserved during evolution and represent key regulators in multiple developmental processes in cell differentiation and morphogenesis.

The homeobox genes share a sequence of approximately 180 bp, the homeobox, that encodes a homeodomain protein of 60 aminoacids. A key feature that allows homeoproteins to fulfill so many and various functions is the presence of various additional DNA binding motifs, such as those found in homeoproteins of the Paired, POU, and LIM class (Gehring, 2004).

Numerous homeobox genes have been discovered both in vertebrates and in invertebrates; in Metazoa they are more than 200. Despite the high degree of conservation of residues within the homeodomain, transcription factors with the homeodomain may recognize target sequences different from each other. Small variations, such as the replacement of a single aminoacid within the homeodomain, are able to alter the binding specificity (Treisman *et al*., 1989; Kuziora and McGinnis, 1989). The interaction with other transcription factors which possess homeodomains or other DNA-binding motifs and post-transcriptional modification, such as acetylation and phosphorylation (Wigle and Eisenstat, 2008), are key determinants for the binding specificity.

1.7 Rx genes

The *Rx* genes (Retinal homeobox) belong to the family of homeobox genes and encode proteins characterized by two domains: the homeodomain and the Paired domain (Mathers *et al*., 1997). The expression of *Rx* genes has been studied during the early embryonic stages of all species which have been identified and it has been shown that in Chordates they are involved in the determination of the cells which will give rise to the retina and, probably, in their proliferation.

In zebrafish three *Rx* genes (Zrx1, Zrx2, Zrx3) were characterized, two in Xenopus (Xrx1 and Xrx2) and chicken, one in humans, in mouse, in planarians *Dugesia japonica* and *Schmidtea mediterranea*, in *Drosophila melanogaster* and in ascidian *C. intestinalis* (Chuang *et al.,* 1999; Furukawa *et al.,* 1997; Casarosa *et al.,* 1997; Voronina *et al.,* 2004; Mathers *et al.,* 1997; Mannini *et al.,* 2008; Eggert *et al.,* 1998 , D'Aniello *et al.,* 2006; D'Aniello *et al.,* 2011).

In the laboratory of Cell and Development Biology of the Zoological Station Anton Dohrn of Naples, in the research group of Dr. Branno, with whom I did my thesis activity, the Rx gene in *C. intestinalis* (*Ci-Rx*) has been identified and isolated, and it started a study to try to understand the role of the Ci-Rx transcription factor in embryonic development and the characterization of the gene cascade in which it is involved (D'Aniello *et al.,* 2006) (D'Aniello *et al.,* 2011).

Ci-Rx has been classified as paired-like gene belonging to the subclass Q50 because it presents a residue of glutamine in position 50 essential for the recognition and for the DNA binding. Whole mount *in situ* hybridization experiments performed on *Ciona* embryos at various stages of development, revealed that *Ci-Rx* transcript appears initially at early tailbud stage in a group of three cells in the anterior part of the head that correspond to the sensory vesicle precursors (Fig. 1.4B).

At middle tailbud stage the transcript continues to be expressed in the two anterior groups of cells and appears also in one more posterior cell along the dorsal midline (Fig. 1.4D).

The *Ci-Rx* expression in the posterior cell is transient and the signal disappears at the late tailbud stage (Fig. 1.4E, F). At this stage, also the frontal left signal becomes weaker and with advancing development it completely disappears (Fig. 1.4G, H). At larva stage, finally, the expression of the gene is restricted in the posterior part of the vesicle sensory, in a small number of cells that surround the ocellus (Fig. 1.4 I).

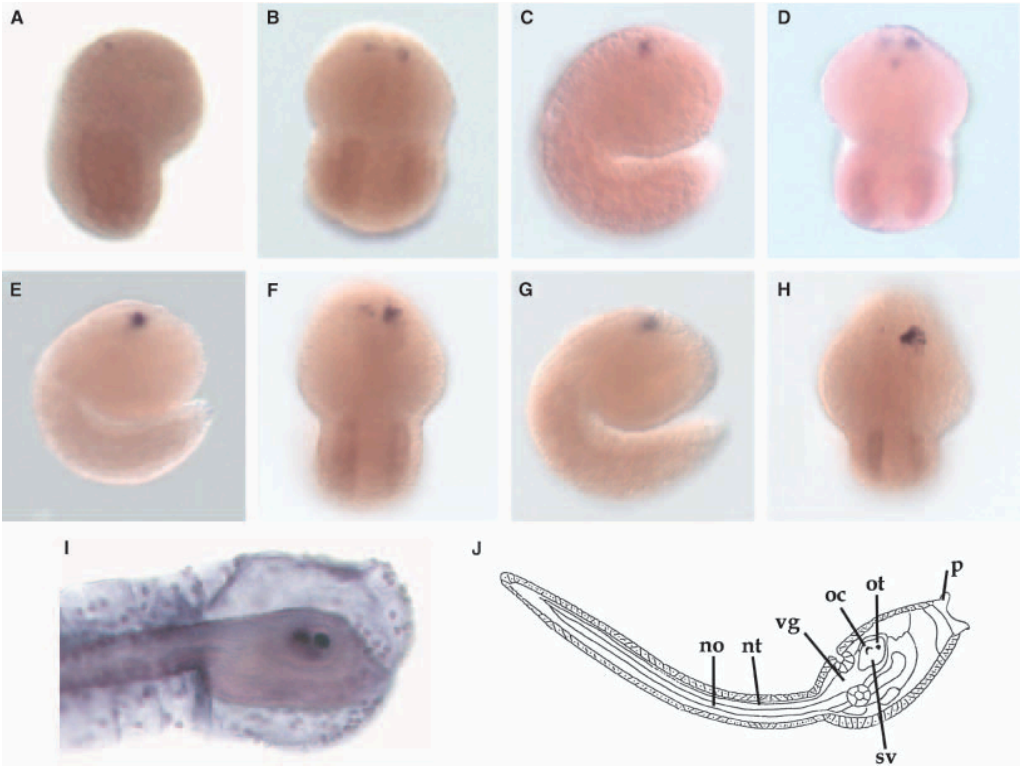


Figure 1.4: ***In situ localization of Ci-Rx transcript in Ciona intestinalis embryos***. Lateral and dorsal views of embryos at early- (A, B), middle- (C, D), and late-tailbud (E–H) stages. *Ci-Rx* is expressed in two groups of cells precursors of the anterior brain at early (A, B) and middle tailbud (C, D) stages. A third signal appears only at middle tailbud (D). At late tailbud, the left signal in the anterior brain precursors tends to disappear (F, H). (I) Larva stage, laterall view, showing a signal in the posterior sensory vesicle around the ocellus. (J) Schematic representation of a *Ciona* larva. no, notochord; nt, neural tube; oc, ocellus; ot, otolith; p, palps; sv, sensory vesicle; vg, visceral ganglion. From D'Aniello *et al*., 2006.

The functional role of the *Ci-Rx* gene during embryogenesis of *C. intestinalis* was analyzed by gene knock-down experiments through microinjection of morpholino antisense oligonucleotide.

The eggs injected with the antisense oligonucleotide develop into larvae which present a single pigmented cell, the otolith, while the ocellus is completely absent. Furthermore by electrophysiological studies it was observed that these larvae have a complete lack of sensitivity to light stimuli. To better clarify the role of *Ci-Rx* in the development and function of the ocellus in *C. intestinalis*, whole mount *in situ* hybridization with two specific photoreceptor markers, *Ci-opsin*1 (ortholog of the vertebrate rhodopsin) and *Ci-arrestin* (ortholog of the visual and β vertebrate *arrestin*), has been

performed on *Ci-Rx* mutant embryos. This analysis showed that the larvae injected with the *Ci-Rx* morpholino completely lack the expressions of *Ci-arrestin* and *Ci-opsin*1 (D'Aniello *et al*., 2006). These data suggest that *Ci-Rx* has a role in the development and differentiation of the photoreceptor cells of *C. intestinalis*.

Furthermore, studies on the promoter sequence of the *Ci-Rx* gene in *Ciona* have allowed the identification of two regions, of approximately 35 kb and 50 kb, respectively, capable of directing the expression of the reporter gene in the same territories of expression of the endogenous gene.

The analysis of these sequences using the bioinformatics software GENOMATIX (http://www.genomatix.de/cgi-bin/eldorado.main.pl) suggested the presence of a putative binding site for the Onecut transcription factor.

1.8 Cut genes

Cut genes are a class of genes that encode for homeobox containing transcription factors with a bipartite DNA binding domain that includes a Cut domain, of about 70 amino acids, and the homeodomain. The bipartite DNA-binding domain, which works independently with the homeodomain or in a cooperative way, is a good way to increase the specificity and affinity in binding to DNA (Ades and Sauer, 1994). Furthermore the homeodomain of this class of proteins has a lower affinity and specificity of DNA binding, unlike the classic homeodomain (Ades and Sauer, 1994). In fact, the DNA binding is mediated primarily by Cut domain, although the sequence specificity and the kinetics of binding is largely influenced by the presence of the homeodomain (Lannoy *et al.,* 1998; Catt *et al.,* 1999). The Cut proteins are classified into three groups depending on the presence of a one (Onecut), two (SATB) or three (CUX) Cut domain repetitions preceding the homeodomain. In agreement with their structural diversity, the Cut proteins perform a variety of functions and they represent an important class of transcriptional regulators involved in the development, both of vertebrates that of invertebrates.

1.9 Onecut genes

Onecut proteins have been identified in humans, mice, in the nematode *C. elegans*, the fruit fly *D. melanogaster*, in the sea urchin *Strangylocentrotus purpuratus*, in the zebrafish *Danio rerio* and in ascidian *Halocynthia roretzi* (Jacquemin *et al*., 2001; Vanhorenbeeck *et al*., 2002; Lannoy *et al*., 1998; Nguyen *et al*., 2000; Matthews *et al*., 2008; Sasakura and Makabe, 2001). The high degree of conservation of the cut domain and of the homeodomain in the Onecut genes during evolution, highlights the key role played by these transcription factors in controlling many cellular

processes. Onecut proteins have two separate domains connected by a long flexible linker (Fig. 1.5A).

The Cut domain is composed of four alpha helices and the hydrophobic core is formed by the interaction of the helix 1 with the helix 5; the helix 3 interacts in the major groove of the DNA and it is important for the recognition. The structure of the homeodomain is highly conserved; it consists of 3 alpha helices (H7-H8-H9) and interacts with the major groove of the DNA by the helix of recognition, H9 (Fig. 1.5B). The interactions with the DNA bases are via the *C*-terminal portion of the helix 3 (H9) of the homeodomain, unlike the other homeodomains that use the aminoterminal portion (Iyaguchi *et al*., 2006). In addition to the DNA binding domain the most bipartite Onecut proteins also share a highly conserved third region, involved in the transcriptional activation, the STP box, a sequence of about 28 amino acids rich in serine/threonine/proline; some aminoacid residues of this sequence were found in all members of the family, including C17H12, the only factor in *C. elegans* to possess such a sequence.

In the Cut domain of the Onecut protein is present a conserved motif, LSDLL, which plays an important role in the stimulation of transcription and allows interaction with other coactivators, such as CREB-binding protein (CBP) (Lannoy *et al*., 2000).

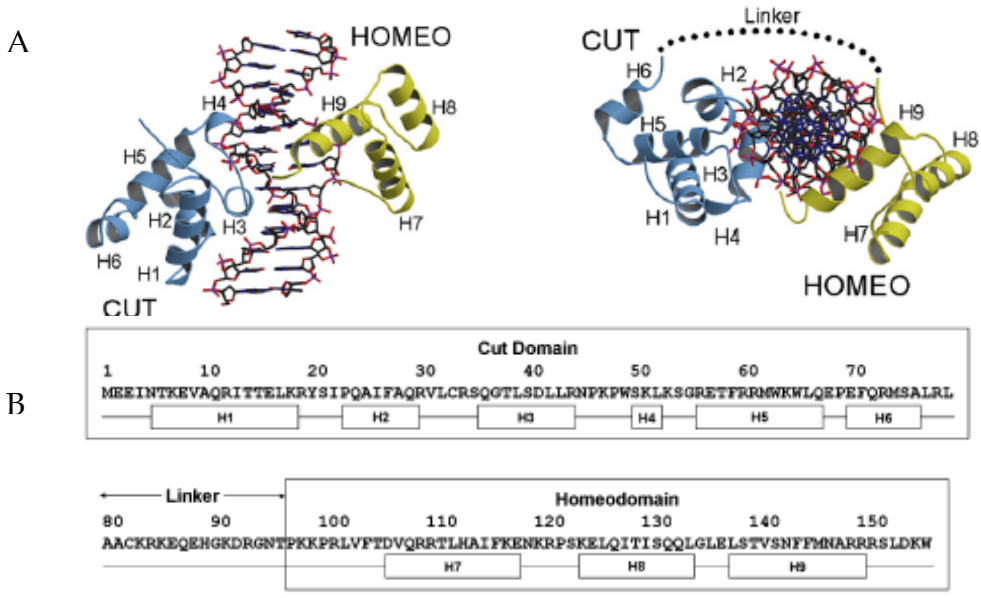


Figure 1.5: ***Schematization of the Ci-Onecut protein.*** (A) Three-dimensional image of the interaction between the Cut rule, the homeodomain and the double helix of DNA. (B). Cut domain amino acid sequence of the Onecut protein and of the homeodomain; the helices shown in the three-dimensional structure (H1-H9), are indicated below the sequence.

1.10 Onecut genes in vertebrates

In mammals three *Onecut* genes have been identified: *HNF-6* (Hepatocyte Nuclear Factor-6, also known as *Onecut1 (OC-1)*, *Onecut2 (OC-2)* and *Onecut3 (OC-3)*. In mice the expression of *HNF-6* and *OC-3* in adults is partially overlapped: both are expressed in the brain and organs derived from the endoderm, respectively, liver and pancreas for *HNF-6* and intestines and stomach for *OC-3*. HNF-6 factor is expressed in the endoderm earlier than OC-3 and it seems that HNF-6 is able to directly activate the expression of *OC-3* (Pierreux *et al*., 2004). The levels of expression are also overlapped for *HNF-6* and *OC-2* during the development of the liver and pancreas, although these two transcription factors control different target genes (Jacquemin *et al*., 1999; Vanhorenbeeck *et al*., 2002). With the exception of the brain, all organs that express the Onecut transcription factors in mammals derived from the endoderm. Moreover, HNF-6 checks the specification of the pancreas starting from the gut endoderm cells front, regulating the expression of *Pdx-1* gene, a transcription factor involved in the regulation of the development of exocrine, endocrine and ductal pancreatic cell precursors (Jacquemin *et al*., 2003a). The pancreas of mice, in which the *HNF-6* gene inactivation was carried *(HNF-6-/-)* is hypoplastic, due to the slowdown of the specification of endodermal pancreatic cells. Furthermore, HNF-6 regulates transcription of genes encoding enzymes involved in the metabolism of glucose and its expression is regulated by growth hormone, Growth Factor (GH) (Lahuna *et al*., 1997). This transcription factor promotes the differentiation of the precursors of the endocrine cells of the pancreas, stimulating the expression of the gene pro-endocrine *Ngn-3*. In fact the pancreas of mutant mice in which *HNF-6* has been inactivated is not properly formed and the islets of Langerhans are absent (Jacquemin *et al*., 2003b).

The HNF-6 and OC-2 factors regulate the differentiation of epatoblasts into hepatocytes and biliary cells, and are both required for the correct morphogenesis of the bile ducts (Clotman *et al*., 2005).

Gene inactivation experiments in mice have shown that in the absence of both *HNF-6* that of *OC-2* there are morphogenetic defects in bile duct nascent. In mice mutated only in *HNF-6* gene was found the same phenotype, less severe than those observed in the double mutant, while the single mutants in the gene *OC-2* have severe changes in the bile ducts (Clotman *et al*., 2005).

Overexpression or ectopic expression of *HNF-6*, instead, is in many cases associated with altered cellular homeostasis or diabetes. The concentration and the expression of *HNF-6* and *OC-2*, in fact, is finely adjusted by some microRNAs, such as miR-495 and miR-218 respectively that regulate *HNF-6* and *OC-2*. This hypothesis was confirmed by the fact that, when in the pancreas and in the liver these two miRs are depleted, *HNF-*

6 and *OC-2* are overexpressed. All these data demonstrate that HNF-6 plays a crucial role in the development of liver and cooperates with OC-2 in the bile specification (Simion *et al*., 2010).

The *Onecut* expression in the nervous system has been found in a few cases, even if the functions performed by these transcription factors have yet to be clarified.

It has been recently demonstrated in mouse that *OC-1* and *OC-2* have very similar expression patterns throughout retinal development and they may regulate the formation of retinal ganglion cells (RGCs) and also have a function in the genesis and manteinance of horizontal cells (HCs) (Fuguo *et al*.,2012). *HNF-6* is expressed in different areas of the central nervous system (Rausa *et al*., 1997), while *OC-2* and *OC-3* are expressed only in the brain.

Experiments performed on mice mutants for *HNF-6* indicate that it controls in motor neurons a genetic program that coordinates the formation of hindlimb neuromuscular junctions (Audouard *et al.,* 2012).

Onecut2 in humans is expressed in melanocytes and regulates the *MITF* gene (Microphthalmia-associated Transcription Factor), which encodes a transcription factor essential in the differentiation of melanocytes (Jacquemin *et al*., 2001). In zebrafish *D. rerio* three Onecut factors were characterized: Onecut1 or HNF-6 (Matthews *et al*., 2004), Neural Onecut (Hong *et al*., 2002) and Onecut3 (Matthews *et al*., 2008).

HNF-6 in zebrafish does not seem to have a major role in the bile specification, unlike its ortholog in mammals (Matthews *et al*., 2004).

Onecut3 shares high sequence homology with the *OC-3* gene in mammals, it is essential for the intrahepatic bile development and is implicated in the specification of the biliary epithelial cells (Matthews *et al*., 2008).

Neural Onecut was the first member of the Onecut family to be identified in the nervous system of vertebrates. This transcription factor is specifically expressed in neural cells and shows highly dynamic expression in primary neurons of the brain and spinal cord during zebrafish embryogenesis (Hong *et al*., 2002).

1.11 Onecut genes in invertebrates

In invertebrates only one gene coding for the Onecut factor was isolated, predominantly expressed in the nervous system. Although the role of *Onecut* genes in vertebrates has not yet been clarified, experiments conducted in different invertebrates suggest the key role of these factors in the formation of the neural tube in vertebrates.

In sea urchin *S. purpuratus* only one gene encoding the Onecut transcription factor has been found, *Sphnf6*, a maternal transcript that until the gastrulation is distributed in a uniform manner (Otim *et al*., 2004). In the

early stages of development its expression is required for the activation of genes involved in the differentiation of primary mesenchyme cells (PMC). After gastrulation Sphnf6 participates in the regulation of ectodermal oral genes (GRN) and in the formation of neural ciliate band. The oral ectoderm adjustment and the formation of the ciliate band are not regular in the absence of the Sphfn6 factor.

In *C. elegans* five *Onecut* genes have been characterized, but their function *in vivo* is still unknown (Lannoy *et al*., 1998). Factors Ceh-21 and Ceh-39 are able to recognize the same binding site for HNF-6 of vertebrate, then these factors may play a role similar to that carried out by HNF-6 in mammals. Another peculiar gene is *Ceh-38*, expressed in different tissues during development, particularly in the endoderm derivatives and in some types of neurons, as happens for the genes in mammals (Cassata *et al*., 1998).

Only one gene of the Onecut family has been identified in *D. melanogaster*, *D-Onecut*, that is expressed exclusively in the nervous system during embryonic development and in the adult it is involved in the specification or differentiation of neural cells (Nguyen *et al*., 2000).

Initially *D-Onecut* was isolated as a potential regulator of the gene coding for the rhodopsin in the photoreceptors cells (R) suggesting its involvement in the regulation of R cell differentiation during the final stages of development of the eye. The overexpression of a dominant negative form of D-Onecut specifically interferes with eye R cell differentiation but not with the determination of their cell fate (Nguyen *et al*., 2000).

Also in the ascidian *H. roretzi* a single *Onecut* gene was isolated, *HrHNF-6*, expressed exclusively in neural cells during embryogenesis where it performs two functions: it is involved in the differentiation of neural cells *via* activation of the *HrTBB2* gene, the homologue of the β-tubulin in mammals, and participates in the specification of the neural tube by adjusting a negative HrPax2/5/8 factor, the corresponding gene Pax-2, Pax-5, Pax-8 in vertebrates (Sasakura and Makabe, 2001). *HrHNF-6* is expressed in the adult stage in two organs that are derived from the endoderm: the endostile and pyloric gland. The endostile is considered to be the counterpart of the thyroid gland in vertebrates (Ogasawara *et al*., 1999), while the pyloric gland could be the homologue of the liver in vertebrates (Ermak 1977). Therefore, the expression of the *Onecut* gene in the endoderm of ascidians could correspond to that of *HNF-6* in hepatocytes of mammals.

1.12 Scientific hypothesis and aim of the thesis

Since ascidians are the closest living relatives of vertebrates, the understanding of their photoreceptor system can be a starting point to discover the origin and evolution of the eye of vertebrates (Delsuc *et al*., 2006). Previous studies have also shown that the photoreceptor cells of the

ocellus sensory organ in *Ciona intestinalis* exhibit morphological, physiological and molecular characteristics similar to those of vertebrates (D'Aniello *et al*., 2006).

The group with whom I did my work has already demonstrated the critical role of the *Ci-Rx* gene in the formation and differentiation of photoreceptor cells and ocellus (D'Aniello *et al*., 2006).

Studies conducted on the *Ci-Rx* promoter sequence have identified two fragments, respectively 35 and 50 bp, able to reproduce the expression of the endogenous *Ci-Rx* gene. These fragments were subjected to bioinformatic analysis using the GENOMATIX software, which identifies known transcription factors binding sites in the context of a nucleotide sequence. The analysis allowed to identify, on both fragments, a binding site for the Ci-Onecut (HNF-6) transcription factor.

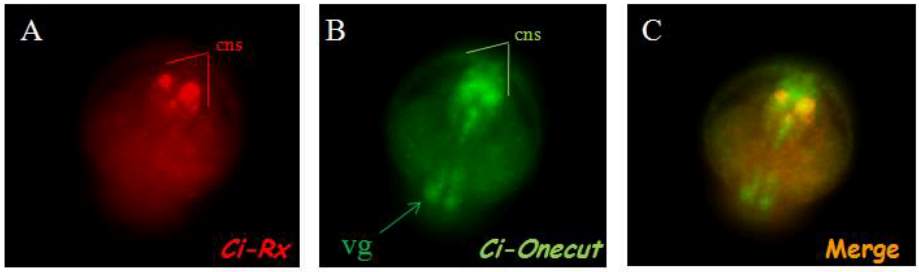


Figure 1.6: ***Positional relationship of Ci-Rx and Ci-Onecut gene expression domains.*** Two colour fluorescent W ISH at tailbud stage (dorsal view). A-C images taken using a fluorescent Zeiss Axio Imager microscope. At the bottom of each panel, *Ci-Rx* and *Ci-Onecut* are indicated in red and green colours respectively, corresponding to the colours of fluorescent staining. In C the two fluorescence images are merged. cns, central nervous system; vg, visceral ganglion.

From these results, it was initially demonstrated the partial spatiotemporal co-localization between the two transcripts through a double whole mount *in situ* hybridization experiments (Fig. 1.6). The effective interaction between Ci-Onecut and the *Ci-Rx* minimum enhancer elements has been demonstrated *in vitro* by an Electrophoretic Mobility Shift Assay (EMSA).

The first aim of my project has been the study of the functional role of *Ci-Onecut* gene in the regulation of the *Ci-Rx* gene cascade leading to the photoreceptor cells and ocellus formation, using *in vivo* approaches.

To this end, I used a promoter-targeted perturbation of *Ci-Onecut*, using a form of constitutive activator and a form of constitutive repressor of the Ci-Onecut protein in the whole nervous system and specifically in the photoreceptor and pigmented cells lineage.

The second step has been the characterization of the *Ci-Onecut* regulatory elements by dissecting the previously isolated 3 kb genomic fragment upstream the translation start site, in a series of partially overlapping and progressively deleted fragments. Once I identified the minimal sequence responsible for *Ci-Onecut* transcription, the following step has been to investigate and characterize the putative trans acting factor/s responsible for its activation using different *in silico* and *in vivo* approaches.

2. MATERIALS AND METHODS

2.1 Ciona eggs and embryos

C. intestinalis adults were collected in the bay of Naples and of Taranto. Eggs and sperm, collected from the gonoducts of several animals, were used for *in vitro* fertilization and fertilized eggs and embryos were used in transgenesis or *in situ* hybridization experiments. Embryos were raised in filtered sea water (FSW) at 18-20° C. I selected for my analysis only the experiments in which at least 60% or more of the embryos developed normally.

2.2 Chemical dechorionation and in vitro fertilization

Before to perform electroporation of the fertilized eggs, it is necessary to deprive the eggs of the chorion. The chemical dechorionation has been effectuated in a Petri dish with a thin layer of 1% agarose in FSW, putting the eggs for 5-6 minutes in a pH 10 solution of Thioglycolic acid (1%) and Proteinase E (0.05%) in MFSW. During this time, the eggs have been shaken continuously in this solution, using a glass pipette, to remove the chorion and the follicular cells surrounding the eggs. After this step the eggs have been washed several times to remove the residual solution and then fertilized *in vitro* with sperm collected from two or more individuals to avoid self-sterility problems. After 10 minutes, fertilized eggs have been washed 2-4 times to eliminate the exceeding sperm and then the embryos have been grown in MFSW at 18-20°C, and directly used or fixed at the suitable stages to perform subsequent experiments: whole-mount *in situ* hybridization (WMISH), immunohistochemistry (IHC).

2.3 Transgenesis by electroporation

In *C. intestinalis* it is possible to obtain transgenic embryos, introducing exogenous DNA into fertilized eggs, by electroporation. In comparison to microinjection technique, that is quite difficult to apply to these animals, due to the hardness of plasmatic membrane, this procedure permits to obtain many more transgenic embryos in one time. Nevertheless, it can occur the mosaicism phenomenon, in which a part of electroporated embryos show the expression of the transgenic DNA only in some regions of the body, due to the unequal incorporation of the transgene in different blastomeres during the first stages of the embryonic development.

❖ *Eggs electroporation*

The fertilized eggs have been immediately transferred in a solution containing 0.77 M Mannitol and 50-100 μg of DNA. The electroporation has

been carried out in Bio-Rad Gene Pulser 0.4 cm cuvettes, using a Bio-Rad Gene Pulser II electroporator , at constant 50 V and 500-800 μF, in order to have a time constant of 14-20 mseconds. The embryos have been allowed to develop until the desired stage on 1% agarose coated Petri dishes, at 18-20°C.

2.4 Embryos' coloration and observation

❖ *Lac-Z reporter gene constructs*

The embryos have been fixed in 1% glutaraldehyde in FSW for 15 minutes at RT, washed twice in 1x PBS and stained at 37°C in a solution containing 3 mM $K_3Fe(CN)_6$, 3 mM $K_4Fe(CN)_6$, 1 mM $MgCl_2$, 0.1% Tween 20, and 250 μg/ml X-gal in 1x PBS, for the time necessary to the staining. Then, they have been washed in 1x PBS, mounted on slides and observed at microscope.

❖ *GFP reporter gene constructs*

In these cases, the embryos and larvae have been observed *in vivo,* for the observation at the microscope of fluorescence and phenotypic abnormalities. To avoid embryos movement, tailbud and larva have been sedated, using menthol crystals in the see water, before the montage on the slides.

2.5 Imaging analysis

The embryos, in PBS buffer or MFSW, have been placed on a microscope slide. A cover slide with some plasticine at the corners have been positioned on the top of the drop of embryos, and pressed until the volume resulted uniformly distributed. DIC and fluorescent images have been taken with a Zeiss Axio Imager M1 microscope equipped with an Axiocam digital camera. For confocal images, embryos have been analyzed with a Zeiss confocal laser scanning microscope LSM 510.

2.6 PCR amplification from DNA

The PCR reactions have been performed in a total volume of 50 μl, using about 100 ng of genomic DNA as template, 0.2 mM of dNTP mix (dATP, dTTP, dCTP, dGTP), 1x PCR buffer (Roche), 0.05 U/μl of Taq DNA Polymerase (Roche) and 2 pmol/μl of each forward and reverse oligonucleotides. PCR amplifications from plasmid DNA have been performed to amplify fragments of interest has been set as follows:

1. First step (one cycle): 5 minutes at 95°C for genomic DNA and 1 minute for plasmid DNA
2. Second step (repeated for 35 cycles):
   DNA denaturation: 1 minute at 95°C.

Oligonucleotides annealing: 2 minutes 50°C for genomic DNA and 1 minute 52-54°C for plasmid DNA (the temperature used in this step has been set at least 8-10°C below the melting temperature of the oligonucleotides).

Polymerization: 72°C for a suitable time (the polymerization time has been calculated considering the desired amplified fragment length and the Taq DNA Polymerase processivity, that is around 1 Kb/minute).

Final elongation step: 10 minutes at 72°C for genomic DNA and 5 minutes at 72°C for plasmid DNA.

The amplified fragments have been separated from the genomic DNA and from the dNTPs and oligonucleotide excess, by gel electrophoresis, using, as fragment length marker, 1x GeneRuler™ 1Kb DNA Ladder (Fermentas). The fragments have been isolated and purified by gel extraction (QIAquick Gel Extraction Kit, QIAGEN; GenElute™ Gel Extraction Kit, Sigma). The concentration has been evaluated by gel electrophoresis, using, as marker, 1x Lambda DNA/HindIII, 2 (Fermentas).

2.7 DNA gel electrophoresis

Preparative and analytic DNA gel electrophoreses experiments have been performed on 1% or 1,5% of agarose gel in 1x TAE buffer (Stock solution 50x TAE buffer: 252 g of Tris base; 57.1 ml glacial acetic acid; 100 ml 0.5M EDTA; H_2O to 1 liter), considering the length of the DNA to be run and adding 0.5 µg/ml Ethidium Bromide (EtBr). Suitable DNA markers have been used to check DNA concentration (1x Lambda DNA/HindIII 2, Fermentas) or length (1x GeneRuler™ 1Kb DNA Ladder; 1x GeneRuler™ 100 bp DNA Ladder, Fermentas). The DNA gel electrophoreses have been performed with constant voltage ranging from 60 to 100 V.

2.8 DNA gel extraction

Digested and PCR amplified fragments have been extracted from gel cutting them with a sterile sharpen blade, using the GenElute™ Gel Extraction Kit (Sigma-Aldrich), following the manufacturer's instructions. After the extraction, the concentration has been estimated by gel electrophoresis.

2.9 DNA digestions with restriction endonucleases

Analytic and preparative plasmid DNA digestions have been performed with the appropriate restriction endonucleases in total volumes of at least 20 times more than the enzyme volume used. The digestion reaction has been prepared as follows: the solution contained the desired amount of DNA, suitable restriction enzyme buffer (1/10) (Roche; New England Biolabs; Amersham), restriction enzyme/s (5 units enzyme per 1 µg of DNA).

Reaction specific temperatures have been used as suggested by manufacturer's instructions.

2.10 Digested plasmids' purification

To eliminate protein contaminations, the DNA has been purified with 1 volume of phenol:chloroform:isoamylic alcohol (25:24:1), vortexed vigorously and centrifuged at 13000 rpm for 5 minutes at 4°C. The soluble phase has been recovered and 1 volume of chloroform:isoamylic alcohol (24:1) has been added; the sample has been vortexed vigorously and centrifuged at 13000 rpm for 5 minutes, at 4°C. The aqueous phase has been recovered and the DNA has been precipitated adding 3 volumes of ethanol 100% and 1/10 volumes of Sodium Acetate 3M pH 5.2. The sample has been mixed and stored over night at -20°C or 1 hour at -80°C. Then, it has been centrifuged at 13000 rpm for 15 minutes, at 4°C. The precipitated DNA has been washed with ethanol 70% (sterile or DEPC-treated), centrifuging at 13000 rpm for 15 minutes at 4°C. The ethanol has been removed and the sample has been air-dried at Room Temperature (RT). At the end, the DNA has been diluted in a suitable volume of H_2O (sterile or DEPC-treated), and its concentration has been evaluated by gel electrophoresis, using 1x Lambda DNA/Hind III, 2 (Fermentas), and using a spectrophotometer (Nanodrop 1000, Thermo SCIENTIFIC), reading the values at the wavelengths of 230, 260 and 280 nm and calculating the ratio between 260/230 nm and 260/280 nm to ascertain respectively the absence of chemical (phenol, ethanol) and protein contamination.

2.11 DNA dephosphorylation

To prevent reclosure of the plasmids after enzyme digestion, DNA dephosphorylation has been conducted as follows: a convenient amount of double stranded linearized DNA has been incubated with 1U of Calf Intestinal Alkaline Phosphatase enzyme (CIP) (Roche) per 1 pmol 5' ends of double stranded linearized DNA, in 1X CIP dephosphorylation buffer (Roche), at 37°C for 30 minutes. After this time, a second aliquot of CIP as been added, and the reaction has been carried on for further 30 minutes, at 37°C. Subsequently, the dephosphorylated DNA has been purified by phenol: chloroform/chloroform extraction (*see* par. 2.10).

2.12 DNA Ligation

Each ligation reaction has been carried out in a final volume of 10 µl containing: 50-100 ng of linearized DNA vector; 3-5 fold vector moles of DNA insert; 1x T4 DNA Ligase buffer (50 mM Tris-HCl pH 7.5, 10 mM MgCl2, 10 mM dithiothreitol, 1 mM ATP, pH 7.5) and 1 unit/µl of T4 DNA

Ligase (New England Biolabs). Generally, the optimal proportion of 1:3, vector:insert, has been used in the reaction. The reaction mix has been incubated at 16 °C overnight or 1,5 hour at RT, and used to transform competent bacteria.

2.13 Bacterial cells electroporation

By this approach it is possible to transform bacterial cells with plasmids containing DNA of interest. Briefly, the circular plasmid and competent bacterial cells, prepared by the Molecular Biology Service of the Stazione Zoologica "A. Dohrn" in Naples, were placed in an electrocuvette. The electrocuvette was subjected to an electric pulse at constant 1.7 V using a Bio-Rad Gene Pulser™ electroporation apparatus. Then the cells were placed in 1 ml of SOC (tryptone 20 g/l, yeast extract 5 g/l, 5 M $NaCl_2$ ml/l, 1 M KCl 2.5 ml/l, 1 M $MgSO_4$ 10 ml/l, 1 M $MgCl_2$ 10 ml/l, 1 M glucose 20 ml/l) shaking at 270 rpm at 37 °C for 1 hour, plated on LB solid medium (NaCl 10 g/l, bactotryptone 10 g/l, yeast extract 5 g/l, agar 15 g/l) in the presence of the specific antibiotic (50 µg/ml) to which the plasmid is resistant and grown at the same temperature overnight.

2.14 PCR Screening

It is possible to carry out a PCR reaction using as template the DNA of a single bacterial colony and in the same time grow the colony. Half of each colony was placed in a PCR mixture described below, and half grown in 3 ml of LB liquid medium in the presence of the suitable antibiotic (50 µg/ml), 8-12 hours shaking at 270 rpm, at 37 ºC.

The PCR reaction have been set in a total volume of 20 µl, with the following composition: 1x PCR buffer (Roche); 0.2 mM dNTP mix (dATP, dTTP, dCTP, dGTP); 1 pmol/µl of each forward and reverse suitable oligonucleotides; and 0.025 U/µl Taq DNA polymerase (Roche; Biogem).
By gel electrophoresis analysis, the samples presenting a band of expected size have been identified and plasmid DNA have been purified from the corresponding bacterial colonies.

2.15 DNA Mini- and Maxi-preparation

A single bacterial colony containing the plasmid of interest was grown in a suitable volume of LB in the presence of the appropriate antibiotic shaking at 37°C overnight. The Sigma-Aldrich Plasmid Purification kit, based on alkaline lyses method, was used to isolate the plasmid DNA from the cells according to the manufacture's instruction.

2.16 Purification on CsCl gradient of DNA constructs

In order to obtain only the circular form of plasmid DNA, plasmids were resuspended in 1xTE and for each ml of DNA 1.2 g of Cesium Chloride (CsCl) was added. After the addition of 100 µl of EtBr (10 mg/ml) for each DNA/CsCl ml, the samples were transferred in Beckman Quick Seal ultracentrifuge tubes and centrifuged in a VTi-65 rotor 16 hours at 55,000 rpm and at 25°C in a Beckman L8-70M ultracentrifuge.

By this technique, circular plasmid DNA was separated by contaminant bacterial DNA and RNA. The separation occurred in virtue of different density acquired by plasmid DNA in the presence of EtBr compared to chromosomal DNA. Two distinct bands were formed on the gradient, the upper one contained nicked bacterial plasmid and chromosomal DNA, the lower one corresponded to supercoiled DNA plasmid. RNA molecules, having higher density, were located in the bottom of the gradient. The band containing the DNA of interest was collected using a 21-gauge needle.

The EtBr was removed by adding 1 volume of isoamyl alcohol and by 10 minutes centrifugation. The extraction was repeated several times until EtBr was completely eliminated from the alcoholic phase. Finally, in order to remove CsCl, plasmid DNA was precipitated 15 minutes on ice after the addition of 3 volumes of H_2O and 2 volumes of 100% ethanol and then centrifuged at 10.000 rpm for 20 minutes at 4°C. The pellet was washed in 70% ethanol and resuspended in sterile H_2O.

2.17 Oligonucleotides' synthesis

All used synthetic oligonucleotides were prepared with a Beckman SM-DNA Synthesizer at the Molecular Biology Service of the Zoological Station "Anton Dohrn" in Naples.

2.18 Sequencing

The DNA sequences have been obtained using the Automated Capillary Electrophoresis Sequencer 3730 DNA Analyzer (Applied Biosystems, Foster City, CA) by the Molecular Biology Service of the Zoological Station "Anton Dohrn" in Naples.

2.19 Ribonucleic probes' preparation

- ***RNA labelling***

The plasmid has been conveniently digested downstream of the template to be transcribed and purified, as described in pars. 2.9 and 2.10. 1 µg of purified, linearized DNA has been used as template for the ribonucleic probe synthesis. At this template, transcription buffer (0.1%) (Roche), Digoxigenin or Fluorescein labelling mix, containing 1 mM of ATP, CTP

and GTP, 0.65 mM UTP and 0.35 mM DIG-11-UTP or 0.35 mM fluorescein-12-UTP (Roche), Sp6 or T7 RNA polymerase (2U/µl) (Roche), and Protector RNase inhibitor (1U/ µl) have been added. The reaction has been performed in a total volume of 20 µl (in DEPC-treated H_2O). The mix has been briefly centrifuged and incubated for 2 hours at 37°C. Then, DNaseI RNase free (0.9U/µl) has been added and the sample has been incubated for 20 minutes at 37°C. Finally, the reaction has been stopped adding EDTA pH 8.0 (16 mM). The synthesized ribonucleic probes have been purified using the mini Quick Spin RNA Columns G-50 Sephadex (Roche), following manufacturer instructions. The ribonucleic probe quality has been checked by gel electrophoresis and the concentration quantification has been evaluated by Dot Blot analysis (see the next paragraph). The recovered samples have been immediately stored at -80°C till the use.

❖ ***Ribonucleic probes' quantification by Dot Blot analysis***

The concentration evaluation of the DIG-11-UTP or fluorescein-12-UTP incorporated in the ribonucleic probes has been estimated making serial dilutions of the sample ribonucleic probes and of a control RNA of reference (Roche), in a buffer containing DEPC-treated H2O, 20x SSC, formaldehyde (5:3:2). 1 µl of each dilution has been blotted on a Hybond-N membrane (Amersham) and air-dried at RT. The RNA has been UV-crosslinked on the membrane with a Stratalinker for 30 seconds. The membrane has been washed in blocking solution (5% BSA in 0.1 M maleic acid pH 7.5), for 30 minutes, shaking at RT. Subsequently, the membrane has been incubated with anti-DIG or anti-Fluo phosphate alkaline antibody (0.15 U/ml) (Roche) in blocking solution for 1 hour, shaking at RT. To remove unbound antibodies, the membrane has been washed twice in a solution containing 0.1 M maleic acid pH 7.5 and 0.15 M NaCl for 15 minutes, at RT. The membrane has been equilibrated in the detection solution (100 mM NaCl; 100 mM Tris pH 9.5; 50 mM $MgCl_2$, in H_2O) for 5 minutes, at RT, and then incubated in the dark in the same solution in which Nitro-Blue Tetrazolium Chloride (NBT) (35 µg/ml) and 5-Bromo-4-Chloro-3'-Indolylphosphate p-Toluidine (BCIP) (175 µg/ml) have been added. The reaction has been monitored every 4-5 minutes and blocked at the suitable moment, putting the membrane under running water. The membrane has been dried on 3MM paper and the concentration of the DIG-11-UTP or fluorescein-12-UTP incorporated in the ribonucleic probes has been calculated from the comparison with the control RNA spots.

2.20 Whole mount immunohistochemistry (IHC)

Control and transgenic *C. intestinalis* larvae have been collected into baskets, provided with a membrane, and fixed in 4%

paraformaldehyde/MOPS (3-(N-morpholino)propanesulfonic acid) buffer overnight, at 4°C. The larvae have been subsequently dehydrated into 70% ethanol and stored at -20 °C, until the use. The primary antibody, rabbit anti-*Ciona* βγ crystallin, used in my experiments, has been kindly provided by Dr Sebastian M. Shimeld (Department of Zoology, University of Oxford, UK). The whole mount immunohistochemistry procedures have been performed as previously described (Shimeld *et al.*, 2005).

2.21 Whole mount in situ hybridization

In situ hybridization was carried out on *C. intestinalis* whole mount embryos at different developmental stages fixed at RT for 90 minutes or at 4°C over night, in a mixture containing: 4% paraformaldehyde, 0.1 M MOPS pH 7.5, 0.5 M NaCl.

The samples were washed 3x15 minutes in 1 ml PBT (PBS + 0.1% tween 20) at RT and incubated 30 minutes at 37°C in 1 ml PBT containing 2 μg/ml protease K for dechorionated embryos or 4 μg/ml for non dechorionated embryos, to increase the permeability to the cells and accessibility to mRNA target. The reaction was stopped by a wash in 2 mg/ml glycine in PBT. After digestion, samples were post-fixed 1 hour at RT in 4% PFA + 0.05% tween-20 in PBS 1x and then washed 3x15 minutes in PBT. Embryos were placed 10 minutes at same temperature in the pre-hybridization solution (50% formamide, 6x SSC, 0.05% tween 20) and finally 2 hour at 55°C in hybridization solution (50% formamide, 1x Denhardt's solution, 6xSSC, 0.05% tween 20, 100 μg/ml Yeast tRNA, 0,005% Heparine). Riboprobe (or Riboprobes for the double *in situ* hybridization) was added up to a final concentration of 0.5 ng/μl and the hybridization occurred overnight at the same temperature.

The following day a series of washes were carried out by varying the temperature and salinity conditions; embryos were, in fact, washed at 55 °C in the following solutions: 2 times for 20 minutes in WB1 (50% formamide, 5x SSC, 0,1% SDS), 2 times for 20 minutes in WB1:WB2 and 2 times for 20 minutes in WB2 (50% formamide, 2x SSC, 0,1% Tween 20). Subsequently they were treated with a 1 ml Solution A (10mM Tris-Cl, pH 8.0, 0.5 M NaCl, 5 mM EDTA, 0.1% Tween 20), 2 times for 5 minutes at RT. To remove aspecific RNA, not bound to the corresponding endogenous mRNA, the embryos were treated with RNase A [20 μg/ml] for 20 minutes at 37 °C in solution A, then washed 1 time in WB3 (2x SSC, 0.1% Tween 20) for 5 minutes at RT and 2 times in WB3 at a temperature of 55°C. Following they were incubated 3 times for 5 minutes in TNT (0.1M Tris, pH 7.5, 150 mM NaCl, 0.1% Tween 20) at RT. At this point the detection of the probes has been carried out.

