UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: **Università degli Studi di Padova**

Dipartimento di Biologia

CORSO DI DOTTORATO DI RICERCA IN: **Bioscienze e Biotecnologie**
CURRICOLO: **Biologia evoluzionistica**
CICLO: **XXIX**

Identification and expression studies of orthologues of vertebrate haematopoietic and neural stem cell markers in the urochordate *Botryllus schlosseri*

Coordinatore: Ch.mo Prof. Paolo Bernardi
Supervisore: Ch.mo Prof. Loriano Ballarin

Dottorando: Francesca Ballin

Summary

All multicellular organisms originate from a small set of totipotent embryonic stem cells that differentiate into a structured body plan during embryogenesis. These cells share the common morphology of undifferentiated cells with a high nucleus/cytoplasm ratio, few cytoplasmic organelles and absence of specific functions, and reach their full functionality through various steps in which they progress towards their specificity. Stem cells with different plasticity persist in tissues during the whole life of the organism and they can be useful in replacing adult cells after their damage or death.

Ascidians are invertebrate chordates, members of the subphylum Tunicata or Urochordata that represents the sister group of vertebrates (Delsuc et al., 2006). They offer the opportunity to investigate and compare the behaviour of both embryonic and adult stem cells (Tiozzo et al, 2008), as adult stem cells play important roles in various processes, such as the regenerative abilities in solitary ascidians and the cyclical renewal of the colony in compound ascidians (Berrill, 1951). In these organisms, it is still debated if there are separate lineages for somatic and germ cell lines or if there is a single population of pluripotent stem cells able to differentiate according to the niche in which they land (Laird et al, 2005). However, it is well known that, in close proximity to the ascidian pharynx, "haematopoietic nodules" are present (Ermak, 1975; Rinkevich et al, 2013) in which haematopoiesis occurs. The endostyle itself has been identified as a novel stem cell niche in the colonial species *B. schlosseri* (Voskoboynik et al, 2008). In addition, morphological data suggest the presence of circulating undifferentiated cells (haemoblasts) able to proliferate and give rise to terminally differentiated cells (Kawamura et al, 2006). Relevant studies were also carried out in the neural lineage of ascidians, in which neural progenitor cells, able to regenerate the brain after extirpation, were found (Bollner et al, 1995). In *B. schlosseri*, during the generation change, there are bud primordial cells, probably deriving from a pool of long-living stem cells (Tiozzo et al, 2009), able to give rise to the neural complex. The latter species offers the possibility to compare the genetic control of neurogenesis during embryogenesis and blastogenesis.

With the aim of better elucidating the process of haematopoiesis and neurogenesis during the colonial blastogenetic cycle (i.e., the interval of time between a generation change and the following one), I screened the *B. schlosseri* genome and transcriptome, looking for genes/transcripts showing similarity to vertebrate molecular markers of haematopoietic and neural stem cells. On these sequences, after an *in silico* translation, I performed the phylogenetic reconstruction that, always, gave tunicate as the vertebrate sister cluster. The four mammalian orthologous genes, used as markers for the recognition of haematopoietic progenitor cells, identified in *B. schlosseri*, are *bsabcg2, bscd133, bsgata1/2/3* and *bsgata4/5/6*. The ISH assay, performed by antisense specific riboprobes, on haemocyte monolayers and colony sections, resulted in a labelling of the sub-endostylar haemolymph lacunae. These results matches previously morphological data that identified the endostyle as a stem cell niche, strengthening our idea to use *bsabcg2, bscd133, bsgata1/2/3* and *bsgata4/5/6* genes for the identification of haematopoietic stem cells in *B. schlosseri*. Quantitative real time PCR (qRT-PCR) highlighted the over-expression of the considered genes in the mid-cycle (MC) phase of

the blastogenetic cycle. During this phase, there is the formation of new secondary buds emerging from the primary buds. The high expression levels of *bsabcg2, bscd133, bsgata1/2/3* and *bsgata4/5/6* genes in the MC phase reflect the presence of undifferentiated cells involved in proliferative and differentiation events required for giving rise to the new blastogenetic generation.

For the neural lineage, I identified and characterised two transcripts identified as orthologues of vertebrate neural stem cell markers Sox2 and Msi2 (BsSox2 and BsMsi2). In addition to these genes, I also examined the expression, during the blastogenetic cycle, of a panel of genes already known to be involved in ascidian larval neurogenesis, i.e., orthologues of Pax2/5/8, Hox1 and Hox3. ISH with riboprobes for BsSox2, BsMsi2, BsPax2/5/8, BsHox1 and BsHox3 revealed a common labelling of the endostyle niche and of the primordia of the branchial basket of the bud. Like the above-considered haematopoietic markers, the presence of *bssox2, bsmsi2, bspax2/5/8, bshox1* and *bshox3* transcripts in the cells of the region known to be a stem cell niche, led us to conclude not only that our probes identified undifferentiated cells but even that in *B. schlosseri* a single population of pluripotent stem cells that can differentiate into haematopoietic or neural cells depending on the microenvironmental signals is probably present. qRT-PCR showed high expression level of all the putative neural markers considered in the MC phase. Again, in this phase new secondary buds are produced from primary buds. Every new bud needs its own neural complex and this requires the proliferation of undifferentiated cells to originate neural gland rudiment and cerebral ganglion. The *bssox2, bsmsi2, bspax2/5/8, bshox1* and *bshox3* increased expression associated with these neurogenesis events support their involvement in neural stem cell identification. Moreover, haematopoietic and neural probes reveal additional labelling in *B. schlosseri* tissues (i.e. oocyte, tissues of primary bud and secondary bud, morula cells). This stimulates additional future studies for better comprehension of the role of these genes in *B. schlosseri*.

Riassunto

Tutti gli organismi pluricellulari originano da un piccolo set di cellule staminali embrionali totipotenti che differenziano in un piano corporeo completo durante l'embriogenesi. Queste cellule mostrano la tipica morfologia delle cellule indifferenziate con un elevato rapporto nucleo/citoplasma, pochi organelli citoplasmatici, assenza di funzioni specifiche e raggiungono il completo differenziamento attraverso stadi successivi. Le cellule staminali persistono nei tessuti durante l'intera vita dell'organismo, in diversi gradi di differenziamento, al fine di fornire cellule di ricambio in seguito a danno o morte cellulare delle cellule tissutali.
Le ascidie sono invertebrati marini appartenenti al phylum dei Cordati ed al sub-phylum dei Tunicati o Urocordati, considerato sister group dei Vertebrati (Delsuc et al, 2006). Esse offrono l'opportunità di investigare e comparare il comportamento di cellule staminali adulte ed embrionali (Tiozzo et al, 2008). La presenza di cellule staminali adulte si evince da vari processi come la rigenerazione in ascidie solitarie e il ciclico ricambio generazionale nelle ascidie coloniali (Berrill, 1951). Nelle ascidie è ancora molto dibattuta la presenza di distinti progenitori staminali multipotenti che portano alla formazione della linea somatica e di quella germinale, o al contrario, di progenitori staminali pluripotenti, in grado di differenziare in entrambe le linee, in base alla nicchia in cui si trovano (Laird et al, 2005). É risaputo che in prossimità del faringe delle ascidie, sono presenti noduli emopoietici (Ermak, 1975) o aggregati cellulari contenenti cellule staminali (Rinkevich et al, 2013). L'endostilo stesso è inoltre stato identificato come una nicchia di staminalità nell'ascidia coloniale *B. schlosseri* (Voskoboynik et al, 2008). Dati morfologici suggeriscono la presenza di cellule circolanti indifferenziate (emoblasti) capaci di proliferare e dare origine a cellule completamente differenziate (Kawamura et al, 2006). Ulteriore evidenza della presenza di cellule staminali adulte nelle ascidie deriva da studi condotti sulla linea neurale che hanno dimostrato la presenza di progenitori neurali capaci di rigenerare il cervello in seguito ad asportazione (Bollner et al, 2009). In *B. schlosseri,* inoltre, si pensa siano presenti pool di cellule staminali longeve che permettono, durante il ciclo blastogenetico, la formazione del complesso neurale nelle nuove gemme (Tiozzo et al, 2009). La presenza di una riproduzione sessuata e asessuata nell'ascidia coloniale *B. schlosseri*, permette di comparare il processo di neurogenesi in due distinti meccanismi di sviluppo: embriogenesi e blastogenesi.
Al fine di chiarire i processi di emopoiesi e neurogenesi durante il ciclo blastogenetico, ho effettuato una ricerca di geni/trascritti di marcatori molecolari di cellule staminali emopoietiche e neurali dei vertebrati da utilizzare per la ricerca di sequenze ortologhe nel genoma e nel trascrittoma di *B. schlosseri*. Una volta individuate queste sequenze, dopo la loro traduzione *in silico*, ho effettuato una ricostruzione filogenetica che, per tutti i trascritti considerati ha restituito un albero in cui gli Urocordati sono compresi in un cluster che risulta sister group dei Vertebrati. I quattro marcatori usati per l'identificazione dei progenitori emopoietici in *B. schlosseri*, ortologhi ai geni trovati nei mammiferi, sono *bsabcg2, bscd133, bsgata1/2/3* e *bsgata4/5/6.* L'ibridazione *in situ,* effettuata mediante specifiche ribo-sonde, su strisci di emociti e su sezioni di

colonie, ha permesso di osservare la presenza dei trascritti emopoietici nelle lacune emolinfatiche della regione sub-endostilare. Questi risultati coincidono con le precedenti analisi morfologiche che identificano l'endostilo come nicchia di staminalità e rinforzano l'idea che *bsabcg2, bscd133, bsgata1/2/3*, *bsgata4/5/6* siano anch'essi marcatori di staminalità emopoietica in *B. schlosseri*. Con la real time-PCR si osserva un'elevata espressione dei suddetti geni nella fase di mid-cycle del ciclo blastogenetico. Durante questa fase si ha la formazione delle gemme secondarie a partire da un ispessimento della parete della gemma primaria. L'elevato livello di espressione di *bsabcg2, bscd133, bsgata1/2/3* e *bsgata4/5/6* nel mid-cycle può essere pertanto messo in relazione con gli eventi di proliferazione e differenziamento di cellule staminali, necessari per la formazione dei tessuti della nuova generazione blastogenetica.
Per la linea neurale, ho identificato e caratterizzato due trascritti ortologhi a Sox2 e Msi2 coinvolti nel riconoscimento di precursori neurali nei vertebrati ed ho indagato, inoltre, l'espressione, durante il ciclo blastogenetico, di un pannello di geni coinvolti nello sviluppo del sistema nervoso della larva delle ascidie (*bspax2/5/8, bshox1 e bshox3*). L'ISH con specifiche ribo-sonde per BsSox2, BsMsi2, BsPax2/5/8, BsHox1 e BsHox3 ha mostrato una marcatura comune nella nicchia endostilare e nella regione di formazione degli stigmi nel cestello branchiale della gemma primaria. Come osservato per i marcatori dei precursori emopoietici, la presenza dei trascritti di *bssox2, bsmsi2, bspax2/5/8, bshox1* e *bshox3* nella regione identificata come nicchia di staminalità, ci permette di concludere che le sonde dei trascritti considerati identificano precursori staminali e che, probabilmente, in *B. schlosseri* sono presenti dei precursori staminali pluripotenti in grado di differenziare in progenitori emopoietici o neurali in base ai segnali ricevuti. La real time-PCR mostra un elevato livello di trascrizione di questi geni nella fase di mid-cycle. Ricordando che in questa fase si sviluppano le gemme secondarie a partire dall'ispessimento della parete delle gemme primarie, la proliferazione di cellule staminali al fine di permettere la formazione di cellule neurali, necessarie per la formazione del primordio della ghiandola neurale e del ganglio cerebrale della nuova gemma, è fenomeno indispensabile. L'aumento di espressione di *bssox2, bsmsi2, bspax2/5/8, bshox1* e *bshox3* nella fase in cui si sa avvenire la formazione del nuovo sistema nervoso della gemma supporta la validità di queste molecole nell'identificazione dei precursori neurali. In aggiunta a questo importante ruolo, i marcatori emopoietici e neurali considerati, rivestono probabilmente ulteriori funzioni nei tessuti di *B. schlosseri* come evidenziato dalla presenza di marcature aggiuntive (ovociti, tessuti di gemme primarie e secondarie, cellule morulari). Quest'ulteriore aspetto verrà approfondito in studi successivi.

What are stem cells?

Stem cells are unspecialised cells that can differentiate in many different cell types giving rise to the entire body of the organism. They can also replace cells lost as a consequence of normal wear and tear, injury, or disease, serving as an internal repair system. Stem cells share an high nucleus/cytoplasm ratio, few cytoplasm organelles, absence of specific functions and are characterised by two important features, self-renewal (the ability to go through numerous cycles of cell division while maintaining the undifferentiated state) and plasticity (capacity to differentiate into specialized cell types). All multicellular organisms contain different type of stem cells, that could be classified on the basis of their plasticity in: totipotent (zygote, able to give rise to all the body cells type), pluripotent (able to give rise to all the cells within a germ layer), multipotent (able to give rise to all the cells of a specific tissues lineage), oligopotent (i.e. myeloid and linfoid progenitors) and unipotent (useful in high turnover tissues, i.e. epidermis). On the basis of their origin, they can be distinguished in embryonic and adult. Embryonic stem cells are pluripotent stem cells derived from the inner cell mass of the blastocyst and can easily proliferate in culture for long periods without differentiating (Yu and Thomson, 2013), whereas adult stem cells are multipotent stem cells, found in various tissues, that act as a repair system for the body. The latter are rare in mature tissues, so isolating these cells from an adult tissue is challenging, and protocols to expand their numbers in cell culture have not yet been worked out. In order to bypass ethical difficulties in the use of embryonic stem cells, in 2006, Takahashi and Yamanaka reprogrammed differentiated cells to an embryonic-like state by transfer of nuclear contents into oocytes or by fusion with embryonic stem (ES) cells and create the so called induced pluripotent stem (iPS) cells. Unspecialised stem cells become more specialised through several steps in a process call differentiation with dramatic changes in cell size, shape, metabolic activity and responsiveness to internal (i.e., DNA-encoded) and external signals (including molecules secreted by other cells, physical contact with surrounding cells, and certain soluble factors in the microenvironment).

How to identify stem cells?

Different types of stem cells are present in small populations (e.g. 1 stem cell out of 100,000 cells in the circulating blood) and look just like any other cell in the tissue under a microscope. Moreover, within the cell lineages cells at different stages of differentiation and activation are present. In order to distinguish these types of cells, scientist looked for a tool that can help to identify them. Specialised proteins, able to bind a variety of ligands, coat the surface of every cell in the body. Each cell type, for example an erythrocyte, has a certain combination of receptors on its surface that makes it distinguishable from other kind of cells. Furthermore, proteins involved in various signal pathways are also known to have important functions in cell fate decision. These molecules can be used as markers. The unique expression pattern and timing of stem cell markers, provide an useful tool to identify and isolate particular stem cell type. The most common approach for the identification and the isolation of stem cell is by use of FACS (fluorescence activated cell sorting) coupled with molecular markers. In recent

years, molecular markers have been reported as features of undifferentiated embryonic stem cells, especially for humans (Zhao et al, 2012).

The crucial characteristic of a molecular marker is its capability to translate an external signal in an internal one thus determining embryonic stem cell fate. Molecular markers can be subdivided in surface molecules and transcription factor. Examples of surface molecular markers are cluster of differentiation (CD) antigens, proteins that belong to various families, such as integrins, adhesion molecules, glycoproteins, and receptors, all recognised by specific antibodies. Transcription factors are also used to identify stem cells. When they bind their recognition sequences, a specific signal transduction event happens. Tracking the expression of the genes of the above proteins can be used as a marker for a specific cell situation.

The colony of *B. schlosseri*

Tunicata or Urochordata evolved from a chordate ancestor ca. 550 million years ago (Laird & De Tomaso, 2005). They include sessile ascidians and pelagic thaliaceans and larvaceans. Ascidians are reliable model organisms for the study of a variety of biological processes, including developmental biology, ageing, immunology, genetics and evolutionary biology (Rinkevich, 2002). They show the chordate features (notochord, dorsal neural tube, pharynx and tail) during the larval stages. After the swimming larval phase, the chordate structures are lost and the metamorphosing larva develops into a sessile organism called oozooid (De Tomaso et al, 2005).

Botryllus schlosseri (Pallas, 1766)

Phylum: Chordata

Subphylum: Tunicata (Urochordata)

Class: Ascidiacea

Order: Pleurogona

Family: Styelidae

Genus: *Botryllus*

Species: *Botryllus schlosseri*

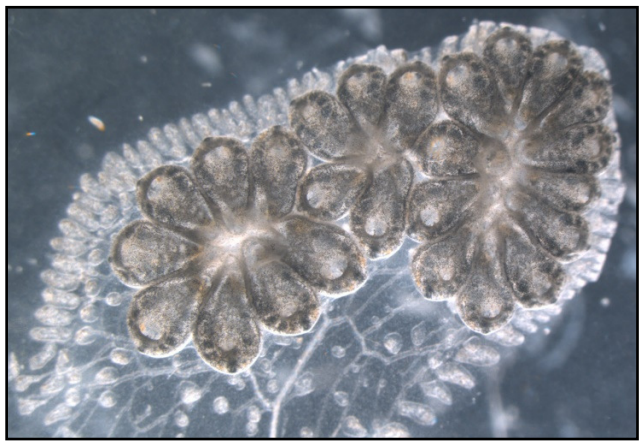


Fig.1: colony of *B. schlosseri*.

Botryllus schlosseri is a colonial ascidian (fig 1), with sexual and asexual reproduction. A colony is a clone, formed by many genetically identical modules (zooids, fig 2C) embedded in a gelatinous matrix, the tunic (fig 2B). The latter is crossed by a common vascular network that, at the periphery, forms many small, blind protrusions called ampullae involved in the adhesion to the substrate. Colonies of the ascidian *B. schlosseri* include three generations of zooids: the mature filter-feeding zooids, the primary buds on zooids and the secondary buds (budlets) emerging from the primary buds, fig 2A (Manni et al, 2007).

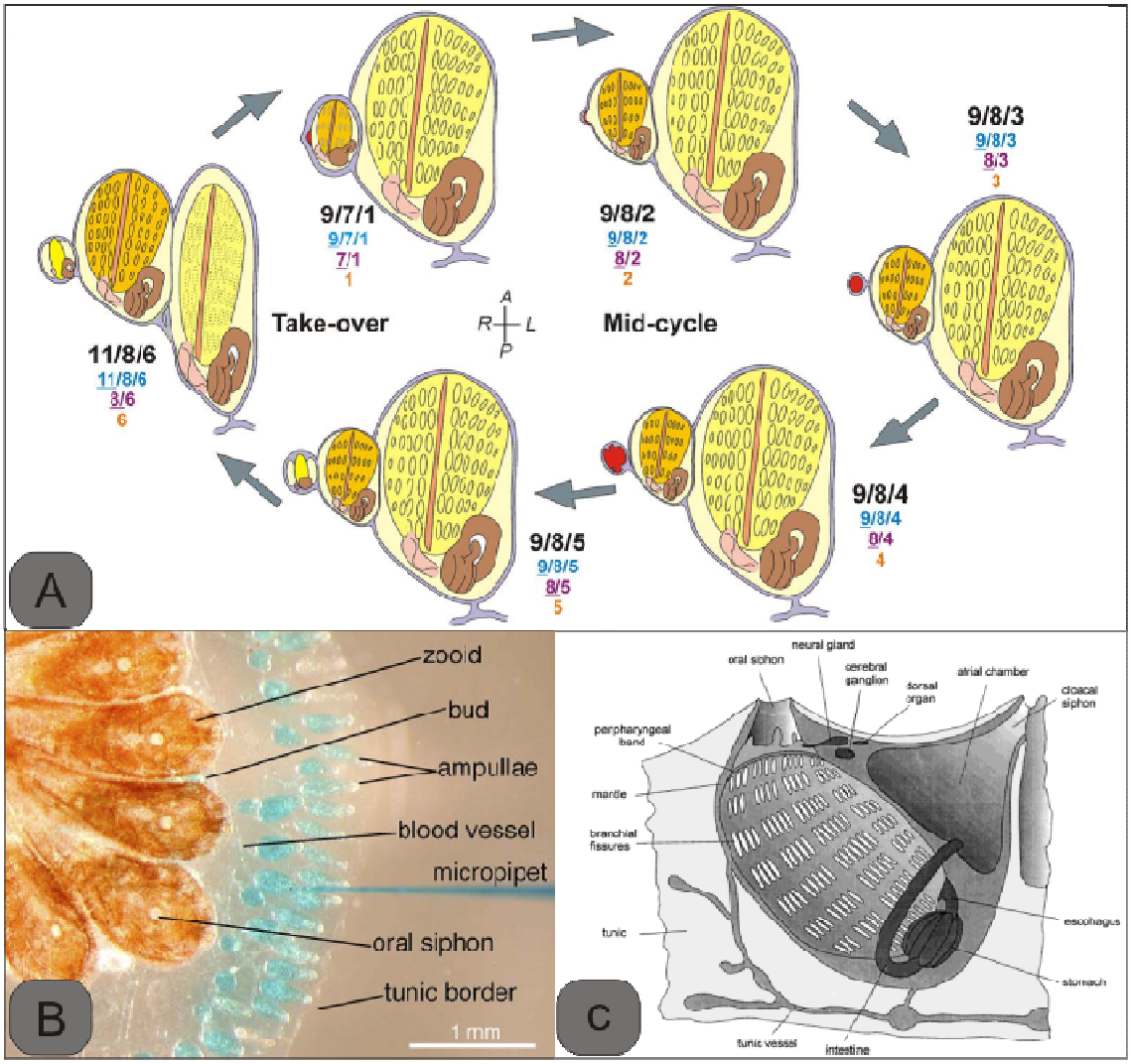


Fig.2: A: blastogenetic cycle of *B. schlosseri.* Colony developmental phases are indicated in black by the combination of three numbers, referring to the developmental stages of zooids, buds and budlets, respectively. Post take-over phase is represented by 9/7/1, mid-cycle cover phases 9/8/2, 9/8/3 and 9/8/4, pre take-over correspond to phase 9/8/5 and take-over to 11/8/6 (modified from Manni et al, 2014). **B:** magnification of a portion of a *Botryllus* colony with zooids, bud and ampullae embedded in the tunic (from Laird et al, 2005). **C:** scheme of the internal organisation of a single zooid of *B. schlosseri* (from Burighel et al, 1998).

Cyclical (weekly at 20°C) generation changes (or take-overs, TOs) occur during which adult zooids are progressively resorbed and replaced by their buds that reach the adult size. In the meantime budlets become buds and a new bud generation appears. A colonial blastogenetic cycle can be defined as the period of time between a TO and the next. Each blastogenetic cycle is characterised by the development and the maturation of primary buds and the appearance of new secondary buds. A bud primordium appears as a thickening of the peribranchial wall and the overlying epidermis in a process called palleal budding: it contains progenitor cells able to differentiate in a variety of tissues (Tiozzo et al, 2008). Vascular budding, i.e. the development of new buds from the vascular system, also occurs when a colony is deprived of all its zooids and buds. Buds

are formed by the vessel ephithelium (deriving from epidermis) and circulating multipotent or pluripotent cells that aggregate at the basis of an ampulla (Sabbadin et al, 1975). Stem cells of colonial organisms like botryllid ascidians are also involved in the cyclical reconstruction of new sets of male and female gonads in newly formed zooids (Rosner et al, 2009). It is still unclear if there are separate lineages for somatic and germ cell lines or if pluripotent stem cells are able to differentiate according to the niche in which they land (Laird et al, 2005). As a consequence, the molecular mechanisms involved in stem cells proliferation and differentiation are scanty known.

Vertebrate haematopoietic stem cell (HSC) markers

Several somatic niches are known in mammals (haematopoietic, skin/hair follicle, mammary gland, incisor teeth, neural and intestine) (Barker et al, 2007; Calvi et al, 2003; Cotsarelis et al, 1990; Doetsch et al, 1999; Harada et al, 1999; Palmer et al, 1997; Potten et al, 2002; Shen et al, 2004; Tumbar et al, 2004; Villadsen et al, 2007; Zhang et al, 2003). HSCs are multipotent, self-renewing progenitor cells responsible for continuous supply of differentiated blood cells of the myeloid and lymphoid lineages. Mammalian HSCs are very rare in bone marrow and blood, they are morphologically very similar to white blood cells and their identification and isolation occurs by cell surface markers and cluster of differentiation antigens (Pazhanisamy, 2013).