❖ For single *in situ* hybridizations embryos were incubated in Blocking TNB buffer (100 mM Tris pH 7, 150 mM NaCl, 1% Blocking Reagent, 0.2% Triton 100x) for 2 hours at RT. At this point embryos were incubated, overnight at 4 °C, with the antibody anti-DIG in the ratio 1:2000 in Blocking TNB Buffer. The next day, the samples were washed at room temperature in TNT with the following modalities: 1 time for 5 minutes, 4 times for 20 minutes, 1 time for 40 minutes and 3 times for 10 minutes in TMN (100mM NaCl, 50mM $MgCl_2$, 100mM Tris-Cl, pH9.5, 0.1% Tween 20).

To identify the localization of the RNA of interest, labeled with DIG and recognized by anti-DIG alkaline phosphatase conjugated, are provided the appropriate substrates that will be converted by the phosphatase in a blue precipitate. The embryos were incubated, therefore, in 1 ml of TMN containing 4.5 µl of NBT (nitroblue tetrazolium) and 3.5 µl of BCIP (5-bromo-4-chloro-3-indolyl-phosphate). The time of formation of the precipitate are conditioned by the amount of antibody bound and, therefore, indirectly on the type of probe used. For this reason, at intervals of 30 minutes a few embryos are taken and observed, after being placed on a microscope slide, with a phase contrast microscope. When positive signal is present the color reaction is stopped using 1x PBT.

❖ For double *in situ* hybridization embryos were first incubated in TNB blocking buffer for 2 hours at RT. At this point samples were incubated over night at 4°C, with anti-DIG Fab Fragments POD HRP diluted 1:400 in Blocking Buffer TNB.

❖ The next day the following series of washes were carried out: 4 times for 15 minutes and 2 times for 5 minutes in TNT at RT.

To identify the localization of the RNA of interest marked with Digoxigenin and recognized by anti-DIG conjugated to alkaline peroxidase (HRP) a substrate is converted by HRP in fluorescent product. The embryos were incubated, therefore, in the 1x Plus Amplification Diluent (Perkin-Elmer) for 15 minutes at RT and subsequently 1:400 Cy3 diluted in the same solution for 1.5 hours at RT. After the reaction, embryos were washed 9 times for 5 minutes in TNT at RT. In order to stop the first antibody reaction an incubation in 50% formammide, 2x SSC, 0,1% tween-20 for 10 minutes at 55°C was performed. The next step consisted in a series of 3 washes of 10 minutes in TNT and incubates, overnight at 4°C, with anti-Fluorescein HRP diluted 1:400 in Blocking Buffer TNB. The next day, the embryos were washed 2 times for 5 minutes in TNT at RT and then incubated in the 1x Plus Amplification Diluent (Perkin-Elmer) for 15 minutes at RT and subsequently treated with 1:400 Cy5 diluted in the same solution for 1.5 hours at RT. After this time the reaction was blocked by carrying out the following washes: 1 time for 5 minutes and 7 times for 10 minutes in TNT at RT. When desired

embryos were treated 3 times for 5 minutes and then over night with 1:10000 DAPI in TNT, in order to highlight the nuclei of the cells.

2.22 Preparation of constructs for transgenesis experiments via electroporation

For the preparation of the constructs used in the functional studies, a starting vector pMESP-Ets (generous gift from Dr. Lionel Christiaen, Levine laboratory, University of California at Berkeley, USA) in two versions was used: one containing a VP16 activator domain, and another one with a WRPW repressor domain (Davidson *et al.*, 2005) (Fig. 2.1).

In order to prepare the constructs it was necessary first of all to delete the sequence of the Ets gene (~ 900 bp) which is located downstream the *Mesp* promoter and replace it with the coding sequence of the *Ci-Onecut* gene. To do this, the vectors were subjected to a double enzymatic digestion with restriction enzymes *Nhe*I/*Spe*I [5U/µg of plasmid DNA] for two hours at 37 °C. The products of the digestions (pMesp-VP16/pMesp-WRPW) were analyzed and extracted from the 1% agarose gel with the gel extraction kit (Sigma-Aldrich).

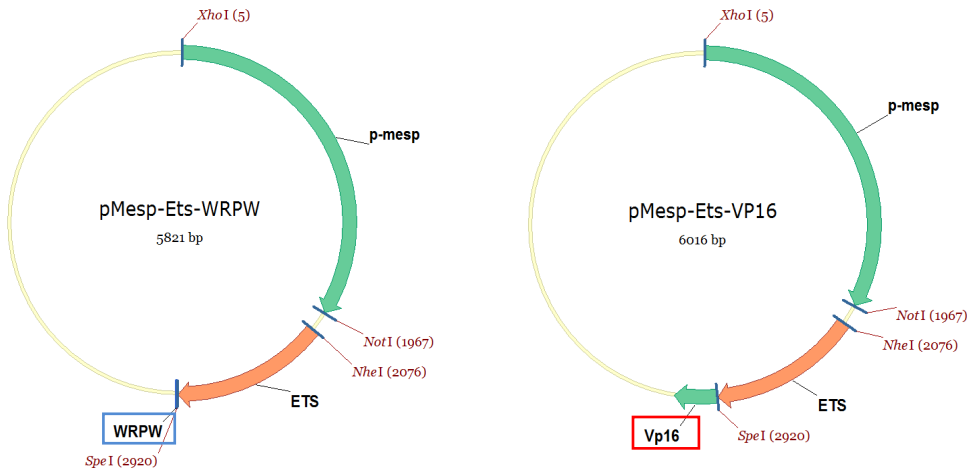


Figure 2.1: ***Map of plasmids pMesp-Ets-WRPW and pMesp-Ets-VP16.*** In the blue rectangle is shown the repressor domain and in red one the activator domain.

*2.23 Preparation of constructs Etr-OC-***VP16** *and Etr-OC-WRPW*

The *Ci-Onecut* coding sequence was obtained by PCR amplification using as template a plasmid containing the *Ci-Onecut* cDNA, already available in the laboratory (D'Aniello *et al.*, 2011). As primers for the PCR reaction two oligonucleotides of about 30 bases in length, complementary to

regions of interest and containing desirable restriction sites, *Spe*I at the 5'end and *Nhe*I at the 3'end, were used.

Onecut *Nhe*I: 5'GACTGCTAGCGTCCCGCACTCAACAGATCGAAG 3'
Onecut *Spe*I: 5'GGACTAGTACGGGGATTCACCAACCCCGTG 3'

PCR reaction was performed as previously described. The amplified fragment was analyzed and extracted from 1% agarose gel as already described. The fragment was digested with *Spe*I and *Nhe*I restriction enzymes, analyzed on 1% agarose gel and then ligated, using a 1:5 ratio, in the vectors pMesp-VP16 and pMesp-WRPW, previously digested with *Spe*I and *Nhe*I.

The reaction mixtures were incubated at 16 °C overnight. The day after the T4 DNA ligase was inactivated at 65 ° C for 10 minutes and microdialized using sterile water for an hour, to remove the salts that could create problems in the subsequent phases.

The obtained circular plasmids (pMesp-OC-VP16 and pMesp-OC-WRPW) were transformed by electroporation in TOP10 electrocompetent bacterial cells, prepared by the Molecular Biology Service of the Zoological Station "Anton Dohrn" of Naples (Dower, 1990).

The bacterial cells were plated on solid medium containing ampicillin and incubated at 37 °C for 18 hours. In order to identify the positive clones a PCR colony using an internal oligo of the plasmid and an oligo specific for the *Ci-Onecut* coding sequence was carried out:

Mespfw1: 5'CTTAAAGGCGATAATGACTTTGCCCG 3'
Onecut7: 5'GTGGAATTGATACAACTCTG 3'

The positive clones resulting from the colony were incubated at 37°C overnight. The next day minipreps were performed and sequences were analyzed by sequencing.

In order to eliminate the sequence of the *MESP* promoter, of about 1900 bp, and replace it with the *Ci-Etr* promoter, the positive clones were subjected to a double digestion *Xho*I/*Not*I. The sequence of the *Ci-Etr* promoter (Etr) was obtained by PCR using as template the plasmid BluescriptII-KS-1229 containing the *Ci-Etr* enhancer, already available in the laboratory, and as primers a pair of specific oligonucleotides with the recognition sites for *Xho*I and *Not*I restriction enzymes. The Etr insert, was double digested, then subjected to electrophoresis on 1% agarose gel, eluted from the gel and used in the reaction of ligase 1:4 in the vectors pMESP-OC-WRPW and pMESP-OC-VP16 previously *Xho*I and *Not*I digested (to eliminate the *MESP* promoter), dephosphorylated and purified. After electroporation of the ligase

reaction, under the conditions previously mentioned, the bacterial cultures were plated and the day after a PCR colony, using a vector internal oligo reverse (Onecut7) and a promoter specific forward oligo (Etr*Not*fw), was performed in order to identify the positive clones for the final constructs Etr-OC-VP16 and Etr-OC-WRPW:

Onecut7: 5' GTGGAATTGATACAACTCTG 3'
Etr*Not*fw: 5' ACATGCGGCCGCGCTCTGCTTCTGGAAGGC 3'

Finally, to increase the efficiency of the proteins expression, the human β-globin minimal promoter was cloned upstream the Ci-Onecut coding sequence. Etr-OC-VP16 and Etr-OC-WRPW constructs were *Not*I linearized and a ligase reaction with a human β-globin was performed. This last one was amplified from a plasmid available in the laboratory using oligonucleotides containing flanking *Not*I restriction enzyme site, digested and purified (as described before).

β-globin*Not*Fw: cgta**gcggccgc**tgcagcccgggctggtgata
(NotI recognition site)
β-globin*Not*Rev: acat**gcggccgc**gctctgcttctgtggaaggc
(NotI recognition site)

ETR-GFP construct, used as control vector in the electroporation experiments, was already available in the lavoratory.

2.24 Electroporation constructs' preparation pOnecut-LacZ

pBlueScript 1230 (gift of R. Krumlauf, Stowers Institute, Kansas City, USA), which contains the LacZ reporter gene and SV40 polyadenylation sequence with the human β-globin basal promoter, was used for all the constructs containing the *Ci-Onecut* promoter fragments. All the fragments were obtained by PCR using specific primers designed according to the sequence of *C. intestinalis* genome (http://genome.jgi-psf.org/ciona4/ciona4.home.html), containing flanking *Kpn*I and *Xho*I restriction sites at the 5′ and 3′ end, respectively (Table 1) and inserted into the pBlueScript 1230 in the 5′–3′ orientation upstream of the LacZ reporter gene, previously digested with the same enzymes.
The fragment K was amplified from genomic DNA and all the other fragments (from D to Fh) were amplified from the K-LacZ plasmid (Figg.3.8, 3.10).

Table 1: ***Oligonucleotides used to amplify Ci-Onecut regulatory region fragments***
Restriction sites are indicated in small

Fragment Name	Oligo name	Sequence 5'=>3'
K	*Kpn*1	gaaggtacCGTGTTCTTGTTGACACCGAATATGC
K	*Xho*1	gatactcgagCGTTATTCTAAACCAGCATTTAAGG
D	*Kpn*1	gaaggtacCGTGTTCTTGTTGACACCGAATATGC
D	*Xho*4	tactcgagCGACTACACCAGATGCTTAGTGG
V	*Kpn*2	gaaggtacCGAGATGTAGCAGGGCAATAAAGG
V	*Xho*1	gatactcgagCGTTATTCTAAACCAGCATTTAAGG
E	*Kpn*5	taggtaccTTGGTGAGGACCAATATGTGC
E	*Xho*5	tactcgagGTGGCATACGGACACGACGCTC
A	*Kpn*2	gaaggtacCGAGATGTAGCAGGGCAATAAAGG
A	*Xho*2	tactcgagATATACGAGTGGTTCGCGTCG
B	*Kpn*3	taggtaccGCGAACCACTCGTATATAAGC
B	*Xho*3	tactcgagCGAACTCCGAAGAATTCAACG
C	*Kpn*4	taggtaccCGTTGAATTCTTCGGAGTTCG
C	*Xho*1	gatactcgagCGTTATTCTAAACCAGCATTTAAGG
AB	*Kpn*2	gaaggtacCGAGATGTAGCAGGGCAATAAAGG
AB	*Xho*3	tactcgagCGAACTCCGAAGAATTCAACG
F	*Kpn*6	taggtaccTCGCCTTGTTTTCGACAGCGGTCC
F	*Xho*6	tactcgagCTTGCGCTATGCGAACTTGAATC
G	*Kpn*7	taggtaccCTAAGATTCAAGTTCGCATAGCG
G	*Xho*3	tactcgagCGAACTCCGAAGAATTCAACG
H	*Kpn*4	taggtaccCGTTGAATTCTTCGGAGTTCG
H	*Xho*8	tactcgagAATGATGCTTCAGATACTCTG
I	*Kpn*8	taggtaccGTGTCCGTATGCCACGCCATATG
I	*Xho*7	tactcgagAGGACCGCTGTCGAAAACAAGGCG
L	*Kpn*9	taggtaccAGAGTATCTGAAGCATCATTG
L	*Xho*1	gatactcgagCGTTATTCTAAACCAGCATTTAAGG
Fg	*Kpn*6	taggtaccTCGCCTTGTTTTCGACAGCGGTCC
Fg	*Xho*12	tactcgagCTGACTAAACTGTTTTATAATTG
Fa	*Kpn*6	taggtaccTCGCCTTGTTTTCGACAGCGGTCC
Fa	*Xho*10	tactcgagATTGTAGATTGTTTGAATTTAAAC
Fb	*Kpn*10	taggtaccTTAAGACAAGTTCAAGCCTTGG
Fb	*Xho*6	tactcgagCTTGCGCTATGCGAACTTGAATC
Fc	*Kpn*6	taggtaccTCGCCTTGTTTTCGACAGCGGTCC
Fc	*Xho*11	tactcgagTGGGCCATCTGCTGCGTTTTTCTG
Fd	*Kpn*11	taggtaccTTGCTTACTCCCATACGAGTACGG
Fd	*Xho*10	tactcgagATTGTAGATTGTTTGAATTTAAAC
Fh	*Kpn*11	taggtaccTTGCTTACTCCCATACGAGTACGG
Fh	*Xho*12	tactcgagCTGACTAAACTGTTTTATAATTG

Each PCR reaction was performed as previously described (*Paragraph 2.6*). The amplified fragments were analyzed and extracted from 1% agarose gel as already described (*Par. 2.7, 2.8*). The fragments and the vector (pBlueScript 1230) were digested with *Kpn*I and *Xho*I, analyzed on 1% agarose gel and then ligated using a 1:4 ratio vector/insert. The recombinant

plasmids were transformed by electroporation, as already described (*Par. 2.14*), in Top10 bacterial competent cells. The bacterial cells were plated and grown at 37 °C overnight.

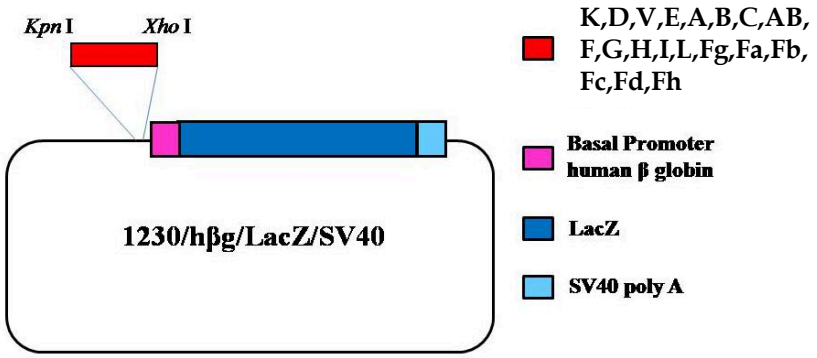


Figure 2.2: ***Cloning strategy in the 1230/hβg/LacZ/SV40 vector.*** The pBlueScript vector, containing the human β-globin basal promoter, the coding sequence of *Lac*Z (reporter gene) and the SV40 polyadenylation signal, was used to clone *Ci-Onecut* promoter fragments (in red) in the *Kpn*I and *Xho*I sites.

2.25 Preparation of the "FG" mutated constructs

The FG(NgnMut), FG(OCaMut), FG(Ngn-OCaMut) and FG(Ngn-OCbMut) constructs have been prepared using the QuikChange® Site-Directed Mutagenesis Kit "Stratagene" from the Fg-*LacZ* construct (called FG, of 262 bp) . The putative Ci-Neurogenin and Ci-Onecut binding sites were replaced by a sequence that reduced the binding affinity by using the mutagenic oligonucleotides of about 44 bp, listed in Table 2, ordered as synthetic single strand oligonucleotides (Primm Company). According to manufacturer's instruction, these mutated oligonucleotides have been used for a PCR reaction using as template the FG 262 bp construct. The PCR reaction has been carried out using the PFU DNA polymerase and these cycling parameters: after a single denaturation step at 95 °C for 30 seconds, 18 amplification cycles were performed as follows: denaturation at 95 °C for 30 seconds, annealing at 55°C for 1 minute, extension at 68 °C for 7 minutes. The presence of amplified product has been checked loading 10 µl of the reaction on 1% agarose gel. At this point to eliminate the FG wild type plasmid, used as template, the mixture has been digested at 37 °C for 1.5 hr with 1 µl of the *Dpn*I restriction enzyme (10 U/µl). This enzyme digests only the methylated supercoiled dsDNA used as template but not the newly synthesized mutated and unmethylated DNA. After the digestion, 2 µl of the reaction has been transformed and grown as previously described (*Par. 2.13*).

Ten clones have been selected and after growth, the isolated plasmids DNAs have been sequenced to check the presence of mutations. Mutated plasmid DNA was purified on CsCl gradient as previously described.

Table 2: ***Oligonucleotides used for mutagenesis***
Mutated nucleotides are indicated in bold-red

Name	Sequence
Ngn MUT up	GCCGAACAGAAAAACGCA**CCCC**A**CC**GCCCATTGCTTACTCCC
Ngn MUT down	GGGAGTAAGCAATGGGC**GG**T**GGGG**TGCGTTTTTCTGTTCGGC
OCa MUT up	GTTTAAATTC**C**AAC**CCC**C**C**ACAATTAAGACAAGTTCAAGCCTTGG
OCa MUT down	CCAAGGCTTGAACTTGTCTTAATTGT**G**G**GGG**GTT**G**GAATTTAAAC
OCb MUT up	GCCTTGGAAAGAAAAATA**CCC**C**CCC**TATAAAACAGTTTAGTCAGTT
OCb MUT down	AACTGACTAAACTGTTTTATA**GGG**G**GG**TATTTTTCTTTCCAAGGC

2.26 In silico analysis of putative trans-acting factors

The Fg sequence, containing *Ci-Onecut* enhancer, were submitted to the GENOMATIX Database using the MatInspector software (http://www.genomatix.de/cgi-bin/eldorado.main.pl) (Cartharius *et al*., 2005). This is a database of transcription factor from many organisms, their genomic binding sites and DNA-binding profiles.

2.27 Isolation of the Ci-Onecut and the Ci-Neurogenin cDNAs

The full-length cDNA of *Ci*-Onecut has been amplified by PCR using as template cDNA from mRNA poly $(A)^+$ isolated at tailbud stage of *C. intestinalis*, as previously described (D'Aniello *et al.,* 2011). The full-length cDNA of *Ci-Neurogenin* was found in Aniseed database (http://www.aniseed.cnrs.fr/): cien83880 (Gateway Gene Collection ID:VES83_M13) and amplified by PCR. The oligonucleotides were designed overlapping the ATG start codon (Ngn up) and overlapping the stop codon (Ngn down) and containing flanking BamHI and MluI restriction sites at the 5′ and 3′ end, respectively (Table 3). The full-length *Ci-Neurogenin* and *Ci-Onecut* PCR fragments were cloned in the pCR®II TOPO Dual Promoter vector (Invitrogen) and sequenced.

Table 3: ***Oligonucleotides used full-length Ci-Neurogenin cDNA*** Restriction sites are indicated in bold and start and stop codon are bold and underlined	
Name	Sequence
Ngn Up *BamH*I	ACGC**GGATCC**TCTTGCTACAAA**ATG**TTGGATT
Ngn DOWN *Mlu*I	GTCG**ACGCG**TGATTTTCCTT**ATC**GGAGTAAATG

2.28 Preparation of co-electroporation constructs

The pBS/pBra700/GFP/SV40 construct (a gift of Dr. A. Spagnuolo, Stazione Zoologica A. Dorhn, Naples) (Corbo *et al*., 1997) was used as starting vector to prepare the construct used for the coelectroporation experiments. It consists of a BlueScript vector containing 700 bp of *Ciona* promoter region of *Brachyury* gene (*Ci-Bra*), the *GFP* as reporter gene and SV40 polyadenylation signal.

❖ ***Preparation of the pBra-Onecut construct***

The pBS/pBra700/GFP/SV40 construct has been digested with *Eco*RI to cut off the *GFP* reporter gene. In parallel, the TOPO-TA vector (Invitrogen), containing the full-length cDNA of *Ci-Onecut*, has been digested with *Eco*RI to excise the *Ci-Onecut* insert. At this point, the two digested vectors have been loaded on the 1% agarose gel and the linearized pBS/pBra700/SV40 vector and the *Ci-Onecut* full length insert were extracted from gel. Then the vector and the insert were ligated in a 1:3 ratio and the recombinant plasmid was transformed by electroporation as previously described (*Par.2.13*). The bacterial cells were plated and grown overnight at 37 °C. Finally, PCR screening was performed to select the clones containing the insert using as primers a forward oligonucleotide, complementary to the vector, and a reverse oligonucleotide, complementary to the *Ci-Onecut* sequence, to discriminate the right orientation of the insert with respect to the promoter (Fig. 2.3). The obtained construct was called pBra-Onecut.

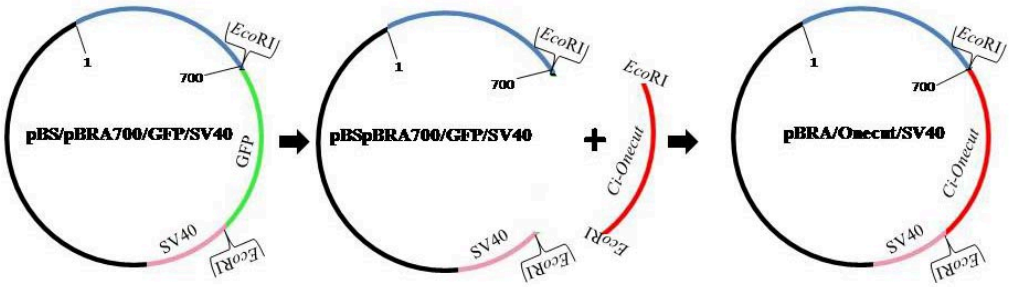


Figure 2.3: ***Cloning strategy of the pBra-Onecut construct.*** The *Ci-Onecut* insert, indicated in red, has been cloned in the *Eco*RI site. In blu the *Ci-Brachyury* promoter region and in pink the SV40 polyadenylation signal.

❖ ***Preparation of the pBra-Neurogenin construct***

The pBS/pBra700/GFP/SV40 construct has been double digested with *BamH*I and *Mlu*I to cut off the *GFP* reporter gene. In parallel, the TOPO-TA vector (Invitrogen), containing the full-length cDNA of *Ci-Neurogenin*, has been digested with *BamH*I and *Mlu*I flanking the *Ci-Neurogenin* coding sequence, to excise the *Ci-Neurogenin* insert. At this point, the two digested vectors have been treated like in the previous experiment in order to obtain the pBra-Neurogenin construct (Fig. 2.4).

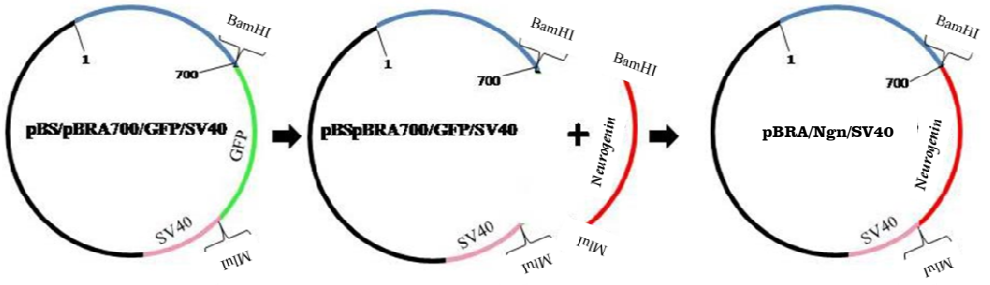


Figure 2.4: ***Cloning strategy of the pBra-Neurogenin construct.*** The *Ci-Neurogenin* insert, indicated in red, has been cloned in the *BamH*I-*Mlu*I sites. In blu the *Ci-Brachyury* promoter region and in pink the SV40 polyadenylation signal.

3. RESULTS

3.1 Studies of the functional connection between Ci-Onecut and Ci-Rx

Previous studies conducted by the research group of Dr. Margherita Branno (Department of Cellular and Developmental Biology, Stazione Zoologica A. Dohrn, Naples), with whom I did my work of thesis, led to the identification of two fragments of the *Ci-Rx* promoter, respectively 35 and 50 bp in size, able to activate the expression of the reporter gene in the same region of the endogenous gene (D'Aniello *et al*., 2011).

Bioinformatics analysis of these fragments identified the presence, on both of them, of a putative binding site for the Ci-Onecut protein.

The complete coding sequence of *Ci-Onecut* (2130 bp) was then obtained by PCR amplification of the cDNA isolated from embryos at tailbud stage. Subsequently, through double *in situ* hybridization experiments, the spatial and temporal co-localization of the *Ci-Rx* and *Ci-Onecut* transcripts had been examined. Furthermore, through an EMSA (Electrophoretic Mobility Shift Assay) experiment, the *in vitro* interaction between the Ci-Onecut transcription factor and the binding sites on the *Ci-Rx* regulatory elements has been demonstrated. From these results, it is therefore inferred that the *Ci-Rx* gene, involved in the differentiation of photoreceptor cells and in the ocellus pigmentation in the ascidian *C. intestinalis*, is regulated by the Ci-Onecut transcription factor.

The first step of my thesis was, therefore, to study the involvement of Ci-Onecut in the *Ci-Rx* cascade and to clarify its functional role.

In order to start the functional analysis it was useful to exploit the regulative region of the *Etr* gene, in order to address the transgenes obtained (Etr-OC-VP16 and Etr-OC-WRPW) throughout the nervous system, then not only in the areas where *Ci-Onecut* is expressed, but also in the precursors of pigmented and photoreceptor cells, starting from gastrula stage (Satou *et al.,* 2001) (Fig. 3.1). These transgenes were in frame with the VP16 activation domain or the WRPW repressor domain (Davidson *et al.*, 2005).

To try to clarify a possible role of Ci-Onecut in the molecular mechanisms that lead to the differentiation of photoreceptor cells and to the pigmented cells formation, I started to analyze its possible involvement in the regulation (direct or indirect) of the *Ci-Rx* expression, a gene expressed later than *Ci-Onecut* during embryogenesis (starting from the early tailbud stage) and known in the literature to be involved in the pathway of photoreceptor cells differentiation and ocellus pigmented cell formation (D'Aniello *et al.*, 2006). I therefore performed electroporation experiments of the Etr-OC-VP16 or Etr-OC-WRPW constructs to investigate the ability of the Ci-Onecut

protein to iperactivate or interfere with the activity of the endogenous *Rx* gene.

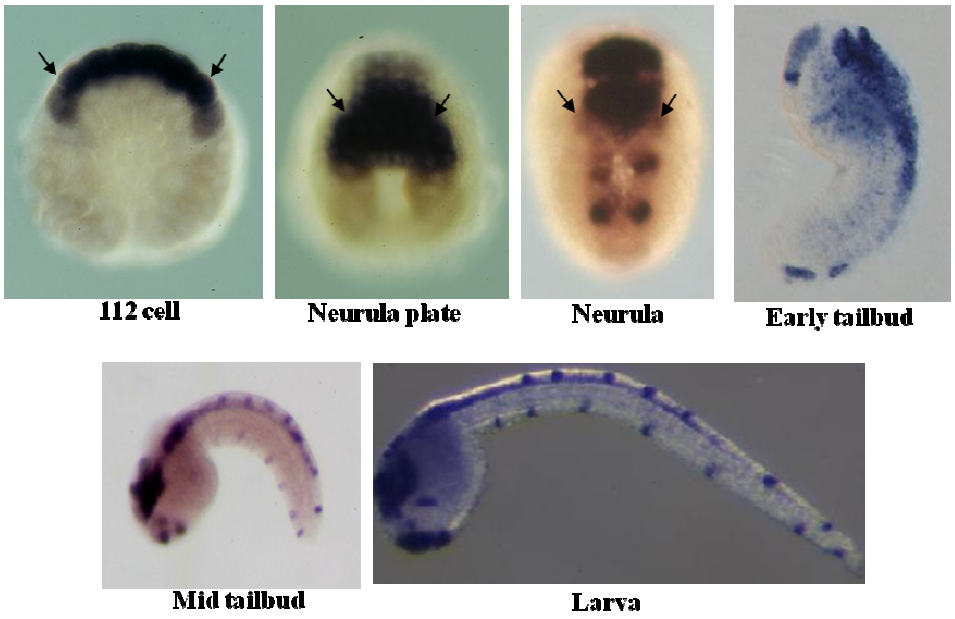


Figure 3.1: ***Wild type embryos probed for Ci-Etr expression.*** (A) Stained region: A8.15 cell pair, A8.16 cell pair, a8.17 cell pair, a8.18 cell pair, a8.19 cell pair, a8.20 cell pair, a8.25 cell pair, pigment cells lineage (black arrows), a8.26 cell pair, A8.7 cell pair, A8.8 cell pair. (B) Stained region: *a line* neural plate, containing the pigment and photoreceptor cells precursors (black arrows), *A line* neural plate. (C) Stained region: *a line* neural plate containing the pigment and photoreceptor cells precursors (black arrows), *A line* neural plate, bipolar tail epidermal neuron precursors (b8.25 line) (D): Stained region: head endoderm, nervous system. (E) Stained region: sensory vesicle, dorsal and ventral caudal epidermal neuron precursors, palps, tail nerve cord, visceral ganglion. (F) Stained region: epidermal neurons, central nervous system, palps. Images obtained from ANISEED.

3.2 pEtr targeted perturbation of Ci-Onecut protein in the nervous system

The Etr-OC-WRPW and Etr-OC-VP16 constructs were used for experiments of transgenesis *via* electroporation into *C. intestinalis* embryo, using Etr-GFP as control construct (Fig. 3.2).

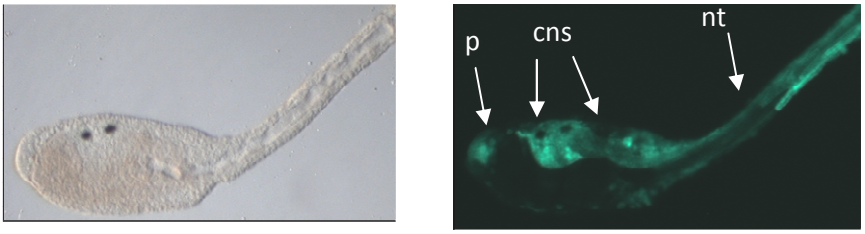


Figure 3.2: ***GFP expression driven by the Etr promoter in Ciona intestinalis larvae electroporated with the Etr-GFP construct.*** The signal is localized in the palps (p), in the central nervous system (CNS) and in the neural tube (nt).

The transgenic larvae for the constitutively repressor Ci-Onecut protein, observed at the microscope, have shown, in 49% of cases, an alteration in the pigmented cells melanization, morphology or number. In some cases, in fact, pigmented cells resutted in one big pigment cell with an altered pigmentation, resembling sometimes an otolith, sometimes an ocellus (Fig. 3.3A, B). In other cases I observed two pigment cells in the sensory vesicle but they sometimes were similar and sometimes one of them presented a very altered shape and pigmentation (Fig. 3.3C).
The remaining 51% of the transgenic larvae showed a wild type similar phenotype with two classic pigmented cells in the sensory vesicle.

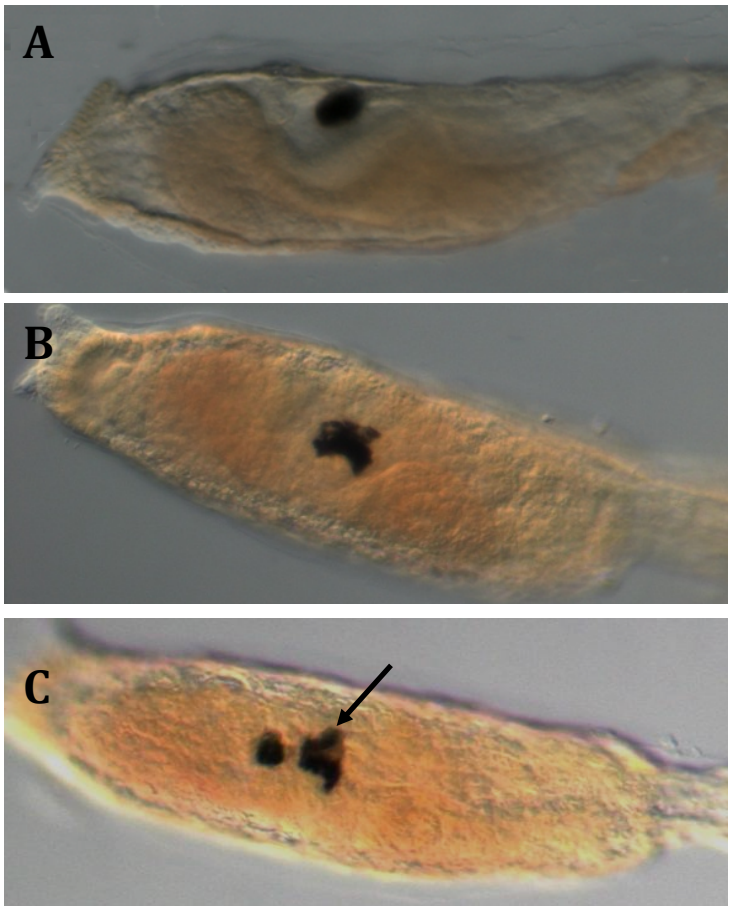


Figure 3.3: ***Phenotypes of transgenic larvae for the construct containing the Etr promoter upstream of Ci-OC-WRPW protein.*** (A, B) Phenotypes of larvae with one pigmented cell;
(C) Phenotype of a larva with two pigmented cells; one of the cells present strange shape (black arrow).
Images obtained with transmitted light microscopy.

In larvae inducted to express the Onecut constitutively active form it was possible to observe various degrees of phenotypic abnormalities regarding pigmented cells. The 81% of the larvae showed, in fact, a greater number of pigmented cells in the sensory vesicle; these cells were in some cases well distinguishable from each other (Fig. 3.4C). In other cases there were one “normal” cell (similar to the wild type one) and others fused to form a single elongated pigmented organ (Fig. 3.4A, B). The remaining 19% of the larvae showed the two pigmented cells as in wild-type ones (data not shown).

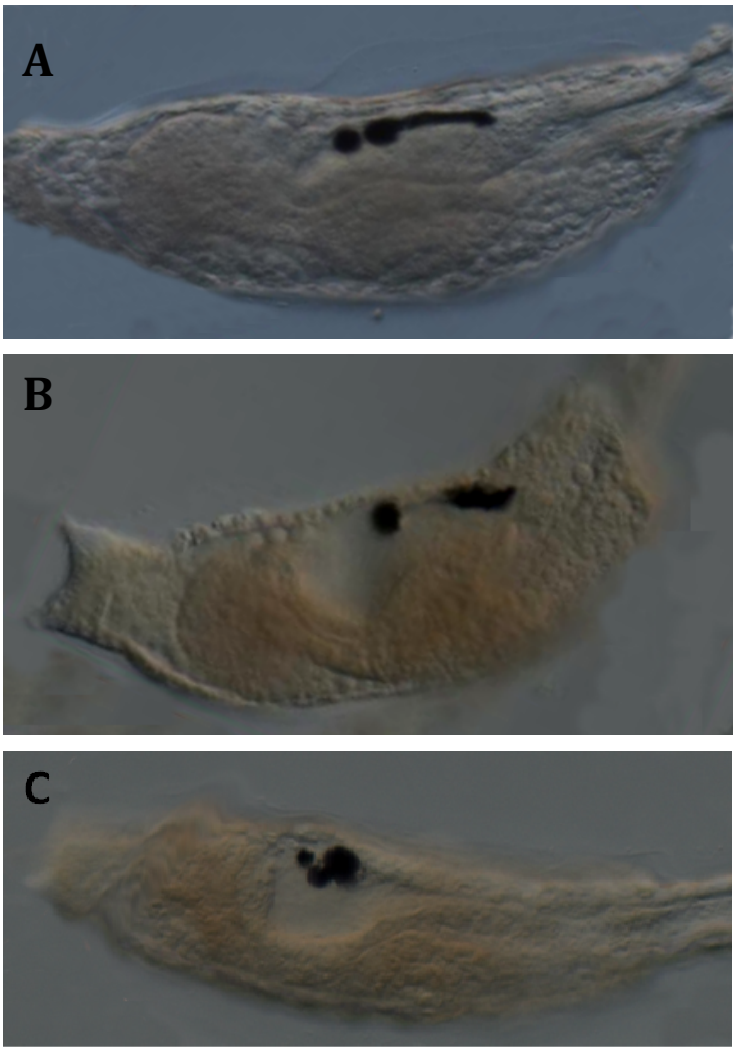


Figure 3.4: ***Phenotypes of the larvae electroporated with the construct containing the Etr promoter upstream to Ci-OC-VP16 protein.***
(A, B) Phenotypes of larvae with fused cells forming an elongated pigmented organ.
(C) Phenotype of a larva with more than two pigmented cells.
Images obtained with transmitted light microscopy.

The irregularity, the altered morphology and the additional number of pigmented cells were not useful elements in the attribution of sensory organ's nature (otolith or ocellus). In order to distinguish the otolith from the ocellus in transgenic larvae that had one, two or more pigmented cells, I have made an immunohistochemical assay with a polyclonal antibody that recognizes the βγ-crystalline, a protein expressed in the otolith and in palps, kindly received by Dr. Sebastian Shimeld (Department of Zoology, University of Oxford, UK) (Shimeld *et al.*, 2005).

The transgenic larvae, electroporated with the Etr-OC-WRPW construct, in which it was present only one pigmented cell (Fig. 3.3A, B), showed expression of βγ-crystalline in correspondence of palps and the pigmented cell, indicating that it was an otolith (Fig. 3.5B). βγ-crystalline signal in larvae showing one "normal" pigmented cell and one altered cell

(Fig. 3.3C) was localized in the "normal" cell, indicating that the altered cell was the ocellus (Fig. 3.5C). In larvae with a wild type phenotype, instead, the expression of the protein was present in palps and in one of two pigmented cells (the otolith), as expected (Fig. 3.5A).

The transgenic larvae for the Etr-OC-VP16 protein that contain more cells, one of which was "normal" and the other ones were fused to form an elongated pigmented organ (Fig. 3.4A, B), showed expression of the βγ-crystalline the palps and in the "normal" cell (Fig. 3.5E), indicating that also in this case the anomalies occurred in the ocellus.

Larvae presenting more than two pigmented cells showed expression of βγ-crystalline in palps and in one of the pigmented cells indicating the presence of a single otolith and more ocelli (Fig. 3.5F). The wild type similar larvae showed expression of βγ-crystalline in the palps and in one pigmented cell (Fig. 3.5D).

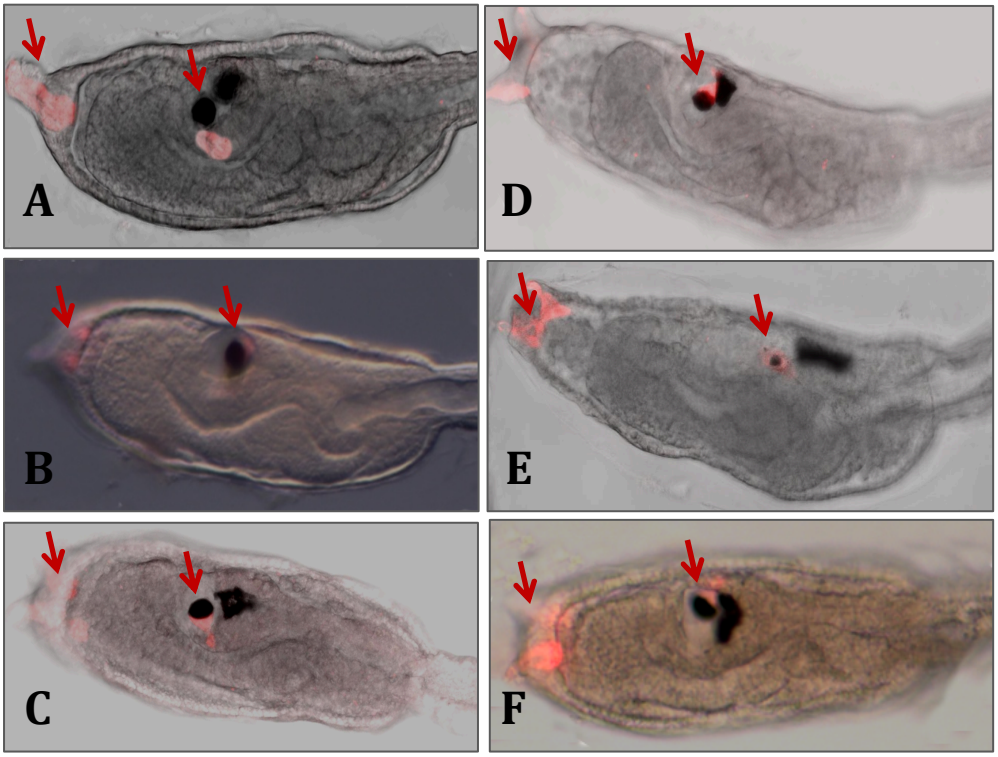


Figure 3.5: ***Immunohistochemical assay with anti-βγ-crystalline in larvae electroporated with the Etr-OC-WRPW (B, C) and Etr-OC-VP16 (E, F) constructs. As control larvae electroporated with the Etr-GFP construct (A) and wild type not electroporated larvae were used (D).*** In red is the visible expression of βγ crystalline. The images (A, B, D, E) are obtained with confocal microscopy; images (C, F) with fluorescence microscope.

As for Ci-Rx, from these results, it is evidenced an involvement of Ci-Onecut in the ocellus formation since the perturbation of the Ci-Onecut protein causes an alteration in the ocellus sensory organ morphology and melanization.

At this point in order to assess the ability of Ci-Onecut protein to influence the *Ci-Rx* endogenous expression, *in situ* hybridization experiments were performed on *Ciona* embryos electroporated with the Etr-OC-WRPW and Etr-OC-VP16 constructs, using *Ci-Rx* as an antisense RNA probe. As control *Ciona* embryos electroporated with the Etr-GFP construct were used.

WMISH experiments for *Ci-Rx* on embryos electroporated with the constitutive active construct (Etr-OC-VP16) revealed the presence of a strong *Ci-Rx* signal in the sensory vesicle (Fig. 3.6C), compared to electroporation of the control Etr-GFP construct that did not affect *Ci-Rx* expression embryos (Fig. 3.6B). Furthermore, 40% of embryos electroporated with the constitutive repressor construct (Etr-OC-WRPW) did not show any *Ci-Rx* endogenous expression, while 45% of the embryos showed a reduced signal (Fig. 3.6D). Therefore, these results demonstrate a functional connection between *Ci-Onecut* and *Ci-Rx* expression and that *Ci-Onecut* acts as a transcriptional activator, being it necessary for *Ci-Rx* expression in the anterior part of the nervous system (D'Aniello *et al.,* 2011).

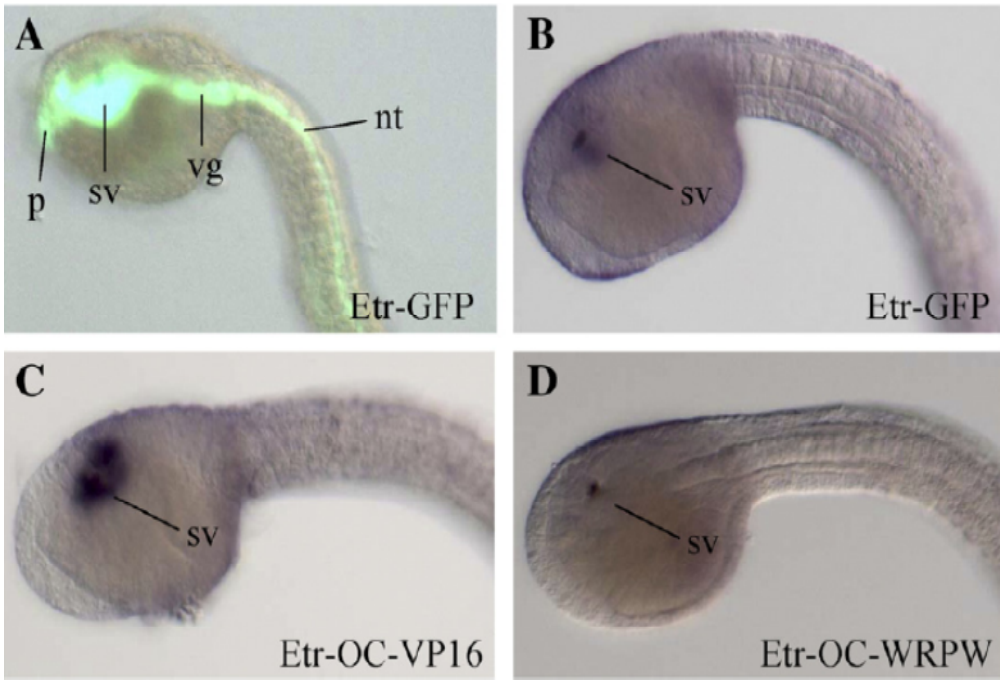


Figure 3.6: ***Mergerd bright-field/fluorescence image of GFP expression driven by the Etr promoter at tailbud stage.*** Transgene expression occurs throughout the nervous system (B-D). *Ci-Rx* expression pattern visualized by whole mount *in situ* hybridization on embryos electroporated with Etr-GFP (B), Etr-OC-VP16 (C) and Etr-OC-WRPW (D) constructs. The Etr-OC-VP16 construct promotes *Ci-Rx* expression in the sensory vesicle, while the Etr-OC-WRPW inhibits *Ci-Rx* expression. nt, neural tube; p, palps; sv, sensory vesicle; vg, visceral ganglion. Images from D'Aniello *et al.,* 2011.

3.3 Studies on the Ci-Onecut regulatory region and its upstream regulators

Ci-Onecut is able to specifically bind two *Ci-Rx* regulatory elements located from -480 to -487 bp and -527 to -534 upstream of the coding sequence and to activate its expression in ocellus and photoreceptor precursors (D'Aniello *et al.*, 2011). In order to characterize the regulatory pathway in which these two genes are implicated, I decided to analyze the *Ci-Onecut* gene and to identify the transcription factor(s) responsible for its tissue specific expression in the nervous system.

3.4 Ci-Onecut expression during embryonic development

The *Ciona Onecut* gene sequence was identified by comparing the *Onecut* nucleotide sequence of the ascidia *Halocynthia roretzi* (HrHNF-6) (AB046937) with the genome sequence of *C. intestinalis* (http://genome.jgi-psf.org/ciona4/ciona4.home.htlm). The complete cDNA clone (2130 bp) was then obtained by PCR amplification from cDNA prepared from embryos at tailbud stage.

To characterize the spatial and temporal expression profile during embryonic development, *in situ* hybridization experiments were carried out from embryos at 110-cell stage to larva stage.

As shown in Fig. 3.7 no signal was detected in embryos at 110-cell stage (A); the expression of *Ci-Onecut* starts at the neural plate stage along the antero-posterior axis of the neural plate (B). In particular, the signal is localized in four cells in the most anterior region of the neural plate, that will give rise to the sensory vesicle (red arrow), and into two pairs of bilateral cells arranged in two rows along the neural folds in the posterior half of the developing nervous system (black arrows). Among these cells, *Ci-Onecut* expression seems to be stronger in the first pair of cells representing the precursors of the visceral ganglion. At neurula and early tailbud stages (C, D) the transcript continues to be present in the same territories observed in the previous stage, although the signal in the most anterior part of the neural plate becomes wider and more intense (red arrow). At middle tailbud stage (E, F), *Ci-Onecut* is expressed in the anterior part of the developing nervous system, at level of the future sensory vesicle (red arrow) and in four cells positioned in the central part and arranged along the neural folds, at level of the visceral ganglion precursors (black arrow). Stressing the staining time of the *in situ*

hybridization experiments, allowed me to appreciate a signal in the posterior part of the tail (white arrow). This signal seems to be localized at level of the dorsal nerve cord. At larva stage (G), *Ci-Onecut* expression is restricted in the rearmost part of the sensory vesicle (red arrow), in an area that surround the ocellus and at level of the visceral ganglion (white arrow) and a faint signal is expressed also in the tail nerve cord (black arrow).

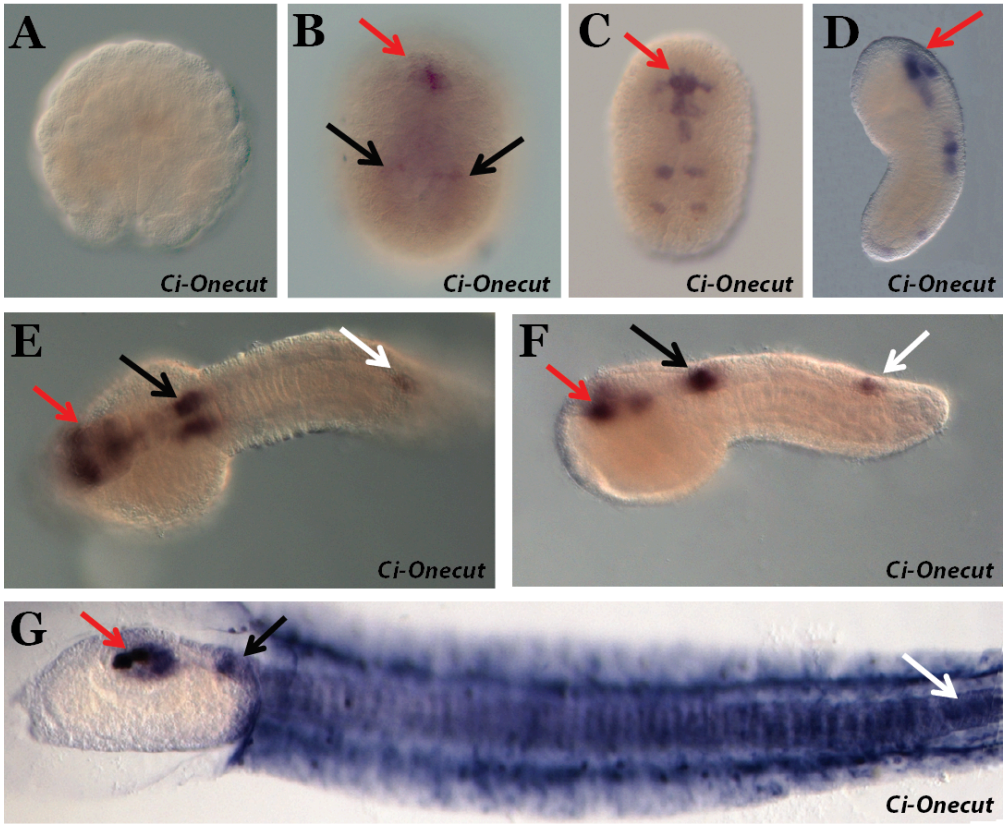


Figure 3.7: ***Whole-mount in situ hybridization of Ci-Onecut.*** (A) 110-cell stage: no signal is detected. (B) Neural plate stage: expression of *Ci-Onecut* is detected in the precursors of the sensory vesicle (red arrow) and visceral ganglion (black arrows). (C) Neurula and (D) early tailbud stage: the signal in the most anterior part of the neural plate becomes wider and more intense (red arrow). (E, F) Middle tailbud, frontal and lateral view: *Ci-Onecut* is expressed in the same territories of the previous stage (red and black arrows) and a signal in the posterior part of the tail seems to be localized at level of the dorsal nerve cord (white arrow). (G) Larva stage: *Ci-Onecut* expression is restricted in the rearmost part of the sensory vesicle (red arrow), in an area that surround the ocellus and at level of the visceral ganglion (black arrow). A faint signal is also revealed in the tail nerve cord (white arrow).

3.5 Ci-Onecut promoter analysis

Ascidians have a very compact genome, in which the majority of minimal promoters are located less than 3 kb upstream of the transcription start site. In order to identify the *trans*-acting factors responsible for *Ci-Onecut* expression, I analyzed the *cis*-regulatory region of this gene. To isolate the regulatory sequence able to activate *Ci-Onecut* tissue specific expression, a 2.6 kb genomic fragment (named K, Fig. 3.8A), extending from the position -2590 to -4 upstream of the translation start site, was isolated by PCR amplification from genomic DNA. The obtained fragment was cloned into the pBluescript-IIKS-1230 vector (construct K) and assayed by transgenesis experiments in *Ciona*. The K construct was able to activate the expression of the LacZ reporter gene in the same territories of the endogenous *Ci-Onecut* transcript, from neurula to larva stage. In particular, at neurula stage the LacZ gene is expressed in the most anterior part of the neural plate and in two pairs of bilateral cells localized in the posterior half of the developing neural plate (data not shown); at tailbud stage the signal is present in three cells in the anterior part of the nervous system, at level of the precursors of the visceral ganglion and a more posterior signal appears in the tail, at level of the epidermis which also seems to affect the neural tube (Fig. 3.9A). In the larvae the LacZ gene is expressed at level of the sensory vesicle around the ocellus, in the visceral ganglion and in the tail (Fig. 3.9B). To better narrow the regulatory elements of *Ci-Onecut* present in the 2.6 kb DNA fragment, a deletion analysis was performed and each construct was tested, repeated in triplicate in order to give a reliably statistical value to the obtained results. In each experiment the K construct was used as control of the electroporation and in the same way, all the obtained fragments were analyzed for their capability to drive the reporter gene expression in the *Ci-Onecut* endogenous territories from neurula to larva stage; here I reported the results obtained at tailbud and larva stages.

The K fragment has been divided into three smaller and overlapping fragments denominated D (Fig.3.8A), extending from the position -2590 to -1740, E (Fig. 3.8A), extending from -1814 to -1293 and V (Fig. 3.8A), from -1377 to – 4.

The D fragment was not able to activate any specific expression of the reporter gene (Fig. 3.8B). The expression of the E construct was observed ectopically in the mesenchyme, from neurula to tailbud stage (Fig. 3.8B, 3.9C); in *Ciona* frequently happens that genomic fragments are activated in a nonspecific manner in these tissues. However a partial specific signal was detected in the sensory vesicle and in the tail at larva stage (Fig. 3.8B, 3.9D). Only the V construct was able to reproduce completely the *Ci-Onecut* endogenous expression profile from neurula to larva stage (Fig. 3.8B e Fig. 3.9E, F).

The results obtained seem to indicate the presence of multiple co-operating elements capable of directing the *Ci-Onecut* gene activation; therefore to better narrow the enhancer sequences responsible for *Ci-Onecut* activation during embryonic development, the region contained in the V construct was further divided into the partially overlapping A, B, C, F, G, H, I and L (Fig. 3.8A) fragments.

At late tailbud stage the A fragment was able to direct the expression of the reporter gene in the anterior nervous system, corresponding to the sensory vesicle and ectopically into the mesenchyme (Fig. 3.8B and Fig. 3.9G). At larva stage the LacZ expression was localized around the ocellus, at level of the mesenchyme and faintly and partially in the visceral ganglion (Fig. 3.8B and Fig. 3.9H).

The LacZ expression driven by the B construct was similar to those obtained with the A construct but in this case, the B fragment was also able to drive the expression of the reporter gene in the tail, both at tailbud and larva stages while at larva stage the signal at level of the sensory vesicle becomes faint and partial (Fig. 3.8B and Fig. 3.9J).

No specific signal was detected when *Ciona* eggs were electroporated with the C construct; only a very strong LacZ expression in the ectopic mesenchyme was observed (Fig. 3.8B).

The F construct is formed by a fragment of 311 bp and is the only one able to direct the expression of the reporter gene in the same territories of the endogenous *Ci-Onecut*, from neurula to larva stage (Fig. 3.8B and Fig. 3.9K, L).

The G, H, I and L constructs did not showed any interesting and specific results, in fact in all analyzed embryonic stages a signal was present only at level of non specific mesenchyme (Fig. 3.8B).

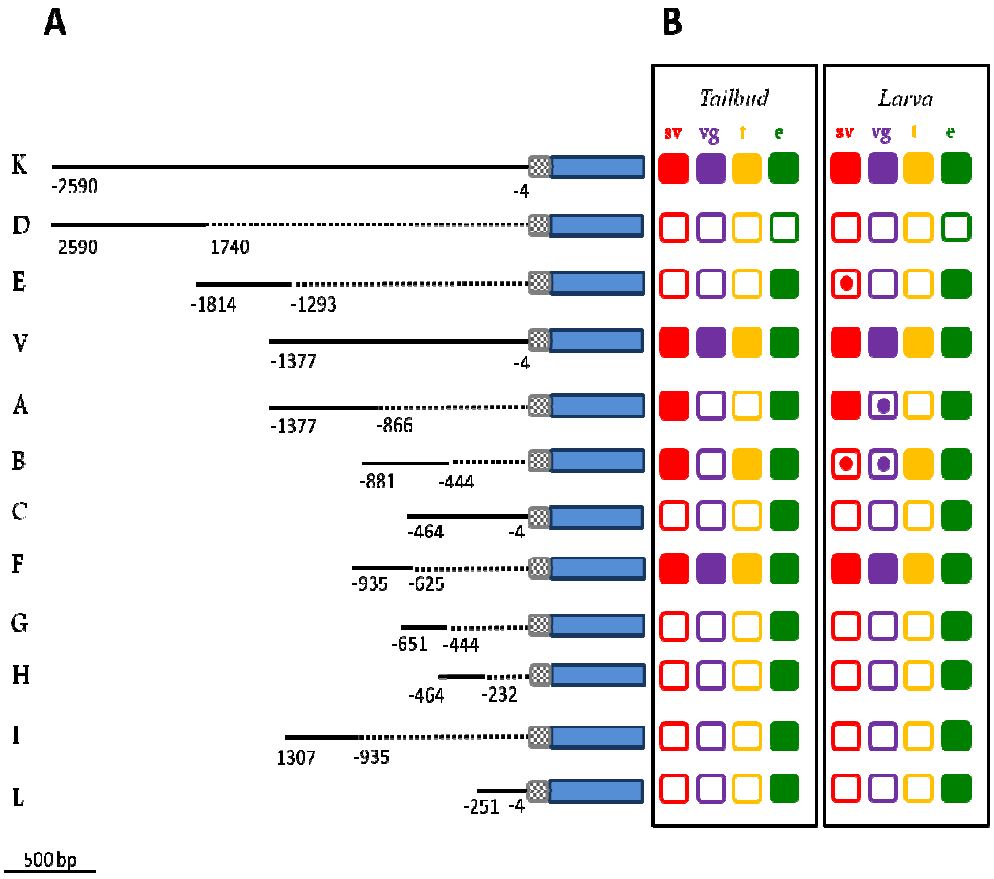


Figure 3.8: ***Summarizing scheme of the results obtained from the Ci-Onecut enhancer region analysis.*** (A) Diagrams of transgenic constructs from K to L. Non coding sequences are represented by black bars. LacZ reporter gene is represented by a blue box. Box with handles indicates the human globin basal promoter. (B) Representation of the results obtained from the reporter gene driving expression at tailbud and larva stage. Full coloured squares indicate reporter gene expression in the coloured-corresponding territory. Framed coloured squares mean absence of reporter gene expression. Framed coloured squares with a dot symbolize a low penetrance of LacZ gene expression in the coloured-corresponding territory. Colour code: e, ectopic mesenchyme, in green; t, tail, in yellow; sv, sensory vesicle, in red; vg, visceral ganglion, in violet.

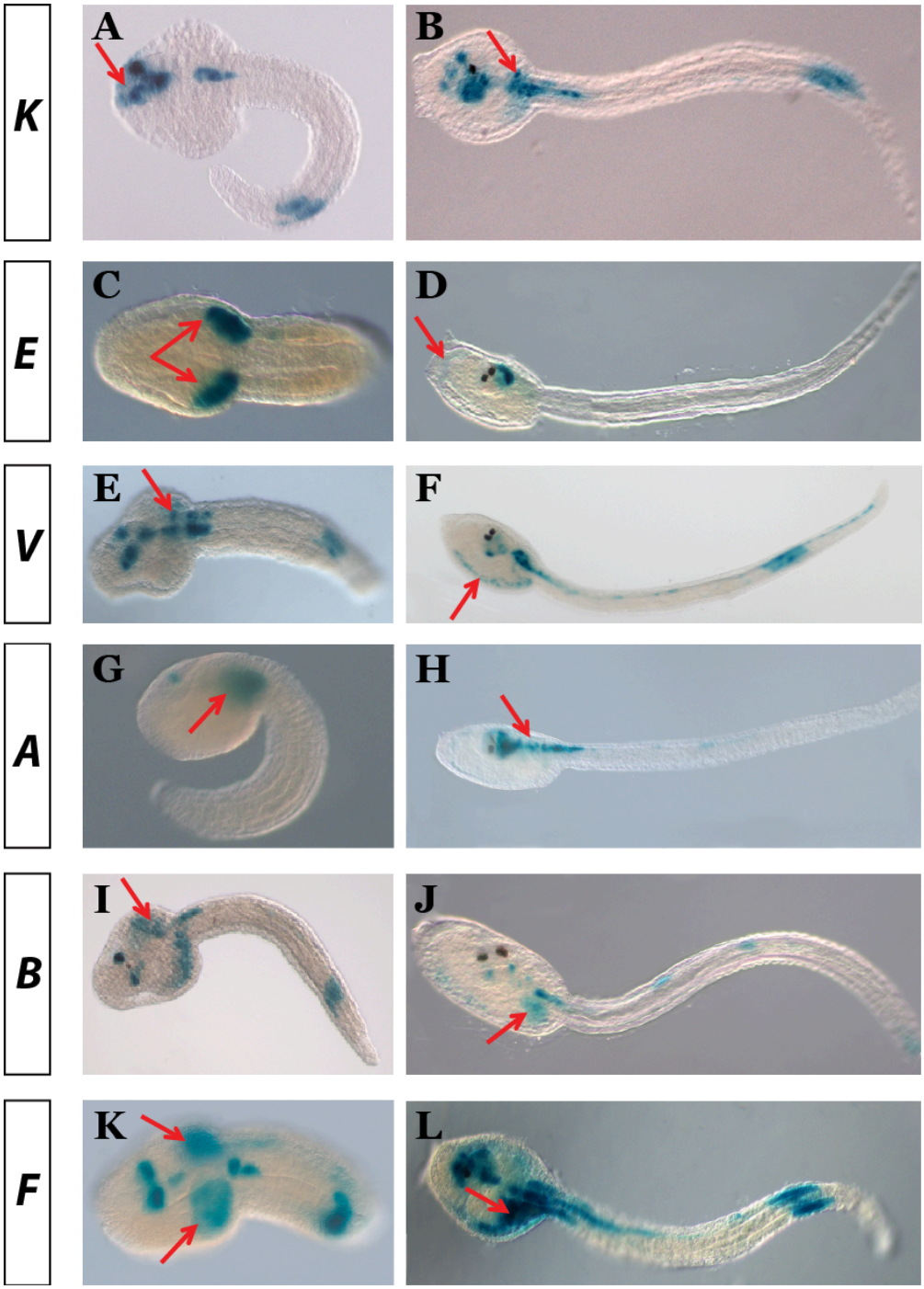


Figure 3.9: ***β-galactosidase histochemical assays of LacZ expression driven by various Ci-Onecut regulatory sequences.*** (A, C, E, G, I, K) Expression of the LacZ reporter gene at tailbud stage in the nervous system for the constructs K, E, V, A, B and E

respectively. (A), (E), (K) Expression of the LacZ reporter gene, for the K, V and F constructs, is detected in the same territories of the endogenous *Ci-Onecut* transcript, respectively. (C) No specific expression of the reporter gene has been detected for the E construct. (G) The expression of the reporter gene for the A construct is localized in the sensory vesicle. (I) LacZ signal for the B construct is visible in the sensory vesicle and in the tail. (B, D, F, H, J, L) Expression of the LacZ reporter gene at larva stage in the nervous system for the constructs K, E, V, A, B and E respectively. (A), (E), (K) Expression of the LacZ reporter gene, for the K, V and F constructs is detected in the same territories of the endogenous *Ci-Onecut* transcript, respectively. (D) Expression of the LacZ reporter gene for the E construct, partially detected in the sensory vesicle. (H) Expression of the reporter gene for the A construct, detected in the sensory vesicle and partially in the visceral ganglion. (J) LacZ signal detected for the B construct, in the tail and partially in the sensory vesicle and visceral ganglion. Red arrows indicate non-specific expression, corresponding to mesenchyme. Lateral (G, H), dorso-lateral (D, E, F, J, K) and dorsal (A, B, C, I, L) view of embryos. Anterior is on the left.

3.6 Identification of Ci-Onecut specific enhancer

The *Ci-Onecut* regulatory sequence contained in the F construct was the only one able to activate the expression of the LacZ gene in the same regions of the endogenous transcript (Fig. 3.10A, B e Fig. 3.9K, L). To isolate the minimal enhancer sequences responsible for *Ci-Onecut* activation, this fragment was divided into the Fa, Fb, Fc, Fd, Fg and Fh smaller fragments (Fig. 3.10A), cloned upstream of the LacZ reporter gene in order to obtain the FA, FB, FC, FD, FG and FH constructs, respectively. Each construct was tested through experiments of electroporation in *Ciona* fertilized eggs, repeated in triplicate in order to give a reliably statistical value to the obtained results. In each electroporation experiment were used as control *Ciona* embryos electroporated with the construct F and the obtained results refer to embryos at tailbud and larva stage.

The FG construct, containing the Fg fragment of 262 bp, extending from position -935 to -674 upstream of the transcription start site of *Ci-Onecut* showed that 90% of the transgenic embryos gave positive results (Fig. 3.10A, B). The reporter gene expression was observed in the sensory vesicle, in particular in the area around the ocellus, in the visceral ganglion, in the tail and also ectopically in the mesenchyme both at tailbud and larva stages (Fig. 3.11A, B).

The 70% of embryos elecroporated with the FA construct, containing the fragment extending from position -935 to -731 upstream of the transcription start site of the *Ci-Onecut* gene showed a strong expression of the LacZ reporter gene in the sensory vesicle around the ocellus and a partial expression in the visceral ganglion and in the posterior part of the tail (Fig. 3.10A, B). Among the positive transgenic embryos, only 20% of them showed a complete expression in all the *Ci-Onecut* endogenous territories (Fig. 3.11C, D).

The Fb fragment extending from position -731 to -625 bp was able to activate the expression of the reporter gene in the ectopic mesenchyme in 67% of observed embryos (Fig. 3.10A, B).

The Fc and Fd fragments, extending from position -935 to -834 bp and from -833 to -731 bp respectively, did not drive any specific or ectopic expression of the reporter gene (Fig. 3.10A, B).

The 73% of transgenic embryos electroporated with the FH construct containing the sequence extending from position -833 to -674 bp showed expression of the reporter gene only in the ectopic mesenchyme while the remaining 27% did not show any signal at all (Fig. 3.10A, B).

In conclusion, this analysis permitted to identify the Fg fragment as the only one able to activate tissue-specific expression of the *Ci-Onecut* gene.

In order to identify possible trans-acting factors able to recognize and activate the Fg sequence, I used the GENOMATIX professional database of vertebrate TFs (http://www.genomatix.de/). Among transcription factors obtained I firstly took into account only those who had a score ≥80. A perfect match to the matrix gets a score of 1.00 (each sequence position corresponds to the highest conserved nucleotide at that position in the matrix), a "good" match to the matrix usually has a similarity ≥ 0.80. Among those with the highest score I selected the ones that had territories of expression similar to those of Ci-Onecut. Therefore for my investigation I have used ANISEED, a database designed to offer a representation of ascidian embryonic development at the level of the genome (*cis*-regulatory sequences, spatial gene expression, protein annotation), of the cell (cell shapes, lineage) or of the whole embryo (anatomy, morphogenesis) (http://www.aniseed.cnrs.fr/). A total of 77 matches were found from the in silico analysis, including binding sites for Neurogenin (Ngn), a proneural basic Helix-Loop-Helix (HLH) protein, and Hnf6 (Onecut, OC). In particular, the Fg sequence contained the presence of one Ci-Neurogenin site, extending from the position -835 to -847 bp upstream of the transcription start site of the *Ci-Onecut* gene and two consensus sequences for Ci-Onecut: the first one (OCa) from -730 to -746 bp and the second one (OCb) extending from -689 to -705 bp (Fig. 3.12A). Most of the other putative factors binding the Fg sequence showed territories of expression arising from the endoderm. Among the factors of greatest interest, that really expressed in the nervous system, I selected for my study Ci-Neurogenin, which presented a score of 0.92, and Ci-Onecut which presented for the first binding site a score of 0.847 and of 0.872 for the second one. The decision to consider these factors is obvious for Ci-Onecut and interesting for Ci-Neurogenin. In fact, Neurogenins are key regulators of neurogenesis even if their function is still not well understood since their primary targets have not been defined. In *Ciona* only one Neurogenin gene was found, that shows territories of expression very

similar to those of Ci-Onecut, starting from earlier stages than when the Ci-Onecut signal starts. A very interesting data, reported on Aniseed, indicates that down regulation experiments, using morpholino oligonucleotide against Neurogenin in *Ciona* cause a loss of *Ci-Onecut* gene expression. Taken together, these data led me to consider this factor as a putative candidate in the pathway of activation of the *Ci-Onecut* gene.

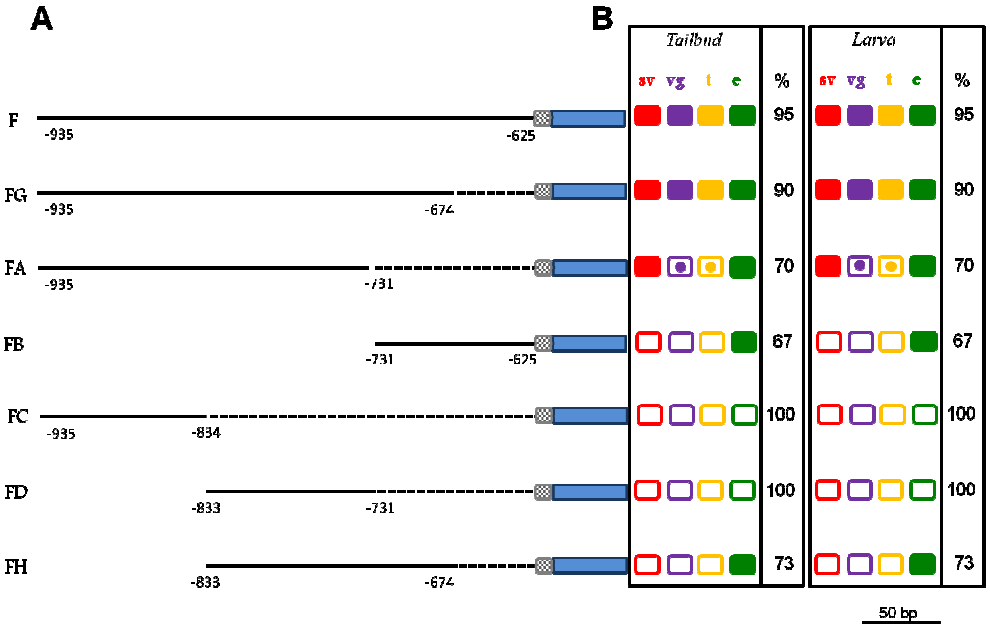


Figure 3.10: ***Summarizing scheme of the results obtained from the Ci-Onecut enhancer region (form F to Fh) analysis***. (A) Diagrams of transgenic constructs from F to FH. Non coding sequences are represented by black bars. LacZ reporter gene is represented by a blue box. Box with handles indicates the human globin basal promoter. (B) Representation of the results obtained about the reporter gene driving expression at tailbud and larva stage. Full coloured squares indicate reporter gene expression in the coloured-corresponding territory. Framed coloured squares mean absence of reporter gene expression. Framed coloured squares with a dot symbolize a low penetrance of LacZ gene expression in the coloured-corresponding territory. Colour code: e, ectopic mesenchyme, in green; t, tail, in yellow; sv, sensory vesicle, in red; vg, visceral ganglion, in violet. "%" indicates percentage of embryos showing expression of the reporter gene.

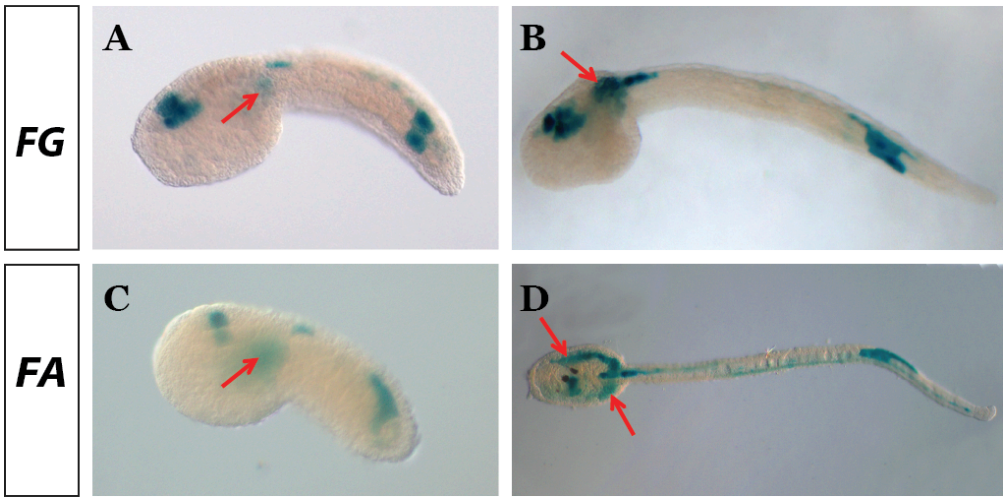


Figure 3.11: ***ß-galactosidase histochemical assays of LacZ expression driven by the Fg and Fa Ci-Onecut regulatory sequences.*** Expression of the LacZ reporter gene for the FG construct, detected in all the *Ci-Onecut* endogenous territories, at tailbud (A) and larva (B) stage, respectively. Expression of the LacZ reporter gene for the FA construct, detected in partial *Ci-Onecut* endogenous territories, at tailbud (C) and larva (D) stage, respectively. Lateral view of tailbud embryos. Lateral (B) and dorsal (D) view of larva embryos. Anterior is on the left. Red arrows indicate non-specific expression, corresponding to mesenchyme.

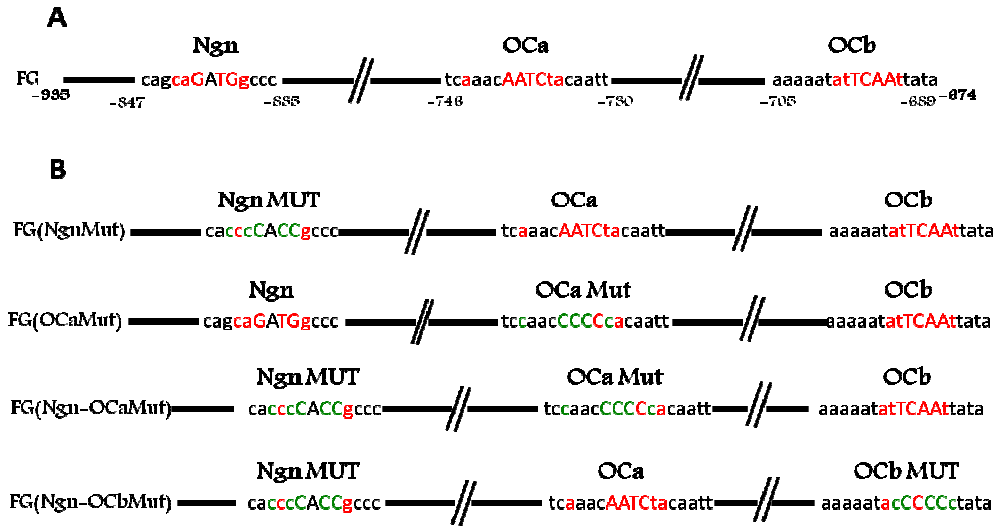


Figure 3.12: ***Schematization of the constructs used for mutational analysis.*** (A) FG construct used as control, in red are indicated the Ci-Neurogenin and the Ci-Onecut binding sites recognized by GENOMATIX database. Core region is shown in capital letters. (B) Constructs mutated in the Ci-Neurogenin and/or in the Ci-Onecut binding sites. In red are represented the nucleotides with a e-value ≥80; in green are indicated the mutated nucleotides in the Ci-Neurogenin and Ci-Onecut binding sites.

3.7 Comparison of Ci-Onecut and Ci-Neurogenin expression

In order to understand if Ci-Neurogenin can be a possible regulator of *Ci-Onecut* expression, I first of all investigated their co-localization in the same territories during embryonic development. I performed double whole mount *in situ* experiments from neurula, when *Ci-Onecut* starts to be expressed, to larva stages (Fig. 3.13A-L). At late neurula stage *Ci-Neurogenin* shows a wide expression in the whole nervous system (Fig. 3.13B), while *Ci-Onecut* shows a more restricted expression profile in only some territories of the future nervous system (Fig. 3.13A). As expected, the merge image obtained by confocal microscopy, demonstrated the co-localization of these two genes in the *Ci-Onecut* territories (Fig. 3.13C). At later stages, from early tailbud to larva stages (Fig. 3.13D-L), this co-localization is still evident although *Neurogenin* expression continues to be wider than that of *Ci-Onecut* and also appear some restricted areas expressing only *Ci-Onecut.* In particular, the merge images demonstrated that these two genes are co-expressed in almost all the *Ci-Onecut* territories at level of the future sensory vesicle, visceral ganglion and part of the posterior tail nerve cord with the exception of two small areas in the most anterior part of the nervous system corresponding to part of the future sensory vesicle and visceral ganglion at tailbud stages (Fig. 3.13F, I, L). Again at larva stage it is

possible to observe a *Ci-Onecut* specific expression in the ventral part of the sensory vesicle and in the visceral ganglion (Fig. 3.13J-L).

To better define the *Ci-Onecut* and *Ci-Neurogenin* co-localization in the sensory vesicle I performed double *in situ* hybridizations with the *Ci-Arrestin*, a specific molecular marker for photoreceptor cells. As shown in figure 3.13M-R, both genes show a clear co-localization with *Ci-Arrestin* at larva stage, thus demonstrating that at level of the sensory vesicle the *Ci-Onecut* and *Ci-Neurogenin* genes are both expressed in the photoreceptor cells.

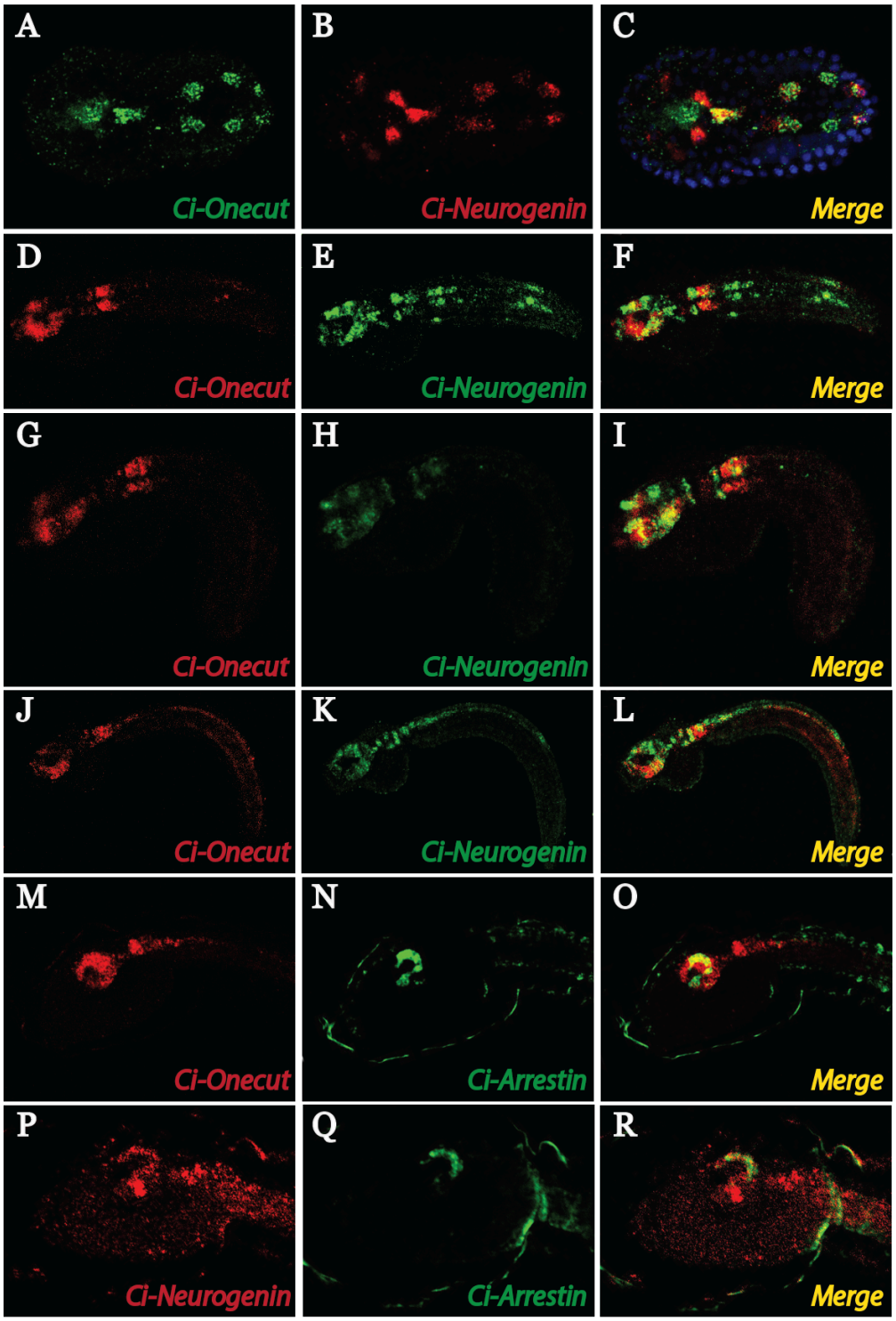

A
Ci-Onecut
B
Ci-Neurogenin
C
Merge
D
Ci-Onecut
E
Ci-Neurogenin
F
Merge
G
CI-Onecut
H
CI-Neurogenin
I
Merge
J
Ci-Onecut
K
Ci-Neurogenin
L
Merge
M
Ci-Onecut
N
Ci-Arrestin
O
Merge
P
Ci-Neurogenin
Q
Ci-Arrestin
R
Merge

Figure 3.13: ***Double whole-mount in situ hybridizations.*** (A-L) Double whole-mount in situ hybridization between *Ci-Onecut* and *Ci-Neurogenin* from neurula to larva stage. (A-C) Neurula stage. Yellow color in the merge images © represents co-expression territories of these two genes in the *Ci-Onecut* territories even if *Ci-Onecut* shows a more restricted expression profile in only some territories of the future nervous system. (D-F) Middle tailbud stage. (G-I) Late tailbud stage. (J-L) Larva stage. The merge images (F, I, L) reveal that these two genes are co-expressed in almost all the *Ci-Onecut* territories at level of the future sensory vesicle, visceral ganglion and part of the posterior tail nerve cord with the exception of two small areas in the most anterior part of the nervous system corresponding to part of the future sensory vesicle and visceral ganglion at tailbud stages. (M-O) Double whole-mount in situ hybridization between *Ci-Onecut* and *Ci-Arrestin* and (P-R) between *Ci-Neurogenin* and *Ci-Arrestin* at larva stage. Merge images (O, R) show that both genes colocalize with the *Arrestin* probe at larva stage thus demonstrating that at level of the sensory vesicle the *Ci-Onecut* and *Ci-Neurogenin* genes are both expressed in the photoreceptor cells. At the bottom of each panel the different genes used are indicated in red and green colors corresponding to the colors of fluorescent staining.

3.8 Ci-Neurogenin and Ci-Onecut bind and activate in vivo Ci-Onecut enhancer

In order to demonstrate that *in vivo* the Ci-Neurogenin and Ci-Onecut transcription factors are able to recognize the binding sites identified on the Fg sequence and consequently to activate *Ci-Onecut* expression, I induced by transgenesis experiments, the ectopic expression of these two TFs in the notochord, a territory where normally they are not expressed. I prepared two constructs, called pBra-Onecut and pBra-Neurogenin, in which the *Ci-Onecut* and *Ci-Neurogenin* coding sequences were cloned under the control of the *Brachyury* promoter (Corbo *et al*., 1997; Yamada *et al*., 2003). Each of these constructs has been co-electroporated together with the FG construct containing the Fg regulatory sequence fused to the *LacZ* reporter gene. I then tested in the transgenic embryos if the ectopic expression of *Ci-Neurogenin* or *Ci-Onecut* in the notochord cells was able to activate the LacZ expression of the FG construct. As a control I used the pBra-RFP construct, in which the coding sequence of the Red Fluorescence Protein (RFP) was under the control of the *Brachyury* promoter (Fig. 3.14A).
As shown in Figure 3.14, both transcription factors induce reporter gene expression in the notochord cells thus demonstrating that these factors recognize binding sequences on the Fg regulatory fragment and are sufficient to activate the gene located under their control (Fig. 3.14B, D, F).

In order to demonstrate that the Ci-Neurogenin and Ci-Onecut binding sites identified by the bioinformatics analysis were effectively the ones responsible for Fg enhancer activation in the notochord cells, I performed a mutational analysis on the Fg fragment. I performed point mutations in five nucleotides of the core sequences for Ci-Neurogenin and Ci-Onecut binding sequences and prepared a series of constructs containing various

combinations of mutated binding sites (Fig. 3.12B). Embryos co-electroporated with the pBra-Neurogenin construct and the FG construct mutated in the Ci-Neurogenin binding site, FG(NgnMut), did not show any LacZ expression in the notochord cells, compared to the control (Fig. 3.14B-C).