2B4/CD244/SLAMF4	EVI-1
ABCG2	Flt-3/Flk-2
Aldehyde Dehydrogenase 1	GATA-2
BMI-1	GFI-1
C1q R1/CD93	Integrin alpha 6/CD49f
CD34	Mcl-1
CD38	MYB
CD44	PLZF
CD45	Podocalyxin
CD45.1	Prominin 2
CD45.2	PTEN
CD48/SLAMF2	PU.1/Spi-1
CD90/Thy1	Sca-1/Ly6
CD117/c-kit	SLAM/CD150
CD133	Spi-B
CDCP1	STAT5a/b

CXCR4	STAT5a
Endoglin/CD105	STAT5b
EPCR	VCAM-1/CD106
Erythropoietin R	VEGF R2/KDR/Flk-1

Table1: human molecular markers associated with haemopoietic stem cells.

In Table 1 a brief list of molecular markers known for human HSCs is presented. In particular, ABCG2 (ATP-Binding Cassette, G2 subfamily) is a membrane protein, member of the ATP-binding cassette superfamily, one of the largest group of transporter proteins. It is considered a haematopoietic marker with a regulatory role in maintaining pluripotency (Zhou et al, 2001). In human bone marrow, ABCG2 is highly expressed in immature haematopoietic precursors and is sharply down-regulated in the subsequent stages of differentiation (Scharenberg et al, 2002). CD133, also called Prominin-1, is a cholesterol-binding, five-transmembrane-domain glycoprotein, also considered a stem cell marker and identified on the surface of both neural (Weigmann et al, 1997) and haematopoietic stem cells (Yin et al, 1997; Shmelkov, 2005). In humans, it is involved in the cell metabolism (Li, 2013) and have a central role in the asymmetric division that is believed to characterise true stemness. Vertebrate GATAs are transcription factors represented by six members, each of which contains a highly conserved DNA binding domain consisting of one or two zinc fingers motif, namely N-terminal zinc finger and C-terminal zinc finger, that bind the consensus DNA sequence GATA (Yue et al., 2014). They are grouped in two subfamilies, the first including GATA1, GATA2 and GATA3, prominently expressed in haematopoietic stem cells, and the second represented by GATA4, GATA5, and GATA6, expressed in various mesoderm- and endoderm- derived tissues (Molkentin, 2000). In particular, GATA2 is relevant in self-renewal and maintenance of developing and adult haematopoietic stem cells (Tsai and Orkin, 1997).

Vertebrate neural stem cell (NSC) markers

Thanks to studies on vertebrate and invertebrate model systems, our understanding of nervous system development and neural stem cell biology has progressed rapidly during the last decade (Brand and Livesey, 2011). NSCs derive from the neural crest and differentiate through a complex combination of signalling pathways in order to form the specific cell types. Commonly used markers for human NSCs identification include Nestin and SOX2, but numerous other NSC markers are already known (Table 2).

ASCL1/Mash1	Noggin
beta-Catenin	Notch-1
BMI-1	Notch-2
Brg1	Nrf2

N-Cadherin	Nucleostemin
Calcitonin R	Numb
CD15/Lewis X	Otx2
CD133	Pax3
CDCP1	Pax6
COUP-TF I/NR2F1	PDGF R alpha
CXCR4	PKC zeta
FABP7/B-FABP	Prominin 2
FABP8/M-FABP	ROR2
FGF R2	RUNX1/CBFA2
FGF R4	RXR alpha/NR2B1
FoxD3	sFRP-2
Frizzled-9	SLAIN1
GCNF/NR6A1	SOX1
GFAP	SOX2
Glut1	SOX9
HOXB1	SOX11
ID2	SOX21
Meteorin	SSEA-1
MSX1	SSEA-4
Musashi-1	TRAF-4
Musashi-2	Vimentin
Nestin	ZIC1

Table 2: human molecular markers associated with neural stem cells.

In particular, Sox2, that belongs to the Sry gene family, contains a DNA-binding domain called high-mobility group (HMG) domain. It is required for inhibiting neuronal differentiation, maintaining proliferative capacity and enhancing neuroectoderm specification that give rise to neural stem cells (Zhang et Cui, 2014). Transcription factor Sox2 is, than, considered one of the earliest markers for vertebrates neural plate cells (Bouzas et al, 2016). Musashi (Msi) genes contain a conserved motif called RRM (ribonucleoprotein-type RNA recognition motifs) with two core domains termed RNP-1 and RNP-2. Vertebrate Msi2 shows sequence homology, similar role and co-expression site with Msi1. An intense expression level of Msi2 was found in proliferating cells of

the central nervous system from postnatal development to adult brain but its expression disappeared in most of the post-mitotic or migrating neurons as neurogenesis proceed (Sakakibara S et al, 2001).

Stemness in *B. schlosseri*

B. schlosseri represents an interesting model organism to study stem cell biology for its continuous and cyclical asexual development in which the colony renews itself, including germline and somatic tissues. In addition, contacting, genetically compatible colonies can fuse and share their tunic and vasculature. In this case, a germline/somatic chimerism/parasitism occurs where undifferentiated circulating cells can proliferate, differentiate and replace the cells of the host (Laird and DeTomaso, 2005; Sabbadin and Zaniolo, 1979, Rinkevich et al, 2013). Furthermore, germline chimerism (the presence of germ cells within gonads, derived from more than a single colony) and germline parasitism (replacement of host gonads by foreign germ cells), following the vascular fusion between genetically distinct colonies, are suggested to involve circulating germline stem or progenitor cells (Rinkevich et al, 2013). In botryllid ascidians, stem cells mediate a myriad of biological events, including blastogenesis, take-over, colonial chimerism, somatic and germline parasitism and diverse regeneration phenomena (Oka and Watanabe, 1957, 1960; Pancer et al, 1995; Rinkevich et al, 1995, 2007b, 2010; Stoner et al, 1999; Lauzon et al, 2002; Voskoboynik et al, 2007; Brown et al, 2009).
Furthermore, botryllid ascidians circulating stem cells are thought to be involved in the process of whole body regeneration (WBR) from (experimentally) fragmented colonies (Oka and Watanabe, 1957, 1960; Sabbadin et al, 1975; Rinkevich et al, 1995; Voskoboynik et al, 2007; Brown et al, 2009).

In *B. schlosseri,* anatomical compartments that harbour and maintain stem cells and regulate their fates *in vivo* via specific spatiotemporal inputs, called stem cell niches, have been recently identified. The anterior ventral area of the endostyle and the subendostylar sinus are regions where stem cells, that can contribute to a wide range of developing tissues, are located (Voskoboynik et al, 2008). Only the cells deriving from the endostyle niche (EN, Voskoboynik et al, 2008) proliferate, migrate in the bud and contribute to the formation of new tissues. In addition to this stem cells harbouring region, in 2013, Rinkevich et al, characterised an additional niche for germline and somatic stem cells in cell aggregates (called cell island, CIs) that line laterally the adult endostyle. These cells were found to cyclically migrate to colonise CIs of developing buds in order to reconstitute both soma and germline in each blastogenetic generation of zooids (Rinkevich et al, 2013). What all the above researches on botryllid stem cells and niche show is that stem cells not reside primarily in the circulation (much like haematopoietic stem cells in mammalian bone marrow) but are sequestered in specific compartments, microenvironment, that are left only after the release of specific signals (Laird et al, 2005). Primordial germ cells (PGC)-like populations is probably the most studied stemness lineage in *B. schlosseri,* with numerous known genes, where male and female germ cells derive from circulating cells. PGC populations survive throughout the life of the colony by repeated weekly migration to newly formed buds. As far as genes involved in stemness are considered, it is commonly accepted that *vasa* gene product is

the best molecular markers for identification of the germline lineage in many organisms (Raz, 2000) and it has been detected in spermatogonia and in the oocytes of *B. schlosseri*; *pl10* is expressed by multipotent bud tissues and by proliferating germline cells (Rosner et al, 2006) including *B. schlosseri*; *Piwi* and *Oct4* are expressed by germ cells and circulating undifferentiated cell even in *B. schlosseri* (Rosner et al, 2009). In *Botryllus primigenus* a gene, called *nanos*, required for primordial germ cells development, is widely expressed in somatic multipotent cells, circulating undifferentiated cells (Kawamura et al, 2011) and germ cells in the gonadal space.

Haematopoietic lineage in *B. schlosseri*

In solitary ascidians, haematopoiesis occurs mainly in close proximity to the pharyngeal vessels, in the so-called "haematopoietic nodules" (Ermak, 1975). The haemolymph of *B. schlosseri* contain a variety of haemocytes. Among the circulating cells, immunocytes are directly involved in immune defense. In *B. schlosseri*, they are represented by phagocytes, that include hyaline amebocytes (*ha*) and macrophage-like cells (*mlc*), and by cytotoxic cells that include granular amebocytes (*ga*) and morula cell (*mc*). Circulating *ha* change their morphology and decrease their frequency at the TO phase, while *mlc* increase their number. This, and the sharing of common enzymatic content, suggest that they probably represent different functional stages of the same cell type (Ballarin et al, 1993, 1994). Furthermore, *mc*, that are the most abundant circulating cell type, have in *ga* their precursors (Ballarin et al, 1993; Cima et al, 2001). Despite all the above-reported suggestions, there is a general lack of molecular data supporting the above assumptions. The differential gene expression patterns that direct haematopoietic precursors towards differentiation are not well characterised although of great interest. Recently, novel stem cell niches have been identified in the endostyle and in the cell islands, the cells of which proliferate and migrate to regenerating organs in developing buds (Voskoboynik et al, 2008). However, the relationships between these stem cells and the circulating cells are still unclear. Haemoblasts are considered undifferentiated circulating cells. They are small (4-6 μm in diameter), round cells with a high nucleus/cytoplasm ratio and a hyaline cytoplasm with few organelles. They rapidly proliferate and are thought to give rise to terminally differentiated cells acting as multipotent somatic cells (Ballarin and Cima, 2005; Kawamura et al, 2006). In *B. schlosseri*, during the generation change, an increase in the number of haemoblasts occurs that will replace the circulating cells died by apoptosis (Ballarin et al, 2008). Only a protein recognised by the antibody anti-CD34, an evolutionally conserved molecular marker for undifferentiated hematopoietic stem cells, has been identified (Ballarin and Cima, 2005).

Neural lineage in *B. schlosseri*

The ascidian central nervous system (CNS) develops from a neural plate, like vertebrates, and consists of less than 100 neurons and 250 glial cells (Meinertzhagen and Okamura, 2001). Despite the numerous efforts in the comparison of ascidian and vertebrates nervous system (fig 3A), the poor anatomical description of the ascidian nervous system and the small number of available marker genes renders it a difficult

task (Lemaire et al, 2002). Its already known that, in the ascidian larva, *otx* gene expressions was located in the sensory vesicle (Lemaire et al, 2002), *pax2/5/8* in the neck, the ascidian equivalent of the vertebrate midhindbrain boundary (Wada et al, 1998) and *hox* genes expression was found in the visceral ganglion (Katsuyama et al, 1995; Locascio et al, 1999) and in the anterior tail nerve cord (Gionti et al, 1998) (fig 3B).

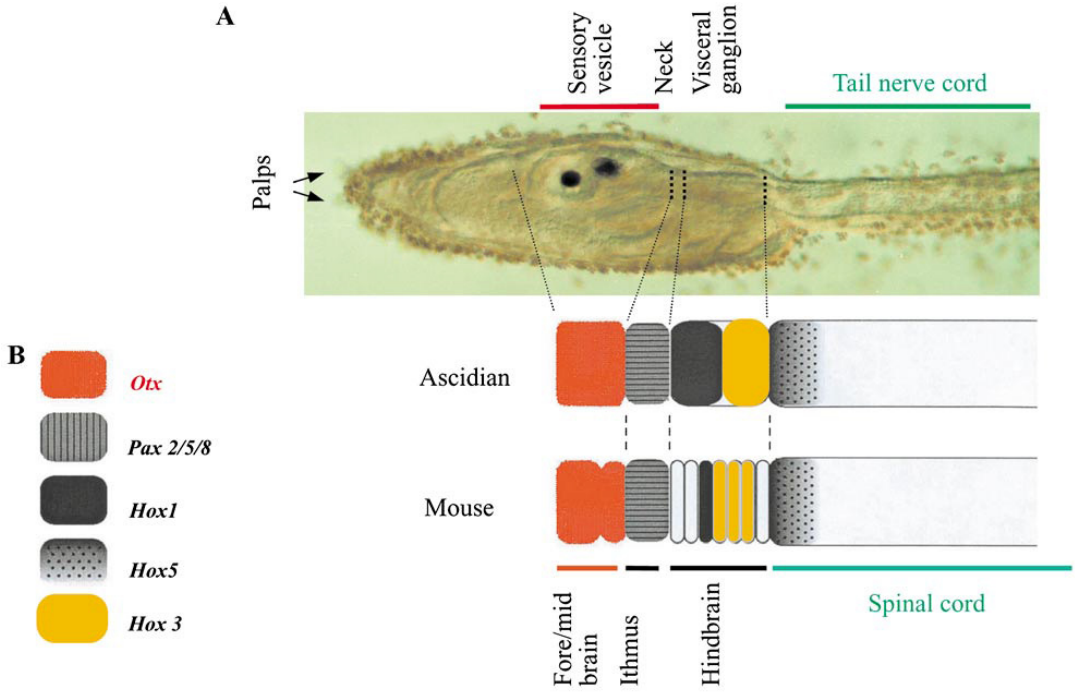


Fig.3: A: lateral view of the anterior section of the *C. intestinalis* larvae with the neural plate regionalization in sensory vesicle, neck, visceral ganglion and tail nerve cord. **B:** similar expression pattern of conserved genes between ascidians and vertebrates, with the homologues vertebrate territories (modified from Lemaire et al, 2002).

The ascidian adult CNS shows a different organisation: a neural complex composed of a cerebral ganglion and a neural gland, the latter homologue to vertebrate hypophysis (Manni et al, 1999). Interestingly, in the solitary ascidian *Ciona intestinalis*, the presence of neuronal progenitor cells able to regenerate the brain after extirpation was demonstrated (Bollner et al, 1995). This process, in addition to the bud primordial cells able to give rise to the neural complex in *B. schlosseri*, appears to be related to the presence of a pool of long-lived stem cells. In the colonial species *B. schlosseri*, the mechanisms of neurogenesis during embryogenesis and blastogenesis are comparable (fig 4). Previous studies (Kawashima et al, 2000), revealed the presence of a Msi homologous sequence in *C. intestinalis* and *Halocynthia roretzi* adult organisms (CiMsi and HrMsi). In subsequent studies, these genes were considered misnamed as the phylogenetic reconstruction of the RRM domain inserted these sequences into the Dazap1 group with strong statistical support (Gasparini et al, 2011).

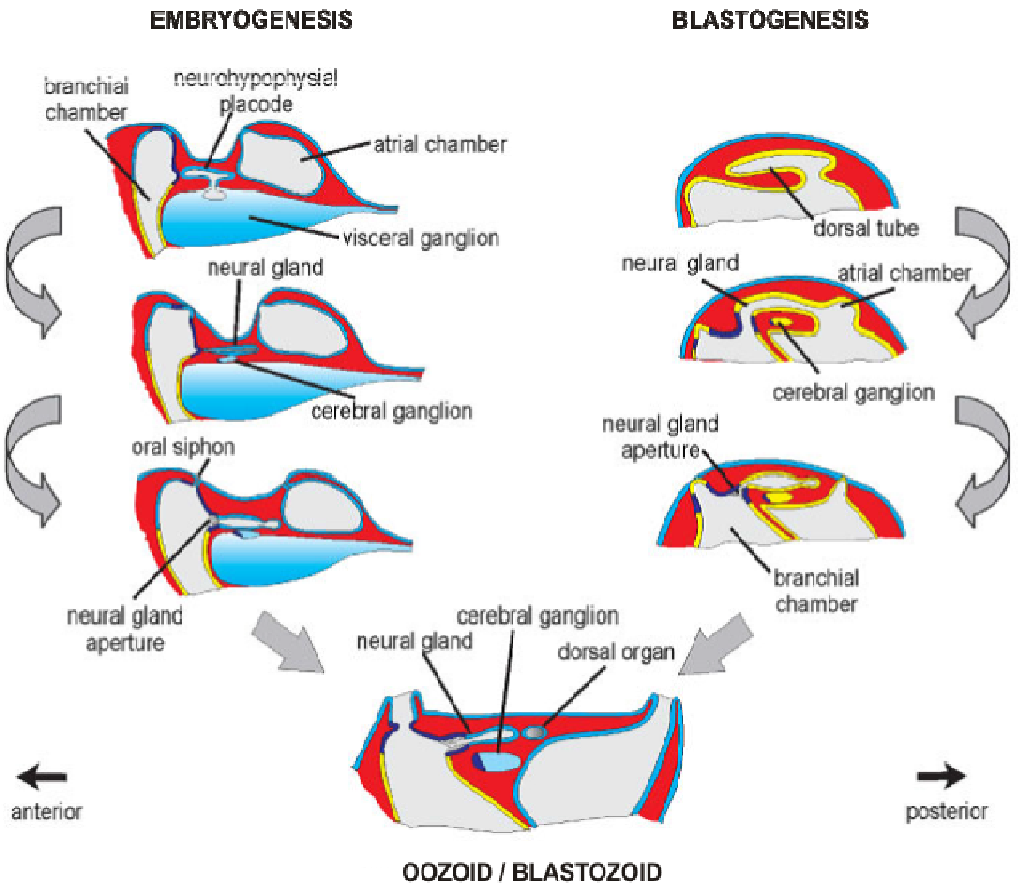


Fig.4: comparison of neurogenesis during embryogenesis and blastogenesis in the zooid of *B. schlosseri.* Germ layer derivation is indicated by different colours (light blue: ectoderm, red: mesoderm, yellow: endoderm). Oozooids and blastozooids present the same organisation in morphology, tissue components and cell types (modified from Manni and Burighel, 2006).

Aim of the study

The aim of this study was the identification of stemness and differentiation molecular markers in the colonial ascidians *B. schlosseri*. We know that in *B. schlosseri*, numerous processes that involve stem cells occurs. During the colonial blastogenetic cycle we can observe palleal budding, a physiological process that cyclically happen, in which undifferentiated cells generate the new secondary bud. After injury, even a vascular budding may happened, with a complete blastozooid regeneration from undifferentiated circulating cells. Despite the numerous morphological evidence of the presence of stem cells, up to now, there are few molecular data in support of this. In the past, the presence of undifferentiated cells able to proliferate and produce all the circulating cells was suggested by the description of cells with high nucleus/cytoplasm ratio, few cytoplasm organelles and absence of specific functions, called haemoblast. Recently, Voskoboynik (2008) and Rinkevick (2013) described the endostyle and the cell island, as

haematopoietic sites. Today there is the need of tools that help in the identification of stem cells in order to better understand the molecular mechanism that control stemness and differentiation. From the vertebrate, and in particular mammalian, stem cell research, we know numerous molecular markers. The availability of an assembled transcriptome of *B. schlosseri*, recently obtained at the Dept of Biology, University of Padova, and of the first draft of the genome, recently published (Voskoboynik et al, 2013) allowed me to start a screening of orthologues of mammalian genes to identify putative haematopoietic and neural progenitors molecular markers in our model organism. With this work I aim to integrate the morphological data with robust molecular analysis. For the haematopoietic progenitor cells we identified and characterised four orthologues genes, called *bsabcg2, bscd133, bsgata1/2/3* and *bsgata4/5/6,* on which we performed a phylogenetic reconstruction, a spatial localisation and a time-expression analysis. As for neurogenesis, we know that orthologues of vertebrate neural genes are expressed in the same nervous territories of ascidian larvae, reflecting a conserved gene expression. Here I aimed to investigate if the orthologues of *pax2/5/8, hox1* and *hox3* are transcribed during the *Botryllus* blastogenetic cycle, in association with bud neurogenesis, and if their expression is maintained even in the adult blastozooid. Moreover, I also searched for orthologues of vertebrate genes used for neural stem cells identification and I focused on Sox2 and Msi2 genes. For all the *B. schlosseri* putative molecular markers (*bspax2/5/8, bshox1, bshox3, bssox2* and *bsmsi2)* I obtained phylogenetic reconstruction and I studied the transcript spatial and temporal expression pattern.

References

Ballarin L, Cima F, Sabbadin A. 1993. Hystoenzymatic staining and characterization of the colonial ascidian *Botryllus schlosseri* hemocytes. Boll Zool 60: 19-24

Ballarin L, Cima F, Sabbadin A. 1994. Phagocytosis in the colonial ascidian *Botryllus schlosseri* (Urochordata: Ascidiacea). Anim Biol 3: 41-48

Ballarin, L., Cima F. 2005. Cytochemical properties of *Botryllus schlosseri* haemocytes: indications for morpho-functional characterisation. Eur J Histochem 49: 255-264

Ballarin L, Menin A, Tallandini L, Matozzo V, Burighel P, Basso G, Fortunato E, Cima F. 2008b. Haemocytes and blastogenetic cycle in the colonial ascidian *Botryllus schlosseri*: a matter of life and death. Cell Tissue Res 331: 555–564

Barker N, Clevers H. 2000. Catenins, wnt signaling and cancer. Bioessays *22*: 961–965

Barker N, vanes JH, Kuipers J, Kujala P, Van den Born M, Cozijnsen M, Haegebarth A, Korving G, Begthel H, Peters PJ, Clevers H. 2007. Identification of stem cell in small intestine and colon by marker gene Lgr5. Nature 449: 1003–1007

Berrill NJ. 1951. Regeneration and budding in Tunicates. Biol Rev 26: 456–475

Bollner T1, Howalt S, Thorndyke MC, Beesley PW. 1995. Regeneration and post-metamorphic development of the central nervous system in the protochordate *Ciona intestinalis*: a study with monoclonal antibodies. Cell Tissue Res 279: 421-32

Bouzas SO, Marini MS, Torres Zelada E, Buzzi AL, Morales Vicente DA, Strobl-Mazzulla PH. 2016. Epigenetic activation of Sox2 gene in the developing vertebrate neural plate. Mol Biol Cell 27: 1921-7

Brand AH, Livesey FJ. 2011. Neural stem cell biology in vertebrates and invertebrates: more alike than different? Neuron 70: 719-729

Brown FD, Keeling EL, Le AD, Swalla BJ. 2009. Whole body regeneration in a colonial ascidian, *Botrylloides violaceus*. J Exp Zool B Mol Dev Evol 312: 885–900

Burighel P, Lane NJ, Zaniolo G, Manni L. 1998. Neurogenic role of the neural gland in the development of the ascidian, *Botryllus schlosseri* (Tunicata, Urochordata). J Comp Neurol 394: 230-41

Calvi LM, Adams GB, Weibrecht KW, Weber JM, Olson DP, Knight MC, Martin RP, Schipani E, Divieti P, Bringhurst FR, et al. 2003. Osteoblastic cells regulate the haematopoietic stem cell niche. Nature 425: 841–846

Cima F, Perin A, Burighel P, Ballarin L. 2001. Morpho-functional characterisation of haemocytes of the compound ascidian *Botrylloides leachi* (Tunicata, Ascidiacea). Acta Zool (Stockh) 82: 261-74

Cotsarelis G, Sun TT, Lavker RM. 1990. Label-retaining cells reside in the bulge area of pilosebaceous unit: implications for follicular stem cells, hair cycle, and skin carcinogenesis. Cell 61: 1329–1337

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

De Tomaso AW, Nyholm SV, Palmeri KJ, Ishizuka KJ, Ludington WB, Mitchel K, Weissman IL. 2005. Isolation and characterization of a protochordate histocompatibility locus. Nature 438: 454-9

Doetsch F, Caille I, Lim DA, Garcia-Verdugo JM, Alvarez-Buylla A. 1999. Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. Cell 97: 703–716

Ermak TH. 1975. An autoradiographic demonstration of blood cell renewal in *Styela clava* (Urochordata: Ascidiacea). Experientia 31: 837-838

Gasparini F, Shimeld SM, Ruffoni E, Burighel P, Manni L. 2011. Expression of a Musashi-like gene in sexual and asexual development of the colonial chordate *Botryllus schlosseri* and phylogenetic analysis of the protein group. J Exp Zool B Mol Dev Evol 316: 562-73

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

Harada H, Kettunen P, Jung HS, Mustonen T, Wang YA, et al. 1999. Localization of putative stem cells indental epithelium and their association with notch and FGF signaling. J Cell Biol 147: 105-120

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

Kawamura K, Takeoka S, Takahashi S, Takeshi S. 2006. In vitro culture of mesenchymal lineage cells established from the colonial tunicate *Botryllus primigenus*. Zoolo Sci 23: 245-254

Kawamura K, Tiozzo S, Manni L, Sunanaga T, Burighel P, De Tomaso AW. 2011. Germline cell formation and gonad regeneration in solitary and colonial ascidians. Dev Dyn 240: 299-308

Laird DJ, De Tomaso AW, Weissman IL. 2005. Stem cells are units of natural selection in a colonial ascidian. Cell 123: 1351-60

Laird DJ and De Tomaso AW. 2005. Predatory stem cells in the non-zebrafish chordate, *Botryllus schlosseri*. Zebrafish 1: 357-61

Lauzon RJ, Ishizuka KJ, Weissman IL. 2002. Cyclical generation and degeneration of organs in a colonial urochordate involves crosstalk between old and new. A model for development and regeneration. Dev Biol 249: 333–348

Lemaire P, Bertrand V, Hudson C. 2002. Early Steps in the Formation of Neural Tissue in Ascidian Embryos. Dev Biol 252: 151-169

Li Z. 2013. CD133: a stem cell biomarker and beyond. Exp Hematol Oncol 2: 17

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

Manni L, Burighel P. 2006. Common and divergent pathways in alternative developmental processes of ascidians. Bioessays 28: 902-912

Manni L, Zaniolo G, Cima F, Burighel P, Ballarin L. 2007. *Botryllus schlosseri*: a model ascidian for the study of asexual reproduction. Dev Dyn 236: 335-352

Manni L, Lane NJ, Sorrentino M, Zaniolo G, Burighel P. 1999. Mechanism of neurogenesis during the embryonic development of a tunicate. J Comp Neurol 412: 527-41

Manni L, Gasparini F, Hotta K, Ishizuka KJ, Ricci L, Tiozzo S, Voskoboynik A, Dauga D. 2014. Ontology for the asexual development and anatomy of the colonial chordate *Botryllus schlosseri*. PLoS One 9: e96434

Meinertzhagen IA and Okamura Y. 2001. The larval ascidian nervous system: The chordate brain from its small beginnings. Trends Neurosci 24: 401–410

Molkentin JD. 2000. The zinc finger-containing transcription factors GATA-4, -5, and -6. Ubiquitously expressed regulators of tissue-specific gene expression. J Biol Chem 275: 38949-38952

NIH Stem Cell Information Home Page. In Stem Cell Information [World Wide Web site]. Bethesda, MD: National Institutes of Health, U.S. Department of Health and Human Services, 2016

Oka H, Watanabe H. 1957. Vascular budding, a new type of budding in *Botryllus*. Biol bull 112: 225-240

Oka H, Watanabe H. 1960. Problems of colony specificity in compound ascidians. Bull Mar Biol Stat Asamushi, Tohoku Univ. 10: 153–155

Palmer TD, Takahashi J, Gage FH. 1997. The adult rat hippocampus contains primordial neural stem cells. Molecular and cellular neurosciences. 8: 389–404

Pancer Z, Gershon H, Rinkevich B. 1995. Coexistence and possible parasitism of somatic and germ cell lines in chimeras of the colonial urochordate *Botryllus schlosseri*. Biol Bull 189: 106–112

Pazhanisamy S. 2013. Adult stem cell and embryonic stem cell markers. Mater methods. 3: 200

Potten CS, Owen G, Booth D. 2002. Intestinal stem cells protect their genome by selective segregation of template DNA strands. Journal of cell science. 115: 2381–2388

Raz E. 2000. The function and regulation of vasa-like genes in germ-cell development. Genome Biol. 1: 1017.1–1017.6

Rinkevich B, Shlemberg Z, Fishelson L. 1995. Whole-body protochordate regeneration from totipotent blood cells. Proc Natl Acad Sci USA. 92: 7695-9

Rinkevich B. 2002. The colonial urochordate *Botryllus schlosseri*: from stem cells and natural tissue transplantation to issues in evolutionary ecology. Bioessays 24: 730-740

Rinkevich Y, Paz G, Rinkevich B, Reshef R. 2007. Systemic bud induction and retinoic acid signaling underlie whole body regeneration in the urochordate *Botrylloides leachi*. PLoS biology 5: e71

Rinkevich Y, Rosner A, Rabinowitz C, Lapidot Z, Moiseeva E, Rinkevich B. 2010. Piwi positive cells that line the vasculature epithelium, underlie whole body regeneration in a basal chordate. Dev Biol 345: 94-104

Rinkevich Y, Voskoboynik A, Rosner A, Rabinowitz C, Paz G, Oren M, Douek J, Alfassi G, Moiseeva E, Ishizuka KJ, Palmeri KJ, Weissman IL, Rinkevich B. 2013. Repeated, long-term cycling of putative stem cells between niches in a basal chordate. Dev Cell 24: 76-88

Rosner A, Paz G, Rinkevich B. 2006. Divergent roles of the DEAD-box protein BS-PL10, the urochordate homologue of human DDX3 and DDX3Y proteins, in colony astogeny and ontogeny. Dev Dyn 235: 1508–1521

Rosner A, Moiseeva E, Rinkevich Y, Lapidot Z, Rinkevich B. 2009. Vasa and the germ line lineage in a colonial urochordate. Dev Biol 331: 113-128

Sabbadin A, Zaniolo G, Majone F. 1975. Determination of polarity and bilateral asymmetry in palleal and vascular buds of the ascidian *Botryllus schlosseri*. Dev Biol 46: 79-87

Sakakibara S, Nakamura Y, Satoh H, Okano H. 2001. Rna-binding protein Musashi2: developmentally regulated expression in neural precursor cells and subpopulations of neurons in mammalian CNS. J Neurosci 21: 8091-8107

Scharenberg CW, Harkey MA, Torok-Storb B. 2002. The ABCG2 transporter is an efficient Hoechst 33342 efflux pump and is preferentially expressed by immature human hematopoietic progenitors. Blood 99: 507-512

Shen Q, Goderie SK, Jin L, Karanth N, Sun Y, Abramova N, Vincent P, Pumiglia K, Temple S. 2004. Endothelial cells stimulate self-renewal and expand neurogenesis of neural stem cells. Science 304: 1338–1340

Shmelkov SV, StClair R, Lyden D, Rafii S. 2005. AC133/CD133/Prominin-1. J Biochem Cell Bio 37: 715-719

Stoner DS, Rinkevich B, Weissman IL. 1999. Heritable germ and somatic cell lineage competitions in chimeric colonial protochordates. Proc Nat Acad Sci U.S.A. 96: 9148–9153

Tiozzo S, Brown FD, De Tomaso AW. 2008. Regeneration and stem cells in ascidians. Stem cells 6: 95-112

Toren A, Bielorai B, Jacob-Hirsch J, Fisher T, Kreiser D, Moran O, Zeligson S, Givol D, Yitzhaky A, Itskovitz-Eldor J, Kventsel I, Rosenthal E, Amariglio N, Rechavia G. 2005. CD133-positive hematopoietic stem cell "stemness" genes contain many genes mutated or abnormally expressed in leukemia. Stem Cells 23: 1142–1153

Tsai FY, Orkin SH. 1997. Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. Blood 89: 3636-3643

Tumbar T, Guasch G, Greco V, Blanpain C, Lowry WE, Rendl M, Fuchs E. 2004. Defining the epithelial stem cell niche in skin. Science 303: 359–363

Villadsen R, Fridriksdottir AJ, Ronnov-Jessen L, Gudjonsson T, Rank F, LaBarge MA, Bissell MJ, Peterson OW. 2007. Evidence for a stem cell hierarchy in the adult human breast. J Cell Biol 177: 87–101

Voskoboynik A, Simon-Blecher N, Soen Y, Rinkevich B, De Tomaso AW, Ishizuka KJ, Weissman IL. 2007. Striving for normality: whole body regeneration through a series of abnormal generations. FASEB J. 21: 1335-44

Voskoboynik A, Soen Y, Rinkevich Y, Rosner A, Ueno H, et al. 2008. Identification of the endostyle as a stem cell niche in a colonial chordate. Cell Stem Cell 3: 456-464

Wada H, Saiga H, Satoh N, Holland PW. 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

Zhou S, Schuetz JD, Bunting KD, et al. 2001. The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. Nat Med 7: 1028-1034

Weigmann A, Corbeil D, Hellwig A, Wuttner WB. 1997. Prominin, a novel microvilli-specific polytopic membrane protein of the apical surface of epithelial cells, is targeted to

plasmalemmal protrusions of non-epithelial cells. Proc Natl Acad Sci U.S.A. 94: 12425–12430

Yin AH, Miraglia S, Zanjani ED, Almeida-Porada G, Ogawa M, Leary AG, Olweus J, Kearney J, Buck DW. 1997. AC133, a novel marker for human hematopoietic stem and progenitor cells. Blood 90: 5002–5012

Yu J, Thomson JA. 2008. Pluripotent stem cell lines. Genes Dev. 22: 1987-1999

Yue F, Zhou Z, Wang L, Wang M, Song L. 2014. A conserved zinc finger transcription factor GATA involving in the hemocyte production of scallop *Chlamys farreri*. Fish Shellfish Immunol 39: 125-35

Zhang J, Niu C, Ye L, Huang H, He X, Tong WG, Ross J, Haug J, Johnson T, Feng JQ, et al. 2003. Identification of the haematopoietic stem cell niche and control of the niche size. Nature 425: 836–841

Zhang S and Cui W. 2014. Sox2, a key factor in the regulation of pluripotency and neural differentiation. World J Stem Cells 6: 305–311

Zhao W, Ji X, Zhang F, Li L, Ma L et al. 2012. Embryonic Stem Cell Markers. Molecules 17: 6196-6236

<u>Research article:</u> Identification and expression studies of four orthologues of vertebrate haematopoietic molecular marker in the urochordate *Botryllus schlosseri*: tissue and temporal analyses of *bsabcg2, bscd133, bsgata1/2/3* and *bsgata4/5/6.*

Introduction

Vertebrate circulating cells, such as erythrocytes, lymphocytes and phagocytes, are examples of cells with specific functions and tissues location. These cells reach their functionality through various steps in which they progress towards their specific fate. Differentiation implies modification in shape and dimension, and the acquisition of the ability to respond to extracellular signals in order to acquire the features of specialised cells, according to the information in their genetic and epigenetic program. The combination of internal (i.e., DNA-encoded) and external signals (including secreted molecules, physical contact with surrounding cells, and certain soluble factors in the microenvironment) induces the differentiation process and the expression of "luxury" molecular patterns (molecular markers), associated with each type of differentiated cell, useful for the identification and the isolation of specific cell populations. The expression and the combination of specific molecular markers are related to the degree of differentiation of the cells and allow the discrimination between a scarcely differentiated progenitor and a terminally differentiated cell (Nemeth and Karpati, 2014).

Botryllus schlosseri is a cosmopolitan colonial ascidian living in shallow waters. A colony derives from the metamorphosis of a single larva into a first zooid which starts to reproduce asexually through palleal budding (Manni et al., 2007). In a colony, adult zooids are grouped in star-shaped systems of 8-12 individuals sharing a common cloacal siphon in the center of the system. Three blastogenetic generations are usually present in a colony, i.e.: adult, filter-feeding zooids, buds on zooids and budlets on buds. A colony performs cyclical (weekly at 20°C) generation changes, or take-overs (TOs), that allow its recurrent rejuvenation (Manni et al., 2007; Ballarin et al., 2010). A colonial blastogenetic cycle can be defined as the period of time between a TO and the next and can be divided in: post-TO (or beginning of a new cycle, developmental phase 9/7/1), mid-cycle (MC, developmental phases 9/8/2, 9/8/3, 9/8/4), pre-TO (developmental phase 9/8/5) and TO (the end of the cycle, developmental phase 11/8/6). The combination of the three numbers corresponding to the developmental stage of the adult zooids, primary and secondary bud, respectively (Manni et al, 2007). In the course of the generation change, tissues of adult zooids undergo diffuse apoptosis and cells and corpses are rapidly ingested by phagocytes having left the circulation and infiltrating the tissues (Cima et al., 2003; Manni et al., 2007; Ballarin et al., 2008a,b). In addition to palleal budding, a colony, deprived of its zooids and buds, can originate new zooids by vascular budding (Sabbadin et al, 1975). The latter is a physiological process in the Japanese species *Botryllus primigenus*, involving the

gathering of circulating cells in differentiation niches at the bases of the peripheral ampullae (Oka and Watanabe, 1957).

In both cases, blastogenesis (by either palleal or vascular budding) requires the progressive differentiation of pluri/multipotent stem cells that can persist in the adult in specific niches (Voskoboynik et al, 2008; Rinkevich et al, 2013) and can be shared by all the zooids and buds through the circulation (Rinkevich et al, 2013). These stem cells are also responsible of the whole body regeneration reported in some species of botryllid ascidians (Rinkevich et al, 1995; Voskoboynik et al, 2007; Brown et al, 2009).

The trans-colonial migration of stem cells, through the fused vasculature of countering, genetically compatible colonies, and their long persistence in heterologous adult *B. schlosseri* colonies, has been clearly demonstrated by various authors (Sabbadin and Zaniolo, 1979; Sabbadin and Astorri, 1988; Laird and DeTomaso, 2005).

As regards haematopoietic cells, we know that new waves of young haemocytes are released in the circulation from haematopoietic sites, presumably the endostyle (Voskoboynnik et al, 2008; Kawamura et al, 2008), at the beginning of a new blastogenetic cycle, to replace the circulating cells that die by apoptosis at TO and are cleared by phagocytes (Ballarin et al, 2008b; Franchi et al, 2016). Up to now, however, few molecular markers have been used to characterise these cells. Laird et al (2005) isolated cells on the basis of their high aldehyde dehydrogenase activity, whereas a CD133 homologue was reported in a subset of circulating cells, assumed to be progenitor cells, by Braden et al (2014).

In the attempt to better characterise the haematopoietic cells of *B. schlosseri*, we examined available databases looking for genes/transcript with similarity in gene /transcript involved in haematopoietic differentiation in vertebrates. We were able to identify and characterise four transcripts showing similarity to mammalian ABCG2, CD133, Gata1/2/3 and Gata4/5/6. We studied the dynamics of transcription during the blastogenetic cycle, carried out a phylogenetic analysis on the putative proteins and searched for the locations of specific mRNAs through *in situ* hybridization (ISH). The domain structure and organisation of all the considered proteins result well conserved during the evolution and our data suggest that such molecules can be considered haematopoietic stem cell markers also in *B. schlosseri*.

Materials and methods

Animals

Colonies of *B. schlosseri* (Tunicata, Ascidiacea) were collected near Chioggia, in the southern part of the Lagoon of Venice. They were reared according to Gasparini et al. (2015), stuck to glass slides (5 x 5 cm), in aerated aquaria filled with 0.45-µm filtered seawater (FSW) changed every other day, at the constant temperature of 19°C, and fed with Liquifry marine (Liquifry Co., Dorking, England). At this temperature, colonies undergo weekly TOs during which old

zooids are resorbed and replaced by their buds in 24-36 h. Colonial developmental phases more than one day from the preceding or the following generation change, are collectively called MC (Manni et al., 2007).

Primer design, RNA extraction, cDNA synthesis, cloning and sequencing

Sequences of interest were identified through a comparison of recent *B. schlosseri* transcriptomes (Campagna et al, 2016) and *Botryllus* genome (Voskoboynik et al., 2013) with the sequences of the vertebrate genes of interest. On that sequences, specific primers (Table 1) for PCR amplification and 5'/3' RACE were designed. We focused our attention on the sequences of orthologues of ABCG2, CD133, GATA1/2/3 and GATA4/5/6, named BsABCG2, BsCD133, BsGATA1/2/3 and BsGATA4/5/6, respectively.

Total RNA was isolated from *B. schlosseri* colonies with the SV total RNA isolation system (Promega); its purity was determined spectrophotometrically by the $A_{260/A280}$ and $A_{260/A230}$ ratio. The integrity of RNA preparation was checked by visualisation of the rRNA in ethidium bromide-stained 1.5% agarose gels. The first strand of cDNA was reverse-transcribed from 1 μg of total RNA according to the Improm II manual (Promega). cDNA amplification was performed with Go-Taq Polymerase (Promega; 5 U/μl) with the following cycling parameters: 94°C for 2 min, 45 cycles of 94°C for 30 s, melting temperature (T_m) for 30 s (T_m for the various primers is indicated in Table 1), 72°C for 1 min, and a last step at 72°C for 10 min. The 5' and 3' rapid amplification of the cDNA ends (RACE) was performed using the 2nd Generation of the 5'/3' RACE Kit (Roche). Specific primers (Table 1) were then designed for nested PCR with anchor forward/reverse primer according to the manufacturer's instruction (Roche).

Amplicons were subjected to electrophoresis and the corresponding bands were purified with ULTRAPrep Agarose Gel Extraction Mini Prep kit (AHN Biotechnologie), ligated in pGEM T-Easy Vector (Promega) and cloned in DH-5α *Escherichia coli* cells (Tang et al, 1994). In order to confirm the sequences and their expression, positively screened clones were sequenced at BMR Genomics (University of Padova) on an ABI PRISM 3700 DNA Analyzer (Applied Biosystems). Gene reconstructions were based on *B. schlosseri* genome database using Spidey's algorithm (http://www.ncbi.nlm.nih.gov/spidey/).

Electrophoresis and immunoblot analysis

Total proteins were obtained from subclones, 2–3 systems in size, of various colonies at different phases of the blastogenetic cycle and from haemocyte samples obtained from colonies at various phases of the blastogenetic cycle. They were homogenised in lysis buffer (50 mM Tris–HCl, 0.25 M sucrose, 1% SDS, 1 mg/ml pepstatin, 1 mg/ml leupeptin, 40 mg/ml PMSF, 2 mM Na orthovanadate, 10 mM NaF, 0.1% NP-40, 5 mM EDTA, 5 mM N-ethylmaleimide) and

centrifuged at 10,000 × g for 10 min. The supernatants of colony lysates were frozen in liquid nitrogen and stored at −80°C until use. Pellets were treated with 6 M Urea and centrifuged at 10,000 x g for 10 min in order to obtain membrane proteins that were frozen as described above. Samples, diluted 1:1 in sample buffer (SB: 0.1 M Tris–HCl, 2% sodium dodecyl sulphate (SDS), 10% glycerol), were boiled for 7 min. Each well of a 4-20% SDS-PAGE minigel received a volume of supernatant equivalent to 10 µg of proteins. Low molecular weight marker proteins (BioRad) were added to a reference well and proteins were run at the constant current of 80 V for approximately 2 h. After electrophoresis, proteins were transferred to 0.2 µM Electran nitrocellulose membrane (BDH), according to Towbin et al. (1979), with 25 mM Tris, 160 mM glycine, 20% methanol and 0.7 mM SDS as transfer buffer. After blotting, membranes were thoroughly washed in Tris-buffered saline (TBS: 50 mM Tris–HCl, 150 mM NaCl, pH 7.4), incubated for 30 min in TBS containing 5% powdered milk and probed overnight with 1 µg/ml of mouse anti-ABCG2 (Abcam) or rabbit anti-CD133 (Abcam) primary antibodies. After further extensive washing in TBS, membranes were overlaid for 1 h with goat anti-mouse IgG or goat anti-rabbit IgG, conjugated with peroxidase (BioRad), diluted 1/1000 in TBS according to the manufacturer's instructions. Immunogenic bands were revealed with 0.63 mM DAB in TBS, containing 0.004% hydrogen peroxide.

Sequence alignment and phylogenetic analyses

Amino acid sequences of the proteins of interest were obtained by *in silico* translation. For sequence alignment and phylogenetic analyses, sequences were compared with those of the corresponding proteins from deuterostomes and protostomes. Alignments were carried out with Clustal W and analyzed with the MEGA 6 program to infer evolutionary relationships among the various orthologous isoforms.

Phylogenetic reconstructions were performed according to neighbor-joining (NJ; Saitou and Nei, 1987) method. The robustness of tree topologies was tested by the nonparametric bootstrap test (Felsenstein, 1985), with 1000 replicates.

Transcripts localisation by in situ hybridisation (ISH)

For the localisation of mRNAs, sense and antisense probes for BsABCG2, BsCD133, BsGATA1/2/3 and BsGATA4/5/6 transcripts were obtained using T7 RNA- and SP6 RNA-polymerase. Probes were further purified with mini-Quick Spin Columns (Roche) and used in both haemocytes and whole colonies.

Haemolymph was collected with a glass micropipette after puncturing, with a fine tungsten needle, the tunic marginal vessels of the colonies. It was diluted 1:1 in 0.38% Na-citrate in FSW (as anti-agglutinating agent), pH 7.5, and then centrifuged at 780 x g for 10 min, at room temperature. The pellet was then re-suspended in FSW to get a final concentration of $5x10^5$

haemocytes/ml. Haemocytes were left to adhere for 30 min on Superfrost Plus (Menzel-Glaser) slides and fixed in freshly prepared MOPS buffer (0.1 M MOPS, 1 mM $MgSO_4$, 2 mM EGTA, 0.5 M NaCl) and 4% paraformaldehyde for 30 min. Whole colonies, at various phases of the blastogenetic cycle, were fixed in 4% paraformaldehyde, 1% glutaraldehyde in 0.2M cacodylate buffer containing 1.6% NaCl, dehydrated and clarified by treatment with xylene. After pre-hybridisation in Hybridisation Cocktail 50% Formamide (Amresco) for 1 h at 55°C, cells and colonies were incubated with sense and antisense probes (2 µg/ml digoxigenin-labelled riboprobe in Hybridisation Cocktail) overnight, at 55°C. Nonspecific single filament RNA transcripts were removed by incubation for 30 min at 37°C in RNAse solution. Haemocytes were then incubated with 5% powdered milk in phosphate-buffered saline (PBS: 8 g/l NaCl, 0.2 g/l KCl, 0.2 g/l KH_2PO_4, 1.15 g/l Na_2HPO_4, pH 7.2, and colonies in 1% powdered milk in TBS, for 1h, to prevent unspecific labelling. Samples were then treated with 1:3000 anti-digoxigenin-AP FAB antibodies (Roche), overnight, at 4°C. Positivity was revealed by the incubation in 3% NBT-BCIP solution (Sigma), substrate for alkaline phosphatase. Haemocytes were finally mounted with Acquovitrex (Carlo Erba) whereas colonies were dehydrated, included in Paraplast Plus Xtra (Sigma), cut into 7-µm sections with a Jung microtome, mounted with Eukitt and finally observed under the light microscope at 1250x.

Quantitative real-time PCR (qRT-PCR)

To estimate the level of transcripts of *bsabcg2*, *bscd133*, *bsgata1/2/3* and *bsgata4/5/6* during the blastogenetic cycle, we performed qRT-PCR with the SYBR green method (KAPA SYBR FAST qPCR Kit Master Mix (2X) Universal, KAPA Biosystems). mRNA was extracted from 3 pools of 10 systems derived from 10 different specimens each, at various phases of the blastogenetic cycle, to evaluate transcription changes in physiological conditions. Considered phases were 9/7/1, 9/8/2, 9/8/3, 9/8/4, 9/8/5, 11/8/6. Forward and reverse primers for *bsabcg2*, *bscd133*, *bsgata1/2/3* and *bsgata4/5/6* and b*s-βactin*, the latter used as housekeeping gene, were synthesized by Sigma Aldrich (Table 1). The stable expression level of *bsβ-actin* (Campagna et al., 2016) explains the choice of cytoplasmic actin as reference gene for qRT-PCR experiments. qRT-PCR analyses were performed using Applied Biosystems 7500 Fast Real-Time PCR System.

cDNA synthesis was carried out as described above and qRT-PCR was performed with the following cycling parameters: 95°C for 3 min, 45 cycles of 95°C for 15 s and 60°C for 1 min. Each set of samples was run three times and each plate contained cDNA from three different biological samples (n=3). The method of $2^{-\Delta\Delta Ct}$ (Livak and Schmittgen, 2001) was used to estimate gene transcription level. The amounts of transcripts in different conditions were normalized to *β-actin* in order to compensate for variations in the amounts of cDNA.