In order to verify if also Ci-Onecut was able to recognize and to bind *in vivo* its two binding sites, called OCa and OCb, I performed co-electroporation experiments of the pBra-Onecut construct (D'Aniello *et al.,* 2011) together with the FG construct mutated in the Ci-Neurogenin binding site and alternatively mutated in the Ci-Onecut sites. In this case I observed a strong reduction in LacZ expression in the notochord cells with respect to the wild type FG construct when pBra-Onecut construct was co-electroporated together with FG(OCaMut) construct (Fig. 3.14D-E). A reduced Lac-Z expression, comparable to that obtained with the construct mutated in the first OC site, was also detected after the co-electroporation of the pBra-Onecut construct together with the FG(OCbMut) (Fig. 3.14F-G). As expected I observed in both cases a reduction of LacZ expression and not its total disappearing because of the presence of only one mutated OC site but another still active. It is to note that the misexpression of Ci-Neurogenin and Ci-Onecut in the notochord induces severe alterations in its normal development. The ability of Ci-Neurogenin and Ci-Onecut to induce ectopic expression of the Lac-Z reporter in the notochord cells suggests that *in vivo* these two TFs are able to bind and to activate *Ci-Onecut* enhancer and the similar results obtained with the FG(OCaMut) and FG(OCbMut) constructs indicate that both these binding sites are active and recognized with the same efficiency.

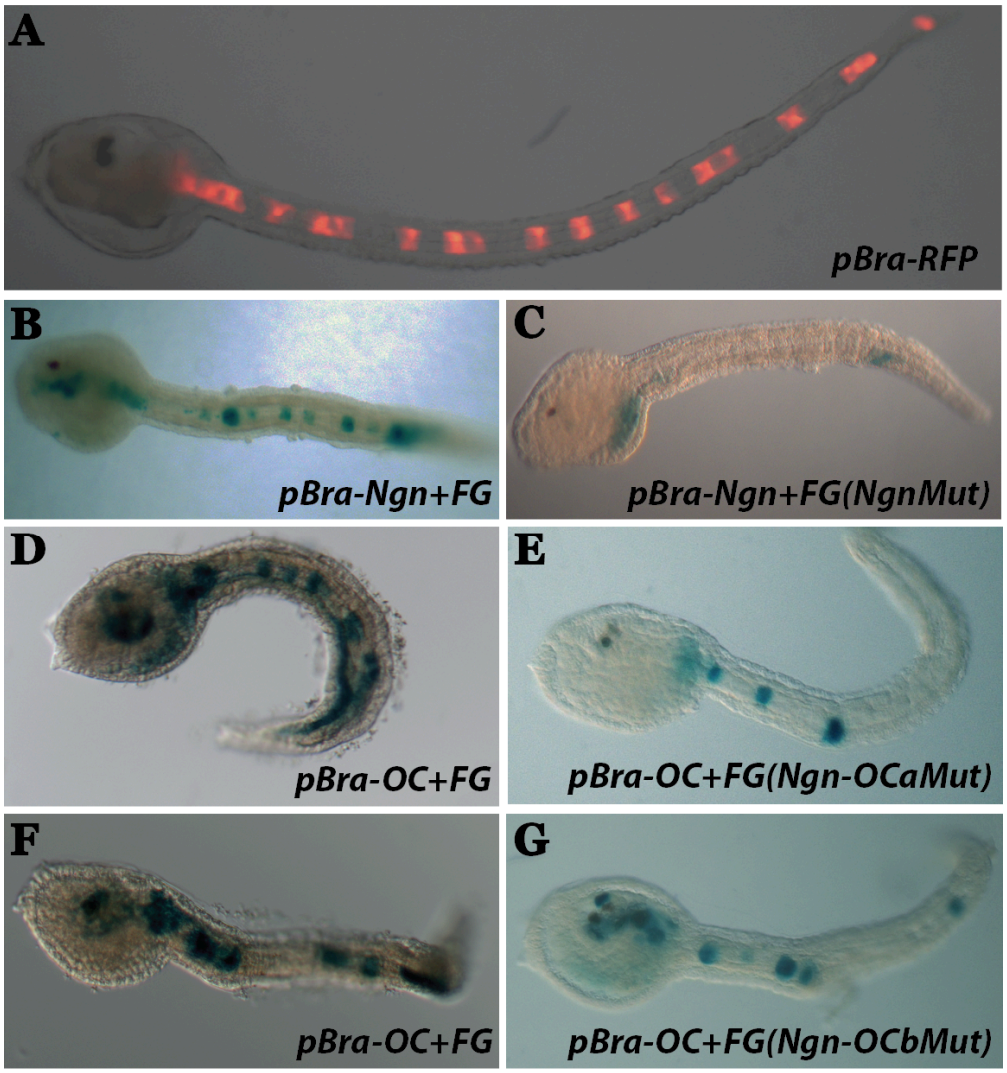


Figure 3.14: ***Ectopic expression of Ci-Neurogenin and Ci-Onecut in Brachyury territories at late tailbud stage.*** (A) Construct used as control showing the expression of the reporter gene driven by the *Ci-Brachyury* promoter in the notochord. (B, D, F) Results of the co-electroporation of the pBra-Onecut or pBra-Neurogenin construct together with the construct containing the wild type Fg regulatory region upstream of the LacZ reporter gene. The expression is visible not only in the *Ci-Onecut* endogenous territories but also in the notochord cells indicating that both Ci-Neurogenin and Ci-Onecut are able to recognize and activate their respectively binding sites on the Fg sequence. (C, E, G) Results of the co-electroporation of the pBra-Onecut or pBra-Neurogenin construct together with the construct containing the Fg regulatory region, mutated in the Ngn, or/and the OC binding sites, upstream of the LacZ reporter gene. The expression is lost in the notochord cells but just partially visible in the *Ci-Onecut* endogenous territories confirming the specific trans-acting binding between the two TFs with the cis-regulatory region.

3.9 Ci-Neurogenin activates Ci-Onecut that in turn controls its maintenance

The same series of experiments with the wild type and mutated FG constructs has been performed in order to analyze in detail the activity of the regulatory elements contained in the Fg sequence in the *Ci-Onecut* endogenous territories of expression. Electroporation experiments with the FG(NgnMut) construct evidenced that only 40% of the transgenic embryos show a positive signal with a low penetrance in the sensory vesicle and the tail in contrast to the 90% of positive embryos obtained with the wild type FG construct (Fig. 3.15A-B, Fig. 3.16A-B). This result indicates that *in vivo* Ngn is responsible for *Ci-Onecut* activation in the nervous system.

To verify if also the Ci-Onecut sites could have a functional role in the regulation of the endogenous *Ci-Onecut* expression, I electroporated the constructs mutated in the first (OCa) or the second (OCb) Onecut binding site and I observed that in both cases the 70% of embryos show only a partial expression of the reporter gene in the sensory vesicle, visceral ganglion and tail with less efficiency with respect to the wild type FG (Fig. 3.15A-B, Fig. 3.16A, C-D). The reduced number of embryos showing expression of the reporter gene in the *Ci-Onecut* endogenous territories, obtained with the two Onecut mutated constructs with respect to the control FG transgenic larvae suggests that, after Ci-Neurogenin activation, both the Ci-Onecut binding sites are important for the maintenance of its gene expression.

To verify this hypothesis I also electroporated the double mutant construct in the sites for Ngn and the first OC (OCa), leaving wild type the second OC site (OCb). In this case the transgenic larvae did not show any expression of the reporter gene in the territories of interest. Only 10% of embryos showed an ectopic mesenchymal expression (Fig. 3.15A-B, Fig.3.16E).

To appreciate putative different efficiency between the two Onecut sites, I conducted the same experiments using the construct double mutant in the sites for Ngn and for the second OC site (OCb). Also in this case was observed a complete absence of the LacZ expression in the *Ci-Onecut* endogenous terrirories (Fig. 3.15A-B, Fig. 3.16F).

The mutational analysis thus indicates not only that all the three identified binding sites (Ngn, OCa and OCb) are necessary for *Ci-Onecut* full activation in its endogenous territories but also that exists a cooperation between the two Ci-Onecut sites for the maintenance of its expression.

All together these results suggest that once activated by Ci-Neurogenin, an autoregulatory mechanism maintain *Ci-Onecut* expression in its endogenous territories.

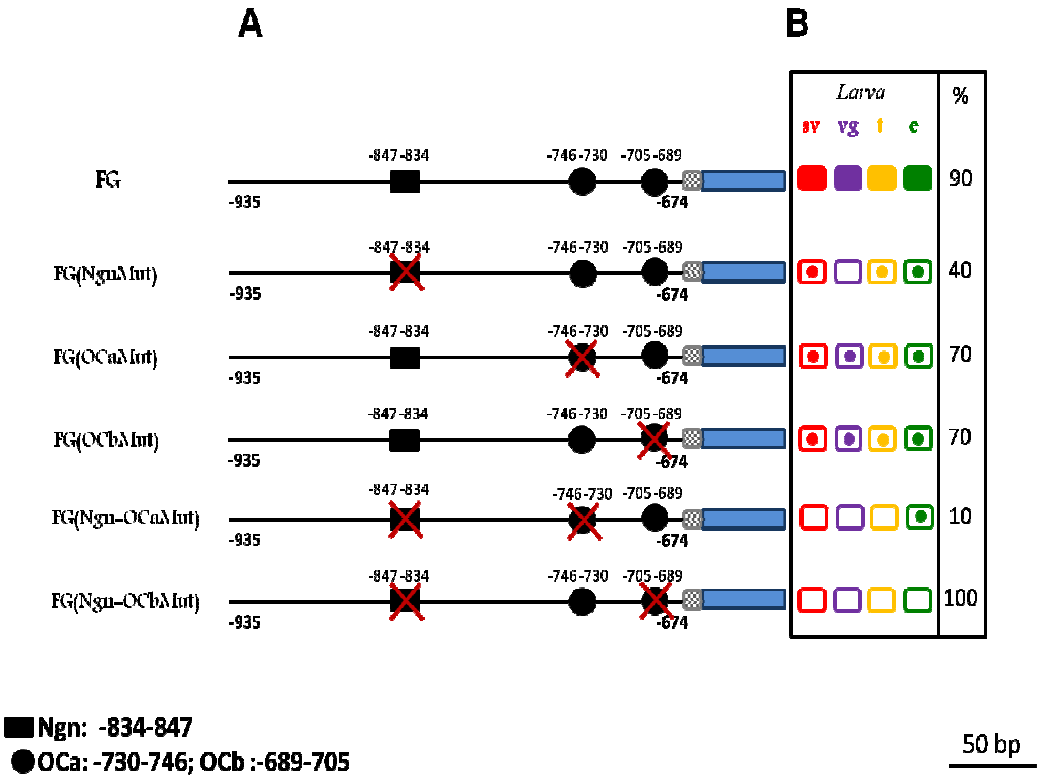


Figure 3.15: ***Summarizing scheme of the results obtained from the Ci-Onecut mutated enhancer region analysis***. (A) Diagrams of transgenic constructs used for mutational analysis. Non coding sequences are represented by black bars. LacZ reporter gene is represented by a blue box. Box with handles indicates the human globin basal promoter. Ci-Neurogenin and Ci-Onecut binding sites are represented by black box and circles, respectively. The red "X" indicates that the corresponding binding site is mutated. (B) Representation of the results obtained about the reporter gene driving expression at larva stage. Full coloured squares indicate reporter gene expression in the coloured-corresponding territory. Framed coloured squares mean absence of reporter gene expression. Framed coloured squares with a dot symbolize a low penetrance of LacZ gene expression in the coloured-corresponding territory. Colour code: e, ectopic mesenchyme, in green; t, tail, in yellow; sv, sensory vesicle, in red; vg, visceral ganglion, in violet . "%" indicates percentage of embryos showing expression of the reporter gene.

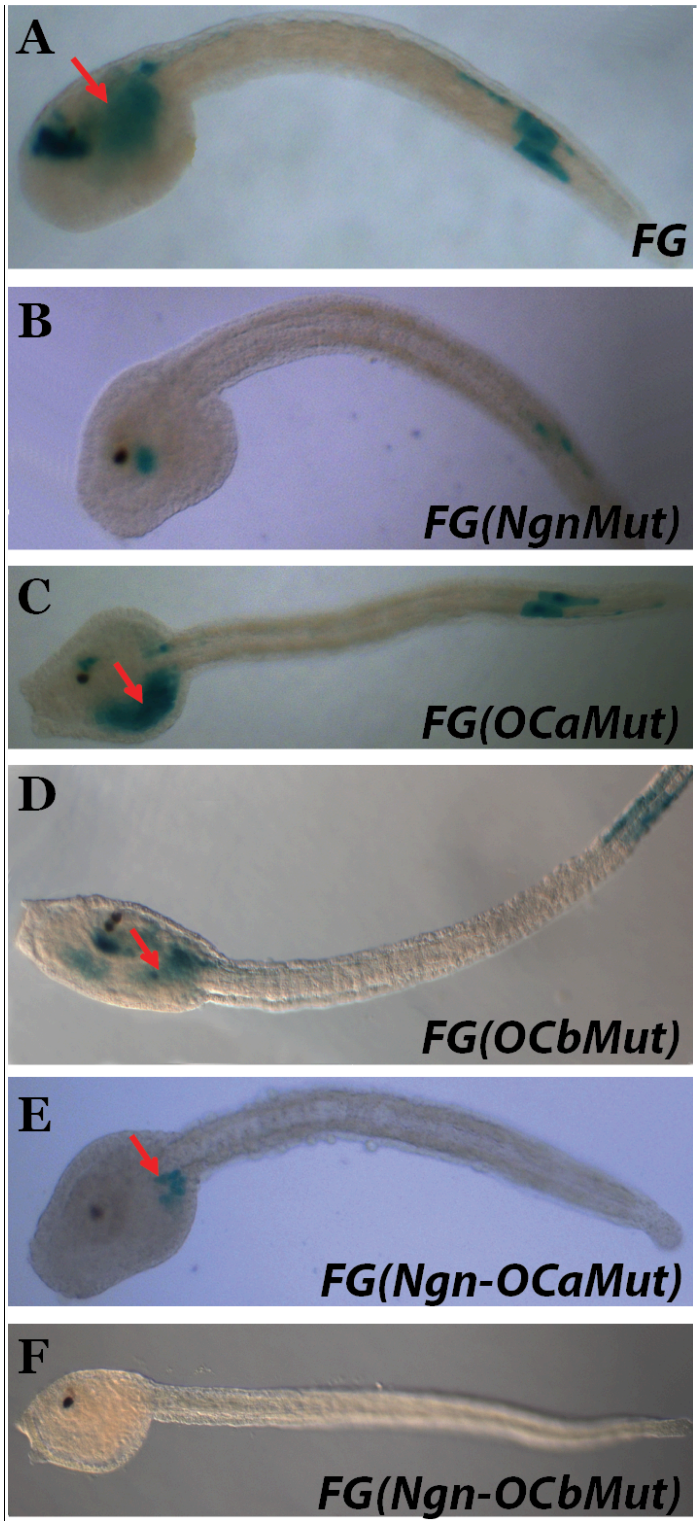


Figure 3.16: ***β-galactosidase histochemical assays of LacZ expression driven by the mutated Fg regulatory sequences.*** (A) Expression of the LacZ reporter gene at late tailbud stage driven by the wild type FG construct, used as control. (B) Transgenic embryo at late tailbud stage for the FG(NgnMut) construct. Positive signal with a low penetrance is present in the sensory vesicle and the tail. (C, D) Embryos at late tailbud and larva stage, respectively, electroporated with the construct mutated in the first (OCa) or the second (OCb) Onecut binding site. Partial expression of the reporter gene is visible in the sensory vesicle, visceral ganglion and tail with less efficiency with respect to the wild type FG (A). (E, F) Double mutants phenotypes obtained at late tailbud stage in the sites for Ngn and the first OC (OCa) or the second OC (OCb) site. Transgenic embryos did not show any expression of the reporter gene in the *Ci-Onecut* endogenous territories. Red arrows indicate non-specific expression, corresponding to mesenchyme.

4. DISCUSSION

4.1 Why to study Ci-Onecut?

Sea squirts are considered the closest relatives of vertebrates and share with this phylum key developmental and physiological characteristics. For these reasons the study of complex phenomena, such as the vision, in a simple model organism, such as the ascidian *C. intestinalis*, could help the understanding of the molecular mechanisms that underlie the complex regulatory networks that coordinate the processes of vertebrate development.

The photoreceptor cells of the ocellus of *C. intestinalis* have morphological, physiological and molecular characteristics that are similar to those of vertebrates (D'Aniello *et al*., 2006), and to shed light on the functioning of the photosensitive system of ascidians can be a great advantage to understand the origin and evolution of the eye of more complex organisms.

The study of the *Ci-Rx* gene expression and regulation in the ascidian *C. ntestinalis*, demonstrated the fundamental role of this gene in the ocellus formation and photoreceptor cells differentiation (D'Aniello *et al*., 2006).

The analysis performed on the *Ci-Rx* promoter sequence, to determine the minimal region responsible for its tissue specific activation, has allowed the identification of two enhancer fragments, of 35 and 50 bp, able to reproduce the *Ci-Rx* endogenous specific expression. These fragments were subjected to a bioinformatic analysis using the GENOMATIX software, that allows to identify on a nucleotide sequence putative binding sites for known transcription factors. The analysis led to the identification on both enhancers of a Onecut (HNF-6) transcription factor binding site.

Similar to *Rx* genes, the expression pattern and function of *Onecut* is highly conserved in invertebrates and vertebrates. *Ci-Onecut* appears conserved within tunicates as *H. roretzi* (Sasakura and Makabe, 2001). In *D. melanogaster*, the *Onecut* homolog has been demonstrated to have a direct role in the central and peripheral nervous system and it has been described to have a role in the formation of photoreceptors (Nguyen *et al*., 2000). In vertebrates various members of the Onecut family have been identified and has been demonstrated the key role played by these factors in the development of the liver and pancreas (Jacquemin *et al*., 1999; Vanhorenbeeck *et al*., 2002). In zebrafish and mammals, *Onecut* genes are expressed in the nervous system, including the retina and pineal gland where *Rx* genes are also expressed (Hong *et al*., 2002; Landry *et al*., 1997). Furthermore, by knock-down gene experiments, it has been shown that Ci-Onecut controls *Chox10* and *Irx* genes (Imai *et al*., 2009). These genes seem to be implicated in retina and photoreceptors development in zebrafish and mouse embryos (Katoh *et al*., 2010; Leung *et al*., 2008).

Although the expression of *Onecut* genes was also detected in various regions of the nervous system, the functions performed by these transcription factors in these areas still remain to be clarified. It is interesting that although it has been widely reported in multiple species that *Onecut* expression and function is in the same territories of *Rx*, such as the nervous system, pineal gland and retina, it had never been implicated in being a direct regulation of the *Rx* genes. *C. intestinalis* diverged from modern vertebrates more than 500 Mya ago. However, the function of *Ci-Rx* gene seems to be conserved in the development of photosensitive structures (D'Aniello *et al*., 2006).

The analysis *in silico* on the *Ci-Rx* promoter sequence, integrated with the data reported in the literature make it possible to advance the hypothesis that Ci-Onecut could be a factor involved in the regulation of *Ci-Rx*, although this supposition has not yet been demonstrated and also suggest that Ci-Onecut could be implicated in a conserved genetic pathway involved in the formation of photosensitive structures.

In order to verify the involvement of Ci-Onecut in the regulation of *Ci-Rx*, was investigated its functional role in the development of the nervous system in *C. intestinalis*, especially in the formation of the ocellus and photoreceptor cells, by analyzing the functional relationship with *Ci-Rx*, and characterizing the genetic cascade in which it is involved. During my research work I have determined the spatio-temporal expression profile of the *Ci-Onecut* transcription factor by *in situ* hybridization experiments during *Ciona* embryonic development (Fig. 1). Previously the research team with whom I worked had shown, through "whole mount" double *in situ* hybridization experiments, the localization of transcripts of *Ci-Onecut* and *Ci-Rx* genes in the same areas at early tailbud and larva stages (D'Aniello *et al*., 2011). However, the *Ci-Onecut* expression includes larger territories; in fact at tailbud stage the transcript is present in the most anterior and central region of the nervous system (Fig. 1E, F, red arrows) while in the larva stage the gene is also expressed at the level of visceral ganglion and in the caudal nerve cord (Fig. 1G, white and black arrows).

Probably, the Ci-Onecut transcription factor, in addition to the development and differentiation of photoreceptor cells, is also involved in the regulation of development of other structures of the central nervous system.

4.2 Studies on the functional connection between Ci-Onecut and Ci-Rx

One of the goals of my PhD work was to clarify the *Ci-Onecut* functional role in the regulation of the *Ci-Rx* gene in the formation of the photosensitive ocellus and in the differentiation of photoreceptor cells. To this aim I conducted perturbation experiments using a specific promoter-driven strategy, using a constitutive activator form and a constitutive

repressor form of the Ci-Onecut protein. I used the sequence of the *Ci-Etr* promoter; *Ci-Etr* is a gene expressed in most cells of the nervous system of the ascidian embryos, including the precursors of the photoreceptor and pigmented cells (Y. Satou *et al.,* 2001) (Fig. 3.1). The transgenic embryos for the Etr-OC-VP16 construct at first glance, seem to suggest that the Ci-Onecut transcription factor is a positive regulator in the genetic cascade in which it is involved. In fact the transgenic larvae for the Etr-OC-WRPW protein showed sometimes two pigmented cells, one of which was "normal" (similar to the wild type one) and another one with altered shape, or sometimes a single pigmented cell, very big and altered in morphology and melanization (Fig. 3.3). The electroporation of the constitutively active form of *Ci-Onecut*, however, gave as a result larvae characterized by the presence of one "normal" and many fused cells forming one elongated pigmented cell or larvae with two or more pigmented cells, of which at least one with a strange morphology or pigmentation (Fig. 3. 4). Considering the variable number of pigmented cells, and the opposite effects of the two forms Etr-OC-WRPW and Etr-OC-VP16 it seems that the Onecut factor could positively regulate genes involved in the formation of pigmented cells.
Since previous knock-down gene experiments through microinjection of antisense morpholino oligonucleotides, have clarified the key role of the *Ci-Rx* gene in the development and differentiation of the ocellus (D'Aniello *et al.,* 2006), it was assumed that also *Ci-Onecut* could be involved in the ocellus formation in *Ciona*. Starting from these considerations I have conducted an immunohistochemical assay, by the use of an antibody against the βγ crystalline protein, a structural protein expressed in the otolith and in the palps, on larvae electroporated with the Etr-OC-VP16 and Etr-OC-WRPW constructs. As expected, in transgenic larvae for the Etr-OC-WRPW protein was observed that the altered cell corresponds to the otolith and that the ocellus is absent (Fig. 3.5B, C). On the contrary, transgenic larvae for the Etr-OC-VP16 construct and containing a supernumerary number of pigmented cells, or one “normal” cell and many others fused one to another to form a single elongated cell, showed the presence of one otolith and various ocelli (Fig. 3.5E, F)

From these results, Ci-Onecut can be regarded as a positive regulator of the *Ci-Rx* gene and involved in the ocellus formation in *Ciona intestinalis.* However, to determine the mechanism of action of *Ci-Onecut* in the ocellus formation, I had to uniquely associate the overexpression of Ci-Onecut to an increase of the *Ci-Rx* endogenous expression and also to the number of ocelli. *Viceversa*, I also had to associate a downregulation of the Ci-Onecut protein to a decrease or absence of the *Ci-Rx* expression and to a reduction or absence of the number of ocelli. First of all, I conducted *in situ* hybridization

experiments on larvae electroporated with the Etr-OC-WRPW and Etr-OC-VP16 constructs using *Ci-Rx* as a probe. As a control *Ciona* embryos electroporated with the construct Etr-GFP were used (Fig. 3.6A, B). This experiment allowed to deduce that the interaction with the Ci-Onecut transcription factor is necessary for *Ci-Rx* gene expression: when in fact Ci-Onecut is in a constitutively active form, the *Ci-Rx* gene expression is visibly increased in its endogenous territories (Fig. 3.6C). On the contrary, in the presence of the Ci-Onecut constitutively repressive form, the *Ci-Rx* expression is reduced or absent in the same territories (Fig. 3.6D). Therefore I provided evidence that Ci-Onecut transcriptional activation is necessary and sufficient to activate *Ci-Rx* endogenous expression in the sensory vesicle (D'Aniello *et al.,* 2011). It would be interesting to investigate if there are conserved *Onecut* binding sites that regulate *Rx* gene expression in vertebrates. It is also possible that the *Onecut* gene was an ancestral regulator of *Rx* expression, but that due to the accumulation of mutational events, vertebrate *Rx* genes acquired new regulatory mechanisms and the dependence on Onecut has been lost. While considering the origin of Onecut and Rx regulation in the sensory vesicle, it is also important to understand the possible homology of the ascidian ocellus, to photosensitive structures in vertebrates (D'Aniello *et al.,* 2006; Horie *et al.,* 2008; Sakurai *et al.,* 2004; Tsuda *et al.,* 2003). Until recently, the prevailing hypothesis was that the ocellus was ancestral to the vertebrate eye. However, in recent years, some authors have proposed that the ascidian ocellus could be homologous to the vertebrate median eye, also called epiphysis or pineal gland. The photoreceptive function of the pineal gland is less conserved and is still present only in some non-mammal vertebrate species. In support of this hypothesis, similar to the ascidian ocellus, the epiphysis of amphibian and teleost larvae triggers the shadow response, when the developing lateral eyes are still not competent to respond to light stimuli (Foster and Roberts, 1982). In addition, the ascidian ocellus and vertebrate epiphysis are both derived from cells located in the lateral part of the embryonic neural plate (Eagleson and Harris, 1990; Nishida, 1987). Together, these similarities between the ocellus and vertebrate epiphysis suggested a possible homology between these two structures. Despite the similarities of the ocellus and the epiphysis, other characteristics do not permit us to distinguish whether or not the ocellus is truly homologous to the eyes or epiphysis. For instance, as in vertebrates, ascidian photoreceptors are ciliary in origin (Eakin and Kuda, 1971), hyperpolarize to light (Gorman et al., 1971), and express *Opsin1*. Furthermore, the sequence of *Ci-Opsin1* is highly homologous to both vertebrate retinal and pineal opsins (Kusakabe *et al.,* 2001). Vertebrate *Rx* genes are expressed in both the pineal organ and the retina (Hong *et al.,*

2002; Mathers and Jamrich, 2000) and play a specific role in the formation of projection neurons of zebrafish pineal gland (Cau and Wilson, 2003). Taking into account a parsimonious theory of evolution, a possible scenario could be that the ocellus represents the ancestral structure of both light sensing organs, which diversified into the vertebrate pineal gland and retinal photoreceptors (Klein, 2006). In this context, it would explain why *Onecut* and *Rx* genes are both expressed in the pineal gland and the retina of many vertebrate species. Further studies in *Ciona* as well as in various vertebrate species will help to elucidate the evolutionary mechanisms that led to the formation of the ascidian ocellus and vertebrate pineal gland and eye.

4.3 Analysis of the Ci-Onecut cis-regulatory region

Given the importance of the Ci-Onecut transcription factor in the study of the molecular mechanisms underlying the development and differentiation of photoreceptor cells in *C. intestinalis*, another important goal of my study was the identification of its minimal elements and the factors responsible for the regulation of *Ci-Onecut* tissue specific expression. To this aim, an analysis of the promoter sequence was performed, using transgenesis by electroporation of constructs consisting of the *Ci-Onecut* promoter regions fused to the LacZ reporter gene cloned into plasmid vectors.

The K construct, which contains a 3 kb fragment of the *Ci-Onecut* promoter, was able to guide the expression of the reporter gene in the same territories of endogenous *Ci-Onecut* transcripts (Fig. 3.8A, B and Fig. 3.9A, B). In order to restrict the isolated region, the K fragment has been divided into progressively smaller and partially overlapping fragments (Fig. 3.8A). Only some of the obtained constructs were able to reproduce the full endogenous *Ci-Onecut* expression. In particular the F construct, constituted by a 311 bp fragment, was the only one capable to activate the expression of the reporter gene in the same areas of the K construct and then, recapitulate, the *Ci-Onecut* full expression (Fig. 3.8A, B and Fig. 3.9K, L).

The E, A and B constructs activates the expression of the reporter gene only in some of the expected territories (Fig. 3.8A, B and Fig. 3.9C-D, G-H, I-J). The A and B constructs were able at larva stage to reproduce almost completely the endogenous *Ci-Onecut* expression although with the A construct the signal is absent in the tail and less intense in the ganglion (Fig. 3.8A, B and Fig. 3.9H), while at tailbud stage it was not detected also in the precursors of the visceral ganglion (Fig. 3.9G). The B construct instead shows no expression in the ganglion at tailbud stage (Fig. 3.8A, B and Fig. 3.9I) and at larva stage it seems to be less efficient in the sensory vesicle (Fig. 3.8A, B and Fig. 3.8J).

Starting from these results, I divided the sequence contained in the F fragment into a series of progressively smaller and partially overlapping fragments, in order to identify the minimal elements responsible for *Ci-Onecut* tissue-specific expression (Fig. 3.10A, B). Among the obtained constructs only FG, constituted by a promoter fragment of 262 bp, was able to reproduce the *Ci-Onecut* full endogenous expression (Fig. 3.10A, B and Fig. 3.11A, B). In order to identify the putative binding sites for known transcription factors responsible for *Ci-Onecut* expression, the sequence of this fragment was subjected to bioinformatic analysis using the GENOMATIX program. From the *in silico* analysis, the more significant results were two binding sites recognized by Ci-Onecut (OC) and one by Ci-Neurogenin (Ngn) (Fig. 3.12A). I analyzed the expression profile of these transcription factors reported in the Aniseed (Ascidian Network for In Situ Expression and Embriological Data) database and Ci-Neurogenin showed an expression profile very interesting in *C. intestinalis*, recalling the *Ci-Onecut* expression, since it is expressed in the anterior sensory vesicle, in the visceral ganglion and in the caudal nerve cord, starting from the 32-cell stage, before the activation of *Ci-Onecut* expression. Furthermore, it was already known in literature that the morpholino against Ci-Neurogenin causes shutdown of the *Ci-Onecut* gene. It was decided to continue the research in this direction in order to verify the possible involvement of Ci-Neurogenin in the *Ci-Onecut* genetic cascade.

The position of the Ci-Onecut binding sites located on the Fg fragment permitted to hypothesize a possible mechanism of autoregulation that only through the site-specific mutagenesis was then confirmed (Fig. 4.1A, B).

In particular, it was noted that the only fragment able to recapitulate Ci-Onecut tissue specific expression was Fg, which is the only one that include all the three considered binding sites. Assuming that Ci-Neurogenin was the enhancer activator, it is also justified that the fragments obtained by deletion of Fg gave absence of expression since they lacked the Ngn binding site (Fig. 4.1FB, FD and FH constructs). On the Fc fragment, instead, the binding site was localized at the 3' end, immediately attached to the β-globin minimal promoter of the vector; this could have created steric problems for the Ci-Neurogenin factor that would probably be unable to efficiently recognize its binding site (Fig. 4.1FC construct).

What made me reflect on the possible mechanism of regulation was the Fa fragment, which was the only one able to direct, although partially, not in all embryos, the tissue-specific expression of the reporter gene (Fig. 3.10A, B and Fig. 3.11C, D). In fact, it presents the binding site for the Ci-Neurogenin and also the first Ci-Onecut binding site (Fig. 4.1FA construct).

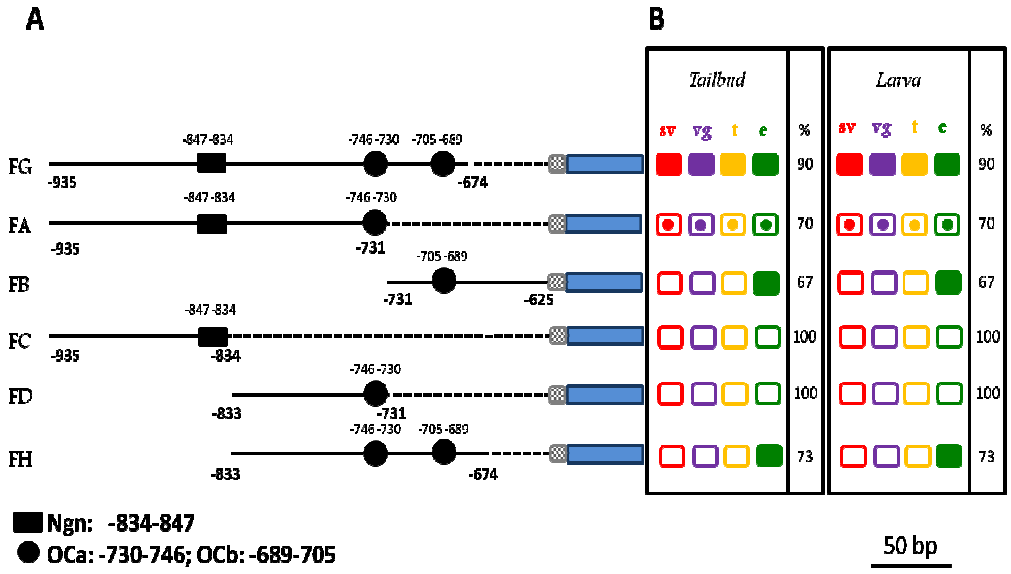


Figure 4.1: ***Summarizing scheme of the structure and of the results obtained from the FG-FH constructs***. (A) Schematization of the Ngn and OC binding sites positions. Non coding sequences are represented by black bars. LacZ reporter gene is represented by a blue box. Box with handles indicates the human globin basal promoter. Neurogenin (Ngn) and Onecut (OC) binding sites are represented by black box and circles, respectively. (B) Representation of the results obtained about the reporter gene driving expression at tailbud and larva stage. Full coloured squares indicate reporter gene expression in the coloured-corresponding territory. Framed coloured squares mean absence of reporter gene expression. Framed coloured squares with a dot symbolize a low penetrance of LacZ gene expression in the coloured-corresponding territory. Colour code: e, ectopic mesenchyme, in green; t, tail, in yellow; sv, sensory vesicle, in red; vg, visceral ganglion, in violet . "%" indicates percentage of embryos showing expression of the reporter gene.

The reduced efficiency and reduced number of positive embryos compared to the control FG could be explained by the absence of the second Ci-Onecut site, which, in cooperation with the first one, might be responsible for the maintenance of the gene expression, after its activation by Ci-Neurogenin.

In order to demonstrate this hypothesis, I performed mutagenesis of the binding sites for Ngn and OC and the analysis of the mutant phenotypes has shown that when the Ngn binding site was mutated, the percentage of positive embryos compared to control FG was drastically reduced (Fig. 3.15A, B and Fig. 3.16A, B). The partial expression of the reporter gene not observed in all territories could be due to the recognition of the *Ci-Onecut* sites by the endogenous Ci-Onecut protein, which would then be able to maintain its own basal expression. The binding specificity between the Ci-

Neurogenin factor and its binding site was analyzed through the over-expression of Ci-Neurogenin in the notochord cells. This experiment highlighted that when Ci-Neurogenin is translated in an ectopic tissue (the notochord) it is able to recognize the binding site present in the Fg sequence and to induce the LacZ expression in the notochord cells (Fig. 3.14B). Furthermore, mutations of the Ngn site present in this sequence abolish ectopic activation of the Fg enhancer in the notochord cells (Fig. 3.14C), recapitulating the phenotype obtained by the electroporation of the FG construct mutated in the Ngn binding site, since there was a partial expression just in the Ci-Onecut endogenous territories (Fig. 3.14C and Fig. 3.16B).

The mutational analysis of the first *Ci-Onecut* binding site confirms the hypothesis suggested by the binding sites positional analysis. In fact the percentage of positive embryos, compared to the mutant phenotypes for Ngn, considerably increases, but is reduced compared to the FG control (Fig. 3.15A, B and Fig 3.16A, C). In addition, the same percentage of embryos obtained from the analysis of the FA phenotype, containing the Ngn and only the first OC sites (Fig.4.1FA construct and Fig. 3.15A, B), clearly suggests that there is a cooperation between the Ngn and both the two OC binding sites. More specifically this result seems to indicate that OC would "potentiate" the expression driven by Ngn and also indicates that the two OC sites are equally efficient. The second OC mutant has in fact confirmed what I have just described, showing that the FA construct was not completely efficient because it lacked one OC site. In addition to demonstrate *in vivo* that also the binding between the Ci-Onecut protein and its binding sites were specific, co-electroporation experiments were conducted by using the FG construct, containing the putative enhancer element of *Ci-Onecut* fused to the LacZ reporter gene and the pBra-Onecut construct. The use of the pBra-Onecut construct allowed to induce ectopic expression of *Ci-Onecut* in the notochord, a territory where normally *Ci-Onecut* is not expressed (Fig. 3.14D, F). The results of these experiments demonstrate that the Ci-Onecut factor, synthesized in the notochord, recognizes the binding site present in FG and is able to activate the expression of the LacZ reporter gene in some cells of this ectopic tissue. Furthermore, the electroporation of the Ci-Onecut protein in the notochord cells, together with the FG construct mutated alternately in the Ngn and in one of the OC binding sites showed identical phenotypes explaining that if one of the two OC sites is mutated, the other one is recognized with the same efficiency, but the expression in the notochord cells is greatly reduced compared to that obtained when both sites are wild type (Fig. 3.14D-E, F-G). These data also support the previous results demonstrating the role of Ci-Onecut in potentiate the Ci-Neurogenin

function to have the full *Ci-Onecut* endogenous expression. The analysis of the double mutant phenotypes has shown that if Ngn and alternatively one of the OC sites are mutated there is no expression of the protein, nor its maintenance, and confirmed that both OC sites are similarly functional and efficient (Fig. 3.15A, B and 3.16E, F) and that Ci-Neurogenin and Ci-Onecut need to be both present to cooperate in enhance the *Ci-Onecut* expression. In particular, the obtained phenotypes showed a LacZ reporter gene expression very similar to those obtained by the co-electroporation of the pBra-Onecut construct together with the FG double mutant constructs (Fig. 3.16E, F and Fig. 3.14E, G), clarifying why in these transgenic embryos the reporter gene expression was observed only in the notochord cells and not in the endogenous *Ci-Onecut* territories.

5. CONCLUSION

5.1 Ci-Neurogenin regulates Ci-Onecut, which controls its expression and directly activates Ci-Rx

All the three identified binding sites are efficient and necessary for *Ci-Onecut* gene regulation and that there is a mechanism of self-regulation of the *Ci-Onecut* gene expression. In particular, Ci-Neurogenin is the necessary and sufficient factor required for the transcription initiation of the *Ci-Onecut* gene and that the Ci-Onecut protein, once expressed, is able to recognize its binding sites located in the Fg enhancer, and to maintain its gene expression in the endogenous territories (Fig. 4.2) (Pezzotti MR *et al*., *in preparation*).

The obtained results demonstrate not only the direct activation by the Neurogenin, but above all show for the first time a mechanism of positive feedback regulation in the family of the CUT genes. The double *in situ* hybridization experiment between the *Ci-Neurogenin* and *Ci-Onecut* genes had already shown that in the early stages in which OC is expressed, the two transcripts present common territories of expression (Fig. 3.13A-C), but at later stages, from early tailbud to larva stages, although this co-localization is still evident, *Ngn* expression continues to be wider than that of *OC* (Fig. 3.13D-L), suggesting that Ngn could be responsible for *Ci-Onecut* activation, which then, at later stages will self-regulate its maintenance. In addition, given my interest in the *Ci-Rx* gene pathway leading to the formation of photoreceptor cells and pigmented ocellus, and shown that Ci-Onecut is a direct activator of the *Ci-Rx* gene *(see Chaper 3, pars.3.1, 3.2, 3.3)*, I performed a double *in situ* hybridization with the *Ci-Arrestin*, showing that the cells of the sensory vesicle in which *Ngn* and *OC* co-localize correspond to photoreceptor cells, the territories in which the *Ci-Rx* gene is also expressed (Fig. 3.13M-R). Ci-Neurogenin is then an excellent candidate in the activation cascade of *Ci-Onecut* (in the regulation of the *Rx* gene) and because data in the literature only indirectly suggested a possible interaction between the two factors, it is interesting to note that the present study highlights for the first time that Ci-Neurogenin in *Ciona* is part of this regulative network, so it is the first time that the three Neurogenin-Onecut-Rx transcription factors are directly associated one to each other (Fig. 5.1).

.

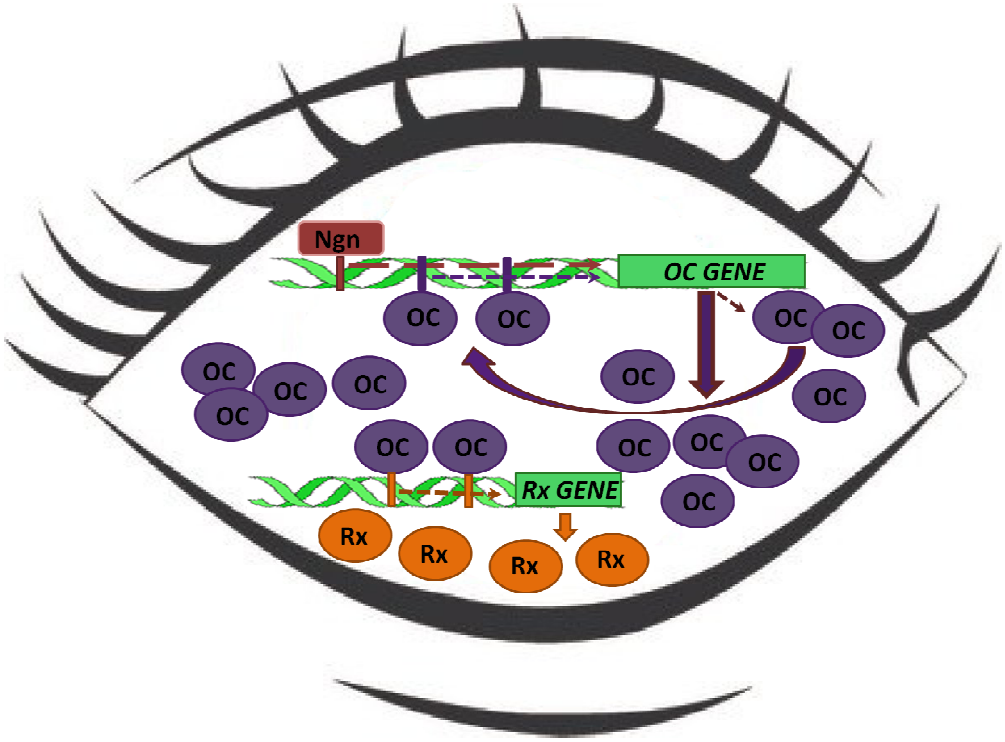


Figure 5.1: ***Proposed Ci-Onecut genetic pathway:*** Ci-Neurogenin (Ngn) is required for the basal *Ci-Onecut* activation (pink arrows). Once expressed, the Ci-Onecut transcription factor is able to recognize its binding sites contained on its promoter (violet arrow) and, by potentiating the Ci-Neurogenin action, is able to give the full expression of the Ci-Onecut protein in its endogenous territories (violet and pink arrow). Among its downstream target genes, Ci-Onecut recognizes its binding sites contained on the *Ci-Rx* promoter and, acting as a direct activator, it regulates the *Ci-Rx* full expression in photoreceptor cells.