Statistical analysis

Each experiment was replicated at least three times with three independent samples (n = 3); data are expressed as mean ± SD. The significance of the qRT-PCR results was assessed with the ANOVA; means were compared with the Duncan's test (Snedecor and Cochran, 1980).

Results

Identification and characterisation of transcripts for putative molecular markers involved in haematopoietic differentiation

The *bsabcg2* transcript is 2018 bp in length and contains a coding sequence (CDS) of 1899 bp with 5' UTR of 119 bp. The gene includes a single exon carrying the ATG start and TAA stop codons.

The CDS of *bscd133* transcript is 1665 bp length, with the corresponding gene organized in 12 exons with the ATG start codon located in the first exon and the TAA stop codon in the last one. All the introns are provided of the canonical GT and AG splicing signal consensus.

As for GATA gene family transcripts, in *B. schlosseri* transcriptome and genome, we found one transcript orthologue of vertebrate *gata1/2/3* and one orthologue of vertebrate *gata4/5/6. bsgata1/2/3* transcript is 2662 bp in length with a CDS of 1905 bp, a 5'UTR of 207 bp and a 3'UTR of 550 bp. The structure of the gene results in 3 exons, with the ATG start codon in the first exon and the TAA stop codon in the last one. BsGATA4/5/6 cDNA results in 2037 bp length and includes a CDS of 1584 bp with a 5' UTR of 342 bp and a 3' UTR of 111 bp. *bsgata4/5/6* has 4 exons with the ATG start codon in the first one and the TGA stop codon in the last one. In the above two transcripts, all the introns are provided of the canonical GT and AG splicing signal consensus.

Protein organization

The *in silico* translation of BsABCG2 results in a putative protein of 671 amino acids with a highly conserved ATP binding cassette (ABC) domain from residue 367 to 578 and a less conserved transmembrane domain (TMD; residues 642 – 664). In the region delimited by residues 76 and 267, an AAA domain (ATPases associated with a variety of cellular activities) is present. This structure parallels the human ABCG2 protein organization. BsABCG2 sequence is highly conserved and displays a typical phosphate-binding loop, Walker A (residues $G^{45, 48, 50}$, K^{51},S^{52}), and a magnesium binding site, Walker B (residues L^{172},F^{173},L^{174},D^{175},E^{176}) with the ABC signature in the middle (I/V^{151},S^{152},G^{153},G^{154},E^{155}).

The sequence of BsCD133 is 554 amino acids in length and contains the specific prominin domain (residues 15-554) that, like the mouse and human orthologues, is a transmembrane domain and shows an N-terminal end exposed to the extracellular space followed by four,

alternating small cytoplasmic (IC) and large extracellular (EC), loops and a cytoplasmic C-terminal end. In *B. schlosseri*, the four transmembrane regions are delimited by residues 28-50, 79-101, 356-378, and 399-421, respectively.

As for GATA family orthologues, the two transcripts found in the transcriptome of *B. schlosseri* show the presence of the zinc finger domain that binds to the DNA consensus sequence [AT]GATA[AG]. BsGata1/2/3 protein present two zinc finger domains (residues 401-451 and 455-505, respectively) in which the zinc ion is coordinated by four cysteines ($C^{407, 410, 428, 431}$; $C^{461, 464, 482, 485}$) that interact with DNA. BsGata4/5/6 transcript, unlike vertebrate GATA5 and other invertebrate GATA4/5/6, presents a single zinc finger domain (residues 196 – 246) with a core in which a zinc ion is coordinated by four cysteines ($C^{202, 205, 223, 226}$).

Phylogenetic analysis

Phylogenetic analysis, carried out on the deduced amino acid sequences of the four proteins, gave similar results, with *B. schlosseri* closely related to vertebrates (Supplementary data). In each NJ tree, tunicate protein were well separated from echinoderms and cephalochordates. From the tree of Gata1/2/3 and Gata4/5/6 we can even observe that, in the cephalochordate *Branchiostoma floridae* and in the considered protostomes, there is a single transcript with similarity to deuterostome Gata4, Gata5 and Gata6. This is in agreement with previous studies that identified single GATA1/2/3 and GATA4/5/6 orthologous genes in basal deuterostomes, including echinoderms, urochordates, and cephalochordates (Gillis et al, 2007).

Western blot analysis

Electrophoretic and immunoblot analysis of haemocyte extracts revealed the presence of a 40-kDa band, recognised by the monoclonal anti-ABCG2 antibody. Also the policlonal anti-CD133 antibody identified a protein of apparent molecular weight of 40 kDa (fig 1). Both the proteins were found in the haemocytes membrane fractions.

ISH

The digoxigenin labelled RNA probes for BsABCG2, when used on colony sections, labelled population of haemocytes close to the endostyle, oocytes, heart primordium of the primary bud and tissues of the secondary buds. Specific riboprobes for BsCD133 where found in population of haemocytes close to the endostyle, oocytes, and haemocytes in peripheral ampullae. BsGATA1/2/3 riboprobes labelled the haemocytes close to the endostyle, haemocytes in growing ampullae and the secondary bud proliferation tissues. ISH with BsGATA4/5/6 show labelling of oocytes and secondary buds proliferation tissues (fig 2). When assayed on haemocyte monolayers, the same riboprobes recognised different fraction of morula cells: 20%

for BsABCG2, 8.5% for BsCD133, 19% for BsGata1/2/3 and 32% for BsGata4/5/6 (fig 2 e, i, m, p). Only in the case of BsGata4/5/6, haemoblasts were also labelled (fig 2q).

qRT-PCR

Specific primers designed on the nucleotide sequences of *B. schlosseri* putative haematopoietic molecular markers (Table 1), allowed me to amplify a short sequence of 150 bp used to evaluate the degree of transcript variation during the blastogenetic cycle. Figure 2 shows that, in phases 9/8/3 and 9/8/4, there is a significant ($p < 0.05$) increase of the extent of transcription of *bsabcg2*, *bscd133*, *bsgata1/2/3* and *bsgata4/5/6* with respect to the other phases (fig 3).

Discussion

In vertebrates, many types of stem cells exist, spread in small populations within the organism (http://stemcells.nih.gov/info/2001report/appendixE.htm). During the last few years, several vertebrate stemness markers, unique of each cell type, were identified (Pazhanisamy, 2013). Conversely, in invertebrate chordates (i.e., members of the subphyla Cephalochordata and Urochordata) these molecules are still poorly investigated. Among invertebrate, chordates Urochordates represents the sister group of vertebrates (Delsuc et al., 2006) and Ascidians are the majority and the best studied urochordates. Their peculiar phylogenetic position explains the persistent interest towards these organisms for evolutionary studies (Rinkevich, 2002).

Colonies of the compound ascidian *B. schlosseri* undergo cyclical events of rejuvenation by the replacements of old adult zooids with their buds that grow to adult size, open their siphons and become fully functional (Manni et al., 2007). In addition, vascular budding (Sabbadin et al., 1975) can assure the survival of injured colonies having lost their functional zooids. These events of blastogenesis are sustained by the presence of a reservoir of stem cells, located in specific niches (Voskoboynik et al., 2008; Rinkevich et al., 2010). A particular type of stem cells, haematopoietic stem cells, is responsible for the production of circulating haemocytes from haematopoietic niches. In *Botryllus*, new haemocytes usually enter the circulation at the beginning of a blastogenetic cycle, when they are required to replace those died by apoptosis during the TO (Ballarin et al.,2008b; Franchi et al., 2016). Unfortunately, these cells are not well characterized and specific haematopoietic markers were not studied in detail.

In the present work, in the attempt to identify haematopoietic markers, we found, in a recently-produced cDNA library (Campagna et al., 2016), four complete transcripts, the predicted amino acid sequences of which show high homology with vertebrate transcripts for ABCG2s, CD133s, members of the Gata1/2/3 subfamily (i.e., Gata1, Gata2 and Gata3) and members of the Gata4/5/6 subfamily (i.e., Gata4, Gata5, and Gata6). The domain structure and organisation of

all the considered proteins result well conserved and all the phylogenetic trees show the close relationship between *B. schlosseri* and vertebrate homologous sequences.

In vertebrates the ATP-Binding Cassette, G2 subfamily (ABCG2), located in the plasma membrane, is a member of the ATP-binding cassette superfamily, one of the largest group of transporter proteins. It is considered a haematopoietic marker with a regulatory role in maintaining pluripotency (Zhou et al, 2001). In human bone marrow, ABCG2 is highly expressed in immature haematopoietic precursors and is sharply down-regulated in the subsequent stages of differentiation (Scharenberg et al, 2002). Its sequence has highly conserved motifs, found in sea urchin, *Ciona* (Annilo et al., 2006; Goldstone et al., 2006) and *C. elegans* where putative homologous have been identified.
In *B. schlosseri,* I identified BsABCG2, a highly conserved vertebrate ABCG2 homologous protein that, in phylogenetic reconstruction, clustered with all chordates as Vertebrate sister group. By immunoblot assay we tested an anti-human ABCG2 monoclonal antibody on cytoplasmic and membrane protein lysate from *B. schlosseri* haemocyte and we found, only in membrane fraction, a 40 kDa single band. Through *in silico* analyses we estimated, for the putative BsABCG2 protein, a molecular weight of 73 kDa. The same difference between detected and predicted dimension was reported for human ABCG2 (http://www.proteinatlas.org) and probably reflects the presence of intra-molecular disulphide bridges that constrain the protein, as documented also for *Botryllus* rhamnose-binding lectin (Gasparini et al., 2008).
It is well known that, in botryllid ascidians, the anterior ventral area of the endostyle is considered a stem cell niche (Voskoboynik et al, 2008). In that area our ISH, with BsABCG2 antisense riboprobe, revealed labelling of haemocytes, in agreement with the above-reported view. The heart primordium in the buds and tissues in the budlets are also labelled, thus reinforcing the idea that BsABCG2 is a stem cell marker.

CD133, also called Prominin-1, is a cholesterol-binding, five-transmembrane-domain glycoprotein, also considered a stem cell marker and identified on the surface of both neural (Weigmann et al, 1997) and haematopoietic stem cells (Yin et al, 1997; Shmelkov, 2005). In humans, it is involved in the cell metabolism (Li, 2013) and have a central role in the asymmetric division that is believed to characterise true stemness; CD133+ cells probably represent more primitive cells than the one identified by the CD34 marker (Toren et al, 2005). CD133 has been found in all the metazoan species analyzed so far (Jászai et al., 2007) and, in mammals, prominin-1 is mainly associated with microvilli and plasma membrane protrusions (Corti et al., 2007). This molecule is expressed also in epithelial cells and peripheral blood, taking part in angiogenetic processes (Li, 2013).

BsCD133 clustered with invertebrate orthologues in our phylogenetic analysis. In Western blot analyses, I detected a single band of 40 kDa recognised by anti-human CD133 polyclonal antibody. Similarly to BsABCG2, and accordingly to its nature of membrane protein, BsCD133 is located in haemocyte membrane fraction. Again, the difference between the predicted protein dimension (60kDa) and the detected one resembles what documented for human CD133 (http://www.proteinatlas.org) and can be ascribed to the presence of disulphide bonds.
On a previous paper on vascular regeneration in *Botryllus,* Braden and coworker (2014) identified the expression of BsCD133 in the resident cells within the peripheral and new budding ampullae. In agreement with that data we detected the expression of BsCD133 in peripheral ampullae haemocytes as well as in the endostyle stem cell niche. The co-localisation of BsCD133 in actively proliferating tissues such as growing ampullae and stem cell niche supports its use as stemness molecular marker also in *B. schlosseri,* as in vertebrates.

Vertebrate GATAs are transcription factors represented by six members, each of which contains a highly conserved DNA binding domain consisting of one or two zinc finger motifs, namely N-terminal zinc finger and C-terminal zinc finger, that bind the consensus DNA sequence GATA (Yue et al., 2014). They are grouped in two subfamilies, the first including GATA1, GATA2 and GATA3, prominently expressed in haematopoietic stem cells, and the second represented by GATA4, GATA5, and GATA6, expressed in various mesoderm- and endoderm- derived tissues (Molkentin, 2000). Briegel et al. (1993) stressed the relevance of GATA family members in vertebrates erythropoiesis. In mammals, GATA1 is essential for the differentiation of erythroid and some myeloid lineages (Patient and McGhee, 2002), whereas GATA2 is expressed in developing and adult haematopoietic stem cells and is required for their self-renewal and maintenance (Tsai and Orkin, 1997). As far as invertebrates are concerned, two GATA factors have been described in the tunicate *Ciona intestinalis*, ascribed to the GATA1/2/3 and the GATA4/5/6 superfamily, respectively (Yamada et al., 2003), the latter involved in cardiomyocyte specification and positioning (Ragkousi et al., 2011). In the nematode *C. elegans* eleven GATA genes have been found, five in *D. melanogaster* and two in *Strongylocentrotus purpuratus* (Gillis et al., 2008). However, as Patient and McGhee (2002) pointed out, the evolutionary homology between invertebrate and vertebrate GATAs are still unclear. For this reason, members of the GATA4/5/6 superfamily, potentially involved in regulating vertebrate cardiogenesis and differentiation of gut epithelium, could play pivotal roles in haematopoietic differentiation of invertebrates.
Like what observed in *Ciona* (Yamada et al., 2013; D'Ambrosio et al., 2013), here, we reported the presence of two GATA transcripts in *B. schlosseri*, one belonging to the GATA1/2/3 superfamily and the other to the GATA 4/5/6 one. BsGATA1/2/3 and BsGATA4/5/6, in their

phylogenetic reconstruction, cluster with *Ciona intestinalis* orthologues as sister group of vertebrates.

As all the other considered molecular marker analysed so far, our *ISH* data on BsGATA1/2/3 indicated the presence of gene transcripts in the secondary bud proliferation tissues and in the endostyle stem cell niche. In addition, labeling was found in resident cells of the growing ampullae. BSGATA4/5/6, conversely, were found in secondary bud proliferation tissues and in oocytes. To reinforce the idea of their connection to stem cell identification BsGATA4/5/6 mRNA, unique among the examined transcripts, was also present in haemoblasts, stressing a possible role in early differentiation of haemocytes.

With qRT-PCR, we analysed the expression of BsABCG2, BsCD133, BsGATA1/2/3 and BsGATA4/5/6 during the *Botryllus* blastogenetic cycle observing a high expression level of all the transcripts in 9/8/3-9/8/4: the mid-cycle phase. That phase of blastogenetic development is characterised by the beginning of tissue differentiation in secondary buds that require new undifferentiated cells for the new tissues formation (Kawamura et al, 2008) and we can hypothesise that the same holds true also for haemocytes.

To date, it is still unclear when and where haematopoietic events occur in *B. schlosseri*. Now we know that the four different transcripts considered, BsABCG2, BsCD133, BsGATA1/2/3 and BsGATA4/5/6, vertebrate haematopoietic orthologous gene, are present in sub-endostylar haemolimph lacunae and are all over-expressed in 9/8/3-9/8/4 phases. The presented data, collectively, allow us to hypothesise that they are haematopoietic molecular markers and that haemocyte turnover occurs during the mid-cycle phase reinforcing the notion that haemopoiesis occurs in endostylar niche.

Moreover, we hypothesise that BsABCG2, BsCD133, BsGATA1/2/3 and BsGATA4/5/6 are involved not only in the specification of haematopoietic stem cells of *B. schlosseri*, but could exert also additional roles in the tissues. The presence of the mRNAs of BsABCG2, BsCD133 and BsGATA4/5/6 inside the oocytes suggests that they are stored inside female gametes in order to assure the early development, when gene transcription has not yet started in the embryo.

As regards the fraction of labelled morula cells, i.e., cytotoxic immunocytes, it can be ascribed to a particular differentiation stage of this cells that can represent either not yet fully differentiated cells or, alternatively, adults cells with some peculiar functions requiring the activity of the genes. For instance, from studies on vertebrates, we know that ABCG2 is also expressed in adult tissues, especially in cells with secretory or barrier functions where it is involved in the efflux of various materials from the cytoplasm to extracellular milieu or intracellular compartments (Mo et al, 2012). Important secretory roles of *Botryllus* morula cells, have been demonstrated. In fact, they are site of synthesis of cytokines (Ballarin et al., 2001; Cima et al., 2004; Menin et al., 2005), complement factors (Franchi and Ballarin, 2014),

antioxidant molecules (Franchi et al., 2016, in press) and cytotoxic enzymes (Ballarin et al., 1995) that can be released in the extracellular milieau. A detailed study of the role of these genes in circulating haemocytes will be the subject of future investigations.

In addition, it is interesting to note that circulating cells defined as haemoblasts, morphologically resembling a differentiated cells and considered, by many authors, precursors of haemocytes, are labelled only by BsGata4/5/6 riboprobes. This indicates that circulating haemoblasts are not true stem cells but already determined, even if not yet fully differentiated, haemocytes.

References

Ballarin L, Cima F. 2005. Cytochemical properties of *Botryllus schlosseri* haemocytes: indications for morpho-functional characterisation. Eur J Histochem 49: 255-264

Ballarin L, Burighel P, Cima F. 2008a. A tale of death and life: natural apoptosis in the colonial ascidian *Botryllus schlosseri* (Urochordata, Ascidiacea). Curr Pharm Des 14: 138–147

Ballarin L, Menin A, Tallandini L, Matozzo V, Burighel P, Basso G, Fortunato E, Cima F. 2008b. Haemocytes and blastogenetic cycle in the colonial ascidian *Botryllus schlosseri*: a matter of life and death. Cell Tissue Res 331: 555–564

Ballarin L, Schiavon F, Manni L. 2010. Natural apoptosis during the blastogenetic cycle of the colonial ascidian *Botryllus schlosseri*: a morphological analysis. Zool Sci 27: 96–102

Braden BP, Taketa DA, Pierce JD, Kassmer S, Lewis DD, et al. 2014. Vascular regeneration in a basal chordate is due to the presence of immobile, bi- functional cells. PLoS ONE 9: e95460

Briegel K, Lim KC, Plank C, Beug H, Engel JD, Zenke M. 1993. Ectopic expression of a conditional GATA-2/ estrogen receptor chimera arrests erythroid differentiation in a hormone-dependent manner. Genes Dev 7: 1097-1109

Campagna D, Gasparini F, Franchi N, Manni L, Telatin A, Vitulo N, Ballarin L, Valle G. 2015. SATRAP: SOLiD Assembler TRAnslation Program. PLoS One 10: e0137436

Campagna D, Gasparini F, Franchi N, Vitulo N, Ballin F, Manni L, Valle G, Ballarin L. 2016. Transcriptome dynamics in the asexual cycle of the chordate *Botryllus schlosseri*. BMC Genomics 17: 275

Cima F, Basso G, Ballarin L. 2003. Apoptosis and phosphatidylserine-mediated recognition during the take-over phase of the colonial life-cycle in the ascidian *Botryllus schlosseri*. Cell Tissue Res 312: 369–376

Corbeil D, Röper K, Hellwig A, Tavian M, Miraglia S, Watt SM, Simmons PJ, Peault B, Buck DW, Huttner WB. 2000. The human AC133 hematopoietic stem cell antigen is also expressed in epithelial cells and targeted to plasma membrane protrusions. J Biol Chem 275: 5512-5520

Corti S, Nizzardo M, Nardini M, Donadoni C, Locatelli F, Papadimitriou D, Salani S, Del Bo R, Ghezzi S, Strazzer S, Bresolin N, Comi GP. 2007. Isolation and characterization of murine neural stem/progenitor cells based on Prominin- 1expression. Exp Neurol 205: 547-562

Franchi N, Ballin F, Manni L, Schiavon F, Basso G, Ballarin L. 2016. Recurrent phagocytosis-induced apoptosis in the cyclical generation change of the compound ascidian *Botryllus schlosseri*. Dev Comp Immunol 62: 8-16

Gasparini F, Franchi N, Spolaore B, Ballarin L. 2008. Novel rhamnose-binding lectins from the colonial ascidian *Botryllus schlosseri.* Dev Comp Immunol 32: 1177-1191

Gasparini F, Manni L, Cima F, Zaniolo G, Burighel P, Caicci F, Franchi N, Schiavon F, Rigon F, Campagna D, Ballarin L. 2015. Sexual and asexual reproduction in the colonial ascidian *Botryllus schlosseri*. Genesis 53: 105–120

Gillis WJ, Bowerman B, Schneider SQ. 2007. Ectoderm- and endomesoderm-specific GATA transcription factors in the marine annelid *Platynereis dumerilli*. Evol Dev 9: 39-50

Kawamura K, Tachibana M, Sunanaga T. 2008. Cell proliferation dynamics of somatic and germline tissues during zooidal life span in the colonial tunicate *Botryllus primigenus*. Dev Dyn 237: 1812-1825

Laird DJ, De Tomaso AW, Weissman IL. 2005. Stem cells are units of natural selection in a colonial ascidian. Cell 123: 1351-60

Lin T, Islam O, Heese K. 2006. ABC transporters, neural stem cells and neurogenesis- a different perspective. Cell R 16: 857-871

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-514 time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 25: 402–408

Lorkowski S, Cullen P. 2002. ABCG subfamily of human ATP-binding cassette proteins. Pure Appl Chem 74: 2021-2081

Manni L, Zaniolo G, Cima F, Burighel P, Ballarin L. 2007. *Botryllus schlosseri*: a model ascidian for the study of asexual reproduction. Dev Dyn 236: 335-352

Molkentin JD. 2000. The zinc finger-containing transcription factors GATA-4, -5, and -6. Ubiquitously expressed regulators of tissue-specific gene expression. J Biol Chem 275: 38949-38952

Nemeth K, Karpati S. 2014. Identifying the stem cell. J Invest Dermatol 134: e26

NIH Stem Cell Information Home Page. In Stem Cell Information [World Wide Web site]. Bethesda, MD: National Institutes of Health, U.S. Department of Health and Human Services, 2016

Oka H, Watanabe H. 1957. Vascular budding, a new type of budding in *Botryllus*. Biol bull 112: 225-240

Patient RK and McGhee JD. 2002. The GATA family (vertebrates and invertebrates). Curr Opin Gen Dev 12: 416-422

Pazhanisamy S. 2013. Adult stem cell and embryonic stem cell markers. Mater methods 3: 200

Rinkevich Y, Voskoboynik A, Rosner A, Rabinowitz C, Paz G, Oren M, Douek J, Alfassi G, Moiseeva E, Ishizuka KJ, Palmeri KJ, Weissman IL, Rinkevich B. 2013. Repeated, long-term cycling of putative stem cells between niches in a basal chordate. Dev Cell 24: 76-88

Ragkousi K, Beh J, Sweeney S, Starobinska E, Davidson B. 2011. A single GATA factor plays discrete, lineage specific roles in ascidian heart development. Dev Biol. 352: 154-163

Rothenberg EV, Pant R. 2004. Origins of lymphocyte developmental programs: transcription factor evidence. Seminars in Immunology 16: 227–238

Sabbadin A, Zaniolo G, Majone F. 1975. Determination of polarity and bilateral asymmetry in palleal and vascular buds of the ascidian *Botryllus schlosseri*. Dev Biol 46: 79-87

Sabbadin A, Zaniolo G. 1979. Sexual differentiation and germ cell transfer in the colonial ascidian *Botryllus schlosseri*. J Exp Zool 207: 289-304

Sabbadin A, Astorri C. 1988. Chimeras and histocompatibility in the colonial ascidian *Botryllus schlosseri*. Devel Comp Immun 12: 737- 747

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406–425

Scharenberg CW, Harkey MA, Torok-Storb B. 2002. The ABCG2 transporter is an efficient Hoechst 33342 efflux pump and is preferentially expressed by immature human hematopoietic progenitors. Blood 99: 507-512

Shmelkov SV, StClair R, Lyden D, Rafii S. 2005. AC133/CD133/Prominin-1. J Biochem Cell Biol 37: 715-719

Taneda Y, Watanabe H. 1982. Studies on colony specificity in the compound ascidian, *Botryllus primigenus* Oka. I. Initiation of "nonfusion" reaction with special reference to blood cells infiltration. Dev Comp Immunol 6: 43-52

Tang X, Nakata Y, Li HO, Zhang M, Gao H, Fujita A, Sakatsume O, Ohta T, Yokoyama K. 1994. The optimization of preparations of competent cells for transformation of *E coli*. Nucleic acids res 22: 2857-2858

Towbin H, Staehelin T, Gordon J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Biotech 24: 145-149

Toren A, Bielorai B, Jacob-Hirsch J, Fisher T, Kreiser D, Moran O, Zeligson S, Givol D, Yitzhaky A, Itskovitz-Eldor J, Kventsel I, Rosenthal E, Amariglio N, Rechavia G. 2005. CD133-positive hematopoietic stem cell “stemness” genes contain many genes mutated or abnormally expressed in leukemia. Stem Cells 23: 1142–1153

Tsai FY, Orkin SH. 1997. Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. Blood 89: 3636-3643

Uchida N, Buck DW, He D, Reitsma MJ, Masek M, Phan TV, Tsukamoto AS, Gage FH, Weissman IL. 2000. Direct isolation of human central nervous system cells. Proc Natl Acad Sci U.S.A. 97: 14720–14725.

Voskoboynik A, Soen Y, Rinkevich Y, Rosner A, Ueno H, et al. 2008. Identification of the endostyle as a stem cell niche in a colonial chordate. Cell Stem Cell 3: 456-464

Voskoboynik A, Neff NF, Sahoo D, Newman AM, Pushkarev D, Koh W, Passarelli B, Fan HC, Mantalas GL, Palmeri KJ, Ishizuka KJ, Gissi C, Griggio F, Ben-Shlomo R, Corey DM, Penland L, White RA, Weissman IL, Quake SR. 2013. The genome sequence of the colonial chordate, *Botryllus schlosseri*. Elife 2: e00569–e00569

Zhou S, Schuetz JD, Bunting KD, et al. 2001. The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. Nat Med 7: 1028-1034

Weigmann A, Corbeil D, Hellwig A, Wuttner WB. 1997. Prominin, a novel microvilli-specific polytopic membrane protein of the apical surface of epithelial cells, is targeted to plasmalemmal protrusions of non-epithelial cells. Proc Natl Acad Sci U.S.A. 94: 12425–12430

Yamada L, Kobayashi K, Degnan B, Satoh N, Satou Y. 2003. A genomewide survey of developmentally relevant genes in *Ciona intestinalis*. Dev Genes Evol 5–6: 245–253

Yin AH, Miraglia S, Zanjani ED, Almeida-Porada G, Ogawa M, Leary AG, Olweus J, Kearney J, Buck DW. 1997. AC133, a novel marker for human hematopoietic stem and progenitor cells. Blood 90: 5002–5012

Yue F, Zhou Z, Wang L, Wang M, Song L. 2014. A conserved zinc finger transcription factor GATA involving in the hemocyte production of scallop *Chlamys farreri*. Fish Shellfish Immunol 39: 125-35

http://www.proteinatlas.org

Table 1. PCR primer used and relative melting temperatures (Tm).

Primer	**Tm(°C)**	**Sequence 5'-3'**
BsABCG2F	54.5	GCCACCGAATCAAAACCC
BsABCG2R	53.9	GTATCTCGGCTGGTGGAT
BsABCG2R-RT	59.7	CGAGATACACGCTTCCAGTGATGC
BsCD133F	55.1	TGCCGCTGTTGCTGTAAT
BsCD133R	55.7	GTCGGGATTGCCATCAGAG
BsCD133R-RT	62.3	CCGAGCACGACTCCAGCGA
BsGATA1/2/3F	63.5	GCACCAGTATGTCCCGACCGAGG
BsGATA1/2/3R	56.7	GCCGTTTGTGTTGCGTCT
BsGATA1/2/3F-RT	57.2	GAGCGACTTCCACTCCTCTG
BsGATA4/5/6F	58	CGAGTCTGTCCCACCCG
BsGATA4/5/6R	57.7	GCCGCCAGAGTGTTGTTACT
BsGATA4/5/6R-RT	58.3	CGCAACAGTGTCGTCCGT
BsACTF-RT	55.6	CTCAACCCGAAGGCTAACC
BsACTR-RT	56.2	GTACAGCGACAGAACGGC

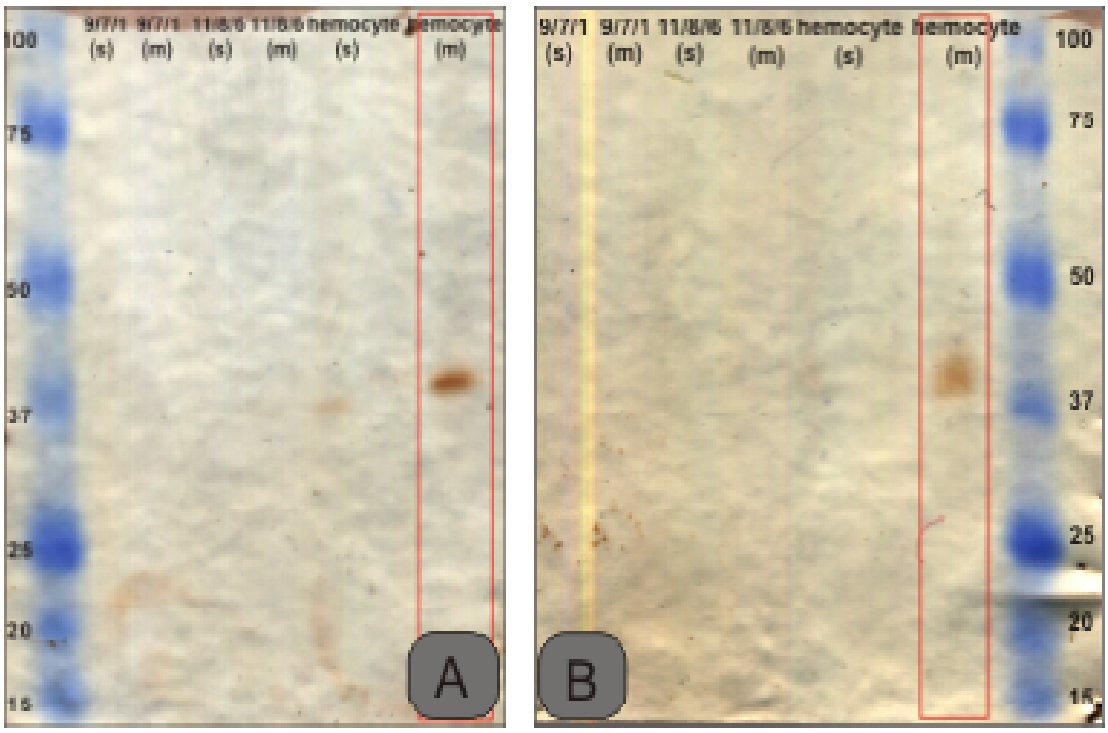


Fig1: Electrophoresis and immunoblot analysis on tissues and haemocyte lysates from colonies at different phases of the blatogenetic cycle, with anti-CD133 and anti-ABCG2 antibody. In blue the reference molecular weight in kDa; (s) soluble proteins; (m) membrane proteins.

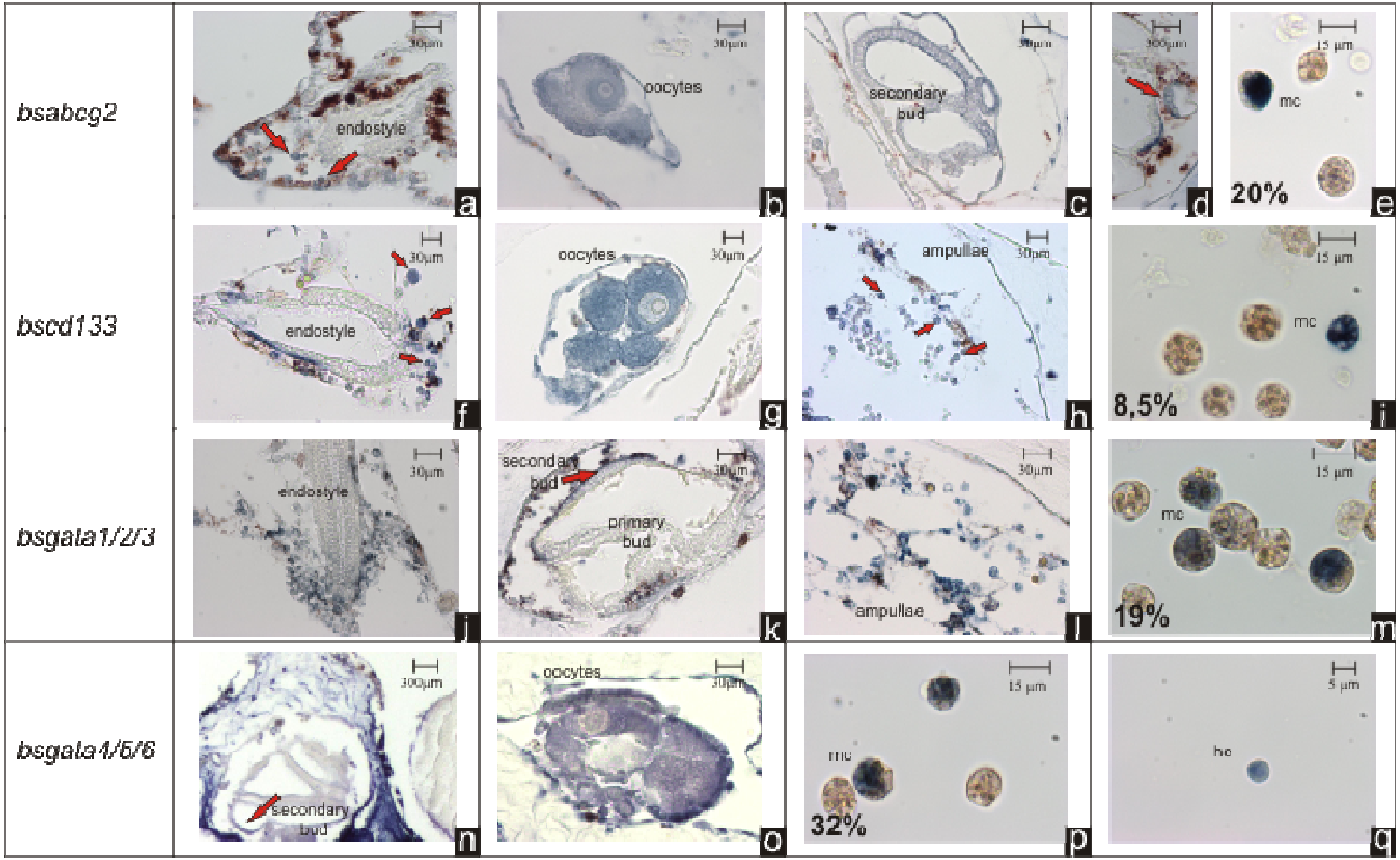

bsabcg2
endostyle
a
oocytes
b
secondary bud
c
d
mc
20%
e
bscd133
endostyle
f
oocytes
g
ampullae
h
mc
8,5%
i
bsgala1/2/3
endostyle
j
secondary bud
primary bud
k
ampullae
l
mc
19%
m
bsgala4/5/6
secondary bud
n
oocytes
o
mc
32%
p
hc
q

Fig2: ISH on tissues sections and haemocytes monolayers stained with antisense riboprobes for *bsabcg2 (a-e), bscd133 (f-i), bsgata1/2/3 (j-m), bsgata4/5/6 (n-q).* (a-d) sections and (e) haemocytes labelled by BsABCG2, red arrow indicates haemocytes in (a) and heart primordia in (d); (f-h) sections and (i) haemocytes labelled by BsCD133, red arrow indicates haemocytes in the endostyle stem cell niche (f) and in the ampullae (h); (j-l) sections and (m) haemocytes labelled by BsGata1/2/3; (n-o) sections and (p-q) haemocytes labelled by BsGata4/5/6. mc= morula cells, he= haemoblast.

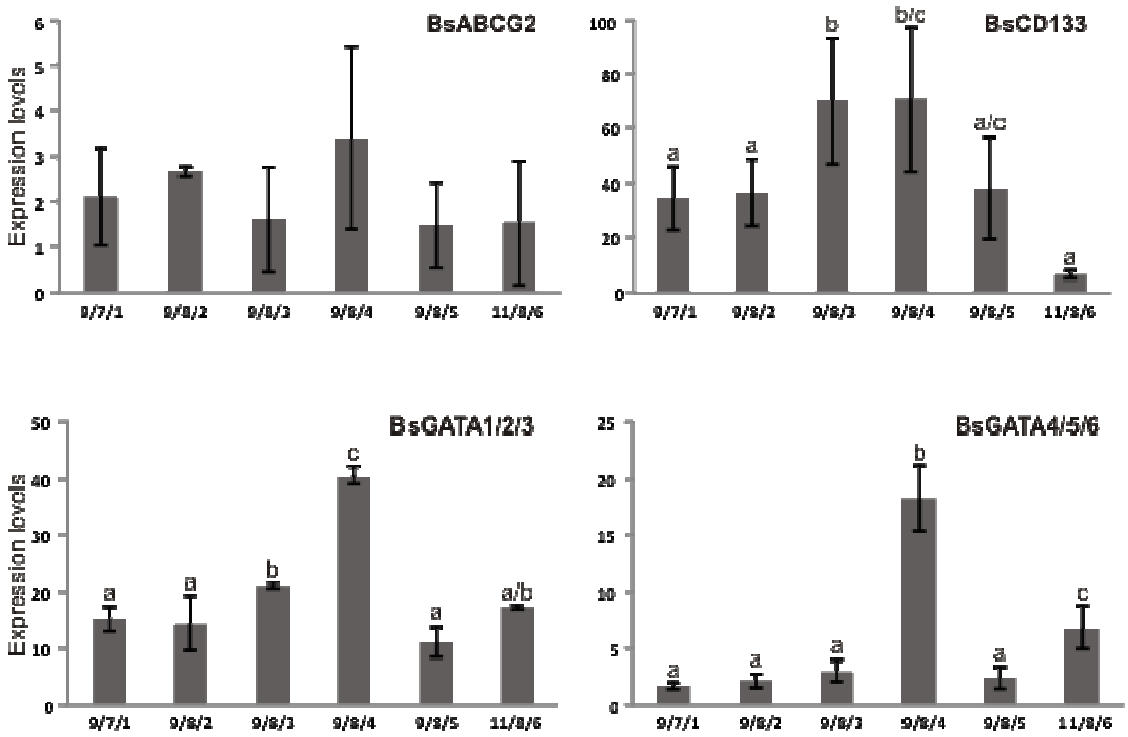
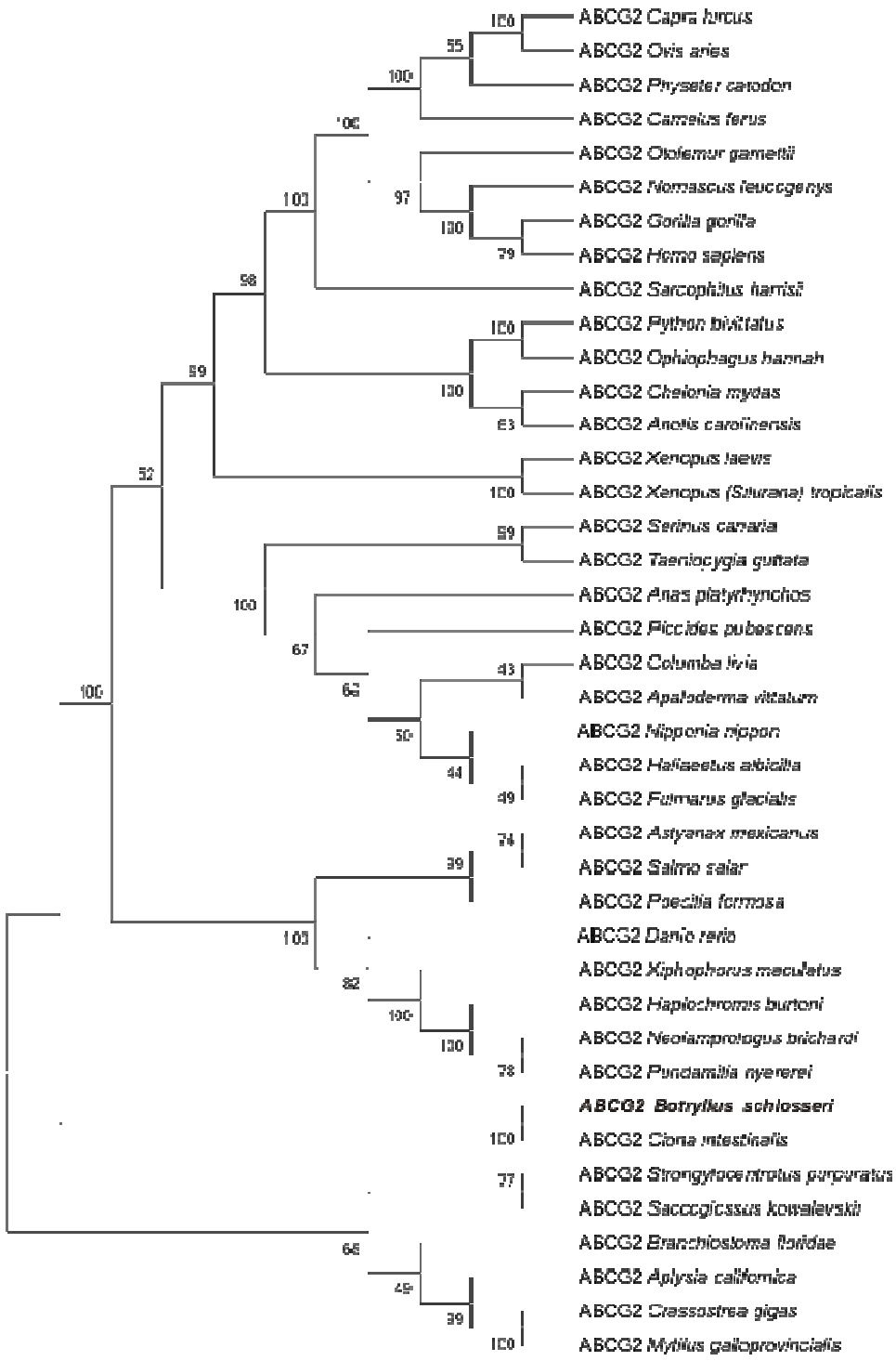
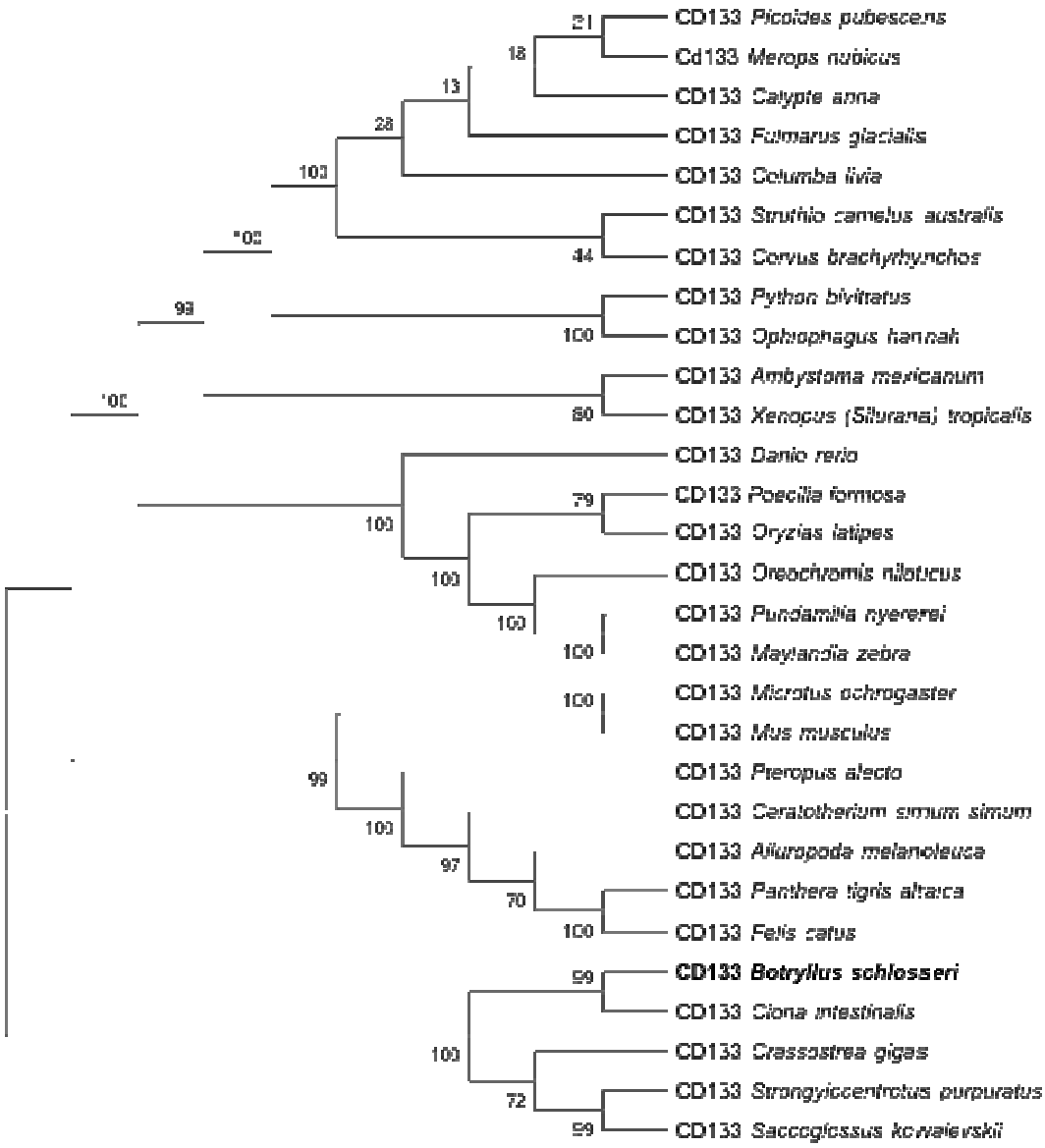


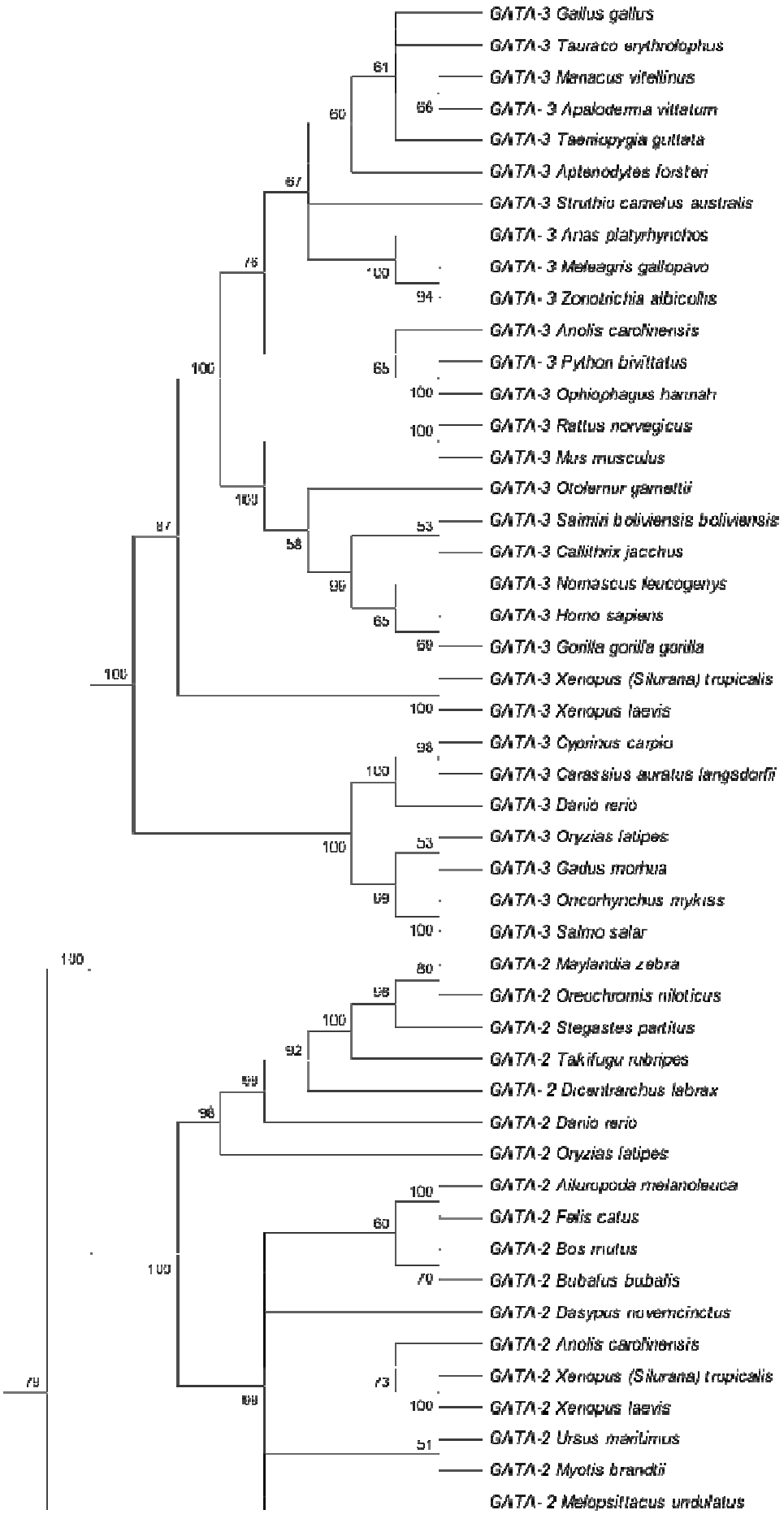
Fig3: Relative expression levels of *bsabcg2, bscd133, bsgata1/2/3, bsgata4/5/6* in different phases of the blastogenetic cycle. Normalisation of expression was achieved using endogenous β-actin as housekeeping gene. Each bar of the histogram corresponds to the average of three independent experiments (n=3) ± SD. Different letters indicate significant differences ($p < 0.05$).