6. REFERENCES

Ades, S. E., Sauer, R. T. (1994). Differential DNA-binding specificity of the engrailed homeodomain: the role of residue 50. *Biochemistry* 33, 9187-9194.

Audouard, E., Schakman, O., René, F., Huettl RE., Huber AB., Loeffler JP., Gailly P., Clotman F. (2012). The Onecut transcription factor HNF-6 regulates in motor neurons the formation of the neuromuscular junctions. *PLoS One*. 7(12).

Bourlat, S. J., Juliusdottir, T., Lowe, C. J., Freeman, R., Aronowicz, J., Kirschner, M., Lander, E. S., Thorndyke, M., Nakano, H., Kohn, A. B., Heyland, A., Moroz, L. L., Copley, R. R., Telford, M. J. (2006). Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. *Nature* 444, 85-88.

Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M., Werner, T. (2005). MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* 21, 2933-2942.

Casarosa, S., Andreazzoli, M., Simeone, A., Barsacchi, G. (1997). Xrx1, a novel Xenopus homeobox gene expressed during eye and pineal gland development. *Mech Dev* 61, 187-198.

Cassata, G., Kagoshima, H., Pretot, R. F., Aspock, G., Niklaus, G. and Burglin, T. R. (1998). Rapid expression screening of *Caenorhabditis elegans* homeobox open reading frames using a two-step polymerase chain reaction promoter-gfp reporter construction technique. *Gene* 212, 127-135.

Catt, D., Luo, W., Skalnik, D. G. (1999). DNA-binding properties of CCAAT displacement protein cut repeats. *Cell Mol Biol* 45, 1149-1160.

Cau, E., Wilson, S. W. (2003). *Ash1a* and *Neurogenin1* function downstream of Floating head to regulate epiphysial neurogenesis. *Development* 130, 2455-2466.

Chuang, J. C., Mathers, P. H., Raymond, P. A. (1999). Expression of three Rx homeobox genes in embryonic and adult zebrafish. *Mech Dev* 84, 195-198.

Clotman, F., Jacquemin, P., Plumb-Rudewiez, N., Pierreux, C. E., Van der Smissen, P., Dietz, H. C., Courtoy, P. J., Rousseau, G. G., Lemaigre, F. P. (2005). Control of liver cell fate decision by a gradient of TGF beta signaling modulated by Onecut transcription factors. *Genes Dev* 19, 1849-1854.

Cole, A. G., Meinertzhagen, I. A. (2004). The central nervous system of the ascidian larva: mitotic history of cells forming the neural tube in late embryonic *Ciona intestinalis*. *Dev Biol* 271, 239-262.

Conklin, E. G. (1905a). Mosaic development in ascidian egg. *J Exp Zool* 2, 145-223.

Conklin, E. G. (1905b). Organ-forming substances in the eggs of ascidians. *Biol Bull* 8, 205-230.

Corbo, J. C., Levine, M., Zeller, R. W. (1997). Characterization of a notochord-specific enhancer from the Brachyury promoter region of the ascidian, *Ciona intestinalis*. *Development* 124, 589-602.

D'Aniello, S., D'Aniello, E., Locascio, A., Memoli, A., Corrado, M., Russo, M. T., Aniello, F., Fucci, L., Brown, E. R., Branno, M. (2006). The ascidian homolog of the vertebrate homeobox gene Rx is essential for ocellus development and function. *Differentiation* 74, 222-234.

D'Aniello, E., Pezzotti, MR., Locascio, A., Branno, M. (2011). Onecut is a direct neural-specific transcriptional activator of Rx in *Ciona intestinalis*. *Dev Biol.* 355(2):358-371.

Darras, S., Nishida, H. (2001). The BMP/CHORDIN antagonism controls sensory pigment cell specification and differentiation in the ascidian embryo. *Dev Biol* 236, 271-88.

Davidson, B., Shi, W., Levine, M. (2005). Uncoupling heart cell specification and migration in the simple chordate *Ciona intestinalis*. *Development* 132, 4811-8.

Dehal, P., Satou, Y., Campbell, R. K., Chapman, J., Degnan, B., De Tomaso, A., Davidson, B., Di Gregorio, A., Gelpke, M., Goodstein, D. M., et al. (2002). The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science* 298, 2157-2167.

Delsuc, F., Brinkmann, H., Chourrout, D., Philippe, H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* 439, 965-968.

Di Gregorio, A., Levine, M. (2002). Analyzing gene regulation in ascidian embryos: new tools for new perspectives. *Differentiation* 70, 132-139.

Dunn, C. W., Hejnol, A., Matus, D. Q., Pang, K., Browne, W. E., Smith, S. A., Seaver, E., Rouse, G. W., Obst, M., Edgecombe, G. D., Sorensen, M. V., Haddock, S. H., Schmidt-Rhaesa, A., Okusu, A., Kristensen, R. M., Wheeler, W. C., Martindale, M. Q., Giribet, G. (2008). Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* 452, 745-749.

Eagleson, G. W., Harris, W. A. (1990). Mapping of the presumptive brain regions in the neural plate of *Xenopus laevis*. *J Neurobiol* 21, 427-440.

Eakin, R. M., Kuda, A. (1971). Ultrastructure of sensory receptors in Ascidian tadpoles. *Z Zellforsch* **112,** 287-312.

Eggert, T., Hauck, B., Hildebrandt, N., Gehring, W. J., Walldorf, U. (1998). Isolation of a Drosophila homolog of the vertebrate homeobox gene Rx and its possible role in brain and eye development. *Proc Natl Acad Sci U S A* 95, 2343-2348.

Ermak, T. H. (1977). Glycogen deposits in the pyloric gland of the ascidian *Styela clava* (Urochordata). *Cell Tissue Res* 176, 47-55.

Ferrier, D. E., Holland, P. W. (2002). *Ciona intestinalis* ParaHox genes: evolution of Hox/ParaHox cluster integrity, developmental mode, and temporal colinearity. *Mol Phylogenet Evol* 24, 412-417.

Foster, R. G., Roberts, A. (1982). The pineal eye in *Xenopus Laevis* embryos and larvae: A photoreceptor with a direct excitatory effect on the behaviour. *J Comp Physiol* 145, 413-419.

Furukawa, T., Kozak, C. A., Cepko, C. L. (1997). Rax, a novel paired-type homeobox gene, shows expression in the anterior neural fold and developing retina. *Proc Natl Acad Sci U S A* 94, 3088-3093.

Gehring, W. J. (2004). Historical perspective on the development and evolution of eyes and photoreceptors. *Int J Dev Biol* 48, 707-717.

Gionti, M., Ristoratore, F., Di Gregorio, A., Aniello, F., Branno, M., Di Lauro, R. (1998). Cihox5, a new *Ciona intestinalis* Hox-related gene, is involved in regionalization of the spinal cord. *Dev Genes Evol* 207, 515-523.

Gorman, A. L., McReynolds, J. S., Barnes, S. N. (1971). Photoreceptors in primitive chordates: fine structure, hyperpolarizing receptor potentials, and evolution. *Science* 172, 1052-1054.

Gruss, P., Walther, C. (1992). Pax in development. *Cell* 69, 719-722.

Holland, L. Z., Gibson-Brown, J. J. (2003). The *Ciona intestinalis* genome: when the constraints are off. *Bioessays* 25, 529-532.

Holland, L. Z., Holland, N. D. (2001). Evolution of neural crest and placodes: amphioxus as a model for the ancestral vertebrate? *J Anat* 199, 85-98.

Hong, S. K., Kim, C. H., Yoo, K. W., Kim, H. S., Kudoh, T., Dawid, I. B. , Huh, T. L. (2002). Isolation and expression of a novel neuron-specific onecut homeobox gene in zebrafish. *Mech Dev* 112, 199-202.

Horie, T., Orii, H., Nakagawa, M. (2005). Structure of ocellus photoreceptors in the ascidian *Ciona intestinalis* larva as revealed by an anti-arrestin antibody. *J Neurobiol* 65, 241-250.

Horie, T., Sakurai, D., Ohtsuki, H., Terakita, A., Shichida, Y., Usukura, J., Kusakabe, T., Tsuda, M. (2008). Pigmented and nonpigmented ocelli in the brain vesicle of the ascidian larva. *J Comp Neurol* **509**, 88-102.

Iyaguchi, D., Yao, M., Watanabe, N., Nishihira, J., Tanaka, I. (2006). Crystallization and preliminary X-ray studies of the DNA-binding domain of hepatocyte nuclear factor-6alpha complexed with DNA. *Protein Pept Lett* 13, 531-533.

Jacquemin, P., Lannoy, V. J., O'Sullivan, J., Read, A., Lemaigre, F. P., Rousseau, G. G. (2001). The transcription factor onecut-2 controls the microphthalmia-associated transcription factor gene. *Biochem Biophys Res Commun* 285, 1200-1205.

Jacquemin, P., Lannoy, V. J., Rousseau, G. G., Lemaigre, F. P. (1999). OC-2, a novel mammalian member of the ONECUT class of homeodomain transcription factors whose function in liver partially overlaps with that of hepatocyte nuclear factor-6. *J Biol Chem* 274, 2665-2671.

Jacquemin, P., Lemaigre, F. P., Rousseau, G. G. (2003a). The Onecut transcription factor HNF-6 (OC-1) is required for timely specification of the pancreas and acts upstream of Pdx-1 in the specification cascade. *Dev Biol* 258, 105-116.

Jacquemin, P., Pierreux, C. E., Fierens, S., van Eyll, J. M., Lemaigre, F. P., Rousseau, G. G. (2003b). Cloning and embryonic expression pattern of the mouse Onecut transcription factor OC-2. *Gene Expr Patterns* 3, 639-644.

Katsuyama, Y., Wada, S., Yasugi, S., Saiga, H. (1995). Expression of the labial group Hox gene HrHox-1 and its alteration induced by retinoic acid in development of the ascidian *Halocynthia roretzi*. *Development* 121, 3197-3205.

Katz, M. J. (1983). Comparative anatomy of the tunicate tadpole, *Ciona intestinalis*. *Biol Bull* 164, 1-27.

Klein, D. C. (2006). Evolution of the vertebrate pineal gland: the AANAT hypothesis. *Chronobiol Int* 23, 5-20.

Kusakabe, T., Kusakabe, R., Kawakami, I., Satou, Y., Satoh, N., Tsuda, M. (2001). *Ci-opsin*1, a vertebrate-type opsin gene, expressed in the larval ocellus of the ascidian *Ciona intestinalis*. *FEBS Lett* 506, 69-72.

Kuziora, M. A., McGinnis, W. (1989). A homeodomain substitution changes the regulatory specificity of the deformed protein in Drosophila embryos. *Cell* 59, 563-571.

Lahuna, O., Fernandez, L., Karlsson, H., Maiter, D., Lemaigre, F. P., Rousseau, G. G., Gustafsson, J., Mode, A. (1997). Expression of hepatocyte nuclear factor 6 in rat liver is sex-dependent and regulated by growth hormone. *Proc Natl Acad Sci USA* 94, 12309-12313.

Lannoy, V. J., Burglin, T. R., Rousseau, G. G., Lemaigre, F. P. (1998). Isoforms of hepatocyte nuclear factor-6 differ in DNA-binding properties, contain a bifunctional homeodomain, and define the new ONECUT class of homeodomain proteins. *J Biol Chem* 273, 13552-13562.

Lemaire, P., Bertrand, V., Hudson, C. (2002). Early steps in the formation of neural tissue in ascidian embryos. *Dev Biol* 252, 151-169.

Lemaire, P., Smith, W. C., Nishida, H. (2008). Ascidians and the plasticity of the chordate developmental program. *Curr Biol* 18, 620-631.

Locascio, A., Aniello, F., Amoroso, A., Manzanares, M., Krumlauf, R. and Branno, M. (1999). Patterning the ascidian nervous system: structure, expression and transgenic analysis of the CiHox3 gene. *Development* 126, 4737-4748.

Mannini, L., Deri, P., Picchi, J., Batistoni, R. (2008). Expression of a retinal homeobox (Rx) gene during planarian regeneration. *Int J Dev Biol* 52, 1113-1117.

Mathers, P. H., Grinberg, A., Mahon, K. A., Jamrich, M. (1997). The Rx homeobox gene is essential for vertebrate eye development. *Nature* 387, 603-607.

Mathers, P. H., Jamrich, M. (2000). Regulation of eye formation by the *Rx* and *pax6* homeobox genes. *Cell Mol Life Sci* **57**, 186-194.

Matthews, R. P., Lorent, K., Pack, M. (2008). Transcription factor onecut3 regulates intrahepatic biliary development in zebrafish. *Dev Dyn* 237, 124-131.

Matthews, R. P., Lorent, K., Russo, P., Pack, M. (2004). The zebrafish Onecut gene hnf-6 functions in an evolutionarily conserved genetic pathway that regulates vertebrate biliary development. *Dev Biol* 274, 245-259.

Meinertzhagen, I. A., Lemaire, P., Okamura, Y. (2004). The neurobiology of the ascidian tadpole larva: recent developments in an ancient chordate. *Annu Rev Neurosci* 27, 453-485.

Meinertzhagen, I. A., Okamura, Y. (2001). The larval ascidian nervous system: the chordate brain from its small beginnings. *Trends Neurosci* 24, 401-410.

Munro, E., Robin, F., Lemaire, P. (2006). Cellular morphogenesis in ascidians: how to shape a simple tadpole. *Curr Opin Genet Dev* 16, 399-405.

Nguyen, D. N., Rohrbaugh, M., Lai, Z. (2000). The Drosophila homolog of Onecut homeodomain proteins is a neural-specific transcriptional activator with a potential role in regulating neural differentiation. *Mech Dev* 97, 57-72.

Nicol, D., Meinertzhagen, I. A. (1991). Cell counts and maps in the larval central nervous system of the ascidian *Ciona intestinalis* (L.). *J Comp Neurol* 309, 415-429.

Nishida, H. (1987). Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. III. Up to the tissue restricted stage. *Dev Biol* 121, 526-541.

Nishida, H. (1997). Cell fate specification by localized cytoplasmic determinants and cell interactions in ascidian embryos. *Int Rev Cytol* 176, 245-306.

Nishiyama, A., Fujiwara, S. (2008). RNA interference by expressing short hairpin RNA in the *Ciona intestinalis* embryo. *Dev Growth Differ* 50, 521-529.

Oda, I., Saiga, H. (2001). Hremx, the ascidian homologue of ems/emx, is expressed in the anterior and posterior-lateral epidermis but not in the central nervous system during embryogenesis. *Dev Genes Evol* 211, 291-298.

Ogasawara, M., Di Lauro, R., Satoh, N. (1999). Ascidian homologs of mammalian thyroid peroxidase genes are expressed in the thyroid-equivalent region of the endostyle. *J Exp Zool* 285, 158-169.

Otim, O., Amore, G., Minokawa, T., McClay, D. R., Davidson, E. H. (2004). SpHnf6, a transcription factor that executes multiple functions in sea urchin embryogenesis. *Dev Biol* 273, 226-243.

Pezzotti, MR., Locascio, A., Racioppi, C., Branno, M. (2013). Auto and cross regulatory elements control Ci-Onecut expression in the ascidian nervous system. *Dev Biol. In preparation*

Pierreux, CE., Vanhorenbeeck, V., Jacquemin, P., Lemaigre, FP., Rousseau, GG. (2004). The transcription factor hepatocyte nuclear factor-6/Onecut-1 controls the expression of its paralog Onecut-3 in developing mouse endoderm. *J Biol Chem.* 279(49):51298-304.

Sasakura, Y., Makabe, K. W. (2001). A gene encoding a new ONECUT class homeodomain protein in the ascidian *Halocynthia roretzi* functions in the differentiation and specification of neural cells in ascidian embryogenesis. *Mech Dev* 104, 37-48.

Sakurai, D., Goda, M., Kohmura, Y., Horie, T., Iwamoto, H., Ohtsuki, H., Tsuda, M. (2004). The role of pigment cells in the brain of ascidian larva. *J Comp Neurol* 475, 70-82.

Satoh, G., Harada, Y., Satoh, N. (2000). The expression of nonchordate deuterostome Brachyury genes in the ascidian Ciona embryo can promote the differentiation of extra notochord cells. *Mech Dev* 96, 155-163.

Satoh, N. (1994). Developmental biology of ascidians, *(ed. C. U. Press). New York.*

Satoh, N., Satou, Y., Davidson, B., Levine, M. (2003). Ciona intestinalis: an emerging model for whole-genome analyses. *Trends Genet* 19, 376-381.

Satou, Y., Imai, K. S., Satoh, N. (2002). Fgf genes in the basal chordate Ciona intestinalis. *Dev Genes Evol* 212, 432-438.

Seo, H. C., Kube, M., Edvardsen, R. B., Jensen, M. F., Beck, A., Spriet, E., Gorsky, G., Thompson, E. M., Lehrach, H., Reinhardt, R., Chourrout, D. (2001). Miniature genome in the marine chordate *Oikopleura dioica*. *Science* 294, 2506.

Shimeld, S. M., Purkiss, A. G., Dirks, R. P., Bateman, O. A., Slingsby, C., Lubsen, N. H. (2005). Urochordate betagamma-crystallin and the evolutionary origin of the vertebrate eye lens. *Curr Biol* 15, 1684-1689.

Simion, A., Laudadio, I., Prévot, P. P., Raynaud, P., Lemaigre, F. P., Jacquemin, P. (2010). MiR-495 and miR-218 regulate the expression of the Onecut transcription factors HNF-6 and OC-2. *Biochem Biophis Res Comm* 391, 293-298.

Spagnuolo, A., Ristoratore, F., Di Gregorio, A., Aniello, F., Branno, M., Di Lauro, R. (2003). Unusual number and genomic organization of Hox genes in the tunicate *Ciona intestinalis*. *Gene* 309, 71-79.

Treisman, J., Gonczy, P., Vashishtha, M., Harris, E., Desplan, C. (1989). A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell* 59, 553-562.

Tsuda, M., Kusakabe, T., Iwamoto, H., Horie, T., Nakashima, Y., Nakagawa, M., Okunou, K. (2003b). Origin of the vertebrate visual cycle: II. Visual cycle proteins are localized in whole brain including photoreceptor cells of a primitive chordate. *Vision Res* **43**, 3045-3053.

Vanhorenbeeck, V., Jacquemin, P., Lemaigre, F. P., Rousseau, G. G. (2002). OC-3, a novel mammalian member of the ONECUT class of transcription factors. *Biochem Biophys Res Commun* 292, 848-854.

Vienne, A., Pontarotti, P. (2006). Metaphylogeny of 82 gene families sheds a new light on chordate evolution. *Int J Biol Sci* 2, 32-37.

Voronina, V. A., Kozhemyakina, E. A., O'Kernick, C. M., Kahn, N. D., Wenger, S. L., Linberg, J. V., Schneider, A. S., Mathers, P. H. (2004). Mutations in the human RAX homeobox gene in a patient with anophthalmia and sclerocornea. *Hum Mol Genet* 13, 315-322.

Wada, H., Saiga, H., Satoh, N., Holland, P. W. (1998). Tripartite organization of the ancestral chordate brain and the antiquity of placodes: insights from ascidian Pax-2/5/8, Hox and Otx genes. *Development* 125, 1113-1122.

Wigle, J. T., Eisenstat, D. D. (2008). Homeobox genes in vertebrate forebrain development and disease. *Clin Genet* 73, 212-226.

Wu F, Sapkota D, Li R, Mu X. (2012). Onecut 1 and Onecut 2 are potential regulators of mouse retinal development. *J Comp Neurol*. 520(5):952-969.

Yamada, L., Shoguchi, E., Wada, S., Kobayashi, K., Mochizuki, Y., Satou, Y., Satoh, N. (2003). Morpholino-based gene knockdown screen of novel genes with developmental function in *Ciona intestinalis*. *Development* 130, 6485-6495.

Developmental Biology 355 (2011) 358–371

Contents lists available at ScienceDirect

Developmental Biology

journal homepage: www.elsevier.com/developmentalbiology

Genomes & Developmental Control

Onecut is a direct neural-specific transcriptional activator of Rx in *Ciona intestinalis*

Enrico D'Aniello *[,1], Maria Rosa Pezzotti, Annamaria Locascio, Margherita Branno *

Cellular and Developmental Biology Department, Stazione Zoologica Anton Dohrn, Villa Comunale, 80121 Napoli, Italy

ARTICLE INFO

Article history:
Received for publication 18 November 2010
Revised 21 April 2011
Accepted 4 May 2011
Available online 12 May 2011

Keywords:
Ciona intestinalis
Ascidian
Sensory organs
Pigment cells
Rx enhancer

ABSTRACT

Retinal homeobox (Rx) genes play a crucial and conserved role in the development of the anterior neural plate of metazoans. During chordate evolution, they have also acquired a novel function in the control of eye formation and neurogenesis. To characterize the Rx genetic cascade and shed light on the mechanisms that led to the acquisition of this new role in eye development, we studied Rx transcriptional regulation using the ascidian, *Ciona intestinalis*. Through deletion analysis of the *Ci*-Rx promoter, we have identified two distinct enhancer elements able to induce *Ci*-Rx specific expression in the anterior part of the CNS and in the photosensory organ at tailbud and larva stages. Bioinformatic analysis highlighted the presence of two Onecut binding sites contained in these enhancers, so we explored the role of this transcription factor in the regulation of *Ci*-Rx. By *in situ* hybridization, we first confirmed that these genes are co-expressed in the same cells. Through a series of *in vivo* and *in vitro* experiments, we then demonstrated that the two Onecut sites are responsible for enhancer activation in *Ci*-Rx endogenous territories. We also demonstrated *in vivo* that Onecut misexpression is able to induce ectopic activation of the Rx promoter. Finally, we demonstrated that *Ci*-Onecut is able to promote *Ci*-Rx expression in the sensory vesicle. Together, these results support the conclusion that in *Ciona* embryogenesis, *Ci*-Rx expression is under the control of the Onecut transcription factor and that this factor is necessary and sufficient to specifically activate *Ci*-Rx through two enhancer elements.

Introduction

Rx genes belong to the paired-like homeobox gene family. The developmental role of Rx genes has been widely studied in both vertebrate and invertebrate species. It has been found that they play a conserved role in the development of the anterior neural plate in metazoans. In chordates, they have also acquired a novel later function in the organogenesis of the vertebrate eye (Mathers et al., 1997). Moreover, a growing series of studies suggest that Rx genes are essential for the proliferation and specification of retinal progenitor cells. Mouse Rx null embryos have no eyes and variable reduction of the forebrain and midbrain structures. A similar phenotype has been found in *Xenopus* embryos injected with a dominant negative form of the XRx1 gene. Therefore, this suggests a conserved and essential role of Rx genes in the anterior neural development of *Xenopus* embryos (Andreazzoli et al., 1999). In zebrafish, the *chock*^{s399} mutant, which contains a nonsense mutation in the homeodomain of Rx3, generates embryos missing eye structures. Specifically, the neuroretina fails to differentiate and the retinal pigmented epithelium is not visible (Loosli et al., 2003). In humans, mutations in the Rx locus cause severe ocular malformations, such as anophthalmia (absence of the eye) and microphthalmia (very small eye) (Voronina et al., 2004).

Analysis of Rx dependent genes in vertebrates is consistent with its roles in anterior neural and eye development. In *Xenopus*, Rx inhibits neural differentiation by repressing X-ngnr-1, XDelta1 and N-tubulin (Andreazzoli et al., 2003; Chuang and Raymond, 2001) and activating Zic2 and Xhairy2 (Zilinski et al., 2004). During retinal development in *Xenopus*, Rx induces cellular proliferation through repressing expression of the p27Xic1 gene, a cell cycle inhibitor gene (Andreazzoli et al., 2003). In Rx mutant mouse embryos, the initial activation of *Otx*2, *Six*3 and *Pax*6 gene expression in the anterior neural plate is Rx dependent, although their downregulation could be explained by the absence or lack of proliferation of the retinal progenitor cells (Zhang et al., 2000; Zilinski et al., 2004). However, the finding that ectopic expression of Rx in mouse embryonic stem cells induces their differentiation into retinal progenitor cells supports the hypothesis that Rx genes have a more direct role in retinal differentiation (Tabata et al., 2004). Furthermore, it has been reported that Rx can activate two specific photoreceptor markers, Arrestin and IRBP, by binding specific and conserved elements (PCE/RET1) present in their promoters (Kimura et al., 2000).

Although the function of Rx genes and downstream consequences has been analyzed in vertebrate and invertebrate species, the genetic networks that regulate Rx expression have not received the same attention. Analysis of the XRx1 promoter in *Xenopus* demonstrated that three distinct upstream regions, which contain a high degree of

* Corresponding authors.
E-mail addresses: enrico.daniello@cchmc.org (E. D'Aniello), margherita.branno@szn.it (M. Branno).
[1] Current address: Molecular Cardiovascular Biology, MLC 7020 Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, Cincinnati, OH 45229, USA.

0012-1606/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2011.05.584

conservation with the Rx upstream region of human, mouse, and opossum, are able to drive expression of a reporter gene in the same territories as the endogenous transcript (Danno et al., 2008). Moreover, the authors demonstrate that *Otx2* and *Sox2* are responsible for recapitulating endogenous Rx gene expression. Therefore, given the little that is known about gene regulatory networks responsible for Rx expression, understanding these networks is likely to tell us about the evolutionary and developmental mechanisms by which Rx genes have come to regulate eye development.

Ascidian larvae and vertebrate embryos share anatomical and molecular homologies, making them an excellent model to study the origin of developmental mechanisms, such as Rx regulation and function. At the larval stage, *Ciona intestinalis* embryos have a CNS with many chordate features. The simple ascidian nervous system has four main regions along the anteroposterior axis: the sensory vesicle, the neck, the visceral ganglion and the caudal nerve cord (Meinertzhagen and Okamura, 2001). The sensory vesicle is thought to be homologous to the forebrain of vertebrates and contains two pigmented cells: the anterior spherical otolith, which has a gravity-sensing function and is on the left ventral side of the sensory vesicle, and the photosensing ocellus, which has a half-moon shape and is in the right posterior wall of the sensory vesicle (Lemaire et al., 2002). The ascidian ocellus is composed of a cup-shaped pigmented cell, around 30 photoreceptors and three lens cells (Horie et al., 2005). Photoreceptor cells in *Ciona* exhibit both morphological (ciliary type) (Gorman et al., 1971) and electrophysiological characteristics (hyperpolarization in response to the light) similar to those of vertebrate photoreceptor cells (unpublished data). Furthermore, the ocellus has a photosensitive function during the larval stage, which has been proposed to be ancestral to vertebrate eyes (D'Aniello et al., 2006; Horie et al., 2008; Sakurai et al., 2004; Tsuda et al., 2003). That the *Ciona* ocellus and its photoreceptors are homologous to the vertebrate eye is also suggested by the similarity of *Ci*-Opsin1 to the vertebrate ciliary opsin subfamily (Kusakabe et al., 2001).

In addition to the homology of the ascidian larvae to vertebrates, the compact organization of genes in the ascidian genome (one gene every 7.5 kb of euchromatic DNA) makes it an ideal organism to study conserved gene regulatory networks. The core promoters and associated enhancers are often located within the first 1.5 kb upstream of the transcription start site (Satoh et al., 2003). Using phylogenetic footprinting between the genomic sequences of the ascidians *C. intestinalis* and *Ciona savignyi* allows rapid identification of *cis*-regulatory DNA sequences that mediate tissue-specific patterns of gene expression during *Ciona* embryogenesis. In addition, the ability to rapidly make transgenic larvae offers the opportunity to verify if *Ciona* regulatory elements have been conserved during chordate evolution (Di Gregorio and Levine, 2002; Locascio et al., 1999; Ristoratore et al., 1999). Altogether, the similarities of the ocellus to the vertebrate eye and the compact genome render *C. intestinalis* an optimal and simplified model to study Rx developmental functions and to identify its *cis*-regulatory sequences in chordate embryos (Passamaneck and Di Gregorio, 2005; Satoh et al., 2003).

We have previously studied the function of *Ciona* Rx during development. Similar to what has been demonstrated in vertebrates, we found that *Ciona* ocellus development is dependent on Rx gene function (D'Aniello et al., 2006). Electrophysiological measurements under variable light conditions indicate these Rx-deficient *Ciona* larvae did not show any light dependent changes in their electrical activity and had a corresponding altered ability to swim spontaneously with respect to control larvae (D'Aniello et al., 2006). Furthermore, knockdown of *Ci*-Rx specifically inhibited the expression of two photoreceptor-specific genes, *Ci*-Opsin1 and *Ci*-Arrestin, suggesting the lack of functional photoreceptor cells underlies their inability to sense light stimuli (D'Aniello et al., 2006). In the present study, we describe the analysis of a 2.9 kb non-coding sequence upstream of the *Ciona* Rx gene. Using deletion analysis, we identified a 197 bp region, (−603 to −407 bp) able to recapitulate the endogenous *Ci*-Rx expression pattern in transient transgenic embryos. Bioinformatic analysis indicated that there are putative *Ci*-HNF6/Onecut binding sites within this region of the promoter. We go on to provide evidence that *Ci*-Onecut is directly involved in *Ci*-Rx activation *in vivo* and *in vitro*. Together, these results extend our understanding of the evolutionary regulatory mechanisms underlying eye development through providing the first evidence of a key and direct role of a Onecut gene in the regulation of chordate Rx transcriptional activation.

Methods

Ascidian eggs and embryos

C. intestinalis adults were collected in the bay of Naples, Italy. Eggs and sperm were collected from the gonoducts of several animals and used for *in vitro* fertilization. Fertilized eggs and embryos were used in electroporation or *in situ* hybridization experiments. Embryos were raised in Millipore-filtered seawater at 18–20 °C. Only the batches in which at least 60% of the embryos developed normally were selected for the experiments. Samples at appropriate stages of development were also collected by low speed centrifugation and used for RNA preparation or fixed for *in situ* hybridization.

Constructs preparation

pBlueScript II KS containing the *GFP* reporter gene and the SV40 polyadenylation sequence (Alfano et al., 2007) was used to prepare the B1.6 and C1.3 constructs. The B1.6 construct contains a 0.2 kb DNA fragment corresponding to the basal promoter of the human β-globin. This insertion was necessary to provide the TATA and CAAT boxes necessary for the basal transcription machinery activation. pBlueScript 1230 (gift of R. Krumlauf, Stowers Institute, Kansas City, USA), which contains the *LacZ* reporter gene and SV40 polyadenylation sequence with and without the human β-globin basal promoter, was used for all other constructs. The desired fragments were amplified using polymerase chain reaction (PCR) and inserted in the 5′–3′ orientation upstream of the reporter genes. PCR primers were designed according to the sequence of *C. intestinalis* genome (http://genome.jgi-psf.org/ciona4/ciona4.home.html) and are listed in Table S1.

The constructs from J-A to J-F and OC2 were prepared using a different strategy. Pairs of synthetic purified oligonucleotides (Primm Company; Table S2) were designed containing flanking *Kpn*I and *Xho*I restriction sites at the 5′ and 3′ end, respectively. Oligonucleotides were annealed by boiling 5 min in a hybridization buffer (1× TE and 150 mM NaCl) and slowly cooling to RT. The annealed products were cloned in the pBlueScript 1230 vector.

The 3 kb Etr promoter was amplified by PCR using primers indicated in Table S1. The Etr-GFP construct was prepared by replacing the B1.6 fragment with the Etr promoter in the B1.6 construct. For Etr-OC-VP16 and Etr-OC-WRPW constructs the Etr promoter fragment and the *Ci*-Onecut coding sequence replaced the MESP promoter and the coding sequence of Ets in the MESP>Ets:VP16 and MESP>Ets:WRPW (gift of Brad Davidson, University of Arizona, Tucson, AZ, USA) respectively.

In silico analysis of putative trans-acting factors

The J-A and the J-F sequences were submitted to the MatInspector software of the Genomatix Database (http://www.genomatix.de/cgi-bin/eldorado.main.pl), which is a database of transcription factors binding sites and DNA-binding profiles from many organisms (Cartharius et al., 2005).

Mutagenized Ci-Rx/LacZ constructs

Two different strategies have been used to prepare the J0.2 MUT and J-A MUT constructs. The first was prepared by site-directed mutagenesis from the J0.2 construct with Quik Change Site-Directed Mutagenesis Kit (Stratagene). The putative Onecut-binding site was replaced by a sequence that reduced the binding affinity by using the mutagenic oligonucleotides listed in Table S3. The second was prepared using a pair of synthetic purified oligonucleotides, containing the flanking *Kpn*I and *Xho*I restriction sites at the 5′ and 3′ end, respectively, and replacing the core sequence of the Clox site with a sequence with a reduced binding affinity (Table S3).

Isolation of the Ci-Onecut cDNA

The full-length cDNA of *Ci*-Onecut has been amplified by PCR using as template cDNA from mRNA poly(A)$^+$ isolated at tailbud stage of *C. intestinalis*. The forward oligonucleotide was designed overlapping the ATG start codon (Onecut up) and the reverse oligonucleotide overlapping the stop codon (Onecut down) (Table S3). The full-length *Ci*-Onecut PCR fragment was cloned in the TOPO TA vector and sequenced.

Electrophoretic mobility shift assay (EMSA)

Electrophoretic mobility shift assays (EMSAs) were carried out under the following conditions. The *Ci*-Onecut full-length protein was synthesized *in vitro* using the TNT Coupled Reticulocyte Lysate System (Promega). Each reaction contained ^{32}P labeled substrate DNA fragments and 2 µl of *Ci*-Onecut *in vitro* translated protein in 20 µl of binding mixture, consisting of 10 mM HEPES pH 7.9, 60 mM KCl, 1 mM EDTA pH 8.0, 1 mM DTT, 20% glycerol and 6 pmol poly [dIdC]. DNA fragments were prepared by annealing complementary oligonucleotides. Briefly, 6 pmol single-strand oligonucleotides (Table S3) were radiolabeled by 5′ termini phosphorylation and annealed with cold complementary strand by boiling 5 min and slowly cooling it up to RT. The mixture was incubated 10 min on ice before the addition of radiolabeled DNA probe (10^5 cpm) in the presence or absence of specific, mutated or random (Table S3) double strand competitor oligonucleotides corresponding to a 100× or 200× molar excess. After addition of the labeled DNA, the binding mixtures were placed 20 min on ice and subjected to electrophoresis on a 5% native polyacrylamide/0.5× TBE gel, at 150 mV for 3 h. The products were visualized by autoradiography of the dried gel.

Preparation of co-electroporation construct

The pBS/pBra700/GFP/SV40 construct (gift of Dr. A. Spagnuolo, Stazione Zoologica A. Dohrn, Napoli, Italy) (Corbo et al., 1997) was used to prepare the construct used for the coelectroporation experiments. It consists of a BlueScript vector containing 700 bp of *Ciona* promoter region of Brachyury gene (*Ci*-Bra), GFP as reporter gene and SV40 polyadenylation signal. The GFP reporter gene was replaced with the full-length *Ci*-Onecut cDNA using *Eco*RI to make the Bra-Onecut construct.

Electroporation

Different fusion gene constructs were electroporated into fertilized eggs as described in Locascio and collaborators (1999). Each electroporation was performed using eggs from several different batches, and each construct was tested in two or more electroporations.

Histochemical detection of β-galactosidase activity

Transgenes expression was visualized by histochemical detection of β-galactosidase activity. In brief, embryos at the desired developmental stage were fixed for 15 min in 1% glutaraldehyde in FSW, washed twice with 1× PBS and stained at 37 °C in staining solution (3 mM $K_3Fe(CN)_6$, 3 mM $K_4Fe(CN)_6$, 1 mM $MgCl_2$, 0.1% Tween 20 and 250 µg/ml X-gal in 1× PBS). After incubation, embryos were washed in 1× PBS and imaging capture was made with a Zeiss Axio Imager M1 microscope. For each construct, a minimum of 100 embryos were analyzed in at least five different electroporations.

In situ hybridization

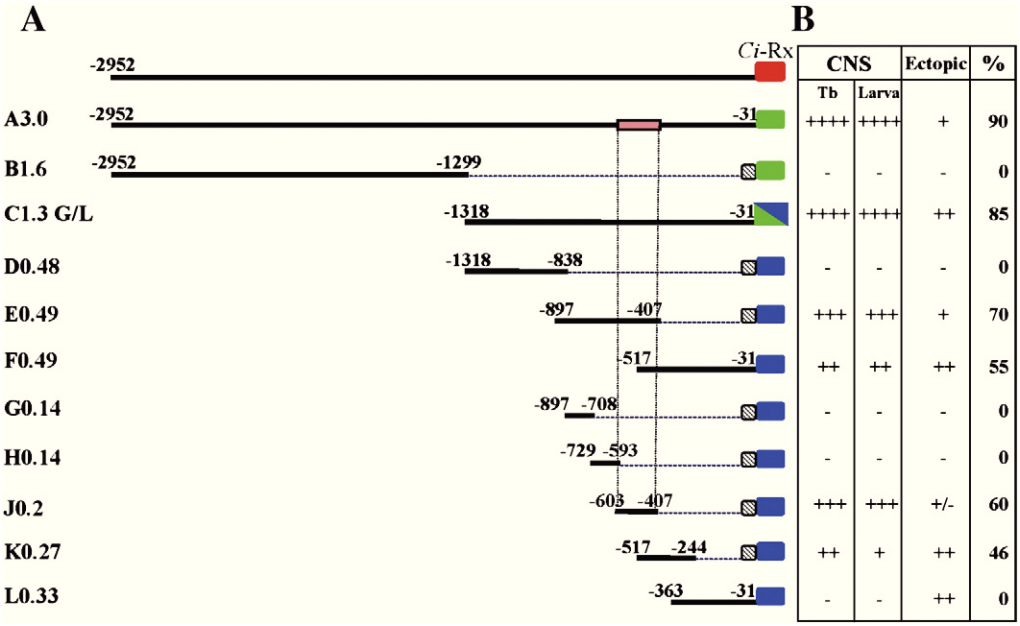
Single whole mount *in situ* hybridization was carried out as previously described (Locascio et al., 1999). Double fluorescent *in situ* hybridization was performed as described by Dufour and collaborators (2006). Embryo imaging capture was made with a Zeiss Axio Imager M1 and a Zeiss LSM 510 META confocal microscopes.

Results

Ci-Rx promoter analysis in electroporated embryos

In order to identify transcription factors (TFs) that control *Ci*-Rx expression in the ocellus and photoreceptor cells, we sought to characterize the regulatory region responsible for *Ci*-Rx tissue specific expression. Previously, we isolated a 2.9 kb genomic fragment (named A3.0, Fig. 1A) extending from the position −2952 to −31 from the translation start site of the *Ci*-Rx gene that was able to reproduce its endogenous expression pattern at tailbud and larva stages (D'Aniello et al., 2006). In order to narrow *Ci*-Rx regulatory elements, we performed a promoter deletion analysis of the 2.9 kb DNA fragment followed by electroporation using GFP or LacZ reporter genes. Initially, we subdivided this fragment into two overlapping fragments: the 1654 fragment, extending from position −2952 to −1299 (B1.6 construct, Fig. 1A), and the 1288 fragment extending from position −1318 to −31 (C1.3 construct, Fig. 1A). The electroporated *Ciona* embryos were allowed to develop until the stage of interest, anesthetized, and were mounted on slides and observed with a fluorescent microscope. The B1.6 construct was unable to induce any specific GFP expression at both tailbud and larva stages (data not shown). However, the C1.3 G construct showed a clear fluorescent signal at the tailbud stage in two anterior groups of cells of the CNS and in one more posterior cell along the dorsal midline of the CNS (Fig. 2A). At the larva stage, the fluorescent signal was detected in the sensory vesicle in the cells surrounding the ocellus (Fig. 2B). The GFP expression in the transgenic embryos completely reproduces the expression pattern of the endogenous *Ci*-Rx transcript (Figs. 2C, D).

In order to identify the regulatory elements responsible for the specific reporter activation, we prepared a series of overlapping DNA fragments covering the entire C1.3 sequence. The C1.3 G construct was subdivided into three smaller and partially overlapping fragments called D0.48, E0.49 and the F0.49 (Fig. 1A). Because of the presence of endogenous GFP auto-fluorescence, particularly evident at the level of the sensory vesicle, we decided to test these new fragments using LacZ as the reporter gene. Moreover, LacZ permits the unambiguous detection of very faint signals due to visualization with histochemical staining. To be able to directly compare the signal intensity of these fragments with that of construct C1.3, we also tested this enhancer using LacZ as reporter gene (called C1.3 L; Figs. 3A, B). This *Ci*-Rx enhancer also induces ectopic expression in the trunk mesenchyme and in some epidermal cells of the tail (Figs. 3A, B). We next investigated the ability of the smaller constructs to induce expression of the reporter gene. We never observed specific expression when embryos were electroporated with the D0.48 construct. However, both



	CNS		Ectopic	%
	Tb	Larva		
A3.0	++++	++++	+	90
B1.6	-	-	-	0
C1.3 G/L	++++	++++	++	85
D0.48	-	-	-	0
E0.49	+++	+++	+	70
F0.49	++	++	++	55
G0.14	-	-	-	0
H0.14	-	-	-	0
J0.2	+++	+++	+/-	60
K0.27	++	+	++	46
L0.33	-	-	++	0

Fig. 1. (A) Diagrams of transgene constructs. Non-coding sequences are represented by black bars. LacZ and GFP reporter genes are represented by a blue and green box, respectively. Red box indicates the Rx gene sequence. The dashed box indicates the human β-globin basal promoter. The pink bar represents a non-coding sequence with high level of homology between *Ciona intestinalis* and *Ciona savignyi*. (B) Scoring chart for expression driven by each transgene in *C. intestinalis* embryos at tailbud and larva stages. Number of "+" symbols denotes relative intensity and penetrance of LacZ and GFP expression; % indicates number of embryos showing expression of the reporter genes. CNS, central nervous system; tb, tailbud stage.

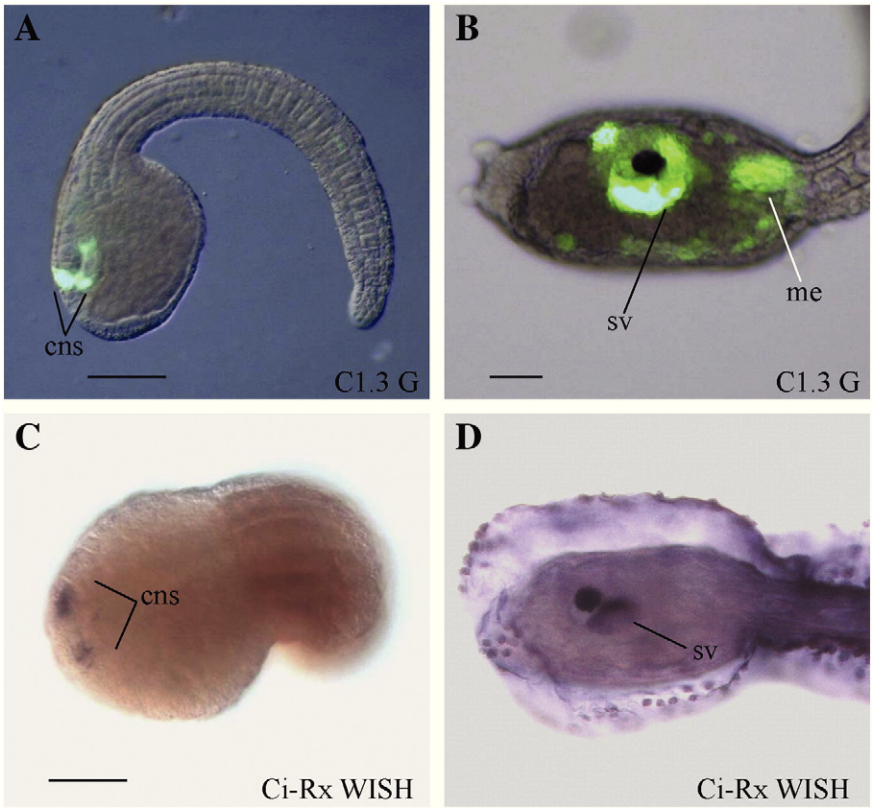


Fig. 2. GFP expression driven by C1.3 G reporter transgene. (A) Tailbud stage, dorsal view. Expression of the reporter gene is visible in the same three groups of cells of the endogenous transcript in the anterior brain. (B) Larva stage, lateral view. Expression of the reporter gene is visible in the sensory vesicle in the cells surrounding the ocellus in the same territory as the endogenous transcript and in some ectopic cells of the mesenchyme (white bar). (C, D) *Ci*-Rx expression pattern visualized by whole-mount *in-situ* hybridization. (C) Tailbud stage, dorsal view. Expression is in three groups of cells in the anterior brain. (D) Larva stage, lateral view. Expression is in the sensory vesicle in the cells surrounding the ocellus. Anterior is to the left in all panels. cns, central nervous system; me, mesenchyme; sv, sensory vesicle. Scale bars indicate 50 μm.