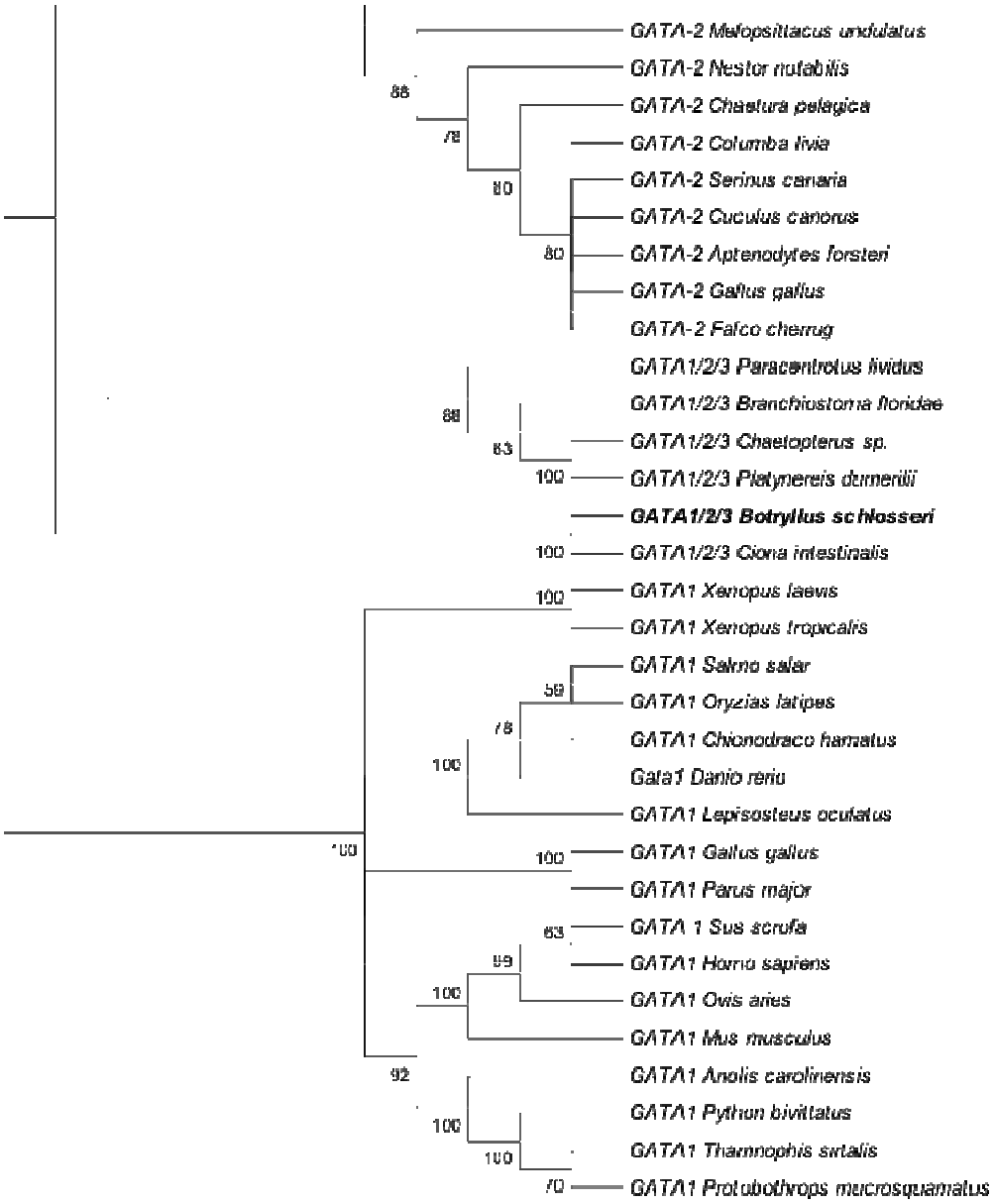


FigS1: Evolutionary relationships (NJ) among proteins *abcg2* with bootstrap confidence values indicated at the left of each branch. BsABCG2 is indicated in bold.

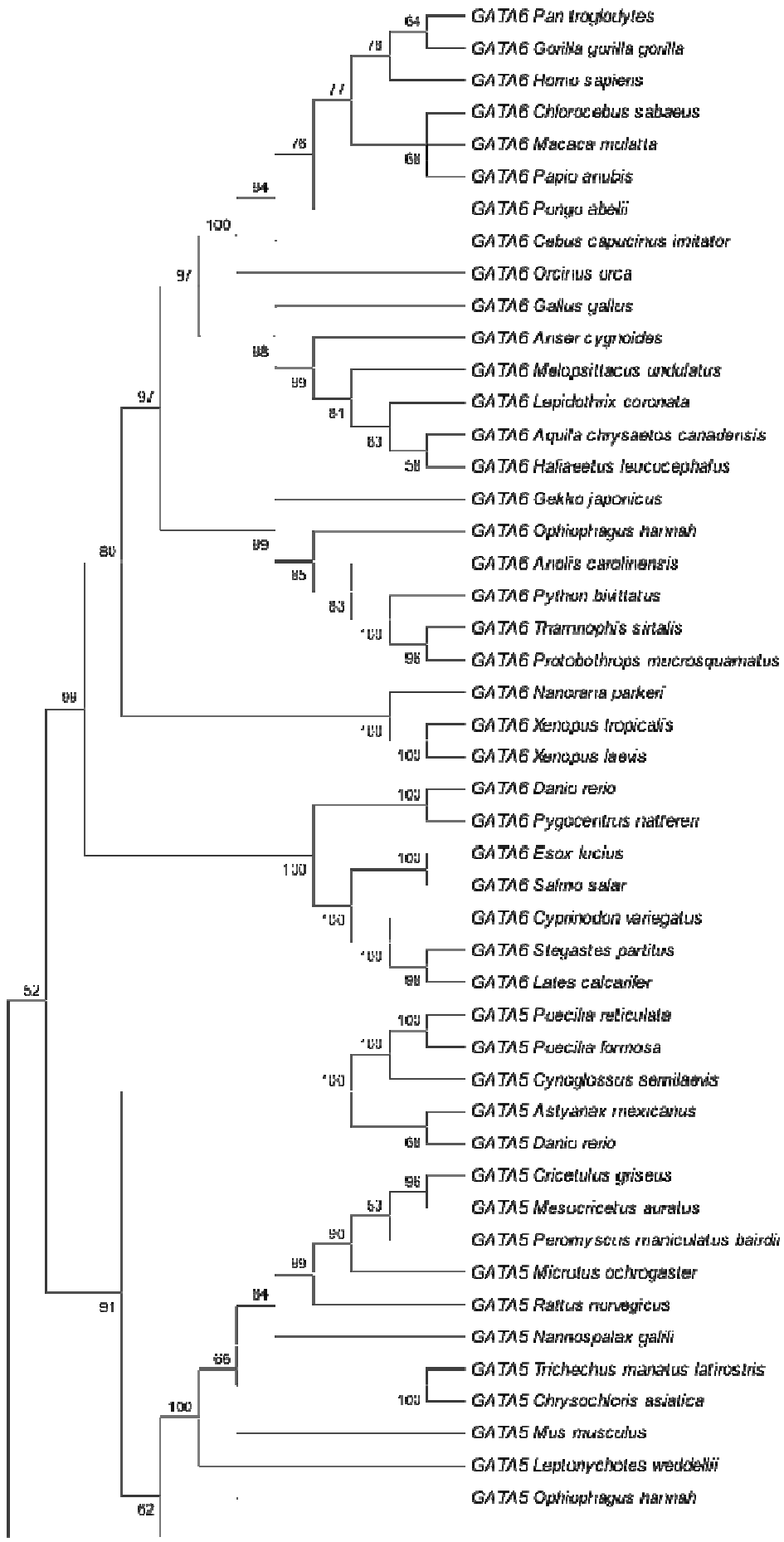


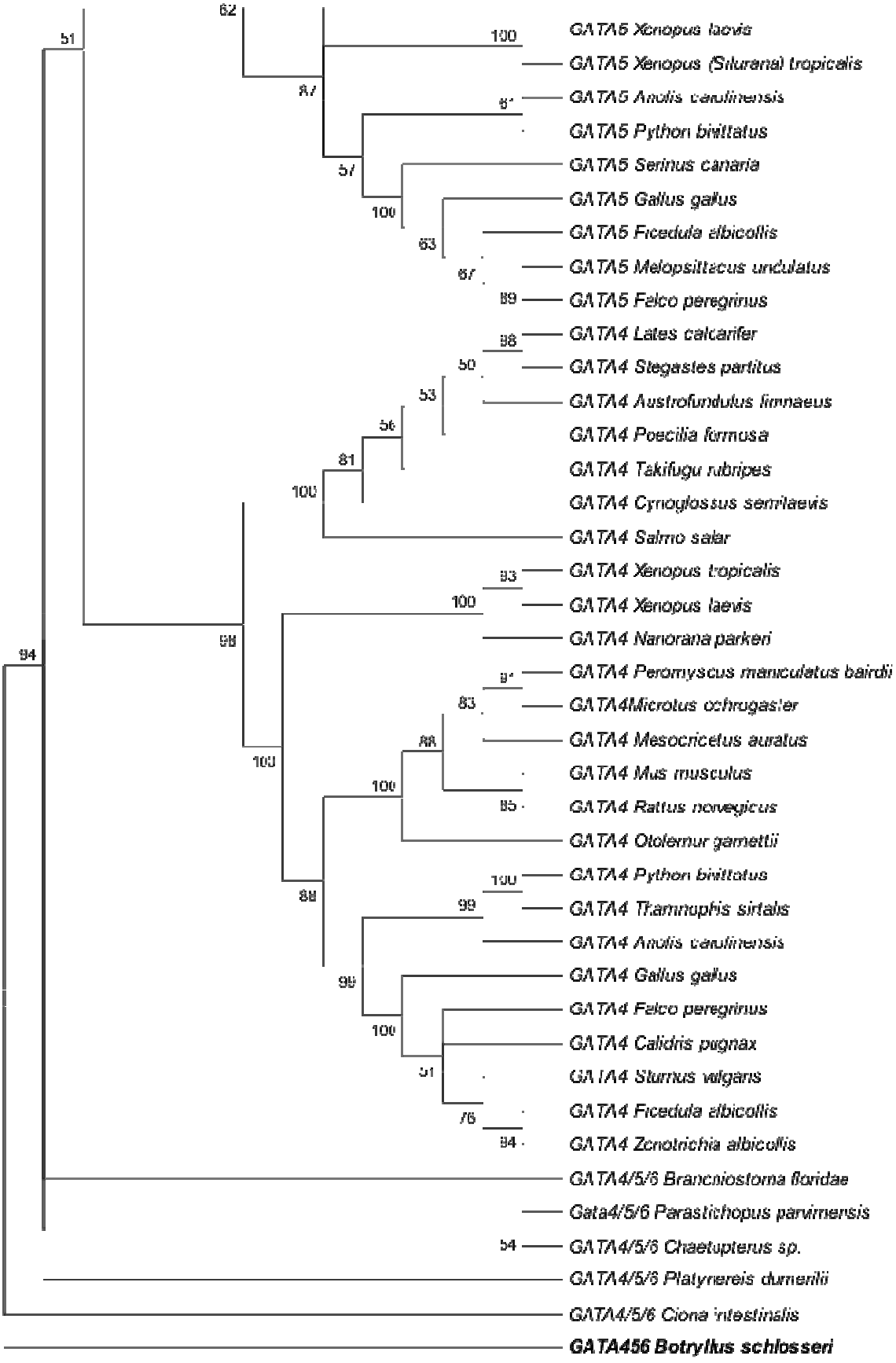
FigS2: Evolutionary relationships (NJ) among *cd133* proteins, with bootstrap confidence values indicated at the left of each branch. BsCD133 is indicated in bold.


GATA-3 Gallus gallus
GATA-3 Tauraco erythrolophus
GATA-3 Menacus vitellinus
GATA-3 Apalopterma vittatum
GATA-3 Taeniopygia guttata
GATA-3 Aptenodytes forsteri
GATA-3 Struthio camelus australis
GATA-3 Anas platyrhynchos
GATA-3 Meleagris gallopavo
GATA-3 Zonotrichia albicollis
GATA-3 Anolis carolinensis
GATA-3 Python bivittatus
GATA-3 Ophiophagus hannah
GATA-3 Rattus norvegicus
GATA-3 Mus musculus
GATA-3 Otolemur garnettii
GATA-3 Saimiri boliviensis boliviensis
GATA-3 Callithrix jacchus
GATA-3 Nomascus leucogenys
GATA-3 Homo sapiens
GATA-3 Gorilla gorilla gorilla
GATA-3 Xenopus (Silurana) tropicalis
GATA-3 Xenopus laevis
GATA-3 Cyprinus carpio
GATA-3 Carassius auratus langsdorfii
GATA-3 Danio rerio
GATA-3 Oryzias latipes
GATA-3 Gadus morhua
GATA-3 Oncorhynchus mykiss
GATA-3 Salmo salar
GATA-2 Maylandia zebra
GATA-2 Oreochromis niloticus
GATA-2 Stegastes partitus
GATA-2 Takifugu rubripes
GATA-2 Dicentrarchus labrax
GATA-2 Danio rerio
GATA-2 Oryzias latipes
GATA-2 Ailuropoda melanoleuca
GATA-2 Felis catus
GATA-2 Bos mutus
GATA-2 Bubalus bubalis
GATA-2 Dasypus novemcinctus
GATA-2 Anolis carolinensis
GATA-2 Xenopus (Silurana) tropicalis
GATA-2 Xenopus laevis
GATA-2 Ursus maritimus
GATA-2 Myotis brandtii
GATA-2 Melopsittacus undulatus



FigS3: Evolutionary relationships (NJ) among *gata1/2/3* proteins, with bootstrap confidence values indicated at the left of each branch. BsGATA1/2/3 is indicated in bold.


GATA6 Pan troglodytes
GATA6 Gorilla gorilla gorilla
GATA6 Homo sapiens
GATA6 Chlorocebus sabaeus
GATA6 Macaca mulatta
GATA6 Papio anubis
GATA6 Pongo abelii
GATA6 Cebus capucinus imitator
GATA6 Orcinus orca
GATA6 Gallus gallus
GATA6 Anser cygnoides
GATA6 Melopsittacus undulatus
GATA6 Lepidothrix coronata
GATA6 Aquila chrysaetos canadensis
GATA6 Haliaeetus leucocephalus
GATA6 Gekko japonicus
GATA6 Ophiophagus hannah
GATA6 Anolis carolinensis
GATA6 Python bivittatus
GATA6 Thamnophis sirtalis
GATA6 Protobothrops mucrosquamatus
GATA6 Nanorana parkeri
GATA6 Xenopus tropicalis
GATA6 Xenopus laevis
GATA6 Danio rerio
GATA6 Pygocentrus nattereri
GATA6 Esox lucius
GATA6 Salmo salar
GATA6 Cyprinodon variegatus
GATA6 Stegastes partitus
GATA6 Lates calcarifer
GATA5 Poecilia reticulata
GATA5 Poecilia formosa
GATA5 Cynoglossus semilaevis
GATA5 Astyanax mexicanus
GATA5 Danio rerio
GATA5 Cricetulus griseus
GATA5 Mesocricetus auratus
GATA5 Peromyscus maniculatus bairdii
GATA5 Microtus ochrogaster
GATA5 Rattus norvegicus
GATA5 Nannospalax galili
GATA5 Trichechus manatus latirostris
GATA5 Chrysochloris asiatica
GATA5 Mus musculus
GATA5 Leptonychotes weddellii
GATA5 Ophiophagus hannah



FigS4: Evolutionary relationships (NJ) among *gata4/5/6* proteins, with bootstrap confidence values indicated at the left of each branch. BsGATA4/5/6 is indicated in bold.

<u>Research article:</u> Identification, characterisation and expression studies of orthologues of vertebrate stemness and differentiation molecular markers of the neural cell lineage in the colonial ascidian *Botryllus schlosseri*

Introduction

Invertebrate chordates are marine, filter-feeding animals including cephalochordates and tunicates or urochordates; with the latter considered the vertebrate sister group (Delsuc et al, 2006).

Ascidians are sessile tunicates representing the richest in species and most studied class of tunicates. Ascidian larvae have a tiny central nervous system (CNS), a miniature chordate nervous system, subject of many studies aiming at understanding the evolutionary developmental processes leading to the complexity of the vertebrate nervous system (Manni et Burighel, 2006). The ascidian CNS develops from a neural plate forming, like vertebrates, a neural tube and consists of less than 100 neurons and 250 glial cells (Meinertzhagen and Okamura, 2001). In the last few years morphological and embryological studies, associated with genetic and molecular data, demonstrated the close relationship between vertebrate and ascidia larva nervous system (Manni et al, 1999, 2005; Tiozzo et al, 2004; Lemaire et al, 2002).At metamorphosis, the neural tube degenerates with the exception of the neurohypophysial duct, that gives rise to the adult neural complex (Willey, 1893). The latter is located between the oral and the cloacal siphon and includes a cerebral ganglion (composed of approximately 1000 neurons, from which nerves of the peripheral nervous system branch off) and a neural gland (a sac-like structure composed of three part, the ciliated duct, the gland and the dorsal organ). Structure homology were suggested between the vertebrate hypophysis and the ascidian neural gland (Julin, 1981). Chordate peripheral nervous system partly derives from neurogenic placodes that, in vertebrates, contribute to the formation of paired sensory organs together with neural crest, a population of migratory cells able to differentiate in various cell types (Manni and Burighel, 2006). These structures probably played a crucial role in vertebrate evolution and were reported also in tunicates (Jeffery et al, 2004; Manni et al, 2004).

In colonial species, such as *Botryllus schlosseri*, in addition to embryogenesis, there is a recurrent budding process during which new tissues cyclically originate. In a colony, adult zooids are grouped in star-shaped systems of 8-12 individuals sharing a common cloacal siphon in the center of the system. Three blastogenetic generations are usually present in a colony, i.e.: adult, filter-feeding zooids, buds on zooids and budlets on buds. A colony performs cyclical (weekly at 20°C) generation changes, or take-overs (TOs), that allow its recurrent rejuvenation (Manni et al. 2007; Ballarin et al. 2010). A colonial blastogenetic cycle can be divided in: post-TO (or beginning of a new cycle, phase 9/7/1), mid-cycle (MC, phases 9/8/2, 9/8/3, 9/8/4), pre-TO (phase 9/8/5) and TO (the end of the cycle, phase 11/8/6). The combination of numbers

corresponding to the developmental phase of the adult zooids, primary and secondary bud, respectively (Manni et al, 2007). In this case, a typical neural plate is absent and the brain develops from neuroblasts that originate as migratory cells from the rudiment of the neural gland (Manni and Burighel, 2006). Manni et al (1999) demonstrated that, even if tissues involved in the neural complex formation are different, the mechanisms of neurogenesis during embryogenesis and blastogenesis can be comparable (Tiozzo et al, 2005; Manni and Burighel, 2006).

Bollner et al. (1995), demonstrated that, in the solitary ascidian *Ciona intestinalis*, there are neuron progenitor cells able to regenerate the brain after extirpation. This process, as well as the formation of the bud neural complex in *B. schlosseri*, appears to be linked to the presence of a pool of long-living stem cells.

Despite the numerous efforts in the comparison of ascidian and vertebrates nervous system, there is still a small number of marker genes useful for molecular studies (Lemaire et al, 2002). With the aim of providing new molecular tools for the study of ascidian neurogenesis, I looked for new orthologous genes/transcript of vertebrate neural stem cell markers in the genome and transcriptome of *B. schlosseri.* In this study I identified and characterised two transcripts recognised as orthologues of vertebrate neural stem cell markers (*bssox2* and *bsmsi2*): I studied the dynamics of their transcription during the colonial blastogenetic cycle. In addition, I examined the expression of a panel of genes already known to be involved in ascidian larval neurogenesis, i.e., orthologues of *Pax2/5/8*, *Hox1* and *Hox3*, during the blastogenetic cycle. In both cases I also studied the localisation of their mRNAs through *in situ* hybridization (ISH) and carried out a phylogenetic analyses on the predicted proteins.

Materials and methods

Animals

Colonies of *B. schlosseri* (Tunicata, Ascidiacea) were collected near Chioggia, in the southern part of the Lagoon of Venice. They were reared according to Gasparini et al. (2015), stuck to glass slides (5 x 5 cm), in aerated aquaria filled with 0.45-µm filtered seawater (FSW) changed every other day, at the constant temperature of 19°C, and fed with Liquifry marine (Liquifry Co., Dorking, England). At this temperature, colonies undergo weekly generation change or take-over (TO) during which old zooids are resorbed and replaced by their buds in 24-36 h. Colonial developmental phases, more than one day from the preceding or the following generation change, are collectively called mid-cycle (MC; Manni et al., 2007).

Primer design, RNA extraction, cDNA synthesis, cloning and sequencing

Comparison of recent *B. schlosseri* transcriptomes (Campagna et al, 2016) and *Botryllus* draft genome (Voskoboynik et al., 2013) with the sequences of the vertebrate genes of interest

allowed us to identify a series of nucleotide sequences and to design specific primers (Table 1) for PCR amplification. We focused our attention on the sequences of the putative transcript orthologues of: SOX2, MSI2, PAX2/5/8, HOX1 and HOX3, named BsSOX2, BsMSI2, BsPAX2/5/8, BsHOX1 and BsHOX3, respectively.

Total RNA was isolated from *B. schlosseri* colonies with the SV total RNA isolation system (Promega); its purity was determined spectrophotometrically by the $A_{260/A280}$ and $A_{260/A230}$ ratio. The integrity of RNA preparation was checked by visualising the rRNA in ethidium bromide-stained 1.5% agarose gels. The first strand of cDNA was reverse-transcribed from 1 μg of total RNA according to the Improm II manual (Promega). cDNA amplification was performed with Go-Taq Polymerase (Promega; 5 U/μl) with the following cycling parameters: 94°C for 2 min, 45 cycles of 94°C for 30 s, melting temperature (T_m) for 30 s (T_m for the various primers is indicated in Table 1), 72°C for 1 min, and a last step at 72°C for 10 min. Amplicons were subjected to electrophoresis and the corresponding bands were purified with ULTRAPrep Agarose Gel Extraction Mini Prep kit (AHN Biotechnologie), ligated in pGEM T-Easy Vector (Promega) and cloned in DH-5α *Escherichia coli* cells (Tang et al, 1994). In order to confirm the sequences and their expression, positively screened clones were sequenced at BMR Genomics (University of Padova) on an ABI PRISM 3700 DNA Analyzer (Applied Biosystems). Gene reconstructions were based on *B. schlosseri* genome database using Spidey's algorithm (http://www.ncbi.nlm.nih.gov/spidey/)

Sequence alignment and phylogenetic analyses

Amino acid sequences of the proteins of interest were obtained by *in silico* translation. For sequence alignment and phylogenetic analyses, sequences of interest were compared with those of the corresponding proteins from deuterostomes and protostomes. Alignments were carried out with Clustal W and analyzed with the MEGA 6 program to infer evolutionary relationships among the various orthologous isoforms.

Phylogenetic reconstructions were performed according to neighbor-joining (NJ; Saitou and Nei, 1987) method. The robustness of tree topologies was tested by the nonparametric bootstrap test (Felsenstein, 1985), with 1000 replicates.

Transcripts localization by in situ *hybridization (ISH)*

For the localization of mRNAs, sense and antisense probes for BsSOX2, BsMSI2, BsPAX2/5/8, BsHOX1 and BsHOX3 transcripts were obtained using T7 RNA- and SP6 RNA-polymerase. Probes were further purified with mini-Quick Spin Columns (Roche) and used on both haemocyte monolayers and whole colonies.

Haemolymph was collected with a glass micropipette after puncturing, with a fine tungsten needle, the tunic marginal vessels of the colonies. It was diluted 1:1 in 0.38% Na-citrate in FSW

(as anti-agglutinating agent), pH 7.5, and then centrifuged at 780 x g for 10 min, at room temperature. The pellet was then re-suspended in FSW to get a final concentration of $5x10^5$ haemocytes/ml. Cells were left to adhere for 30 min on Superfrost Plus (Menzel-Glaser) slides and fixed in freshly prepared MOPS buffer (0.1 M MOPS, 1 mM $MgSO_4$, 2 mM EGTA, 0.5 M NaCl) and 4% paraformaldehyde for 30 min. Whole colonies, at various phases of the blastogenetic cycle, were fixed in 4% paraformaldehyde, 1% glutaraldehyde in 0.2 M cacodylate buffer containing 1.6% NaCl, dehydrated and clarified by treatment with xylene. After pre-hybridisation in Hybridization Cocktail 50% Formamide (Amresco) for 1 h at 55°C, cells and colonies were incubated with sense and antisense probes (2 µg/ml digoxigenin-labelled riboprobe in Hybridization Cocktail) overnight, at 55°C. Nonspecific single filament RNA transcripts were removed by incubation, for 30 min at 37°C, in RNase solution. Haemocytes were then incubated with 5% powdered milk in phosphate-buffered saline (PBS: 8 g/l NaCl, 0.2 g/l KCl, 0.2 g/l KH_2PO_4, 1.15 g/l Na_2HPO_4, pH 7.2, and colonies in 1% powdered milk in TBS, for 1h, to prevent unspecific labelling, and samples were then treated with 1:3000 anti-digoxigenin-AP FAB antibodies (Roche), overnight, at 4°C. Positivity was revealed by the incubation in 3% NBT-BCIP solution (Sigma), substrate for alkaline phosphatase. Haemocytes were finally mounted with Acquovitrex (Carlo Erba) whereas colonies were dehydrated, included in Paraplast Plus Xtra (Sigma), cut into 7-µm sections with a Jung microtome, mounted with Eukitt and samples were finally observed under the light microscope at 1250x.

Quantitative real-time PCR (qRT-PCR)

To estimate the level of transcripts of *bssox2*, *bsmsi2*, *bspax2/5/8, bshox1* and *bshox3* during the blastogenetic cycle, we performed qRT-PCR with the SYBR green method (KAPA SYBR FAST qPCR Kit Master Mix (2X) Universal, KAPA Biosystems). mRNA was extracted from pools of 10 systems, from different colonies at various phases of the blastogenetic cycle, to evaluate transcription changes in physiological conditions. Considered phases were 9/7/1, 9/8/2, 9/8/3, 9/8/4, 9/8/5, 11/8/6 (see Manni et al., 2007), the latter corresponding to the generation change or take-over (TO). The three numbers separated by slashes correspond to the developmental stages of the adult zooids, primary and secondary buds, respectively (Manni et al., 2007). Forward and reverse primers for the considered genes and and b*s-βactin*, the latter used as housekeeping gene, were synthesised by Sigma Aldrich (Table 1). The stable expression level of *bsβ-actin* (Campagna et al., 2016) explains the choice of cytoplasmic actin as reference gene in qPCR experiments. Analyses were carried out using Applied Biosystems 7500 Fast Real-Time PCR System. cDNA synthesis was carried out as described above. qRT-PCR was performed with the following cycling parameters: 95°C for 3 min and then 45 cycles of 95°C for 15 s and 60°C for 1 min. Each set of samples was run three times and each plate contained cDNA from three different biological samples and negative controls. The method of $2^{-\Delta\Delta Ct}$

(Livak and Schmittgen, 2001) was used to estimate gene transcription level. The amounts of transcripts in different conditions were normalised to *β-actin* in order to compensate for variations in the amounts of cDNA.

Statistical analyses

Each experiment was replicated at least three times with three independent samples(n = 3); data are expressed as mean ± SD. The significance of the qRT-PCR results was assessed with the ANOVA; means were compared with the Duncan's test (Snedecor and Cochran, 1980).

Results

Identification and characterisation of transcripts and protein organization

The BsSOX2 transcript is 2138 bp in length and contains a coding sequence (CDS) of 1170 bp with 5' UTR of 95 bp and a 3' UTR of 873 bp. The gene includes two exons carrying the ATG (start) and TAA (stop) codons, respectively. The *in silico* translation of BsSOX2 results in a putative protein of 389 amino acids and contain a DNA-binding domain, the high-mobility group (HMG) box domain from residue 367 to 578. This domain usually consists of three α–helices in an irregular array; the presence of a single HMG-box motif justifies the inclusion of BsSOX2 in the group of sequence-dependent DNA recognition proteins (O'Flaherty and Kaye, 2003).

The gene for BsMSI2 is organized in 7 exons with the ATG start codon located in the first exon and the TAA stop codon in exon number 5. All the introns are provided of the canonical GT and AG splicing signal consensus. It gives a transcript of 2547 bp length, with a 876 bp CDS, a 410 bp 5'UTR and a 1261bp 3'UTR. The sequence of BsMSI2 protein is 291 amino acids in length and contains a conserved motif called RRM (ribonucleoprotein-type RNA recognition motifs) from residue 7 to 79. Unlike vertebrate MSI2, that contain two RRMs, BsMSI2 has only a single motif with two core conserved amino acid sequences, termed RNP-1 ($F^{9}V^{10}G^{11}G^{12}L^{13}$) and RNP-2 ($R^{47}G^{48}F^{49}G^{50}F^{51}V^{52}T^{53}F^{54}$).

The transcript of *bspax2/5/8* is 2333 bp in length, with a 5' UTR of 89 bp and a CDS of 2244 bp. The *bspax2/5/8* gene is composed of 11 exons, all with the canonical GT and AG splicing signal consensus and the ATG start codon in the first exon and the TGA stop codon in the last one. The BsPax2/5/8 predicted protein results of 747 amino acids with the conserved paired box genes (PAX) domain from residue 29 to 153 and the conserved octapeptide ($Y^{194}S^{195}I^{196}N^{197}G^{198}I^{199}L^{200}G^{201}$), tipical of PAX2/5/8 subfamily.

The BsHox1 transcript is 1027 bp with a 612 bp CDS and a 415 bp 5' UTR with the corresponding gene composed of 2 exons. The *in silico* translation results in a 203-amino acid protein with the homeodomain (residues 10 to 72) that presents a well conserved amino acid sequence ($N^{16}F^{17}T^{19}T^{23}A^{38}N^{50}T^{52}Q^{65}R^{68}$). The BsHox3 transcript presents a CDS of 1863 in

length with the corresponding gene split in 3 exons. Its *in silico* translation gives a putative protein of 620 amino acids with the specific homeodomain (residue 497-559) and the pentapeptide $Y^{469}P^{470}W^{471}M^{472}K^{473}$ in the N terminal region of the homeodomain.

Phylogenetic analysis

Phylogenetic analysis, carried out on the deduced amino acid sequences of the considered proteins, gave similar results, with the *B. schlosseri* protein, closely related to the vertebrate ones (Supplementary data). In each NJ tree, tunicates were well separated from echinoderms and cephalochordates. From the tree of Pax2/5/8 we can also observe that cephalochordates and urochordates orthologues clusters as sister group of the Pax2, Pax5 and Pax8 group of vertebrates.

ISH

The specific riboprobes, when used on colony sections, labelled different tissues and cells (fig 1). Using riboprobes for BsSOX2 we observed labelling of haemocytes close to the endostyle and of the oocytes. ISH performed by BsMsi2 labelling of endostyle and haemocytes close to it, oocytes and primary bud branchial basket. BsPax2/5/8 was found in the prebranchial zone close to the ciliated funnel of the adult, in the branchial chamber and endostyle cells of the primary bud, in phase 9/8/2, in the peribranchial chamber of the secondary buds, as well as in oocytes. BsHox1 labelled the branchial chamber of the secondary bud and the oocytes, whereas BsHox3 showed labelling of the endostyle, of the nearby haemocytes, and of oocytes; label was also present in the prebranchial zone close to the ciliated funnel of the adult and in the neural complex. When assayed on haemocyte monolayers, the five digoxigenin labelled RNA probes recognised a fraction of morula cells: 3% for BsSox2, 23% for BsMsi2, 13.4% for BsPax2/5/8, 35% for BsHox1 and 29.4% for BsHox3 (fig 1 d, i, n, q, v).

qRT-PCR

Use of specific primers, designed on the nucleotide sequence of *B. schlosseri bssox2, bsmsi2, bspax2/5/8, bshox1 and bshox3* transcripts, allowed me to evaluate a significant ($p<0.05$) increase of transcription of all the considered genes in late mid cycle phase (9/8/4) of the blastogenetic cycle, with respect to the other considered phases (fig 2).

Discussion

In the colonial ascidian *B. schlosseri* there are two developmental strategies, embryogenesis and blastogenesis, leading to the same end-product, i.e., a functional zooids with a complete neural complex. This allows to compare two apparently different neurogenetic mechanisms that converge to a common final product: the adult neural complex. Neurogenesis is usually dependent on the presence of stem cells; therefore, the possibility to compare two neurogenesis

mechanisms (Manni and Burighel, 2006) stresses the relevance of *B. schlosseri* for neural stem cells research.

In this study I identified and characterised, in *B. schlosseri* transcriptome and genome, two orthologues of vertebrate neural stem cell markers (*bssox2* and *bsmsi2*) and analysed the transcripts of three genes known be involved in ascidian larvae neurogenesis (*Pax2/5/8*, *Hox1* and *Hox3).*

Vertebrate Sox1, Sox2 and Sox3 are classified as member of the subfamily Sox B1, with Sox1 involved in central nervous system development and in particular in eye development, Sox2 with a central role in maintaining self-renewal, or pluripotency, of undifferentiated embryonic and neural stem cells and Sox3 involved in the regulation of embryonic brain development. Previous functional studies on SoxB1 subfamily genes suggest their involvement in consolidating early neural fate and their expression have been observed during development and in proliferating neural progenitors (Graham et al, 2003). Sox2 contains a DNA-binding domain called high-mobility group (HMG) domain and is considered one of the earliest markers for vertebrate neural plate cells (Bouzas et al, 2016). In *B. schlosseri,* I found a BsSox2 transcript with the conserved HMG domain that, in phylogenetic reconstruction, clusters with the other tunicate and cephalocordate Sox2 orthologues as sister group of vertebrates. The anterior ventral area of the endostyle and the region beneath it, are considered stem cell niche in botryllid ascidians (Voskoboynik et al, 2008). The use of BsSox2 specific antisense riboprobe, showed its presence in the haemocytes close to the endostyle.

In the vertebrate adult brain, neural stem cells are located in the subventricular zone where they undergo self-renewal and perform neurogenesis (Shimozaki, 2014). In addition to Sox2, these cells show the expression of Musashi 2 (Msi2), a neural RNA-binding protein. Msi genes contains a conserved motif called RRM (ribonucleoprotein-type RNA recognition motifs) with two core domains termed RNP-1 and RNP-2 (Gasparini et al, 2011). An intense expression level of Msi2 was found in proliferating cells of central nervous system from postnatal development to adult brain of mammals. In this study, I identified a BsMsi2 transcript with a conserved domain structure that cluster with chordates as vertebrate sister group. The expression studies, with BsMsi2 antisense specific riboprobes, reveal labelling in the branchial epithelium primordia: probably related to the role of Msi2 in the regulation of cell proliferation (DeAndrès-Aguayo et al, 2012). The additional labelling of cells in the endostyle and in the ventral island, well-known stem cell niches (Voskoboynik et al, 2008; Rinkevich, 2013), support the relation of BsMsi2 with stem cells. In the last few years, several studies suggested a supplementary role of Msi2 in haematopoietic system with important functions in normal and malignant haematopoiesis (DeAndrès-Aguayo et al, 2012; Kharas et al, 2010; Hope et al, 2010).

A family of genes expressed during neural development in chordates are the paired box genes (PAX) of which four subfamilies are known (PAX1/9, PAX2/5/8, PAX3/7, PAX4/6).

Vertebrate *Pax2, Pax5* and *Pax8* genes were identified in the *HrPax2/5/8* of *H. roretzi* and *CiPax2/5/8* of *C. intestinalis* orthologous. Ascidian Pax2/5/8 expression was found in the anterior neural plate of the early neurulae and in a restricted region of the posterior brain (the neck region) of the larva that represents an important source of adult neurons (Dufour et al, 2006). Here I identified a BsPax2/5/8 transcript that clusters with other invertebrate chordate Pax2/5/8 orthologues. An adult important region of the central nervous system is the prebranchial zone close to the cerebral ganglion. Our *ISH* results performed by BsPax2/5/8 show the presence of the relative mRNA in the ciliated funnel cells. In addition, labelling was localised in the endostyle, known to be a stem cell niche, and in the well active proliferating cells of the secondary bud, the peribranchial chamber.

The ascidian genome, small and simple, presents a single Hox cluster, corresponding to the mammalian HOX D cluster, that might represent the primordial chordate cluster (Di Gregorio et al, 1995). *H. roretzi Hox1*, orthologue of vertebrate *Hox1* gene, was observed in a small number of cells in the left side of sensory vesicle and in the anterior part of the larval visceral ganglion, posterior to the once of HrPax2/5/8 expression. The expression of the gene dropped to lower levels after metamorphosis (Wada et al, 1998). In the posterior region of the visceral ganglion of the larvae, another Hox gene expression was documented. This gene presents an high homology level with vertebrates Hox3 gene (Locascio et al, 1999).

In my work on *B. schlosseri*, I identified two Hox genes called *bshox1* and *bshox3*. In the phylogenetic reconstruction, BsHox1 clusters with other urochordates; BsHox3 localise in the vertebrate sister group position. Again, BsHox1 mRNA is located in the branchial chamber of the secondary bud, an active proliferating tissue, rich in interstigmatic nerve (Burighel et al, 2001; Zaniolo et al 2002). In botryllid ascidians, adult neural complex localize between the oral and cloacal siphons. Antisense riboprobe for BsHox3 shows the presence of the transcript in this region, suggesting a role in neural cell differentiation. The additional labelling of both *hox* genes in the endostyle niche supports our view on BsHox1 and BsHox3 stem cell molecular markers identity.

In addition to the study of the transcripts in the *Botryllus* tissues and haemocytes, we performed even an expression study, comparing the different blastogenetic phases of the cycle. qRT-PCR on BsSox2, BsMsi2, BsPax2/5/8, BsHox1 and BsHox3 reveal a significantly ($p<0.05$) higher expression level of all the transcripts in the late mid-cycle phase (9/8/4). During that period, neuroblasts begin to proliferate in the secondary bud and continue to proliferate even in the primary bud, in order to form cerebral ganglion. The neuroblasts proliferation matches with over expression of putative neural stem cell markers, straightening, their role in the identification of neural precursors.

All our considered transcript, BsSox2, BsMsi2, BsPax2/5/8, BsHox1 and BsHox3, vertebrate neural gene orthologues, co-localise in the endostyle niche and in the epithelial primordia of the

branchial basket, with an elevated expression in the 9/8/4 phase. These observation togheter with previous results that identified Pax2/5/8, Hox1 and Hox3 in other urochordates, allow me to hypothesise that they may be good neural stem cells molecular markers also in ascidians.

Furthermore, the presence of the mRNAs of BsSox2, BsMsi2 and BsHox3 inside the oocytes suggests that they are stored inside female gametes in order to assure the early development, when gene transcription has not yet started in the embryo.

As regards the fraction of labelled morula cells, it can be ascribed to a particular differentiation stage of this cells that can represent either not yet fully differentiated cells or, alternatively, adults cells with some peculiar functions requiring the activity of the genes.

References

Bollner T, Howalt S, Thorndyke MC, Beesley PW. 1995. Regeneration and post-metamorphic development of the central nervous system in the protochordate *Ciona intestinalis*: a study with monoclonal antibodies. Cell Tissue Res. 279: 421-32

Bouzas SO, Marini MS, Torres ZE, Buzzi AL, Morales VDA, Strobl-Mazzulla PH. 2016. Epigenetic activation of Sox2 gene in the developing vertebrate neural plate. Mol Biol Cell 27: 1921-7

Campagna D, Gasparini F, Franchi N, Vitulo N, Ballin F, Manni L, Valle G, Ballarin L. 2016. Transcriptome dynamics in the asexual cycle of the chordate *Botryllus schlosseri*. BMC Genomics 17: 275

DiGregorio A, Spagnuolo A, Ristoratore F, Pischetola M, Aniello F, Branno M, Cariello L, DiLauro R. 1995. Cloning of ascidian homeobox genes provides evidence for a primordial chordate cluster. Gene 156: 253-257

Gasparini F, Shimeld SM, Ruffoni E, Burighel P, Manni L. 2011. Expression of a Musashi-like gene in sexual and asexual development of the colonial chordate *Botryllus schlosseri* and phylogenetic analysis of the protein group. J Exp Zool B Mol Dev Evol 316: 562-73

Gasparini F, Manni L, Cima F, Zaniolo G, Burighel P, Caicci F, Franchi N, Schiavon F, Rigon F, Campagna D, Ballarin L. 2015. Sexual and asexual reproduction in the colonial ascidian *Botryllus schlosseri*. Genesis 53: 105–120

Graham V, Khudyakov J, Ellis P, Pevny L. 2003. SOX2 functions to maintain neural progenitor identity. Neuron 39: 749-65

Jeffery WR, Strickler AG, Yamamoto Y. 2004. Migratory neural crest-like cells form body pigmentation in a urochordate embryo. Nature. 431: 696-699

Julin C. 1981. Recherches sur l'organisation des ascidies simples. Sur l'hypophyse et quelques organes qui s'y attachent, dans les genres *Corella, Phallusia* et *Ascidia.* Arch Biol Paris. 2: 59-126

Katsuyama Y, Wada S, Yasugi S, Saida 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

Kawashima T, Murakami AR, Ogasawara M, Tanaka K, Isoda R, Sasakura Y, Nishikata T, Okano H, Makabe KW. 2000. Expression patterns of musashi homologs of the ascidians, *Halocynthia roretzi* and *Ciona intestinalis*. Dev Genes Evol 210: 162-5

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

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-514 time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 25: 402–408

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

Mackie GO, Burighel P, Caicci F, Manni L. 2006. Innervation of ascidian siphons and responses to stimulation. Can J Zool 84: 1146-1162

Manni L, Lane NJ, Sorrentino M, Zaniolo G, Burighel P. 1999. Mechanism of neurogenesis during the embryonic development of a tunicate. J Comp Neurol. 412: 527-41

Manni L, Lane NJ, Joly JS, Gasparini F, Tiozzo S, Caicci F, Zaniolo G, Burighel P. 2004. Neurogenic and non-neurogenic placodes in ascidians. J Exp Zool 302B: 483-504

Manni L, Agnoletto A, Zaniolo G, Burighel P. 2005. Stomodeal and neurohypophysial placodes in *Ciona intestinalis:* insights into the origin of the pituitary gland. J Exp Zool B 304B: 324-339

Manni L, Burighel P. 2006. Common and divergent pathways in alternative developmental processes of ascidians. Bioessays 28: 902-912

Manni L, Zaniolo G, Cima F, Burighel P, Ballarin L. 2007. *Botryllus schlosseri*: a model ascidian for the study of asexual reproduction. Dev Dyn 236: 335-352

Meinertzhagen IA, Okamura Y. 2001. The larval ascidian nervous system: the chordate brain from its small beginnings. Trends in neuroscience. 24: 401-410

O'Flaherty E, Kaye J. 2003. TOX defines a conserved subfamily of HMG-box proteins. BMC Genomics. 4: 13

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406–425

Shimozaki K. 2014. Sox2 transcription network acts as a molecular switch to regulate properties of neural stem cells. World J Stem Cells. 6: 485-90

Snedecor GW, Cochran WG. 1980. Statistical Methods (7th ed.) Iowa State University Press, Ames, IA

Tang X, Nakata Y, Li HO, Zhang M, Gao H, Fujita A, Sakatsume O, Ohta T, Yokoyama K. 1994. The optimization of preparations of competent cells for transformation of *E coli*. Nucleic acids res 22: 2857-2858

Tiozzo S, Christiaen L, Deyts C, Manni L, Joly JS, et al. 2004. Embryonic versus blastogenetic development in the compound ascidian *Botryllus schlosseri* insights from Pitx expression patterns. Dev Dyn 232: 468-478

Voskoboynik A, Neff NF, Sahoo D, Newman AM, Pushkarev D, Koh W, Passarelli B, Fan HC, Mantalas GL, Palmeri KJ, Ishizuka KJ, Gissi C, Griggio F, Ben-Shlomo R, Corey DM, Penland L, White RA, Weissman IL, Quake SR. 2013. The genome sequence of the colonial chordate, *Botryllus schlosseri*. Elife 2: e00569–e00569

Wada H, Saiga H, Satoh N, Holland PW. 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

Willey A. 1893. Studies on the Protochordata. II. The development of neuro-hypophyseal system in *Ciona intestinalis* and *Clavelina lepadiformis*, with an account of the origin of sense organs in *Ascidia mentula*. Q J Microsc Sci 35: 295-334

Zhang S, Cui W. 2014. Sox2, a key factor in the regulation of pluripotency and neural differentiation. World J Stem Cells. 6: 305–311

Table 1. PCR primer used and relative melting temperatures (Tm).