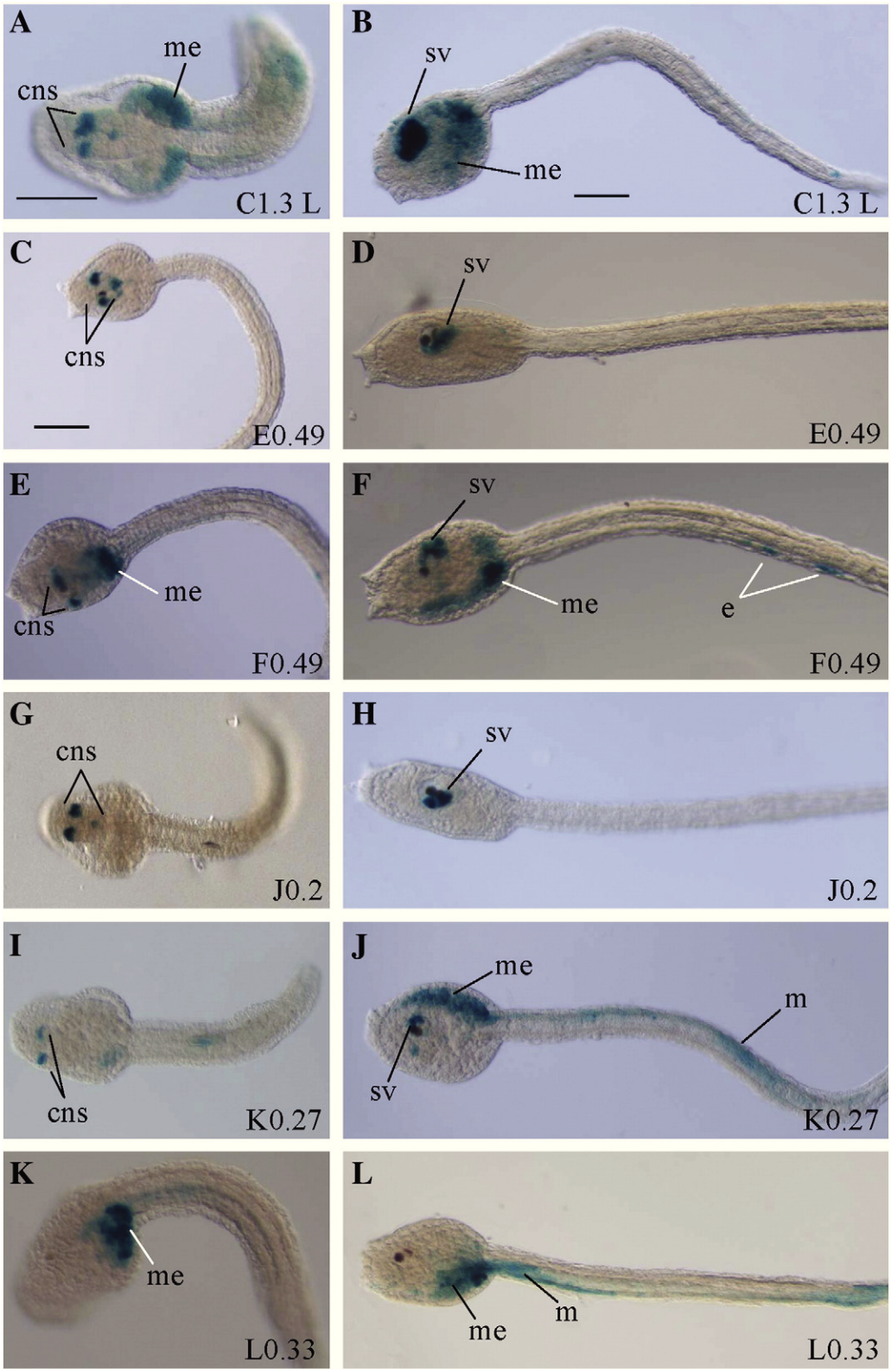


Fig. 3. β-galactosidase histochemical assays of LacZ expression driven by various reporter transgenes. (A, C, E, G, I) Expression of the LacZ reporter gene at tailbud stage in the anterior brain for the constructs, C1.3 L, E0.49, F0.49, J0.2 and K0.27, respectively. Lateral view of tailbud embryos. Anterior is on the left. (K) No specific expression of the reporter gene has been detected for the L0.33 construct at tailbud stage. (B, D, F, H, J) Expression of the LacZ reporter gene at larva stage in the sensory vesicle for the constructs, C1.3 L, E0.49, F0.49, J0.2 and K0.27, respectively. (L) No specific expression of the reporter gene has been detected for the L0.33 construct at larva stage. White bars indicate non-specific expression. cns, central nervous system; m, muscle; me, mesenchyme; sv, sensory vesicle. Scale bars indicate 100 μm.

the E0.49 (Figs. 3C, D) and the F0.49 (Figs. 3E, F) constructs were able to induce the expression of LacZ in the same territories as the endogenous *Ci*-Rx at tailbud and larva stages. Comparing these results with those obtained with the C1.3 L construct, it seems that the E0.49 fragment is slightly more efficient than the F0.49 with respect to induction in the CNS and sensory vesicle, but that both are less efficient

than the previously analyzed C1.3 L fragment (Figs. 3A, B). Moreover, it is evident that the F0.49 fragment also showed ectopic expression in the trunk mesenchyme and in the tail epidermis in a high percentage of transgenic embryos (Figs. 3E, F).

C. intestinalis vs C. savignyi: comparative sequence analysis

Having identified two fragments capable of driving appropriate expression in the sensory vesicle, we decided to carry out a comparative sequence analysis between *C. intestinalis* and *C. savignyi* to highlight the presence of putative conserved sequences between these two closely related species. For this comparison, we used the new *C. intestinalis* genome database (http://genome.jgi-psf.org/Cioin2/Cioin2.home.html/), which already contains a DNA sequence comparison between these two *Ciona* species performed with the VISTA bioinformatic tool (Mayor et al., 2000; http://genome.lbl.gov/vista/index.shtml); *C. savignyi* sequence was taken from the Broad Institute database (http://www.broad.mit.edu/annotation/ciona/). We examined 3 kb upstream of the exon 1 and 1.2 kb downstream of the exon 5 for a total sequence length of 11.2 kb of the *Ci*-Rx genomic locus. This analysis revealed the presence of three conserved non-coding sequences (Fig. 4). The first extends from position −2559 to −2440. The second, which is the only region with more than 75% conserved identity, is in between position −603 and −407. The third conserved sequence, which is positioned between the third and fourth exons, extends from position +3349 to +3473.

Based on these results we decided to investigate *in vivo* the ability of these three conserved regions to specifically activate reporter gene expression in the *Ci*-Rx territories. The first conserved sequence is located within the B1.6 construct, which already has been assayed by electroporation experiments and did not recapitulate endogenous expression (Fig. 1). Consistent with this observation, this reporter construct (M0.1) is unable to activate any specific expression, despite the high level of homology with *C. savignyi* (Fig. 4). In addition, the conserved intronic region located between the third and fourth exons of the *Ci*-Rx gene (INTR 0.12; Fig. 4) did not show any reporter gene activation, indicating that this sequence also lacks a positive regulatory element in this context. However, the last conserved sequence (J0.2), which is located within the E0.49 construct and is partially included in the F0.49 construct (Fig. 1A), was able to drive CNS and sensory vesicle specific expression. Therefore, sequence within J0.2 is likely at least in part responsible for the CNS and sensory vesicle specific expression.

Identification of a Ci-Rx specific enhancer

While J0.2 was able to drive neural specific expression, both E0.49 and F0.49 constructs, which together cover the region from −897 to −31 of the *Ci*-Rx promoter, were also able to drive specific expression. To determine if other sequences within this region are also responsible for CNS and sensory vesicle expression, we made five constructs to cover this region. The E0.49 construct was covered by three partially overlapping constructs G0.14, H0.14 and J0.2, which was the conserved sequence from the Vista analysis analyzed above (Figs. 1A and 4). The F0.49 region, which covered from −517 to −31, was subdivided into two constructs, K0.27 and L0.33 (Fig. 1A). Of these constructs, G0.14 and H0.14 did not show any specific expression (data not shown), while L0.33 drove only ectopic expression not in the CNS or sensory vesicle (Figs. 3K, L). However, the J0.2 and K0.27 constructs drove the reporter gene expression in the nervous system of electroporated embryos at both tailbud and larva stages (Figs. 3G–J). Therefore, these results suggest that these two overlapping sequences contain the element(s) necessary to activate *Ci*-Rx transcription.

To further isolate the minimal sequence responsible for *Ci*-Rx activation, the J0.2 sequence was subdivided into a series of six smaller partially overlapping fragments, J-A through J-F (Fig. 5A). Electroporation experiments showed that the J-B, J-C, J-D and J-E constructs were unable to activate the expression of the reporter gene (Fig. 5B; data not shown). However, the J-A and J-F constructs were able to specifically activate LacZ reporter gene in the sensory vesicle (Fig. 5B). Moreover, the J-A and J-F constructs were able to closely recapitulate the endogenous *Ci*-Rx expression in the sensory vesicle at both the tailbud and larva stages (Figs. 6A–D). Of these two constructs, the J-F fragment produced stronger expression, in terms of signal intensity, with respect to the J-A construct (Figs. 6A–D). However, both fragments were less efficient than the entire J0.2 (Figs. 3G, H).

Putative Onecut DNA-binding sites are found in the conserved Ci-Rx promoter sequence

In order to identify possible binding sites recognized by known transcription factors present in the J-A and J-F sequences, we used the Genomatix professional database of vertebrate TFs (http://www.genomatix.de/cgi-bin/eldorado.main.pl). The Genomatix analysis indicated the presence of multiple potential TF binding sites on both sequences, including those for HNF6 (Onecut). The Cut proteins contain a bipartite DNA-binding domain formed by the Cut domain and a homeodomain and are classified into groups containing one, two or three Cut domains.

Specifically, the J-F sequence contained the presence of a Onecut site (Fig. 6G), while the J-A fragment contained the consensus sequence of a Cut-like gene, called Clox (Fig. 6G). Vista comparison of the J0.2 enhancer between *C. intestinalis* and *C. savignyi* indicated that both the Clox and Onecut binding sites were perfectly conserved (Fig. 8A).

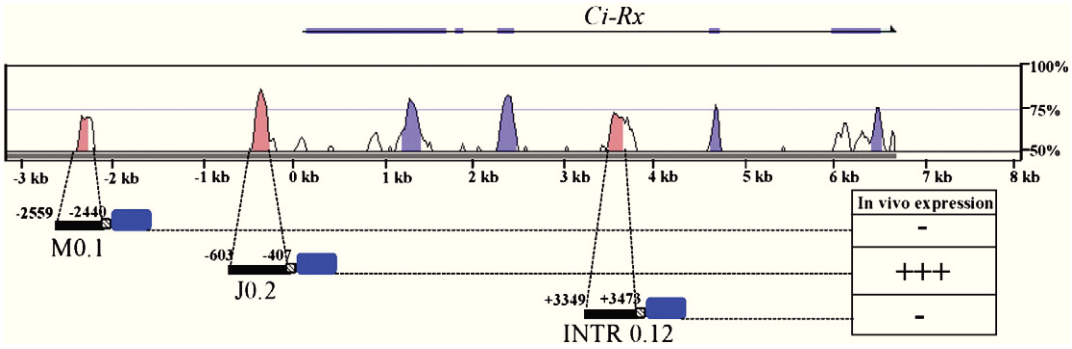
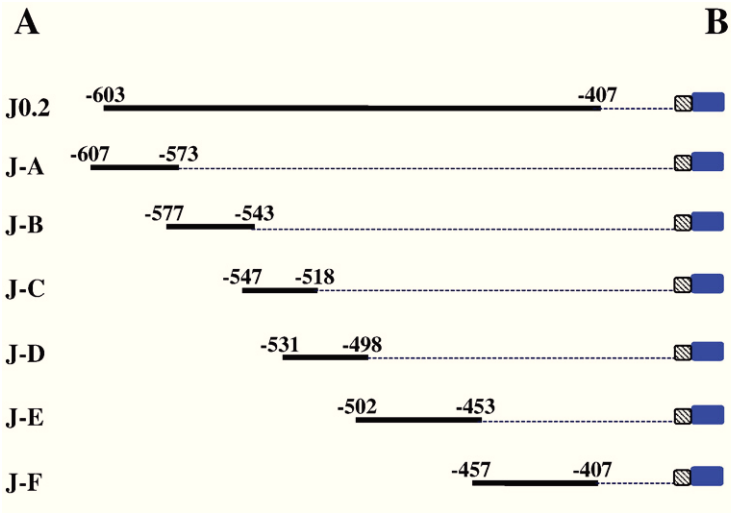


Fig. 4. mVista sequence alignment plot between *C. intestinalis* and *C. savignyi*, with *Ci*-Rx exons shown as blue boxes. Curve represents levels of sequence identity in a 50 bp window. Blue peaks are exons, while pink peaks are non-coding sequences. Below the alignment are the constructs used in this study that have conserved non-coding sequences. The scoring chart indicates expression driven by each transgene in *C. intestinalis* embryos at tailbud and larva stages. "+" symbols denote relative intensity and penetrance of LacZ.



	CNS		Ectopic	%
	Tb	Larva		
J0.2	+++	+++	+/-	60
J-A	+	+	+	22
J-B	-	-	-	0
J-C	-	-	-	0
J-D	-	-	-	0
J-E	-	-	-	0
J-F	++	++	+	30

Fig. 5. (A) Diagrams of deleted transgene constructs of the J0.2 non-coding sequence. Black bars indicate non-coding sequences. Blue box represents the LacZ reporter gene. Dashed box indicates human β-globin basal promoter. (B) Scoring chart for expression driven by each transgene in *C. intestinalis* embryos at tailbud and larva stages. Number of "+" symbols denotes relative intensity and penetrance of LacZ expression. % indicates number of embryos showing expression of the reporter gene. CNS, central nervous system; tb, tailbud stage.

Comparison of Ci-Rx and Ci-Onecut expression

In order to assess if *Ci*-Onecut is a viable candidate to regulate *Ci*-Rx expression, we carried out whole mount *in situ* hybridization to determine if their expression patterns overlap. At the tailbud stage *Ci*-Onecut expression was detected along the antero-posterior axis of the central nervous system in the precursors of the sensory vesicle of the visceral ganglion and in some cells of the spinal cord (Figs. 7A, B). Later at larva stage, *Ci*-Onecut expression in the nervous system is restricted to the sensory vesicle and the visceral ganglion (Fig. 7C). This pattern of *Ci*-Onecut expression is consistent with what has been reported previously in *Halocynthia roretzi* and *C. intestinalis* (Moret et al., 2005; Sasakura and Makabe, 2001). Importantly, comparing the expression of *Ci*-Rx and *Ci*-Onecut at tailbud and larva stages suggested that some of the cells marked by *Ci*-Onecut could overlap with cells in the anterior nervous system that express *Ci*-Rx. Therefore, we then investigated the colocalization of these two genes using double fluorescent *in situ* hybridization. Using confocal microscopy, we found that the most anterior cells expressing *Ci*-Rx consistently also co-express *Ci*-Onecut (Figs. 7E–G). Furthermore, we observed *Ci*-Onecut co-localization in the two more anterior groups of cells that express *Ci*-Rx. Unfortunately, it was not been possible to establish if *Ci*-Onecut is expressed in the more posterior cell marked by *Ci*-Rx, because *Ci*-Rx expression in this cell is too faint and transient to be detected with this technique.

Ci-Onecut binding is necessary for anterior neural expression of Ci-Rx

To study the role of *Ci*-Onecut in the activation of the *Ci*-Rx enhancer, we used the J0.2 fragment for mutational analyses. To test its dependence on the putative Onecut DNA binding elements, we made point mutations in five nucleotides of the core Onecut binding sequence (Fig. 8B). Electroporation experiments of the J0.2 MUT construct, carrying the 5 point mutations, (n>150 embryos analyzed in at least three independent experiments) showed only 29% of the embryos with reporter gene expression in contrast to the 65% of positive embryos electroporated with the wild type J0.2 construct (Fig. 8C). This result supports the hypothesis that the Onecut binding site is involved in the activation of *Ci*-Rx expression. To further verify this hypothesis, we created a new synthetic construct, called OC2, which contains 60 nucleotides with a tandem repetition of the Onecut binding site (Fig. 6G). The OC2 construct was also able to direct expression of the reporter gene in the specific territories of *Ci*-Rx when electroporated in *Ciona* eggs (Figs. 6E, F). To determine if the Clox binding site found in the J-A fragments is also involved in the activation of the *Ci*-Rx enhancer, the Clox binding site in the J-A construct was mutated with 5 nucleotides changes (J-A MUT; Fig. 8B). When electroporated into *Ciona* eggs, the J-A MUT construct did not show any expression of the reporter gene, while the J-A construct was able to drive expression of the LacZ in the 25% of the electroporated embryos (Fig. 8C). Together, these results reveal that at least two Onecut binding sites are involved in Rx transcriptional activation.

Ci-Onecut binds and activates the Ci-Rx enhancers

To address the ability of the Onecut protein to directly bind the *Ci*-Rx enhancer, we next carried out electrophoretic mobility shift assays (EMSA). *In vitro Ci*-Onecut protein was synthesized and combined with either the oligonucleotides encoding the Onecut binding site (Onecut wt) or the Clox binding site (Clox wt), respectively (Table S3). Both the oligonucleotides were able to form specific complexes with *Ci*-Onecut (Fig. 9 lanes 1, 7). In both cases, the retarded band disappeared in the competition with 100–200-fold excess of specific unlabeled oligonucleotide (Fig. 9, lanes 2–3, 8–9), but was not affected by 200-fold excess of the random (R) oligonucleotide (Fig. 9, lanes 4, 10). To further confirm the retarded bands were due to the specific binding of *Ci*-Onecut to these oligonucleotides, the competition assay was performed using mutated oligonucleotides (called Onecut MUT and Clox MUT), which were made by changing the same five nucleotides as used in the analysis above (Fig. 8B). The mutated oligonucleotides were not able to compete with WT oligonucleotides for *Ci*-Onecut (Fig. 9, lanes 5–6, 11–12), nor were the mutated oligonucleotides able to be retarded by *Ci*-Onecut alone (data not shown). Therefore, these results suggest that *Ci*-Onecut is directly involved in the *Ci*-Rx enhancer activation and is able to recognize the Onecut binding sites.

To investigate *in vivo* the ability of the *Ci*-Onecut protein to bind and active the two *Ci*-Rx enhancers, we next determined if *Ci*-Onecut is able to induce the ectopic expression of the reporter constructs in the notochord, a territory where it is normally not expressed. For this purpose, we prepared a construct in which the coding sequence of *Ci*-Onecut has been cloned downstream of the *Ci*-Brachyury promoter sequence (Bra-Onecut). Brachyury is specifically expressed in the notochord and its promoter has already been characterized in *Ciona* (Corbo et al., 1997; Yamada et al., 2003). To verify its ability to induce *Ci*-Onecut ectopic expression in the notochord cells, Bra-Onecut electroporated embryos were analyzed by *in situ* hybridization

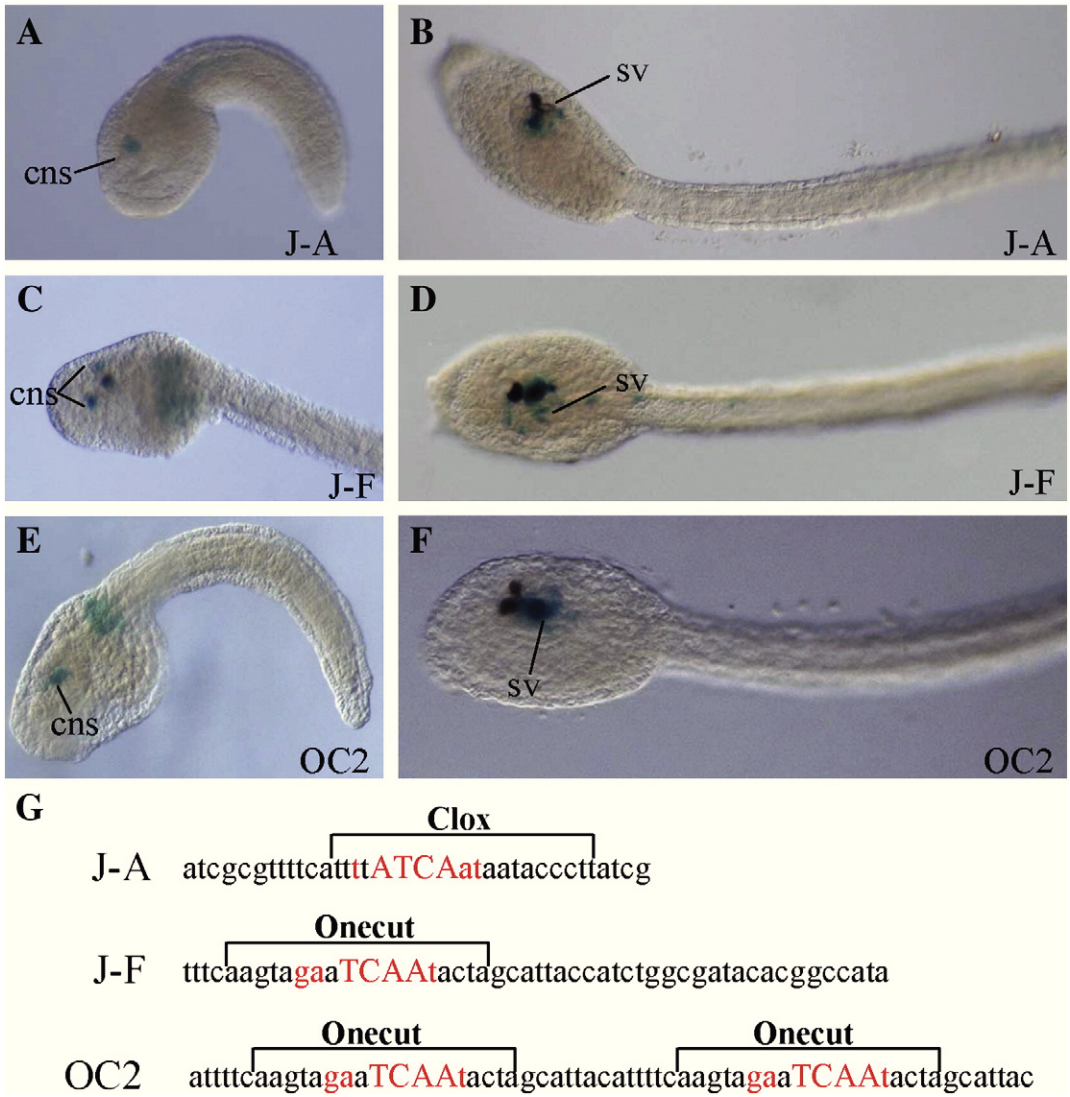


Fig. 6. β-galactosidase histochemical assays of LacZ expression driven by various reporter transgenes. Embryos are in lateral view with anterior to the left. (A, C, E) Expression of the LacZ reporter gene at tailbud stage in the anterior brain for the constructs, J-A, J-F and OC2 respectively. (B, D, F) Expression of the LacZ reporter gene at larva stage in the sensory vesicle for the constructs J-A, J-F and OC2, respectively. (G) Schematization of the J-A fragment containing the Clox (Cut-like) binding site, the J-F fragment containing the Onecut binding site, and the OC2 fragment containing a tandem repetition of the Onecut binding site. cns, central nervous system; m, muscle; me, mesenchyme; sv, sensory vesicle.

experiments using *Ci*-Onecut probe. *Ci*-Onecut transcript was detected in the sensory vesicle and visceral ganglion where it is normally expressed and in various notochord cells (Fig. 10A). *Ci*-Onecut is not expressed in all the notochord cells due to the mosaic incorporation of the transgene. In 40% of the electroporated embryos, we observed an alteration in the normal development of the notochord, likely due to the misexpression of *Ci*-Onecut in this tissue. Having established that Bra-Onecut drives *Ci*-Onecut expression in the notochord, we then electroporated the *Ciona* eggs with Bra-Onecut together with the wild type or mutated J0.2 and J-A constructs. 40% of the embryos coelectroporated with the J0.2 or the J-A constructs express the LacZ reporter genes in the notochord cells as well as the endogenous territories of *Ci*-Rx (Figs. 10B, C, E). However, the J0.2 MUT construct showed strongly reduced expression in the notochord cells compared to the J0.2 construct, likely because of the presence of a mutated Onecut site but a still active Clox site (Fig. 10D). The J-A MUT construct, which only contains the mutated Clox site, did not show any expression in the notochord cells (Fig. 10F). We also coelectroporated the OC2 construct together with Bra-Onecut, which produce ectopic expression in the notochord cells with very high efficiency (Fig. 10G). Therefore, the ability of *Ci*-Onecut to induce ectopic expression of the *Ci*-Rx LacZ reporters in the notochord cells, where none of the constructs of the *Ci*-Rx promoter has never been active, strongly suggests that *in vivo Ci*-Onecut is able to bind and activate *Ci*-Rx expression through both Onecut and Clox sites.

Ci-Onecut activates Ci-Rx gene expression

In order to assess the ability of *Ci*-Onecut protein to influence Rx endogenous expression, we overexpressed a constitutive activator (OC-VP16) or a constitutive repressor (OC-WRPW) form of *Ci*-Onecut. To induce ectopic OC-VP16 or OC-WRPW, we used an enhancer of *Ci*-Etr, which is specifically expressed in the nervous system up to the larval stage and starts to be expressed in the CNS lineage from the 110 cells stage (aniseedV3_2090 on the ANISEED web site; http://aniseed-ibdm.univ-mrs.fr/). The *Ci*-Etr enhancer is able to recapitulate its endogenous expression pattern throughout the CNS of *Ciona* embryos (Fig. 11A). Whole mount *in situ* hybridization experiments for *Ci*-Rx on embryos electroporated with the constitutive activation construct (Etr-OC-VP16) revealed the presence of a strong *Ci*-Rx signal in the sensory vesicle (Fig. 11C), compared to electroporation of control Etr-GFP construct that did not

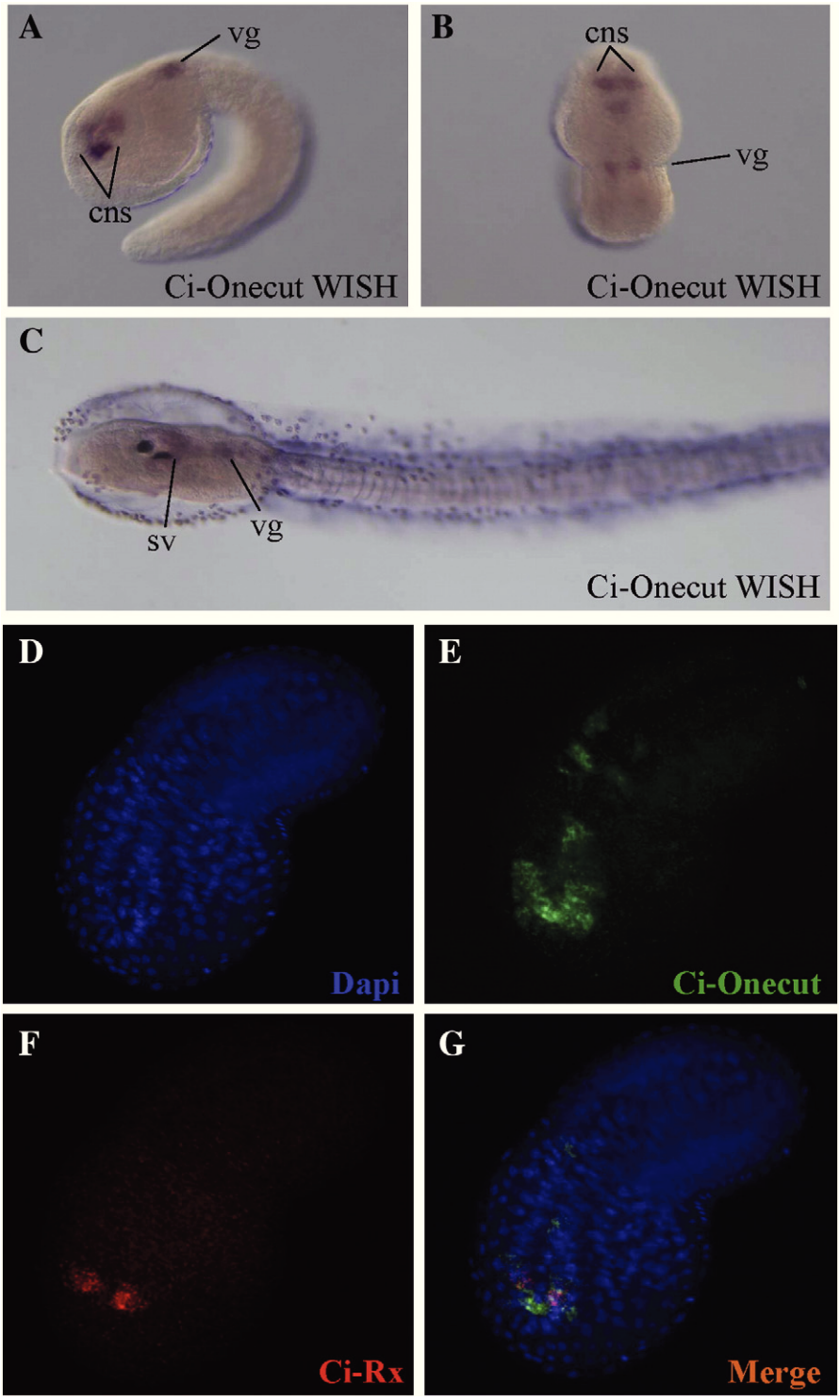


Fig. 7. Whole-mount *in situ* hybridization of *Ci*-Onecut and *Ci*-Rx. Anterior is to the left in all panels. (A, B) Tailbud stage lateral and dorsal view. Expression of *Ci*-Onecut is detected along the antero-posterior axis of the CNS. (C) Larva stage, lateral view. *Ci*-Onecut expression is detected in the sensory vesicle and in the visceral ganglion. (D–G) Positional relationship of *Ci*-Rx and *Ci*-Onecut gene expression domains. (D) DAPI image. (E) *Ci*-Onecut expression at tailbud stage (dorsal view). (F) *Ci*-Rx expression at the same stage. (G) Merge image of D, E, and F. Orange color in G indicates co-expression territory of *Ci*-Rx and *Ci*-Onecut. cns, central nervous system; sv, sensory vesicle; vg, visceral ganglion.

affect *Ci*-Rx expression embryo (Fig. 11B). Furthermore, 40% of embryos electroporated with the constitutive repressor construct (Etr-OC-WRPW) did not show any *Ci*-Rx endogenous expression (Fig. 11D), while 45% of the embryos showed a reduced signal (data not shown).

Therefore, these results demonstrate a functional connection between *Ci*-Onecut and *Ci*-Rx expression and that *Ci*-Onecut is not only able to activate *Ci*-Rx enhancer, but its ability to act as a transcriptional activator is also necessary for *Ci*-Rx endogenous expression in the anterior part of the nervous system. Because ectopic expression is

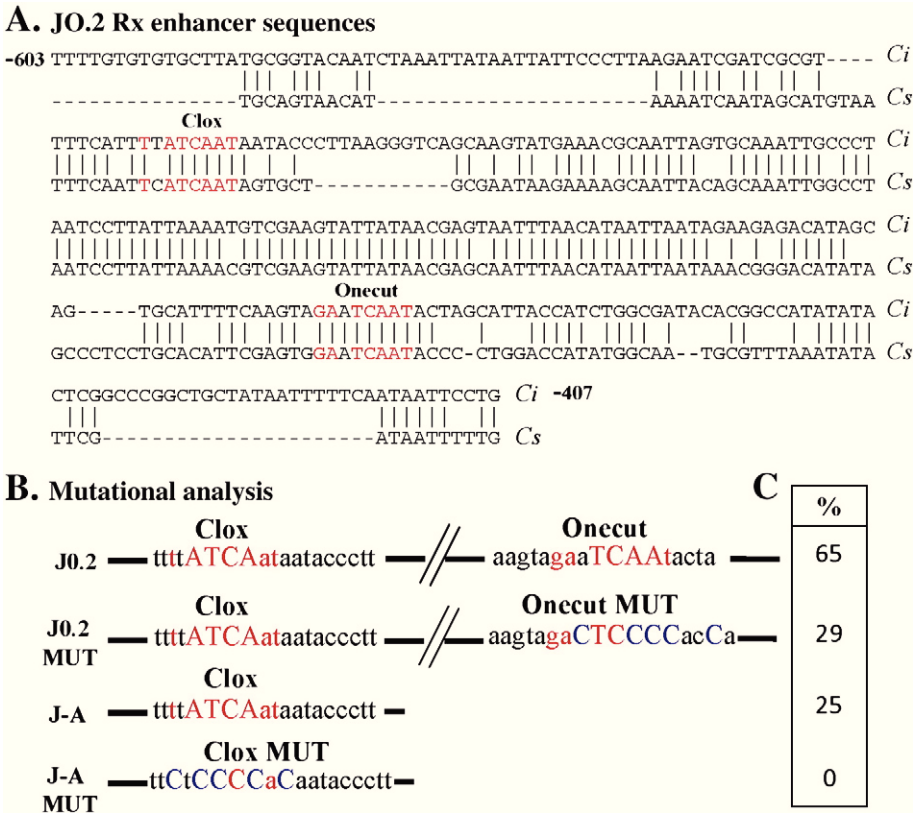


Fig. 8. (A) Phylogenetic footprinting of the Onecut elements (red letters) in the J0.2 sequence between *C. intestinalis* (*Ci*) and *C. savignyi* (*Cs*). Vertical bars indicate conserved nucleotides. (B) Schematization of the constructs used for mutational analysis. Blue indicates mutated nucleotides in the Onecut and Clox binding sites. (C) Scoring chart for expression driven by each transgene in *C. intestinalis* embryos at tailbud and larva stages.

restricted to the sensory vesicle, this suggests other limiting factors are also necessary to promote ectopic *Ci*-Rx expression in the CNS.

Discussion

In all vertebrate species in which Rx genes have been studied, the expression pattern is highly conserved and it has been demonstrated that they are involved, not only in the patterning of the nervous system but also in the development of the eye fields. In invertebrate species, although Rx genes are expressed in the nervous system, they do not share any expression in the eye structures (Arendt et al., 2004; Eggert et al., 1998; Mannini et al., 2008). Therefore, it seems that *Rx* gene function was likely restricted to the brain in the common ancestor of vertebrate and invertebrate, while during Deuterostome evolution Rx genes acquired a new role in the development of the eye structures.

In *C. intestinalis*, we previously demonstrated that the Rx gene has a key function in the development of photosensitive organ, since knockdown of this gene resulted in a larva that lacks the ocellus and the photoreceptor cells and has completely lost the ability to respond to the light stimuli variations (D'Aniello et al., 2006). These data clearly show that the new function of Rx gene in the development of the photosensitive structure has been acquired during chordate evolution before tunicate divergence. In this context, studying Rx gene in an experimental model such as *C. intestinalis* represents a great opportunity to clarify the genetic cascade of Rx genes and to shed light on the evolutionary mechanisms that led to the acquisition of a new function in eye development. In the attempt to answer these intriguing questions, we characterized the regulatory mechanism underlying *Ci*-Rx specific activation during *C. intestinalis* embryogenesis.

Previously, we identified a 2.9 kb genomic region located upstream of the *Ci*-Rx coding sequence responsible for *Ci*-Rx tissue specific activation (D'Aniello et al., 2006). In the present study, we build on this initial observation and perform analysis of this promoter. First, we performed a detailed deletion analysis of the 2.9 kb sequence and tested the ability of these fragments to reproduce *Ci*-Rx expression profile. Computational analysis of the Rx gene locus between two ascidian species, *C. intestinalis* and *C. savignyi*, revealed the presence of three conserved non-coding sequences. Among them, only one (J0.2) was functionally active and able to activate reporter gene expression in the *Ci*-Rx territories consisting of three groups of cells of the developing brain and at the larva stage in the cells surrounding the ocellus (Figs. 3G, H). Interestingly, this region of the promoter was the most conserved and the only one showing higher than 75% conservation. Recently, it has been shown that Nodal controls *Ci*-Rx negative regulation on the left side of the embryo and gives rise to the proper right sided localization of the ocellus and photoreceptor cells (Yoshida and Saiga, 2011). Unfortunately, *Ciona* rapid development and LacZ stability do not permit to observe the switch on and off of the enhancer at the different stages. This hitch does not allow to establish if the Rx enhancers that we have identified are involved in the response to signaling.

Subsequently, to better narrow the *Ci*-Rx enhancer(s), this J0.2 positive fragment has been subdivided and allowed the positive identification of two fragments, called J-A and J-F, able to direct the expression of the reporter gene in the same territories of the endogenous transcript (Figs. 6A–D). Computational search for binding sites revealed that both sequences contain binding sites for the Onecut TF, also called HNF6 in vertebrates. Interestingly, the expression profile of this transcription factor in *C. intestinalis* (Moret et al., 2005) and in *H. roretzi* (Sasakura and Makabe, 2001) appeared similar to that of *Ci*-Rx in the CNS at tailbud stage and in the sensory vesicle at larva stage.

In literature the role of these genes in the development of the nervous system is reported. In *D. melanogaster* a direct role of Onecut

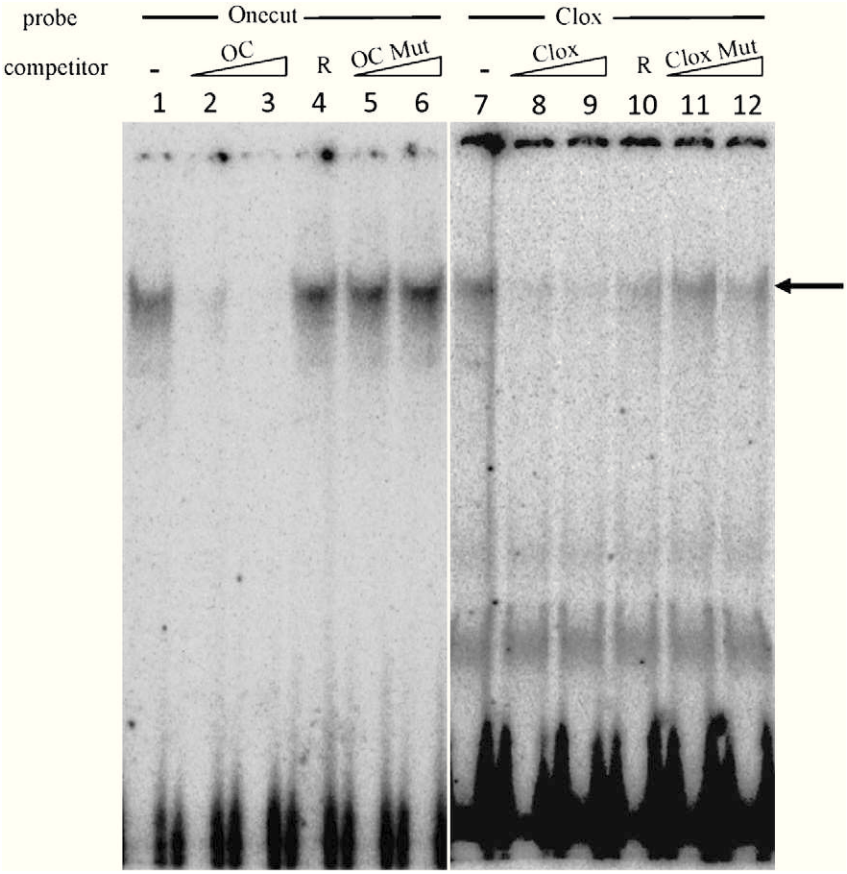


Fig. 9. Gel-shift analysis with *Ci*-Onecut protein and wild-type and mutated oligonucleotides. One shifted band is observed in lanes 1 and 7 where the *in vitro* translated protein was incubated with the OC (Onecut wt, Table S3) and Clox (Clox wt, Table S3) labeled oligonucleotides, respectively. The specific band disappears in the presence of the unlabeled oligonucleotides (lanes 2–3, 8–9), while is not affected by the random one (lanes 4, 10) or by the mutated ones (lanes 5–6, 11–12). Arrow indicates specific retarded band. Increasing amount of cold probe and mutated oligo are represented by open triangles. OC Mut and Clox Mut indicate Onecut MUT and Clox MUT oligonucleotides (Table S3).

as a transcriptional activator in the central and peripheral nervous system and, in particular, in the formation of photoreceptors has been demonstrated (Nguyen et al., 2000). In zebrafish and in mammals these genes are expressed in various parts of the nervous system and in particular in the retina and pineal gland where Rx genes are also expressed (Hong et al., 2002; Landry et al., 1997). We verified by double *in situ* hybridization that Onecut and Rx genes are expressed in the same territories in the anterior developing brain of *C. intestinalis* tailbud embryos (Fig. 7). We then used *in vivo* and *in vitro* approaches to demonstrate that *Ci*-Onecut binds and activates the two predicted *Ci*-Rx enhancer elements. First, we demonstrated *in vivo* that once J-F and J-A are mutated in the Onecut binding site, the expression of the reporter gene is dramatically reduced (Fig. 8). Analysis of these two sequences by EMSA assay revealed that, they are able to form a specific retarded complex with the *Ci*-Onecut protein *in vitro* synthesized (Fig. 9). Furthermore, to demonstrate that Onecut is not only necessary but also sufficient to activate *Ci*-Rx enhancers we overexpressed *Ci*-Onecut in the notochord cells (Fig. 10). This experiment highlighted that when *Ci*-Onecut is translated in an ectopic tissue (the notochord) it is able to recognize the binding sites present in the J-A, J-F and OC2 sequences and induce LacZ expression in the notochord cells. Furthermore, mutations of the Onecut sites present in these sequences abolish ectopic activation of the J-A and J-F Rx enhancer in the notochord cells (Figs. 10D, F). Finally, we provided evidence that Onecut transcriptional activation is necessary and sufficient to activate *Ci*-Rx endogenous expression in the sensory vesicle (Fig. 11).

Similar to Rx genes, the expression pattern and function of Onecut is highly conserved in analyzed invertebrates and vertebrates. *Ci*-Onecut (presented here and Moret et al., 2005) appears conserved within tunicates as *H. roretzi* (Sasakura and Makabe, 2001). In *D. melanogaster*, the Onecut homolog has been demonstrated to have a direct role in the central and peripheral nervous system and it has been described to have a role in the formation of photoreceptors (Nguyen et al., 2000). In zebrafish and mammals, Onecut genes are expressed in the nervous system, including the retina and pineal gland where Rx genes are also expressed (Hong et al., 2002; Landry et al., 1997). Furthermore, by morpholino experiments, it has been shown that *Ci*-Onecut controls Chox10 and Irx genes (Imai et al., 2009). These genes seem to be implicated in retina and photoreceptors development in zebrafish and mouse embryos (Katoh et al., 2010; Leung et al., 2008). These results suggest that Onecut could have a conserved genetic pathway involved in the formation of photosensitive structures.

It is interesting that although it has been widely reported in multiple species that Onecut expression and function is in the same territories of Rx, such as the nervous system, pineal gland and retina, it had never been implicated in being a direct regulator of the Rx genes. To date the only detailed study on Rx regulatory elements was performed in *Xenopus*. Although conserved binding sites for *Otx2* and *Sox2* have been identified in *Xenopus* Rx enhancer elements (Danno et al., 2008), we did not find Otx2 or Sox2 binding sites in the J0.2 sequence or any of the consensus sequences. *C. intestinalis* diverged from modern vertebrates more than 500 Ma ago. However, the function of Rx gene seems to be conserved in the development of photosensitive structures (D'Aniello et al., 2006). Our data suggest that Onecut directly regulates Rx enhancer elements. Therefore, it will be interesting to investigate if there are conserved Onecut enhancers that also regulate Rx gene expression in vertebrates. However, it is also possible that the Onecut gene was an

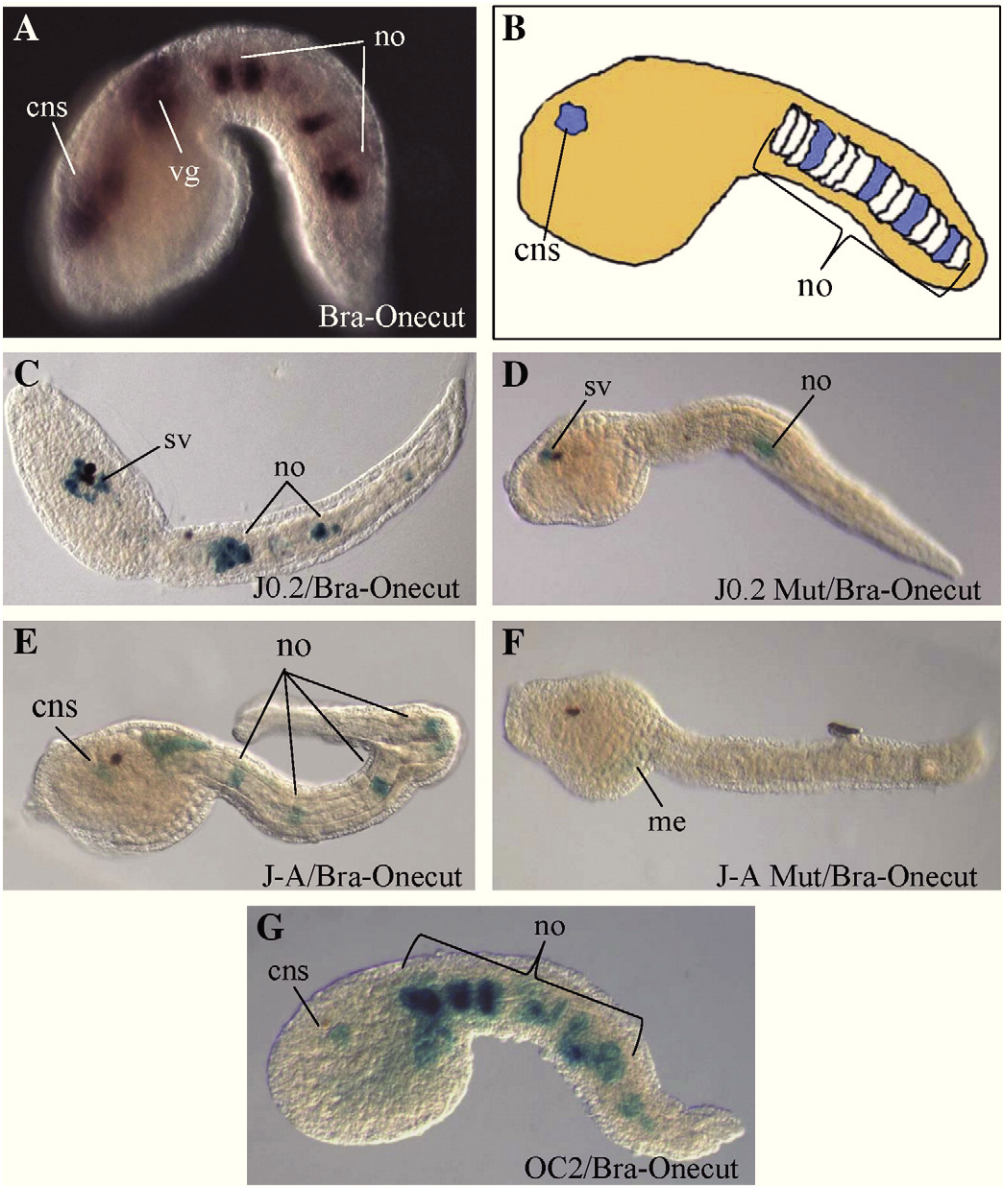


Fig. 10. (A) Spatial expression of *Ci*-Onecut in embryos electroporated with the Bra-Onecut construct at tailbud stage. The expression is visible not only in the *Ci*-Onecut endogenous territories but also in the notochord cells. (B) Predicted result of co-electroporation of the Bra-Onecut construct together with the constructs containing *Ci*-Rx non-coding sequence upstream of the LacZ reporter gene. (C, D) Larva embryos co-electroporated with Bra-Onecut together with J0.2 construct (C) and J0.2 Mut construct (D), respectively. (E, F) Larva embryos co-electroporated with Bra-Onecut together with J-A construct (E) and J-A Mut construct (F), respectively. Expression of the reporter gene has been detected not only in the *Ci*-Rx endogenous tissue in the anterior brain, but also in the notochord cells when the wild type constructs are used, while no signals are detected with the mutated constructs. (G) Embryo at late tailbud stage co-electroporated with Bra-Onecut construct together with the OC2 construct. LacZ expression is visible in the central nervous system and in the notochord. All images are lateral view with anterior to the left. cns, central nervous system; no, notochord; sv, sensory vesicle.

ancestral regulator of Rx expression, but that due to the accumulation of mutational events vertebrate Rx genes acquired new regulatory mechanisms and the dependence on Onecut has been lost.

While considering the origin of Onecut and Rx regulation in the sensory vesicle, it is also important to understand the possible homology of the ascidian ocellus, to photosensitive structures in vertebrates (D'Aniello et al., 2006; Horie et al., 2008; Sakurai et al., 2004; Tsuda et al., 2003). Until recently, the prevailing hypothesis was that the ocellus was ancestral to the vertebrate eye. However, in recent years, some authors have proposed that the ascidian ocellus could be homologous to the vertebrate median eye, also called epiphysis or pineal gland. The photoreceptive function of the pineal gland is less conserved and is still present only in some non-mammal vertebrate species. In support of this hypothesis, similar to the ascidian ocellus, the epiphysis of amphibian and teleost larvae triggers the shadow response, when the developing lateral eyes are still not competent to respond to light stimuli (Foster and Roberts, 1982). In addition, the ascidian ocellus and vertebrate epiphysis are both derived from cells located in the lateral part of the embryonic neural plate (Eagleson and Harris, 1990; Nishida, 1987). Together, these similarities between the ocellus and vertebrate epiphysis suggested a possible homology between these two structures.

Despite the similarities of the ocellus and the epiphysis, other characteristics do not permit us to distinguish whether or not the ocellus is truly homologous to the eyes or epiphysis. For instance, as in vertebrates, ascidian photoreceptors are ciliary in origin (Eakin and Kuda, 1971), hyperpolarize to light (Gorman et al., 1971), and express Opsin1. Furthermore, the sequence of *Ci*-Opsin1 is highly homologous to both vertebrate retinal and pineal opsins (Kusakabe et al., 2001). With respect to Rx genes, vertebrate Rx genes are expressed in both the pineal organ and the retina (Hong et al., 2002; Mathers and Jamrich, 2000) and

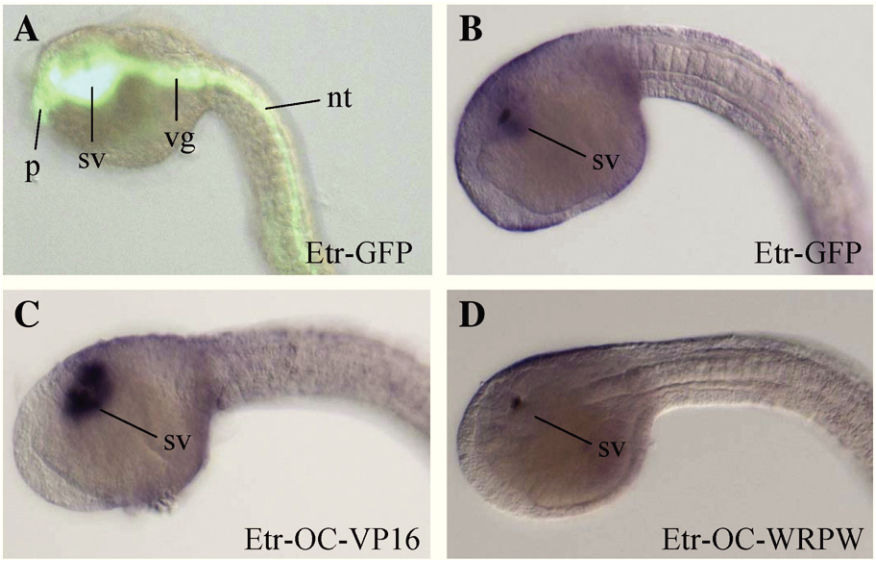


Fig. 11. (A) Merged bright-field/fluorescent image of GFP expression driven by the Etr promoter at tailbud stage. Transgene expression occurs throughout the nervous system. (B–D) *Ci*-Rx expression pattern visualized by whole-mount *in-situ* hybridization on embryos electroporated with Etr-GFP (B), Etr-OC-VP16 (C) and Etr-OC-WRPW (D) constructs. (C) The Etr-OC-VP16 construct promotes *Ci*-Rx expression in the sensory vesicle, while (D) the Etr-OC-WRPW construct inhibits *Ci*-Rx expression. nt, neural tube; p, palps; sv, sensory vesicle; vg, visceral ganglion.

play a specific role in the formation of projection neurons of zebrafish pineal gland (Cau and Wilson, 2003). Taking into account a parsimonious theory of evolution, a possible scenario could be that the ocellus represents the ancestral structure of both light sensing organs, which diversified into the vertebrate pinealocyte and retinal photoreceptors (Klein, 2006). In this context, it would explain why Onecut and Rx genes are both expressed in the pineal gland and the retina of many vertebrate species. Further studies in *Ciona* as well as in various vertebrate species will help to elucidate the evolutionary mechanisms that led to the formation of the ascidian ocellus and vertebrate pineal gland and eye.

Supplementary materials related to this article can be found online at doi:10.1016/j.ydbio.2011.05.584.

Acknowledgments

We are deeply grateful to Joshua Waxman (Cincinnati Children's Hospital Medical Center, Ohio, USA) for a critical reading of the manuscript and for the final English revision of the paper. Special thanks to Nunzia Pastore, Laura Gaveglia and Claudia Racioppi for their help in preparing constructs and performing electroporation during their undergraduate thesis. We thank Alessandro Amoroso for technical assistanceand the SZN Molecular Biology Service members for sequencing and technical assistance. We acknowledge Giovanna Benvenuto of the SZN Confocal Microscopy Service for confocal microscopy assistance.

References

Alfano, C., Russo, M.T., Spagnuolo, A., 2007. Developmental expression and transcriptional regulation of *Ci-Pans*, a novel neural marker gene of the ascidian, *Ciona intestinalis*. Gene 406, 36–41.

Andreazzoli, M., Gestri, G., Angeloni, D., Menna, E., Barsacchi, G., 1999. Role of *Xrx1* in *Xenopus* eye and anterior brain development. Development 126, 2451–2460.

Andreazzoli, M., Gestri, G., Cremisi, F., Casarosa, S., Dawid, I.B., Barsacchi, G., 2003. *Xrx1* controls proliferation and neurogenesis in *Xenopus* anterior neural plate. Development 130, 5143–5154.

Arendt, D., Tessmar-Raible, K., Snyman, H., Dorresteijn, A.W., Wittbrodt, J., 2004. Ciliary photoreceptors with a vertebrate-type opsin in an invertebrate brain. Science 306, 869–871.

Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M., Werner, T., 2005. MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics 21, 2933–2942.

Cau, E., Wilson, S.W., 2003. *Ash1a* and *Neurogenin1* function downstream of Floating head to regulate epiphysial neurogenesis. Development 130, 2455–2466.

Chuang, J.C., Raymond, P.A., 2001. Zebrafish genes *rx1* and *rx2* help define the region of forebrain that gives rise to retina. Dev. Biol. 231, 13–30.

Corbo, J.C., Levine, M., Zeller, R.W., 1997. Characterization of a notochord-specific enhancer from the *Brachyury* promoter region of the ascidian, *Ciona intestinalis*. Development 124, 589–602.

D'Aniello, S., D'Aniello, E., Locascio, A., Memoli, A., Corrado, M., Russo, M.T., Aniello, F., Fucci, L., Brown, E.R., Branno, M., 2006. The ascidian homolog of the vertebrate homeobox gene *Rx* is essential for ocellus development and function. Differentiation 74, 222–234.

Danno, H., Michiue, T., Hitachi, K., Yukita, A., Ishiura, S., Asashima, M., 2008. Molecular links among the causative genes for ocular malformation: *Otx2* and *Sox2* coregulate *Rax* expression. Proc. Natl. Acad. Sci. U. S. A. 105, 5408–5413.

Di Gregorio, A., Levine, M., 2002. Analyzing gene regulation in ascidian embryos: new tools for new perspectives. Differentiation 70, 132–139.

Dufour, H.D., Chettouh, Z., Deyts, C., de Rosa, R., Goridis, C., Joly, J.S., Brunet, J.F., 2006. Precraniate origin of cranial motoneurons. Proc. Natl. Acad. Sci. U. S. A. 103, 8727–8732.

Eagleson, G.W., Harris, W.A., 1990. Mapping of the presumptive brain regions in the neural plate of *Xenopus laevis*. J. Neurobiol. 21, 427–440.

Eakin, R.M., Kuda, A., 1971. Ultrastructure of sensory receptors in Ascidian tadpoles. Z. Zellforsch. 112, 287–312.

Eggert, T., Hauck, B., Hildebrandt, N., Gehring, W.J., Walldorf, U., 1998. Isolation of a *Drosophila* homolog of the vertebrate homeobox gene *Rx* and its possible role in brain and eye development. Proc. Natl. Acad. Sci. U. S. A. 95, 2343–2348.

Foster, R.G., Roberts, A., 1982. The pineal eye in *Xenopus laevis* embryos and larvae: a photoreceptor with a direct excitatory effect on the behaviour. J. Comp. Physiol. 145, 413–419.

Gorman, A.L., McReynolds, J.S., Barnes, S.N., 1971. Photoreceptors in primitive chordates: fine structure, hyperpolarizing receptor potentials, and evolution. Science 172, 1052–1054.

Hong, S.K., Kim, C.H., Yoo, K.W., Kim, H.S., Kudoh, T., Dawid, I.B., Huh, T.L., 2002. Isolation and expression of a novel neuron-specific *onecut* homeobox gene in zebrafish. Mech. Dev. 112, 199–202.

Horie, T., Orii, H., Nakagawa, M., 2005. Structure of ocellus photoreceptors in the ascidian *Ciona intestinalis* larva as revealed by an anti-arrestin antibody. J. Neurobiol. 65, 241–250.

Horie, T., Sakurai, D., Ohtsuki, H., Terakita, A., Shichida, Y., Usukura, J., Kusakabe, T., Tsuda, M., 2008. Pigmented and nonpigmented ocelli in the brain vesicle of the ascidian larva. J. Comp. Neurol. 509, 88–102.

Imai, K.S., Stolfi, A., Levine, M., Satou, Y., 2009. Gene regulatory networks underlying the compartmentalization of the *Ciona* central nervous system. Development 136, 285–293.

Katoh, K., Omori, Y., Onishi, A., Sato, S., Kondo, M., Furukawa, T., 2010. Blimp1 suppresses Chx10 expression in differentiating retinal photoreceptor precursors to ensure proper photoreceptor development. J. Neurosci. 30, 6515–6526.

Kimura, A., Singh, D., Wawrousek, E.F., Kikuchi, M., Nakamura, M., Shinohara, T., 2000. Both PCE-1/RX and OTX/CRX interactions are necessary for photoreceptor-specific gene expression. J. Biol. Chem. 275, 1152–1160.

Klein, D.C., 2006. Evolution of the vertebrate pineal gland: the AANAT hypothesis. Chronobiol. Int. 23, 5–20.

Kusakabe, T., Kusakabe, R., Kawakami, I., Satou, Y., Satoh, N., Tsuda, M., 2001. *Ci-opsin*1, a vertebrate-type opsin gene, expressed in the larval ocellus of the ascidian *Ciona intestinalis*. FEBS Lett. 506, 69–72.

Landry, C., Clotman, F., Hioki, T., Oda, H., Picard, J.J., Lemaigre, F.P., Rousseau, G.G., 1997. *HNF-6* is expressed in endoderm derivatives and nervous system of the mouse embryo and participates to the cross-regulatory network of liver-enriched transcription factors. Dev. Biol. 192, 247–257.

Lemaire, P., Bertrand, V., Hudson, C., 2002. Early steps in the formation of neural tissue in ascidian embryos. Dev. Biol. 252, 151–169.

Leung, Y.F., Ma, P., Link, B.A., Dowling, J.E., 2008. Factorial microarray analysis of zebrafish retinal development. Proc. Natl. Acad. Sci. U. S. A. 105, 12909–12914.

Locascio, A., Aniello, F., Amoroso, A., Manzanares, M., Krumlauf, R., Branno, M., 1999. Patterning the ascidian nervous system: structure, expression and transgenic analysis of the *CiHox3* gene. Development 126, 4737–4748.

Loosli, F., Staub, W., Finger-Baier, K.C., Ober, E.A., Verkade, H., Wittbrodt, J., Baier, H., 2003. Loss of eyes in zebrafish caused by mutation of *chokh*/*rx3*. EMBO Rep. 4, 894–899.

Mannini, L., Deri, P., Picchi, J., Batistoni, R., 2008. Expression of a retinal homeobox (*Rx*) gene during planarian regeneration. Int. J. Dev. Biol. 52, 1113–1117.

Mathers, P.H., Jamrich, M., 2000. Regulation of eye formation by the Rx and pax6 homeobox genes. Cell. Mol. Life Sci. 57, 186–194.

Mathers, P.H., Grinberg, A., Mahon, K.A., Jamrich, M., 1997. The *Rx* homeobox gene is essential for vertebrate eye development. Nature 387, 603–607.

Mayor, C., Brudno, M., Schwartz, J.R., Poliakov, A., Rubin, E.M., Frazer, K.A., Pachter, L.S., Dubchak, I., 2000. VISTA: visualizing global DNA sequence alignments of arbitrary length. Bioinformatics 16, 1046–1047.

Meinertzhagen, I.A., Okamura, Y., 2001. The larval ascidian nervous system: the chordate brain from its small beginnings. Trends Neurosci. 24, 401–410.

Moret, F., Christiaen, L., Deytes, C., Blin, M., Joly, J.S., Vernier, P., 2005. The dopamine-synthesizing cells in the swimming larva of the tunicate *Ciona intestinalis* are located only in the hypothalamus-related domain of the sensory vesicle. Eur. J. Neurosci. 21, 3043–3055.

Nguyen, D.N., Rohrbaugh, M., Lai, Z., 2000. The *Drosophila* homolog of *Onecut* homeodomain proteins is a neural-specific transcriptional activator with a potential role in regulating neural differentiation. Mech. Dev. 97, 57–72.

Nishida, H., 1987. Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. III. Up to the tissue restricted stage. Dev. Biol. 121, 526–541.

Passamaneck, Y.J., Di Gregorio, A., 2005. *Ciona intestinalis*: chordate development made simple. Dev. Dyn. 233, 1–19.

Ristoratore, F., Spagnuolo, A., Aniello, F., Branno, M., Fabbrini, F., Di Lauro, R., 1999. Expression and functional analysis of Cititf1, an ascidian NK-2 class gene, suggest its role in endoderm development. Development 126, 5149–5159.

Sakurai, D., Goda, M., Kohmura, Y., Horie, T., Iwamoto, H., Ohtsuki, H., Tsuda, M., 2004. The role of pigment cells in the brain of ascidian larva. J. Comp. Neurol. 475, 70–82.

Sasakura, Y., Makabe, K.W., 2001. A gene encoding a new ONECUT class homeodomain protein in the ascidian *Halocynthia roretzi* functions in the differentiation and specification of neural cells in ascidian embryogenesis. Mech. Dev. 104, 37–48.

Satoh, N., Satou, Y., Davidson, B., Levine, M., 2003. *Ciona intestinalis*: an emerging model for whole-genome analyses. Trends Genet. 19, 376–381.

Tabata, Y., Ouchi, Y., Kamiya, H., Manabe, T., Arai, K., Watanabe, S., 2004. Specification of the retinal fate of mouse embryonic stem cells by ectopic expression of *Rx*/*rax*, a homeobox gene. Mol. Cell. Biol. 24, 4513–4521.

Tsuda, M., Kawakami, I., Shiraishi, S., 2003. Sensitization and habituation of the swimming behavior in ascidian larvae to light. Zoolog. Sci. 20, 13–22.

Voronina, V.A., Kozhemyakina, E.A., O'Kernick, C.M., Kahn, N.D., Wenger, S.L., Linberg, J. V., Schneider, A.S., Mathers, P.H., 2004. Mutations in the human *RAX* homeobox gene in a patient with anophthalmia and sclerocornea. Hum. Mol. Genet. 13, 315–322.

Yamada, L., Shoguchi, E., Wada, S., Kobayashi, K., Mochizuki, Y., Satou, Y., Satoh, N., 2003. Morpholino-based gene knockdown screen of novel genes with developmental function in *Ciona intestinalis*. Development 130, 6485–6495.

Yoshida, K., Saiga, H., 2011. Repression of Rx gene on the left side of the sensory vesicle by Nodal signalling is crucial for right-sided formation of the ocellus photoreceptor in the development of *Ciona intestinalis*. Dev. Biol. 354, 144–150.

Zhang, L., Mathers, P.H., Jamrich, M., 2000. Function of *Rx*, but not *Pax6*, is essential for the formation of retinal progenitor cells in mice. Genesis 28, 135–142.

Zilinski, C., Brownell, I., Hashimoto, R., Medina-Martinez, O., Swindell, E.C., Jamrich, M., 2004. Expression of *FoxE4* and *Rx* visualizes the timing and dynamics of critical processes taking place during initial stages of vertebrate eye development. Dev. Neurosci. 26, 294–307.

OPEN ACCESS Freely available online

PLoS one

New Insights into the Evolution of Metazoan Tyrosinase Gene Family

Rosaria Esposito[◕], Salvatore D'Aniello[◕], Paola Squarzoni[¤], Maria Rosa Pezzotti, Filomena Ristoratore, Antonietta Spagnuolo*

Cellular and Developmental Biology Department, Stazione Zoologica Anton Dohrn, Villa Comunale, Napoli, Italy

Abstract

Tyrosinases, widely distributed among animals, plants and fungi, are involved in the biosynthesis of melanin, a pigment that has been exploited, in the course of evolution, to serve different functions. We conducted a deep evolutionary analysis of tyrosinase family amongst metazoa, thanks to the availability of new sequenced genomes, assessing that tyrosinases (tyr) represent a distinctive feature of all the organisms included in our study and, interestingly, they show an independent expansion in most of the analyzed phyla. Tyrosinase-related proteins (tyrp), which derive from tyr but show distinct key residues in the catalytic domain, constitute an invention of chordate lineage. In addition we here reported a detailed study of the expression territories of the ascidian *Ciona intestinalis tyr* and *tyrps*. Furthermore, we put efforts in the identification of the regulatory sequences responsible for their expression in pigment cell lineage. Collectively, the results reported here enlarge our knowledge about the *tyrosinase* gene family as valuable resource for understanding the genetic components involved in pigment cells evolution and development.

Citation: Esposito R, D'Aniello S, Squarzoni P, Pezzotti MR, Ristoratore F, et al. (2012) New Insights into the Evolution of Metazoan Tyrosinase Gene Family. PLoS ONE 7(4): e35731. doi:10.1371/journal.pone.0035731

Editor: Hector Escriva, Laboratoire Arago, France

Received July 18, 2011; **Accepted** March 24, 2012; **Published** April 20, 2012

Funding: No current external funding sources for this study.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: nietta.spagnuolo@szn.it

◕ These authors contributed equally to this work.

¤ Current address: Ecole Normale Supérieure, Institut de Biologie de l'ENS, IBENS; INSERM, U1024, Avenir Team, Paris, France

Introduction

In vertebrates three types of melanin-producing pigment cells are known, that have distinct, even if related, embryonic origins: melanocytes of the inner ear, skin, hair-bulbs and uvea, which derive from the neural crest; retinal pigment epithelium (RPE) cells of the eye derived from the neural tube; and pigment cells of the pineal organ, which also arise from the neural tube [1,2,3]. All these cells share the capacity to produce melanins, a class of polymeric pigments whose biosynthesis is mainly governed by evolutionarily conserved enzymes of the tyrosinase family: tyrosinase (tyr), tyrosinase related protein-1 (tyrp1) and tyrosinase related protein-2 (tyrp2) also called DOPAchrome tautomerase (dct). Amongst them, tyr plays the initial and crucial role for melanin production, by converting the amino acid tyrosine to 3,4-dihydroxyphenylalanine (DOPA), while tyrp1 and tyrp2 function in subsequent steps, since they influence the quantity and the quality of the synthesized melanins [4,5]. Furthermore, both tyrps are known to stabilize the tyr enzyme [6,7,8] and to function in melanocyte survival and maintenance of melanosomal structures [9].

The genetic programs leading to the development of the three types of vertebrate pigment cells, although different, thus converge at a certain point to allow the expression of members of the *tyrosinase* family, in order to produce melanin pigments. It is noteworthy that many human genetic inheritable pathologies, as multiple forms of albinism, vitiligo and deafness, are linked to genetic mutations in one or more genes responsible for melanin biosynthesis [10]. These genes therefore represent a good paradigm to answer questions regarding the evolution, genetics, and developmental biology of pigment cells, as well as to approach human disorders associated with defects in their synthesis, regulation or function.

The three tyrosinase family proteins, besides showing extensive similarities at the amino acid level, share many key structural characteristics (see [8,11] for detailed reviews). The first one consists of the presence of two highly conserved metal binding domains, MeA and MeB, that are involved in the proper folding of the active site and in the binding of metal cofactors (copper for tyr, zinc for tyrp2 and unknown for tyrp1). Few differences exist, consisting in four amino acid substitutions, which might be responsible for the switch of affinity from phenolic substrates, typical of tyr enzymes, to indolic substrates, observed in tyrps. A further interesting common trait of tyr and tyrps is the presence of three cysteine clusters, two at the N-terminal and one located between MeA and MeB, likely involved in correct protein folding [8].

Data collected so far have suggested that the *tyr* and *tyrp* gene family has clearly evolved from a common ancestral *tyrosinase* gene [12,13] that was first duplicated before the divergence of urochordates (ascidians) and vertebrates [14], leading to *tyrosinase* (*tyr*) and a *tyrosinase-related protein* (*tyrp*). The *tyrp* was then duplicated early in vertebrate lineage, before the divergence of teleost fishes [15], giving rise to *tyrp1* and *tyrp2* (or *dct*).

However, a survey of the protochordate ascidian *Ciona intestinalis* genome revealed the presence of three *tyrosinase* family genes, one *tyr* (*Ci-tyr*) and two *tyrps* (*Ci-tyrp1/2a* and *Ci-tyrp1/2b*) [16], thus indicating that *tyr* family evolution might be much more complex than previously thought. As a model system for understanding chordate development, ascidians, such as *C. intestinalis*, offers important experimental advantages, compared to vertebrate species. They produce a large number of embryos, have external development, are small in size and have a fixed cell lineage. Furthermore, they have two pigmented sensory organs in the sensory vesicle: the otolith, composed of one pigmented cup cell, which functions in geotactic responses, and the ocellus, involved in photoreception, which is composed of three lens cells, 30 photoreceptor cells and one pigment cell [17]. It is noteworthy that the cell-lineage of the pigment cells has been fully documented [18]; furthermore in ascidians every blastomere of the embryo is distinguishable, so that it is easy to precisely identify cells expressing genes of interest, when gene expression is initiated and lineage in which gene expression is inherited [19]. This peculiarity coupled with the possibility, in *C. intestinalis*, to easily isolate the promoter regions of the gene of interest, by using electroporation of chimeric reporter genes [20], makes *Ciona* a model system ideal to identify marker genes, specific for each lineage, and study the genetic cascades in which they are involved.

In the present study, as a first approach, we have exploited the growing number of sequenced genomes, from different taxa, for a deeper evolutionary analysis in order to shed light on the origin of *tyrosinase* family genes. We have then devoted our attention to the *C. intestinalis tyrosinase* family members, by conducting a detailed characterization of the expression profiles of the two *Ci-tyrps*, in comparison with *Ci-tyr*. Furthermore, analyses of their transcriptional regulation led to the identification of regulatory regions responsible for their spatio-temporal expression during *Ciona* embryogenesis. These enhancers have been successfully used as tools to study the genetic circuits controlling pigment cell differentiation during *Ciona* embryogenesis [21]. These enhancers will be also instrumental to look for modules responsible for the expression patterns of *tyrosinase* family genes in *Ciona*.

Results

Tyrosinase family evolution

To study the evolutionary history of tyrosinase family we conducted a phylogenetic analysis by using deduced protein sequences from eumetazoan available genomes. Among bilaterians we included sequences from deuterostomes, as vertebrates, urochordates (*C. intestinalis* and *Ciona savignyi*) [22,23], cephalocordates (*Branchiostoma floridae*) [24], hemichordates (*Saccoglossus kowalevskii*), and from protostomes, as nematodes (*Caenorhabditis elegans*) [25] and molluscs (*Sepia officinalis, Loligo vulgaris, Pinctada fucata*). Among radiates, tyrosinases from cnidarian genomes (*Nematostella vectensis* and *Hydra magnipapillata*) [26] were also included, while no ctenophore's representatives were found. A tyrosinase-like sequence from sponges (*Suberites domuncula*), which are historically considered to be the earliest diverging metazoan phylum, was used as outgroup. We were unable to identify any putative sequence related to the tyrosinase family in available echinoderm, annelid and arthropod genomes. It is already known that arthropods use phenoloxidases, enzymes that belong, as tyrosinases, to the Type3 Copper protein family, for melanin biosynthesis [27], and there are evidences indicating that also annelids and echinoderms could exploit phenoloxidases, in place of tyrosinases, for this cellular process [28,29].

The topology of our phylogenetic reconstruction revealed the clustering of four distinct groups of proteins (Fig. 1A): 1. cnidarian and protostome tyrs (green box), 2. chordate "canonical" tyrs (pink box), 3. chordate tyrps (blue box) and 4. a group of tyrs, present in cephalochordates and hemichordates, that branched independently and that we called tyrs-like (orange box), given the lack of any functional information. In this phylogenetic reconstruction the cnidarian tyrs grouped with protostome tyrs and not at the base of bilaterian tyrs, as it could be expected from phylogenetic lineage relationships. Notably, tyr and tyr-like independent expansions were observed in most of the analyzed metazoan phyla (Fig. 1A), whose functional significance is still unknown, thus opening the evolutionary history of this gene family to new perspectives.

In order to gain insight into evolutionary phylogenesis of the expanded tyrosinases, we analyzed, when available, the chromosomal distribution of all the *tyrosinase* expanded genes in chordates (*B. floridae*), hemichordates (*S. kowalevskii*), nematodes (*C. elegans*) and cnidarians (*N. vectensis*). The data showed that only *S. kowalevskii* and *C. elegans* expanded *tyrosinases* are contained in pair on two scaffolds or chromosomes (Fig. S1) and this indicates tandem duplication events, but we cannot exclude that future chromosomal reconstructions in other genome models would give a similar layout.

The present survey assessed that tyrosinase-related proteins (tyrp) are present exclusively in chordates; however ascidian and cephalochordate tyrps showed no clear phylogenetic relationships with vertebrate tyrp1 and tyrp2. In an effort to gain more insights into the evolutionary history of these genes, we thus studied the tyrp synteny conservation in amphioxus, ascidian and human genomes and we mapped three independent gene duplications. In amphioxus *tyrp1/2a* and *tyrp1/2b* came from a tandem duplication event, since they lay close on scaffold 61, but we could not establish any synteny conservation with *Ciona* and human *tyrps*, possibly due to the short length of the scaffold (Fig. 2). On the other hand, we detected synteny conservation for *Ci-tyrp1/2a*, on chromosome 5, and human *TYRP1*, on chromosome 9, indicating that these genes are clearly orthologous. No shared genes around the locus of *Ci-tyrp1/2b*, on chromosome 8, and human *TYRP2*, on chromosome 13, were instead identified (Fig. 2), so we could not infer or exclude any orthology in this case.

A detailed analysis of conserved metal binding domains (MeA and MeB), based on previous work [8], was conducted on all tyrosinase family members included in our study. We confirmed that few key residues within the metal binding domains (MeA and MeB) are clearly archetypal of tyr or tyrp proteins [8]. These residues thus represent an important tool to easily distinguish between tyr and tyrps and allocate family memberships (Fig. 1B and Fig. S2). These residues were instrumental, in our analysis, to assign protostome and cnidarian sequences to the tyrosinase group in support of our phylogenetic tree (Fig. 1B and Fig. S2).

A further known characteristic of *tyrosinase* gene family is the presence of cysteine clusters that are probably responsible for correct protein folding. We detected an high degree of cysteine conservation, both at the N-terminal and between MeA and MeB, in the deuterostome proteins (Fig. S3). In the protostome lineage the cysteine clusters appeared conserved at the N-terminal, although with a lower number of cysteine residues; no cysteine cluster was detected between MeA and MeB domains whereas, interestingly, a specific cluster was present at the C-terminus in both nematodes and molluscs (Fig. S3).

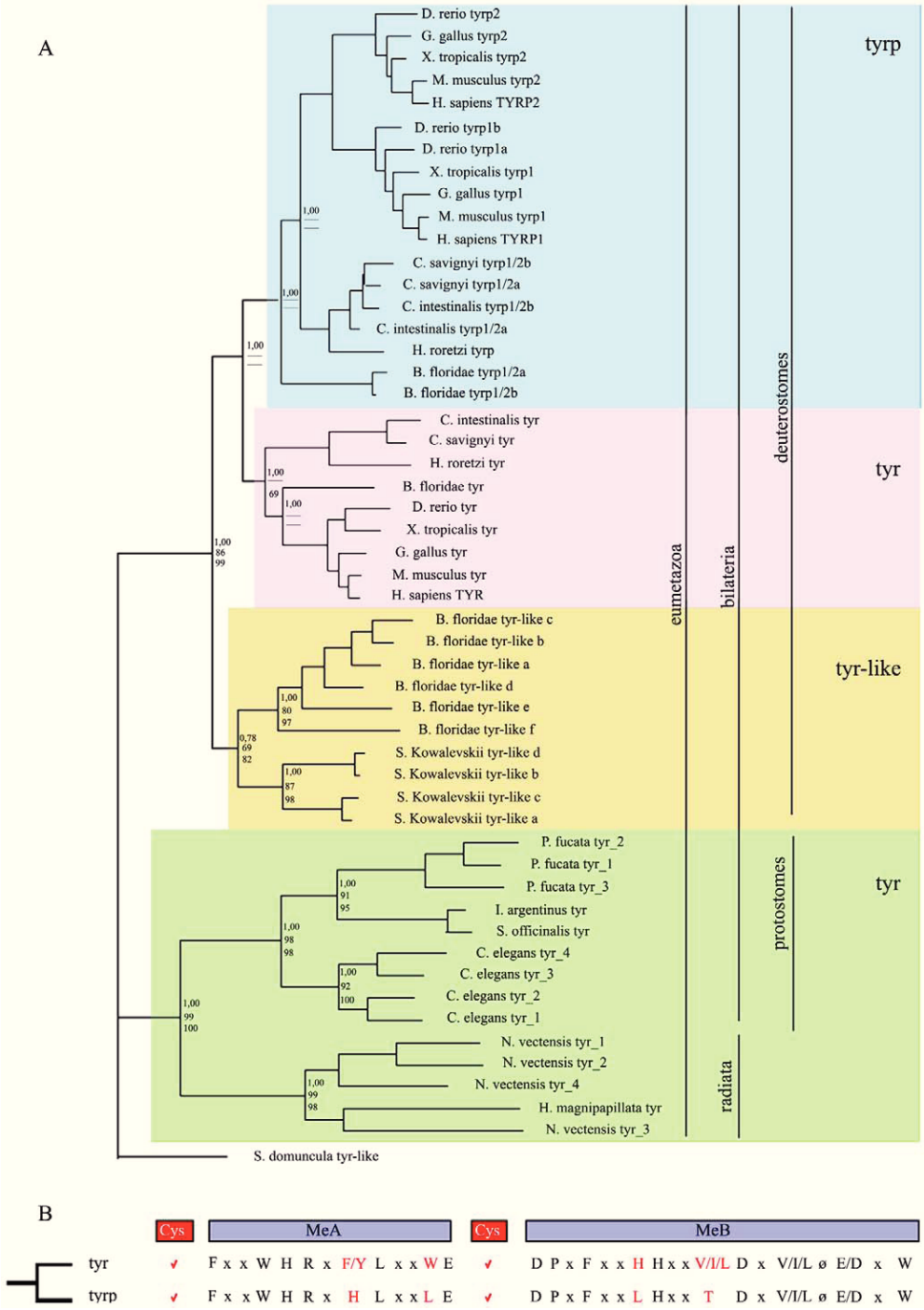


Figure 1. Evolution of tyr and tyrps in metazoa. A) Phylogenetic study of tyrosinase family proteins. Numbers at the branches indicate bootstrap values obtained with MrBayes, Maximum Likelihood and Neighbour Joining methods, respectively. Vertical bars highlight the classification of the analysed eumetazoa in radiata and bilateria, which are in turn subdivided in protostomes and deuterostomes. Colored boxes highlight four groups of proteins: cnidarian and protostome tyrosinases (green box), cephalochordate and hemichordate tyrs-like (orange box), chordate "canonical" tyrosinases (red box), and chordate tyrps (blue box). B) Conserved aminoacid residues in the two metal binding domains (MeA and MeB) of tyrosinase family members, derived from a multiple sequence alignment (see Fig. S2). Key aminoacid positions, probably involved in the change of affinity for phenolic substrates, are reported in red. ø indicates aromatic residues (F, Y or W), x indicates any aminoacid. The position of cysteine clusters (red boxes) refers to deuterostomes (see also Fig. S3).
doi:10.1371/journal.pone.0035731.g001

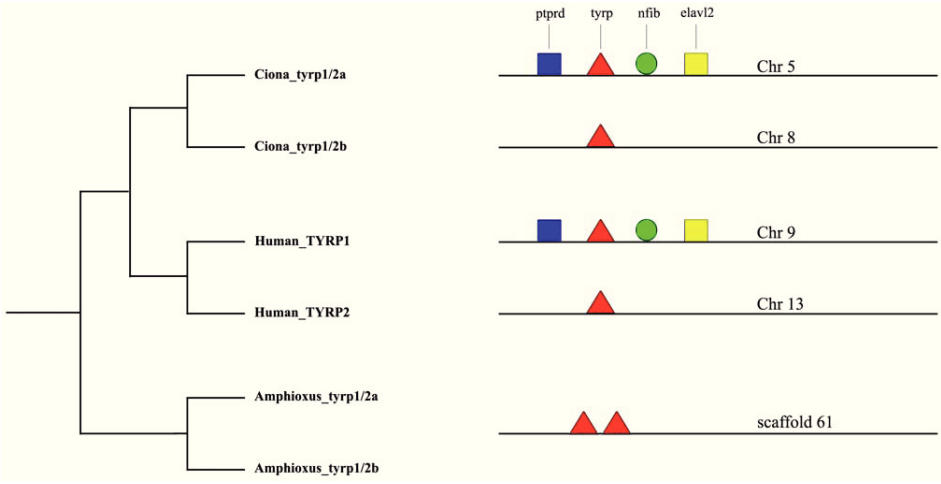


Figure 2. Tyrosinase-related proteins evolution in chordates. Schematic representation of synteny analysis of tyrosinase-related proteins in Ciona (*C. intestinalis*), amphioxus (*B. floridae*) and human (*H. sapiens*). Tyrps are represented by the red triangle. Our analysis revealed syntenic conservation shared by *Ci-tyrp1/2a* on chromosome 5 and *Hs-TYRP1* on chromosome 9. Neighbouring genes *ptprd* (blue square), *nflB* (green circle) and *elavl2* (yellow square) are indicated.
doi:10.1371/journal.pone.0035731.g002

Expression pattern of *tyrosinase* family genes in *Ciona intestinalis*

Expression patterns of *Ci-tyr*, *Ci-tyrp1/2a* and *Ci-tyrp1/2b* were examined through whole mount *in situ* hybridization experiments on *Ciona* embryos at different developmental stages. No signal was detected up to the late gastrula stage.

Ci-tyrp1/2a was the first to be expressed, from the late gastrula stage, in the a9.49 blastomere pair which corresponds to the pigment cell precursors (Fig. 3A). A clear and specific signal was then inherited in both a9.49 progeny (the a10.97 and a10.98 pairs) appearing much stronger in the posterior a10.97, compared to the anterior a10.98 pairs, at middle and late neurula stages (Figs. 3B and 3C). The expression persisted, with the same intensity up to the tailbud stage, in these four blastomeres that line up along the dorsal midline of the developing neural tube (Fig. 3D). The posterior a10.97 cells then differentiate into the otolith and ocellus pigment cells, where the *Ci-tyrp1/2a* mRNA remained localized at the larval stage (Fig. 3E).

Ci-tyr and *Ci-tyrp1/2b* expression territories were superimposable with that of *Ci-tyrp1/2a*. The only difference was that their hybridization signals were first detected at a slightly delayed developmental time, the middle (*Ci-tyrp1/2b*, Figs. 3F, 3G, 3H, 3I) and late (*Ci-tyr*, Figs. 3J, 3K, 3L) neurula stages, compared to *Ci-tyrp1/2a*. These results confirm previous data on *tyrosinase* expression in *C. intestinalis* [30] and strengthen the evidence that *tyrosinase* family members are specific markers of *C. intestinalis* pigment cell lineage from the late gastrula stage.

In vivo cis-regulatory regions analysis

To test the transcription driving activity of *Ci-tyr*, *Ci-tyrp1/2a* and *Ci-tyrp1/2b*, the 5′ genomic regions of each gene were isolated by PCR on *C. intestinalis* genomic DNA. Each upstream fragment was mapped, between the ATG of the transcript and the contiguous 5′ gene, and the length corresponded to 0.9 kb for *pCi-tyr*, and 1.5 kb for both *pCi-tyrp1/2a* and *pCi-tyrp1/2b*. These putative promoters were cloned upstream of a *mCherry* reporter gene (constructs *pCi-tyr>mChe pCi-tyrp1/2a>mChe*, and *pCi-tyrp1/2b>mChe*) and tested, by transgenesis *via* electroporation, for the capability to direct pigment cell lineage-specific expression of the reporter at the larval stage. The data indicated that the *pCi-tyr*, *pCi-tyrp1/2a* and *pCi-tyrp1/2b* all behave like specific enhancers, since the larvae showed a fluorescent signal in the otolith and/or ocellus pigment cells in a high proportion of the electroporated embryos (80–85% for *pCi-tyrp1/2a>mChe*, 60% for *pCi-tyr>mChe* and 50% for *pCi-tyrp1/2b>mChe* constructs) (Figs. 4D, 4G and 4J). The *pCi-tyrp1/2a* appeared the strongest, since many larvae showed a robust signal in the two pigment cells and, in a lower percentage, also in one or two accessory cells in the brain vesicle, that could represent the a10.97 sister cells, the a 10.98 pair, given the long half-life of mCherry protein (Fig. 4D). Furthermore, *pCi-tyrp1/2a* activity was the first to be detected as fluorescent protein product from early tailbud stage (data not shown), compared to the late tailbud (Fig. 4C) stage when mCherry protein signal, driven by *pCi-tyr* or *pCi-tyrp1/2b*, started to appear (Figs. 4F and 4I), confirming the timing of the *in situ* hybridization signal. To check for reporter expression at earlier developmental stages, whole mount *in situ* hybridization experiments, using *mCherry* antisense RNA, were performed on embryos electroporated with *pCi-tyr>mChe*, *pCi-tyrp1/2a>mChe*, and *pCi-tyrp1/2b>mChe* constructs, at late gastrula and neurula stages. The presence of *mCherry* mRNA in the territories where the endogenous genes are expressed (Figs. 4A, 4B, 4E and 4H compared with the Fig. 3) confirmed that the three promoters we have isolated contain the *cis*-regulatory information required for a correct spatial and temporal expression of the corresponding genes.

Comparisons between orthologous *C. intestinalis* and *C. savignyi* sequences have already indicated that these two species are at sufficient evolutionary distance to permit efficient identification of conserved regulatory sequence information [21,31,32]. Phylogenetic footprinting of the three corresponding promoters, *pCi-tyr*, *pCi-tyrp1/2a* and *pCi-tyrp1/2b*, pointed to the presence of a 400 bp highly conserved fragment only in the *pCi-tyrp1/2a* 5′ regulatory sequence. A 600 bp region, including this fragment and extending up to the ATG of *Ci-tyrp1/2a*, named *pCi-tyrp1/2a-0.6*, was cloned upstream of *mCherry* reporter and tested *in vivo* by transgenesis through electroporation. The results indicated that *pCi-tyrp1/2a-0.6* had similar activity to *pCi-tyrp1/2a* (Fig. S4), thus permitting an initial dissection of the regulatory region directing *Ci-tyrp1/2a* transcription.

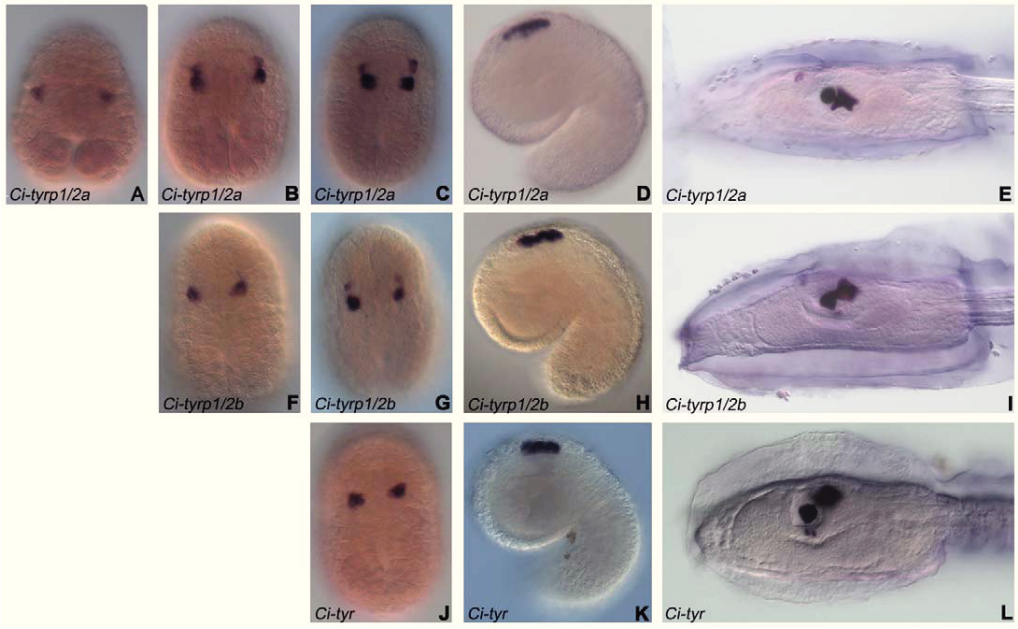


Figure 3. *Ci-tyr-tyrps* expression in pigment cell precursors. Expression pattern of *Ci-tyrp1*/2a (A–E), *Ci-tyrp1/2b* (F–I) and *Ci-tyr* (J–L) in *C. intestinalis* embryos at different developmental stages, detected through whole mount *in situ* hybridization experiments. The three genes are specifically expressed in a9.49 blastomeres (pigment cell precursors) and in their descendant cells from late gastrula (*Ci-tyrp1/2a*, in A), middle neurula (*Ci-tyrp1/2b*, in F) or late neurula (*Ci-tyr*, in J) stages up to the larval stage (E, I, L). A: late gastrula stage, vegetal view. B, F: middle neurula stage, dorsal view. C, G, J: late neurula stage, dorsal view. D, H, K: mid-tailbud stage, lateral view. E, I, L: larval stage, lateral view. A, B, C, F, G, J: anterior is up; D, E, H, I, K, L: anterior is on the left.
doi:10.1371/journal.pone.0035731.g003

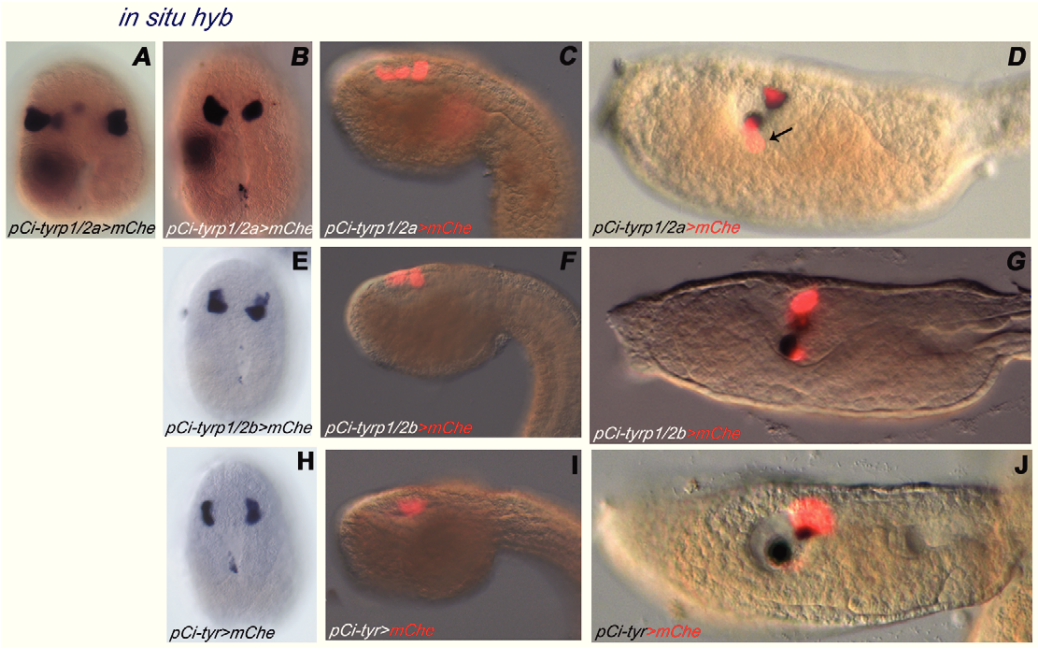


Figure 4. Identification of pigment cell *Ci-tyr* and *tyrps* enhancers. *In vivo* analysis of *pCi-tyrp1/2a>mChe* (A–D), *pCi-tyrp1/2b>mChe* (E–G) and *pCi-tyr>mChe* (H–J) constructs performed through transgenesis *via* electroporation experiments. Merged bright- field/fluorescent images of mCherry expression driven by the three enhancers at the tailbud (C,F,I) and larval stages (D,G,J). Lateral view, anterior is on the left. Note transgene expression in the pigment cell lineage. Arrow in D indicates the accessory cell in which a fluorescent signal is present. *In situ* hybridization experiments, using mCherry antisense RNA as probe, were performed in order to reveal the reporter expression at earlier developmental stages, when fluorescence is still undetectable (A: late gastrula stage; B, E, H: neurula stage, dorsal view). Reporter gene expression perfectly mirrors endogenous genes expression profiles (compare with Fig. 3).
doi:10.1371/journal.pone.0035731.g004

This *pCi-tyrp1/2a-0.6* fragment was then subjected to bioinformatic analyses, in comparison with the corresponding region of *C. savignyi*. This initial approach was mostly focused on the search for consensus motifs, conserved between *Ci-tyrp1/2a* and *Cs-tyrp1/2a* corresponding enhancers, for representatives of families already demonstrated, in vertebrates, to act as important players in pigmentation processes, such as Pax, Oct/Pou, Sox-TCF (HMG family), Mitf-TFE (bHLH-LZ family binding E-box motif) (for a comprehensive review see [33,34]). In this analysis we also included the promoter fragment *Hr-tyrp-333N*, previously identified in the ascidian *Halocynthia roretzi*, and demonstrated to be sufficient for *Hr-tyrp* expression in pigment cell precursors [35].

The software we used for this analysis (http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB = TF_8.3) identified different putative binding sites for Pax, Oct/Pou and Sox family genes. Interestingly, no canonical Mitf binding sites were identified in these regions. The study revealed also that Pax, Oct/Pou and Sox binding sites were organized in modules that are well conserved between *C. intestinalis* and *C. savignyi* and partially conserved also with *H. roretzi* (Fig. 5).

Discussion

Melanins are formed *in vitro* from L-tyrosine in the presence of tyrosinase alone; this led to the deduction that melanogenesis is a simple process requiring a single enzyme, the tyrosinase. This is the case in bacteria, sponges and plants; in the course of animal evolution, however, the situation has become more and more complex up to the mammals where the process is very sophisticated and has to be tightly regulated in terms of amount, type of melanin produced and the environment in which the synthesis takes place. Thus we move from a simple to a complex multi-enzymatic process (as in mammalian melanocytes), where new family members, the tyrps, have been added to finely tune the whole pathway. Sequence comparison of tyr and tyrps reveals that these proteins share many key structural features, indicating their common origin from an ancestral *tyrosinase* gene able to catalyse the critical rate-limiting hydroxylation of L-tyrosine to L-DOPA [11].

Melanogenic toolkit in animal evolution

Our phylogenetic analysis led to the conclusion that the tyrosinase family is divided in four distinct branches: 1. cnidarian and protostome tyrs (green box), 2. cephalochordate and hemichordate tyrs-like (orange box), 3. chordate "canonical" tyrs (pink box) and 4. chordate tyrps (blue box) (Fig. 1A).

The most parsimonious evolutionary scenario, based on the available data of sequenced genomes and gene predictions, as well as taking into account insights from enzymatic activity studies, is that one tyrosinase was present in the ancestor of eumetazoa. Indeed protostomes and radiates possess only one representative *tyr* gene that, in some cases, has been subjected to a lineage specific expansions (green box, Fig. 1A).

In the ancestor of deuterostomes then occurred a duplication that produced two *tyrosinase* genes. One of them was lost in tunicates and vertebrates and was retained only by cephalochordates and hemichordates. We named this group of genes as *tyr-like* (pink box), since no functional assays attesting their tyrosinase activity are yet available. The second *tyrosinase* gene was lost in hemichordates and was further amplified in the ancestor of chordates, giving rise to the canonical tyr (blue box) and tyrps (orange box).

An alternative evolutionary hypothesis is that two tyrosinase genes existed in the ancestor of metazoa. This scenario would imply that one representative (gene 1) has been retained in deuterostomes and lost in protostomes and radiates, while the second representative (gene 2) has been lost in deuterostomes and retained in protostomes and radiates. Thought we do not have extensive data to support either one of the two hypotheses, novel sequencing data, from animal taxa in key positions in the three of life, will help clarifying the phylogenetic history of this gene family.

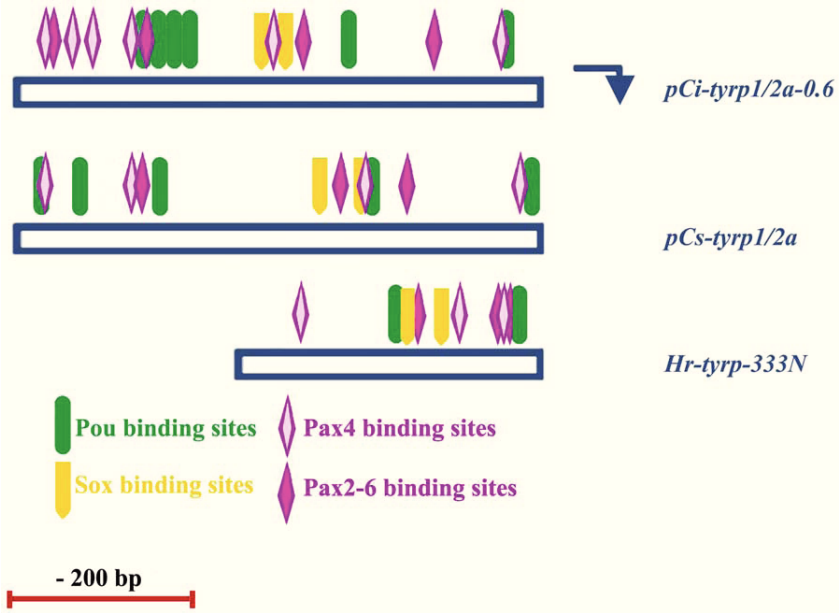


Figure 5. Putative transcription factors binding sites on *pCi-tyrp1/2a-0.6*, *pCs-tyrp1/2a* and *Hr-tyrp-333N* fragments. Bars indicate the analyzed *pCi-tyrp1/2a-0.6*, *pCs-tyrp1/2a* and *Hr-tyrp-333N* fragments. Transcription starting site is indicated by the blue arrow. Putative binding sites for Pou, Pax4, Pax 2–6 and Sox are taken into consideration and respectively represented by colored shapes, as indicated on the bottom.
doi:10.1371/journal.pone.0035731.g005

We were unable to identify any putative sequence related to the *tyrosinase* family in available arthropod, echinoderm, and annelid genomes. Previous studies have demonstrated that arthropods use the phenoloxidases for melanin byosinthesis, in place of tyrosinases [27] and that in these organisms the melanin, apart from providing pigmentation, is involved in other important processes, as wound healing, sclerotization and immunity [36,37]. It is intriguing to note that there are evidences indicating that annelids and echinoderms use phenoloxidases for both melanin byosinthesis and immune defence [28,29], thus supporting the absence of proteins phylogenetically related to tyrosinases in these organisms.

Our study put in light also a novel and intriguing example of independent family expansion in metazoa, with at least one species in all analyzed phyla (Fig. 1). In invertebrates, duplicates in the tyrosinase family are present in *N. vectensis* (cnidarians), in *P. fucata* (molluscs), in *C. elegans* (nematodes), in *S. kowalevskii* (hemichordates) and in *B. floridae* (cephalochordates). It is tempting to speculate that the multiple duplicates, present in these organisms, can play diverse functions, besides being involved in melanogenic processes, as phenoloxidases in arthropods that are used also for sclerotization and primary immune response [27,36,37]. One can suppose that melanin and its intermediates, in these species, are exploited as a defence mechanism to compensate the lack of a complex and specialized immune system. Notably, amongst all metazoa, amphioxus is the only organism that still retains representative genes of tyr, tyrp and tyr-like groups, firstly described in this paper. The multiple tyrosinases present in *B. floridae* can thus be framed in this perspective, since amphioxus has a simple humoral immune system with lymphocyte-like cells in gills [38], but without clearly identified free, circulating blood cells [39,40]. Conversely, the absence of tyr-like proteins in *Ciona* could be related to the presence, in this organism, of a well-developed vascular system, in which defined cell types, such as granular amoebocytes and granulocytes, are known to be involved in immunity reactions [39].

Our study confirmed also that the tyrps represent an invention of chordate lineage (Fig. 1A) where they contribute to a more complex melanogenic process. Concerning the topology of the phylogenetic tree for tyrps, we suppose that a rapid evolution, during their appearance in chordates, make it difficult to establish their exact phylogenetic relationship. It seems clear that the duplication events, leading to two copies of tyrps, happened independently in *C. intestinalis*, amphioxus and vertebrates, and that the two copies of vertebrate tyrps derive from a whole genomic duplication event that took place at the base of this lineage.

Taking all this into consideration, we then studied synteny conservation, surrounding the *tyrps* genes in *Ciona* and humans, in order to establish their orthology (Fig. 2). We excluded amphioxus from this analysis because of the short length of the scaffolds. The data indicated a clear orthology between *Ci-tyrp1/2a* and human *TYRP1*, while the absence of shared genes around *Ci-tyrp1/2b* and human *TYRP2* (Fig. 2) did not permit to infer or exclude any orthology.

From the enzymatic point of view, vertebrate tyrp2 is a DOPAchrome tautomerase, responsible for the conversion of L-DOPAchrome into DHICA [41], while the enzymatic role played by tyrp1 is still controversial, since tyrp1 has been attributed either tyrosinase, as well DHICA oxidase, or DOPAchrome tautomerase function but with a low specific activity [42,43,44,45]. Moreover biochemical data indicate that the two vertebrate tyrps, besides acting as enzyme, play also important functions in modulating tyrosinase activity, in the assembly of the melanogenic apparatus and in the detoxification processes taking place within melanosomes [7,9,46,47,48]. The invention of tyrps could then be related to the need of accessory enzymatic activities to finely tune the melanogenic process. In addition the tyrps could function as structural elements by providing a robust scaffold able to shield tyrosinase from potentially toxic intermediate products of the melanogenic pathway and permit the synthesis of melanin.

Expression profile and transcriptional regulation of *tyr* and *tyrps* in *Ciona*

Based on the expression profiles, in *Ciona*, *Ci-tyrp1/2a* appears first, at the late gastrula stage, and is copiously expressed up to the tadpole stage in the pigment cell lineage. This could be in agreement with a need to accumulate a large amount of tyrp protein in order to exert its protective and scaffolding function toward tyrosinase in the melanogenic complex. A further interesting feature, that ascidians share with vertebrates, is that *tyrosinase* mRNA (Fig. 3) (and, in *Ciona*, also the corresponding protein product [49]) appears well before pigment synthesis begins [49,50]. This could be related to the need of a large quantity of enzyme requested at the time of initial melanin synthesis, in order to catalyze the production of sufficient L-DOPA cofactor to maintain a rapid and sustained tyrosine oxidase activity. Interestingly, the expression territories and the timing of *H. roretzy tyrosinase* messenger RNA overlap those of *Ci-tyr* [51]. *H. roretzy tyrp* (*Hr-tyrp*) messenger RNA instead appears earlier (at 110 cell stage) than *Ci-tyrp1/2a*, being localized in blastomeres other than pigment cell lineage [52]. The functional significance of this early expression is not known, but it is important to note that *Hr-tyrp* messenger RNA becomes exclusive of pigment cell lineage from neurula stage onward.

In amphioxus *tyr*, *tyrp1/2a* and *tyrp1/2b* are coexpressed, throughout the epidermal ectoderm, in gastrula and neurula stages and eventually become localized in the rudiment of the primary pigment spot of Hesse organ located in the amphioxus neural tube [53]. The Hesse ocelli represents a characteristic amphioxus trait, not observed in vertebrates, consisting in the presence of bicellular photoreceptor organs (one receptor and one pigment cell), distributed throughout most of the spinal cord and numbered in hundreds in mature animals [54]. Besides Hesse ocelli, amphioxus has another pigment structure, the frontal eye, that differentiates early in the larva and is implicated in controlling orientation to light [54]. The finding that the first spot of Hesse ocelli coexpresses *tyr*, *tyrp1/2a* and *tyrp1/2b* [53] indicates that this pigment cell lineage utilize the same input, as vertebrates, at least for the first pigment spot melanization. Further *in situ* experiments will be instrumental to clarify potential involvement of tyr, tyrp1/2a and tyrp1/2b in melanogenic processes of the frontal eye pigment cell and of all the ocelli that develop in the mature animals. Furthermore, in-depth examination of the expression pattern of the supernumerary *tyr-like* genes, at different developmental stages, will enormously contribute to our understanding of how and when these genes are used in the melanogenic processes in amphioxus.

In the present study we also identified the regulatory regions responsible for the spatio-temporal expression of the three *tyrosinase* family members in *C. intestinalis* (Fig. 4). The ascidian genome is compact, compared with vertebrate genomes, and intergenic regions as well as introns are relatively small [22]. In the case of the three *tyrosinase* family genes, the upstream intergenic regions we have analyzed range between 0.8 and 1.5 kb, which is small enough to easily isolate and clone the whole intergenic region in a reporter expression vector. Similar to transcript levels, the *pCi-tyrp1/2a* enhancer revealed to be the strongest element in driving a robust reporter expression, from the late gastrula stage up to the

tadpole stage, in all pigment cell lineage descendants. *pCi-tyrp1/2b* and *pCi-tyr*, instead, although specific, appeared weaker in terms of the number of embryos showing transgene expression. It is likely that these regulatory regions contains elements necessary, but not sufficient, for a robust activation in the endogenous territories. Probably other elements located outside the area we have tested, maybe in the introns, may fill this gap. Thus we have identified relatively short upstream intergenic regions, which are lineage specific and are capable to switch on their activity in a concerted way and at precise developmental times of pigment cell formation. In *C. intestinalis*, lineage restricted enhancers are often used as tool for targeted interference with lineage restricted developmental genes [55,56,57]. Our enhancers, given their specificity in labelling the pigment cell lineage, with *pCi-tyrp1/2a* being active from late gastrula and *pCi-tyr* from neurula stages, have been already successfully exploited to interfere with the function of two factors involved in pigment cell differentiation at two consecutive developmental stages [21]. On the other hand, these promoters can give important clues to the network controlling pigment cell development, since tyrosinase family members are typical markers of this lineage.

The bioinformatic analyses we have conducted on these promoters revealed that *C. intestinalis* and *C. savignyi tyrp1/2a* genes share, in their promoter regions, some motifs specific for Pax, Oct/Pou, and Sox family transcription factor bindings (Fig. 5). It is intriguing to note that in *C. intestinalis* the density of specific binding sites in promoter fragments, coupled with their conservation in the corresponding regions of *C. savignyi*, is often indicative of their functional relevance. The distribution of this motif sites is also partially conserved in *Hr-tyrp* promoter. Interestingly, these specific DNA motifs and the corresponding transcription factors have been already demonstrated to be involved in the activation of tyr family members in different vertebrate species [34], thus indicating a certain grade of conservation, among chordates, in the molecular mechanisms controlling pigmentation machinery. Our analysis did not reveal any canonical binding motif for Mitf, a factor fundamental in the development of vertebrate melanin producing cells, both in *C. intestinalis* or *C. savignyi tyr* and *tyrps* 5′ regulatory regions, paralleling the previously findings on *H. roretzi tyr* and *tyrp* promoters [35,58]. In *Halocynthia*, *Hr-Mitf* messenger RNA becomes localized in the pigment cell lineage from neurula stage onward, as in *C. intestinalis* (http://www.aniseed.cnrs.fr/), and *Hr-Mitf* overexpression is able to activate ectopically *Hr-tyr* [59]. Because canonical Mitf binding sites are absent on ascidian *tyr* and *tyrps* promoters it is possible that Mitf action is mediated through additional transcription factors. Alternatively, we can suppose that ascidian Mitf factors are able to recognize and interact with a non-canonical E-box binding motif.

Collectively the data we have presented here revealed that *tyrosinase* gene family phylogeny is much more painted than previously thought, thus opening new perspectives on the way by which the organisms synthesize and exploit melanin during evolution. Additional expansion of the comparative analysis of *tyr* and *tyrps*, by exploiting the "daily" new sequenced genomes, combined with experiments of *in situ* hybridization and studies on transcriptional regulation of *tyr* and *tyrps* in different phyla, will be crucial for accessing further aspects of pigment cell biology and revealing important mechanisms about their evolution.

Materials and Methods

Phylogenetic analysis

Sequences used in phylogenetic analysis were retrieved from the NCBI database and are listed in Table S1. The protein set was aligned using ClustalW [60,61] and Mega5 [62] with default parameters and the amphioxus tyr protein was taken as reference from its Q19 to the E458 residual aminoacid to set the length of the datasets, in order to avoid regions of unreliable alignment along the extremities of the molecule. *Suberites domuncula* tyr-like was used as outgroup.

Phylogenetic tree reconstructions were carried out using the bayesian (MrBayes), neighbor joining (NJ) and the maximum likelihood (ML) methods. Bayesian trees were inferred using MrBayes 3.1.2 [63,64]. Two independent runs of 1 million generations each were performed, each with four chains. For convention, convergence was reached when the value for the standard deviation of split frequencies stayed <0.01. NJ and ML analyses were performed using Mega5 [62] and robustness of the obtained tree topologies was assessed with 1000 Bootstrap replicates.

All phylogenetic methods gave similar tree topology, Fig. 1 shows the consensus tree in which each node reports the bootstrap value for MrBayes, ML and NJ respectively.

Synteny conservation analysis

We analyzed the presence of synteny conservation using the Synteny Database developed by Catchen et al. [65] in addition to manual searches in amphioxus (*B. floridae*, JGI v2.0), *Ciona* (*C. intestinalis*, JGI v2.0) and human (*H. sapiens*, Ensembl release 64) genomes. In Fig. 2 we only show a few representative genes among many that are conserved around the *tyrp* locus. *Ciona* chromosome 5 and human chromosome 9 share an high degree of conservation in the genomic neighborhoods surrounding the *Ci-tryp1/2a* and *Hr-Tyrp1*. Amphioxus *tyrp1/2a* and *tyrp1/2b* lay on scaffold Bf_V2_61 that is most probably not long enough to establish synteny.

Animals and embryos

Adult *C. intestinalis* were collected from the Gulf of Naples. Animal handling and transgenesis *via* electroporation have been carried out as previously described [66,67]. Embryo imaging capture was made with a Zeiss Axio Imager M1.

Whole-mount *in situ* hybridization

Three cDNA clones, presumably encoding tyrosinase family members, were found in *Ciona* genomic database (http://genome.jgi-psf.org/Cioin2/Cioin2.home.html): citb41l04 (N. Satoh Gene Collection 1 ID:CiGC33c19), citb030d10 (N. Satoh Gene Collection 1 ID: CiGC44b23) and cilv069a04 (N. Satoh Gene Collection 1 ID: CiGC31h05). They have been named respectively *Ci-tyr*, *Ci-tyrp1/2a* (previously named *Ci-tyrp1a* [21]) and *Ci-tyrp1/2b*. The corresponding RNA probes were used for whole-mount *in situ* hybridization experiments, performed as previously described [67].

Construct preparation

pCi-tyr>mChe and *pCi-tyrp1/2a>mChe* (originally named *ptyr-p1a>*mChe) constructs were previously prepared [21]. Approximately 1.5 kb of the *Ci-tyrp1/2b* 5′ flanking region was PCR-amplified from genomic DNA using the primers: *Ci-ptyrp1/2bF* (GTAGTATAAACAAACTACCGATAACCTGC) and *Ci-ptyrp1/2bR* (AGAACGAAGAAATAGATGTATGCTTGG). The fragment *pCi-tyrp1/2b* was cloned into pCR®II vector (TOPO® TA Cloning Dual Promoter Kit, Invitrogen), following the manufacturer's indications, and then excised through digestion with the unique restriction sites in pCR®II plasmid polylinker (*Hind*III-*Not*I for cloning upstream of mCherry). The digested

fragment replaced *pCi-tyr* in *pCi-tyr>mChe* vector (previously digested with *Hind*III-*Not*I to eliminate *ptyr*) to create the construct *pCi-tyrp1/2b>mChe*.

pCi-tyrp1/2a-0.6>mChe: the construct *pCi-tyrp1/2a->mChe* was digested with compatible ends generating enzymes, *Spe*I-*Xba*I, to eliminate a fragment of about 0.9 kb at the 5′ end of *pCi-tyrp1a*; the resulting linearized vector was then re-ligated.

In silico analysis for putative *trans*-acting factors

pCi-tyrp1/2a-0.6, the corresponding region of *pCs-tyrp1/2a* and *Hr-tyrp-333N* were analyzed using PROMO program (http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB = TF_8.3). Only dissimilarity values ≤2% (divergence percentage of the given sequence from the consensus matrix) have been taken into consideration.

Supporting Information

Figure S1 Chromosomal distribution of the expanded tyrosinases in *B. floridae*, *S. kowalevskii*, *N. vectensis* and *C. elegans*. The tandem duplicates lying on the same scaffold or chromosome are highlighted using the same color code.
(PDF)

Figure S2 Schematic representation of two metal binding domains (MeA and MeB) and sequence consensus. A) Key aminoacid positions, probably involved in the change of affinity to phenolic substrates, are reported in pink for tyrosinases and yellow for tyrosinase-related proteins. ø indicates aromatic residues (F, Y or W), x indicates any aminoacid. B) Multiple sequence alignment was obtained using Mega5 software. Conserved residues in all proteins analyzed, highlighted in green, define a robust sequence consensus. Aminoacid changes between tyr and tyrps are represented in pink and yellow, respectively.
(PDF)

Figure S3 Cysteine residues conservation in Tyrosinase family proteins. Cysteine (Cys) residues number and positions are reported from deuterostome and protostome organisms. A significant conservation is detectable in deuterostomes; protostomes share with deuterostomes only the N-terminal cluster and present an additional stretch of cysteine residues at the C-terminal.
(PDF)

Figure S4 *In vivo* analysis of *pCi-tyrp1/2a-0.6*>mChe construct. Transgenesis was performed *via* electroporation experiments. Merged bright-field/fluorescent images of mCherry expression driven by *pCi-tyrp1/2a.0-6* region at early (A) and middle (B) larval stages (lateral view, anterior is on the left). Note that the transgene expression in pigment cell lineage corresponds to *pCi-tyrp1/2a* full-length enhancer (compare with Fig. 4D).
(PDF)

Table S1 Accession number of sequences used in the phylogenetic tree of Figure 1.
(DOC)

Acknowledgments

We are deeply grateful to Alison Cole for the final English revision of the paper. We thank Giampiero Lanzotti (SZN) for helping in figure preparation.

Author Contributions

Conceived and designed the experiments: AS FR SDA. Performed the experiments: RE SDA PS MRP. Analyzed the data: SDA RE AS FR. Wrote the paper: RE SDA AS.

References

1. Eakin RM (1973) The third eye. Berkeley: University of California Press.
2. Silvers WK (1979) The coat colors of mice - A model for mammalian gene action and interaction:. New York: Springer-Verlag.
3. LeDouarin NM (1982) The neural crest:. Cambridge: Cambridge University Press.
4. Tsukamoto K, Jackson IJ, Urabe K, Montague PM, Hearing VJ (1992) A second tyrosinase-related protein, TRP-2, is a melanogenic enzyme termed DOPAchrome tautomerase. Embo J 11: 519–526.
5. Luo D, Chen H, Jimbow K (1994) Cotransfection of genes encoding human tyrosinase and tyrosinase-related protein-1 prevents melanocyte death and enhances melanin pigmentation and gene expression of Lamp-1. Exp Cell Res 213: 231–241.
6. Manga P, Sato K, Ye L, Beermann F, Lamoreux ML, et al. (2000) Mutational analysis of the modulation of tyrosinase by tyrosinase-related proteins 1 and 2 in vitro. Pigment Cell Res 13: 364–374.
7. Kobayashi T, Imokawa G, Bennett DC, Hearing VJ (1998) Tyrosinase stabilization by Tyrp1 (the brown locus protein). J Biol Chem 273: 31801–31805.
8. Garcia-Borron JC, Solano F (2002) Molecular anatomy of tyrosinase and its related proteins: beyond the histidine-bound metal catalytic center. Pigment Cell Res 15: 162–173.
9. Hearing VJ (2000) The melanosome: the perfect model for cellular responses to the environment. Pigment Cell Res 13 Suppl 8: 23–34.
10. Goding CR (2007) Melanocytes: the new Black. Int J Biochem Cell Biol 39: 275–279.
11. Olivares C, Solano F (2009) New insights into the active site structure and catalytic mechanism of tyrosinase and its related proteins. Pigment Cell Melanoma Res 22: 750–760.
12. Jackson IJ (1994) Evolution and expression of tyrosinase-related proteins. Pigment Cell Res 7: 241–242.
13. Sturm RA, O'Sullivan BJ, Box NF, Smith AG, Smit SE, et al. (1995) Chromosomal structure of the human TYRP1 and TYRP2 loci and comparison of the tyrosinase-related protein gene family. Genomics 29: 24–34.
14. Sato S, Yamamoto H (2001) Development of pigment cells in the brain of ascidian tadpole larvae: insights into the origins of vertebrate pigment cells. Pigment Cell Res 14: 428–436.
15. Camacho-Hubner A, Richard C, Beermann F (2002) Genomic structure and evolutionary conservation of the tyrosinase gene family from Fugu. Gene 285: 59–68.
16. Takeuchi K, Satou Y, Yamamoto H, Satoh N (2005) A genome-wide survey of genes for enzymes involved in pigment synthesis in an ascidian, Ciona intestinalis. Zoolog Sci 22: 723–734.
17. Tsuda M, Sakurai D, Goda M (2003) Direct evidence for the role of pigment cells in the brain of ascidian larvae by laser ablation. J Exp Biol 206: 1409–1417.
18. Nishida H (1987) Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. III. Up to the tissue restricted stage. Dev Biol 121: 526–541.
19. Satoh N (2001) Ascidian embryos as a model system to analyze expression and function of developmental genes. Differentiation 68: 1–12.
20. Di Gregorio A, Levine M (2002) Analyzing gene regulation in ascidian embryos: new tools for new perspectives. Differentiation 70: 132–139.
21. Squarzoni P, Parveen F, Zanetti L, Ristoratore F, Spagnuolo A (2011) FGF/MAPK/Ets signaling renders pigment cell precursors competent to respond to Wnt signal by directly controlling Ci-Tcf transcription. Development 138: 1421–1432.
22. Dehal P, Satou Y, Campbell RK, Chapman J, Degnan B, et al. (2002) The draft genome of Ciona intestinalis: insights into chordate and vertebrate origins. Science 298: 2157–2167.
23. Vinson JP, Jaffe DB, O'Neill K, Karlsson EK, Stange-Thomann N, et al. (2005) Assembly of polymorphic genomes: algorithms and application to Ciona savignyi. Genome Res 15: 1127–1135.
24. Putnam NH, Butts T, Ferrier DE, Furlong RF, Hellsten U, et al. (2008) The amphioxus genome and the evolution of the chordate karyotype. Nature 453: 1064–1071.
25. (1998) Genome sequence of the nematode C. elegans: a platform for investigating biology. Science 282: 2012–2018.
26. Putnam NH, Srivastava M, Hellsten U, Dirks B, Chapman J, et al. (2007) Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science 317: 86–94.
27. Sugumaran M (2002) Comparative biochemistry of eumelanogenesis and the protective roles of phenoloxidase and melanin in insects. Pigment Cell Res 15: 2–9.

28. Porchet-Hennere E, Vernet G (1992) Cellular immunity in an annelid (Nereis diversicolor, Polychaeta): production of melanin by a subpopulation of granulocytes. Cell Tissue Res 269: 167–174.
29. Wang T, Sun Y, Jin L, Xu Y, Wang L, et al. (2009) Enhancement of non-specific immune response in sea cucumber (Apostichopus japonicus) by Astragalus membranaceus and its polysaccharides. Fish Shellfish Immunol 27: 757–762.
30. Caracciolo A, Gesualdo I, Branno M, Aniello F, Di Lauro R, et al. (1997) Specific cellular localization of tyrosinase mRNA during Ciona intestinalis larval development. Dev Growth Differ 39: 437–444.
31. Bertrand V, Hudson C, Caillol D, Popovici C, Lemaire P (2003) Neural tissue in ascidian embryos is induced by FGF9/16/20, acting via a combination of maternal GATA and Ets transcription factors. Cell 115: 615–627.
32. Johnson DS, Davidson B, Brown CD, Smith WC, Sidow A (2004) Noncoding regulatory sequences of Ciona exhibit strong correspondence between evolutionary constraint and functional importance. Genome Res 14: 2448–2456.
33. Martinez-Morales JR, Rodrigo I, Bovolenta P (2004) Eye development: a view from the retina pigmented epithelium. Bioessays 26: 766–777.
34. Murisier F, Beermann F (2006) Genetics of pigment cells: lessons from the tyrosinase gene family. Histol Histopathol 21: 567–578.
35. Toyoda R, Kasai A, Sato S, Wada S, Saiga H, et al. (2004) Pigment cell lineage-specific expression activity of the ascidian tyrosinase-related gene. Gene 332: 61–69.
36. Gillespie JP, Kanost MR, Trenczek T (1997) Biochemical mediators of insect immunity. Ann Rev Entomol 42: 611–643.
37. Sugumaran M (1998) Unified mechanism for sclerotization of insect cuticle. Adv Insect Physiol 27: 229–334.
38. Huang G, Xie X, Han Y, Fan L, Chen J, et al. (2007) The identification of lymphocyte-like cells and lymphoid-related genes in amphioxus indicates the twilight for the emergence of adaptive immune system. PLoS One 2: e206.
39. Harrison FW, Ruppert EE (1991) Microscopic anatomy of invertebrates. New York: Wiley-Liss.
40. Zhang S, Liang Y, Guangdong J, Zhuang Z (2009) Protochordate amphioxus is an emerging model organism for comparative immunology. Progress in Natural Science 19: 923–929.
41. Aroca P, Garcia-Borron JC, Solano F, Lozano JA (1990) Regulation of mammalian melanogenesis. I: Partial purification and characterization of a dopachrome converting factor: dopachrome tautomerase. Biochim Biophys Acta 1035: 266–275.
42. Jimenez M, Tsukamoto K, Hearing VJ (1991) Tyrosinases from two different loci are expressed by normal and by transformed melanocytes. J Biol Chem 266: 1147–1156.
43. Jimenez-Cervantes C, Garcia-Borron JC, Valverde P, Solano F, Lozano JA (1993) Tyrosinase isoenzymes in mammalian melanocytes. 1. Biochemical characterization of two melanosomal tyrosinases from B16 mouse melanoma. Eur J Biochem 217: 549–556.
44. Jimenez-Cervantes C, Solano F, Kobayashi T, Urabe K, Hearing VJ, et al. (1994) A new enzymatic function in the melanogenic pathway. The 5,6-dihydroxyindole-2-carboxylic acid oxidase activity of tyrosinase-related protein-1 (TRP1). J Biol Chem 269: 17993–18000.
45. Kobayashi T, Urabe K, Winder A, Jimenez-Cervantes C, Imokawa G, et al. (1994) Tyrosinase related protein 1 (TRP1) functions as a DHICA oxidase in melanin biosynthesis. Embo J 13: 5818–5825.
46. Winder A, Kobayashi T, Tsukamoto K, Urabe K, Aroca P, et al. (1994) The tyrosinase gene family–interactions of melanogenic proteins to regulate melanogenesis. Cell Mol Biol Res 40: 613–626.
47. Orlow SJ, Zhou BK, Chakraborty AK, Drucker M, Pifko-Hirst S, et al. (1994) High-molecular-weight forms of tyrosinase and the tyrosinase-related proteins: evidence for a melanogenic complex. J Invest Dermatol 103: 196–201.
48. Sarangarajan R, Boissy RE (2001) Tyrp1 and oculocutaneous albinism type 3. Pigment Cell Res 14: 437–444.
49. Whittaker JR (1973) Tyrosinase in the presumptive pigment cells of ascidian embryos: tyrosine accessibility may initiate melanin synthesis. Dev Biol 30: 441–454.
50. Steel KP, Davidson DR, Jackson IJ (1992) TRP-2/DT, a new early melanoblast marker, shows that steel growth factor (c-kit ligand) is a survival factor. Development 115: 1111–1119.
51. Sato S, Masuya H, Numakunai T, Satoh N, Ikeo K, et al. (1997) Ascidian tyrosinase gene: its unique structure and expression in the developing brain. Dev Dyn 208: 363–374.
52. Sato S, Toyoda R, Katsuyama Y, Saiga H, Numakunai T, et al. (1999) Structure and developmental expression of the ascidian TRP gene: insights into the evolution of pigment cell-specific gene expression. Dev Dyn 215: 225–237.
53. Yu JK, Meulemans D, McKeown SJ, Bronner-Fraser M (2008) Insights from the amphioxus genome on the origin of vertebrate neural crest. Genome Res 18: 1127–1132.
54. Lacalli TC (2004) Sensory systems in amphioxus: a window on the ancestral chordate condition. Brain Behav Evol 64: 148–162.
55. Davidson B, Shi W, Levine M (2005) Uncoupling heart cell specification and migration in the simple chordate Ciona intestinalis. Development 132: 4811–4818.
56. Davidson B, Shi W, Beh J, Christiaen L, Levine M (2006) FGF signaling delineates the cardiac progenitor field in the simple chordate, Ciona intestinalis. Genes Dev 20: 2728–2738.
57. Beh J, Shi W, Levine M, Davidson B, Christiaen L (2007) FoxF is essential for FGF-induced migration of heart progenitor cells in the ascidian Ciona intestinalis. Development 134: 3297–3305.
58. Toyoda R, Sato S, Ikeo K, Gojobori T, Numakunai T, et al. (2000) Pigment cell-specific expression of the tyrosinase gene in ascidians has a different regulatory mechanism from vertebrates. Gene 259: 159–170.
59. Yajima I, Endo K, Sato S, Toyoda R, Wada H, et al. (2003) Cloning and functional analysis of ascidian Mitf in vivo: insights into the origin of vertebrate pigment cells. Mech Dev 120: 1489–1504.
60. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680.
61. Higgins DG, Thompson JD, Gibson TJ (1996) Using CLUSTAL for multiple sequence alignments. Methods Enzymol 266: 383–402.
62. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731–2739.
63. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17: 754–755.
64. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.
65. Catchen JM, Conery JS, Postlethwait JH (2009) Automated identification of conserved synteny after whole-genome duplication. Genome Res 19: 1497–1505.
66. Corbo JC, Levine M, Zeller RW (1997) Characterization of a notochord-specific enhancer from the Brachyury promoter region of the ascidian, Ciona intestinalis. Development 124: 589–602.
67. Ristoratore F, Spagnuolo A, Aniello F, Branno M, Fabbrini F, et al. (1999) Expression and functional analysis of Cititf1, an ascidian NK-2 class gene, suggest its role in endoderm development. Development 126: 5149–5159.