Primer	**Tm(°C)**	**Sequence 5'-3'**
BsSOX2F	55.5	TCAGCACAGAGCCAAATGAC
BsSOX2R	60.5	CAGCAAGTAGCCCACATCCCG
BsSOX2R-RT	57.8	ATGGTGTGGTCTCGTGGC
BsMSI2F	59.6	GTTGGCGGTCTATCTTCGGACAC
BsMSI2R	59.8	GGAAGACCCGTGGCTGGAT
BsMSI2F-RT	59.2	GACCCTGCGACAGTGATGTATCTTG
BsPAX2/5/8F	53.4	CTGCTAAAAGGGAAGGCAG
BsPAX2/5/8R	56.5	TACCTATTCCGTCTGCGAGC
BsPAX2/5/8F-RT	53.7	GGGAATGATAATGCCACACC
BsPAX2/5/8R-RT	57.4	ACTGAGCAGAACAGCGGAAG
BsHOX1F	49.0	GACTGAACTGGAGAAGG
BsHOX1R	63.7	GCAGCGACATCCAACGGCAGG
BsHOX1F-RT	54.5	CCCTGAATGAAACGCAAGTG
BsHOX1R-RT	59.5	GAAAGCCAACGCCACTCAGC
BsHOX3F	58.8	CCATTTCAATAGATACCTCTGCCGACC
BsHOX3R	60.9	GGCGAGGGACGATGGCTTTTTC
BsACTF-RT	55.6	CTCAACCCGAAGGCTAACC
BsACTR-RT	56.2	GTACAGCGACAGAACGGC

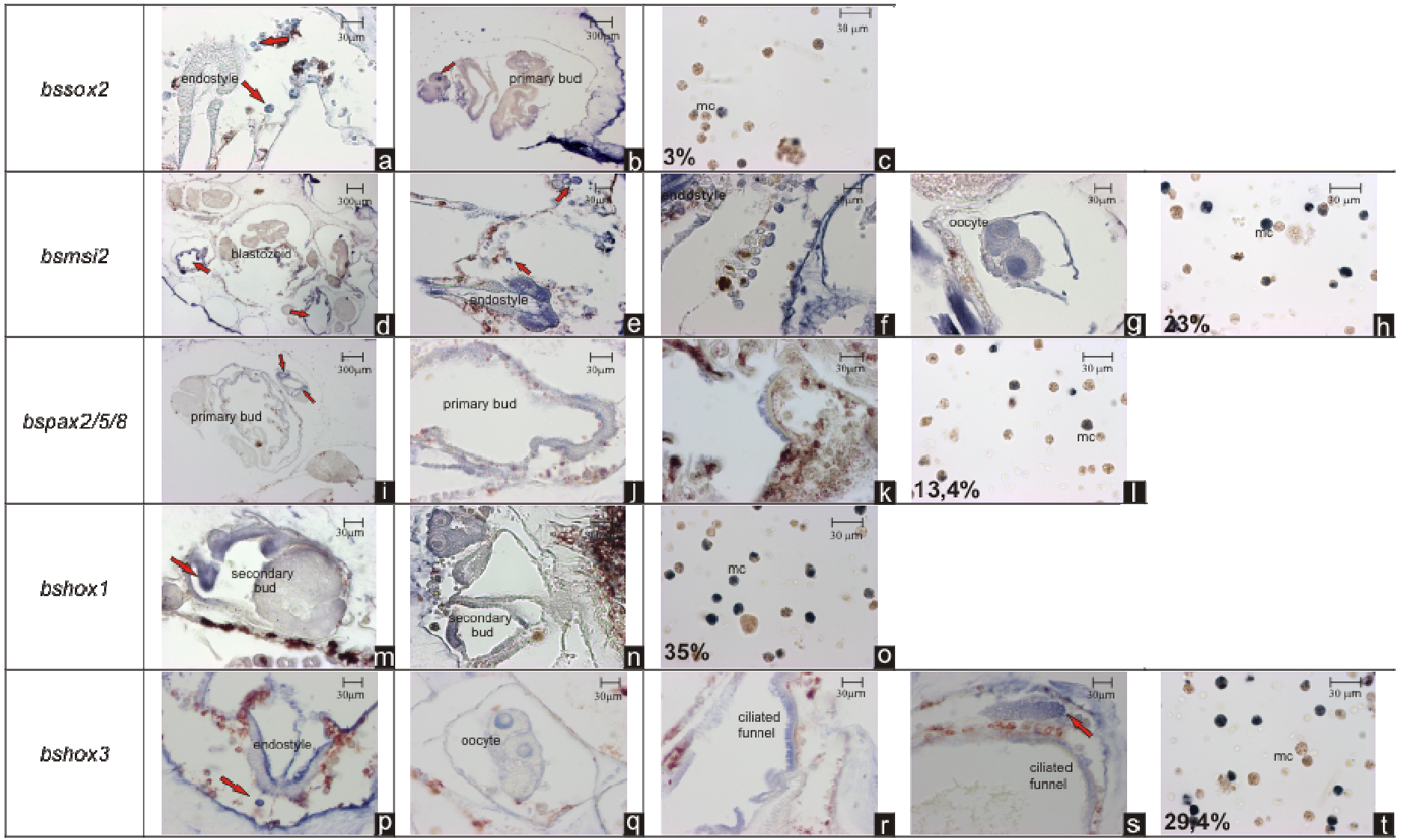

bssox2
endostyle
a
primary bud
b
mc
3%
c
bsmsi2
blastozoid
d
endostyle
e
endostyle
f
oocyte
g
mc
23%
h
bspax2/5/8
primary bud
i
primary bud
j
k
mc
13,4%
l
bshox1
secondary bud
m
secondary bud
n
mc
35%
o
bshox3
endostyle
p
oocyte
q
ciliated funnel
r
ciliated funnel
s
mc
29,4%
t

Fig1: ISH of tissues sections and haemocytes monolayers stained with antisense riboprobes for *bssox2, bsmsi2, bspax2/5/8, bshox1, bshox3.* (a-b) sections and (c) haemocytes labelled by BsSOX2, red arrow indicates haemocytes in (a) and oocyte (b); (d-g) sections and (h) haemocytes labelled by BsMSI2, red arrow indicates primary bud branchial basket primordia (d) and haemocytes in the endostyle stem cell niche (e,f); (i-k) sections and (l) haemocytes labelled by BsPax2/5/8, red arrow indicates secondary bud peribranchial chambers (i); sections (m-n) and (o) haemocytes labelled by BsHox1, red arrow indicate branchial chamber of the secondary bud (m); sections (p-s) and (t) haemocytes labelled by BsHox3, red arrow indicate haemocyte in the stem cell niche (p) and neural complex (s). mc= morula cells.

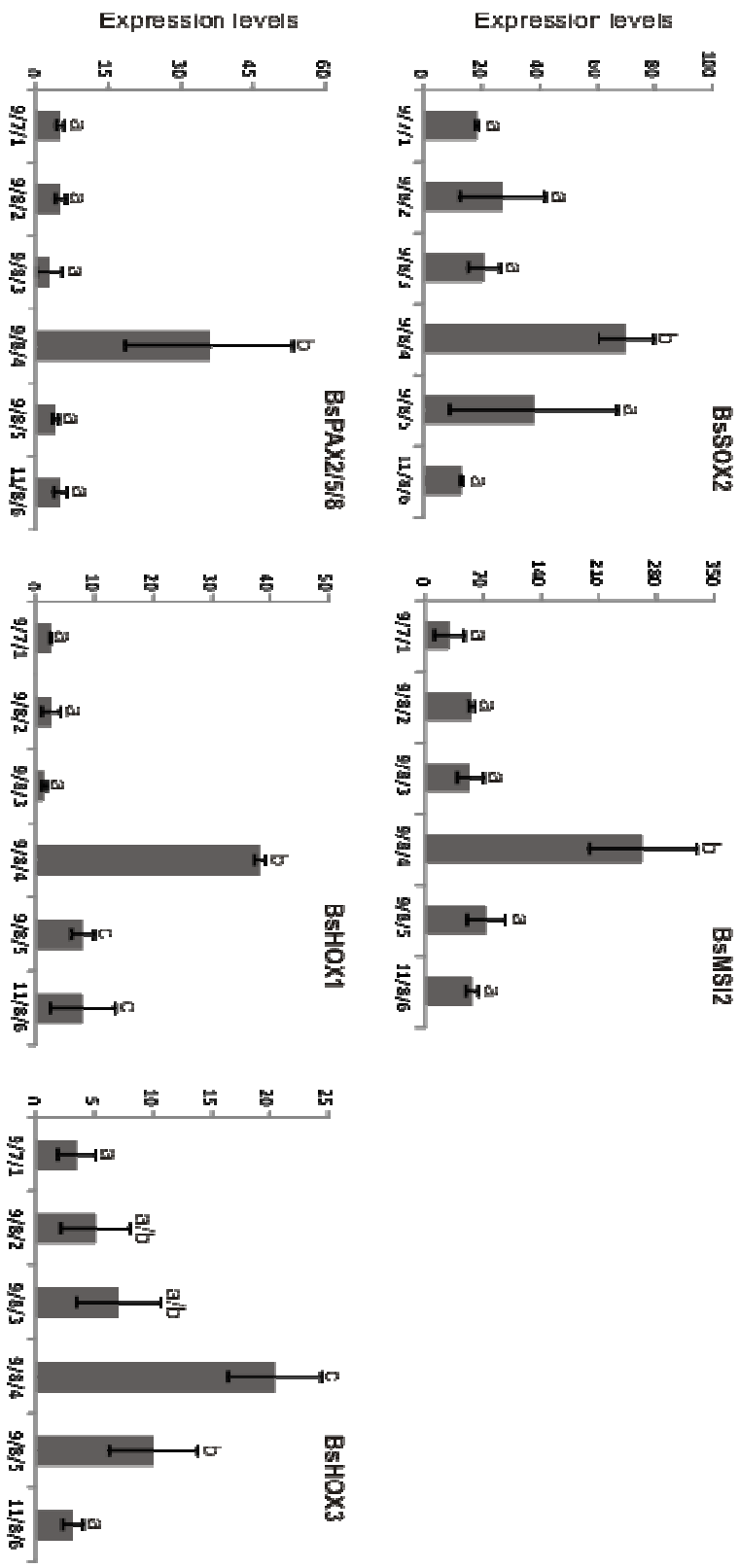
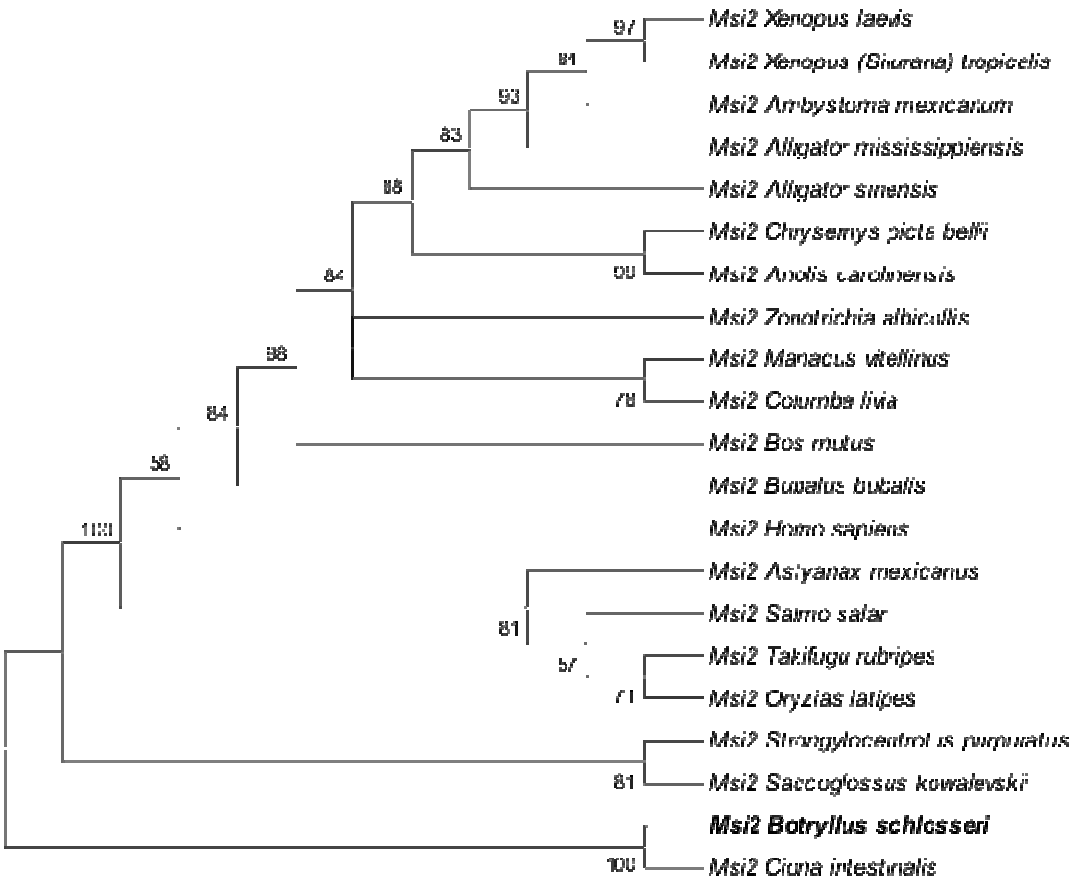


Fig2: Relative expression levels of *bssox2, bsmsi2, bspax2/5/8, bshox1, bshox3* in different phases of the blastogenetic cycle. Normalisation of expression was achieved using endogenous β-actin as housekeeping gene. Each bar of the histogram corresponds to the average of three independent experiments ± SD. Different letters indicate significant differences ($p < 0.05$).

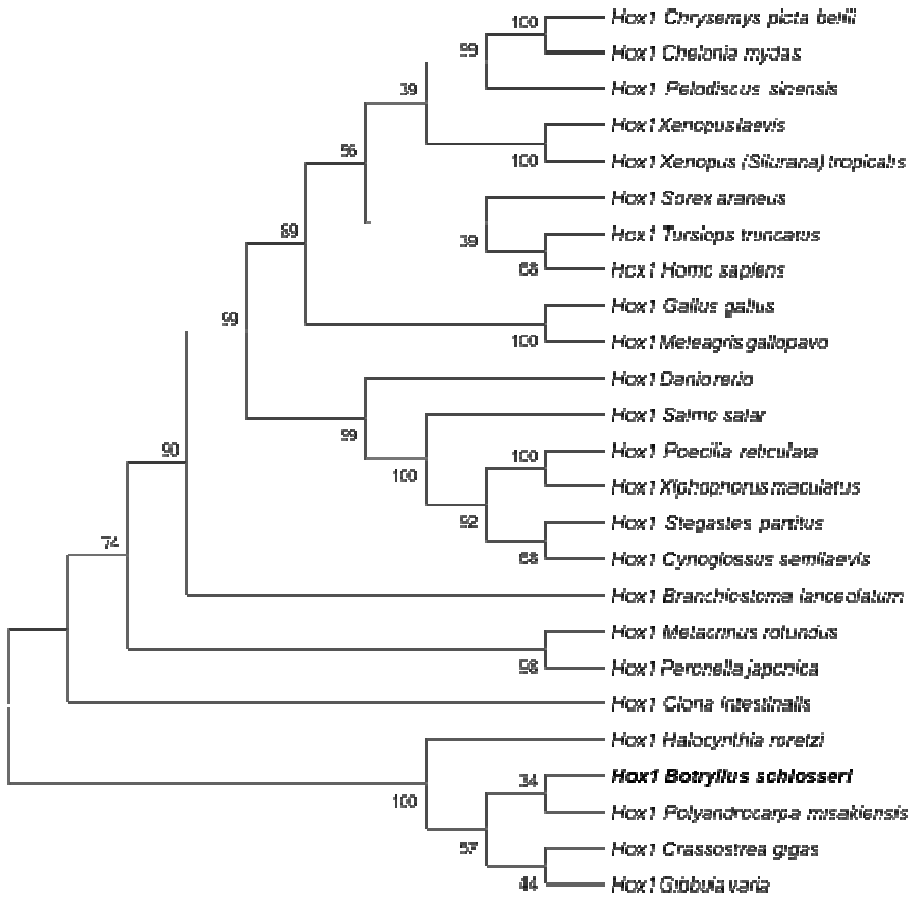
FigS1: Evolutionary relationships (NJ) among *sox2* proteins with bootstrap confidence values indicated at the left of each branch. BsSox2 is indicated in bold.



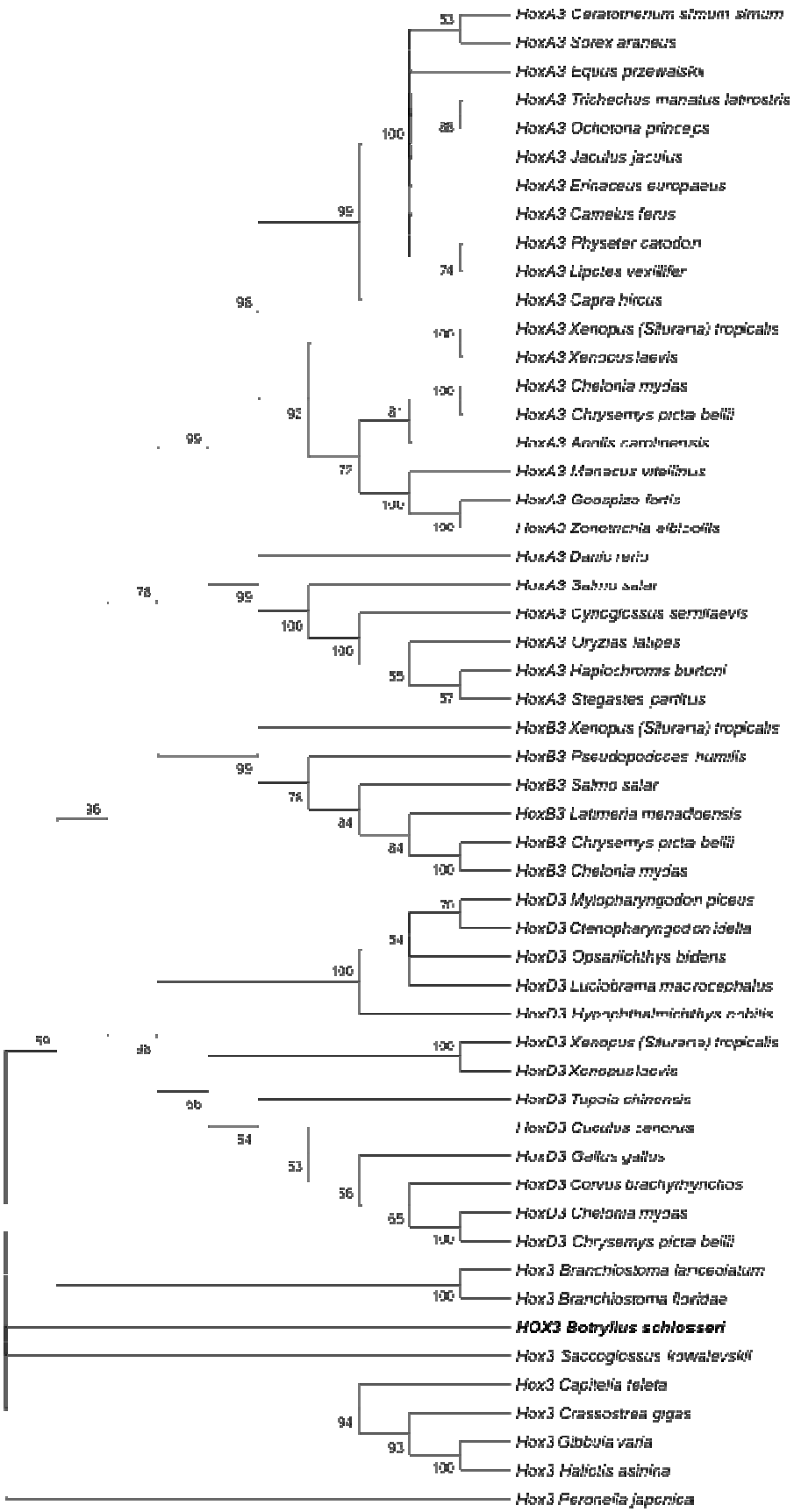
FigS2: Evolutionary relationships (NJ) among *msi2* proteins with bootstrap confidence values indicated at the left of each branch. BsMsi2 is indicated in bold.


Pax-2 Chelonia mydas
Pax-2 Pseudaspis cinerea
Pax-2 Chrysemys picta bellii
Pax-2 Picadula abbicilis
Pax-2 Taeniopygia guttata
Pax-2 Aenigma fortis
Pax-2 Menacus vrellicus
Pax-2 Columba livia
Pax-2 Gallus gallus
Pax-2 Python bivittatus
Pax-2 Anolis carolinensis
Pax-2 Orcinus gamelli
Pax-2 Macaca fascicularis
Pax-2 Oryctenopus cuniculus
Pax-2 Chrysochloris asiatica
Pax-2 Mus musculus
Pax-2 Sus scrofa
Pax-2 Canis familiaris
Pax-2 Xenopus laevis
Pax-2 Danio rerio
Pax2 Haplodactylus zebra
Pax-2 Stegastes partitus
Pax-2 Oryzias latipes
Pax-2 Dicentrarchus labrax
Pax-5 Pygocentrus nattereri
Pax-5 Haplochromis burtoni
Pax-5 Poecilia reticulata
Pax-5 Danio rerio
Pax-5 Carassius auratus langsdorfii
Pax-5 Xenopus (Silurana) tropicalis
Pax-5 Xenopus laevis
Pax-5 Struthio camelus australis
Pax-5 Acanthisitta chloris
Pax-5 Taeniopygia guttata
Pax-5 Geospiza fortis
Pax-5 Buceros rhinoceros silvestris
Pax-5 Melopsittacus undulatus
Pax-5 Anas platyrhynchos
Pax-5 Pan paniscus
Pax-5 Microtus ochrogaster
Pax-5 Columba livia
Pax-5 Homo sapiens
Pax-5 Anolis carolinensis
Pax-5 Papio anubis
Pax-5 Ectopistes hilfaki
Pax-5 Macaca fascicularis
Pax-5 Pteropus alecto
Pax-9 Tetraodon nigroviridis
Pax-9 Oryzias latipes
Pax-9 Oncorhynchus mykiss
Pax-9 Clarias batrachus
Pax-9 Astyanax mexicanus
Pax-9 Xenopus (Silurana) tropicalis
Pax-9 Xenopus laevis
Pax-9 Chrysemys picta bellii
Pax-9 Chelonia mydas
Pax-9 Orcinus orca
Pax-9 Tursiops truncatus
Pax-9 Physeter catodon
Pax-9 Rattus norvegicus
Pax-9 Microtus ochrogaster
Pax-9 Mesocricetus auratus
Pax-9 Cricetulus griseus
Pax2/5/8 Branchiostoma floridae
Pax2/5/8 Saccoglossus kowalevskii
Pax2/5/8 Hydra magnipapillata
Pax2/5/8 Petromyzon marinus
Pax2/5/8 Latimeria chalumnae
Pax2/5/8 Danio rerio
Pax2/5/8 Ciona intestinalis
Pax2/5/8 Strongylocentrotus purpuratus
Pax2/5/8 Lytechinus variegatus

FigS3: Evolutionary relationships (NJ) among *pax2/5/8* proteins with bootstrap confidence values indicated at the left of each branch. BsPax2/5/8 is indicated in bold.



FigS4: Evolutionary relationships (NJ) among *hox1* proteins with bootstrap confidence values indicated at the left of each branch. BsHox1 is indicated in bold.



FigS4: Evolutionary relationships (NJ) among *hox3* proteins with bootstrap confidence values indicated at the left of each branch. BsHox3 is indicated in bold.

Conclusion

In the recent years, the awareness of the importance of ascidians as model organisms for the studying of different biological processes has growly increased. The close evolutionary relation to vertebrates makes urochordates useful organisms in the study of developmental biology, ageing, immunology, genetics and evolutionary biology (Rinkevich, 2002). After the discovery of stem cell niches and the observation of hypothetical undifferentiated cells, the colonial ascidian *B. schlosseri,* is becoming an interesting animal even for studies on stem cells cause it allows to elucidate the mechanism involved in their self-renewal and differentiation. Brand and Livesey (2011) have already demonstrated the importance of invertebrate model system in the understanding of stem cell biology. Moreover, colonial ascidians are the only chordates, with the possibility to perform both sexual and asexual reproduction. The larval structures, originate in the embryonic development, whereas the bud structures, originate in the blastogenetic development: they will give rise, after metamorphosis and growth to an adult zooid (either oozooid or blastozooid) with the same organization and body morphology. The two developmental strategies begin from different populations of stem cells, totipotent in the sexual reproduction and multi- or pluripotent in the asexual one, probably controlled by different signal molecules during their differentiation process. The ability of undifferentiated stem cells to respond at different signals is linked to the presence of specific receptors able to bind them: this interaction acts as a guide to translate an external signal in an activated internal pathway. These receptors can be used as stem cell markers in order to recognise stem cells, and characterise stem cell niches. To this aim, I searched for orthologues of know vertebrate haematopoietic and neural markers in *B. schlosseri* transcriptome and genome. In this research, I identified four sequences with similarity to genes expressed in vertebrate haematopoietic progenitor cells, called *bsabcg2, bscd133, bsgata1/2/3* and *bsgata4/5/6*. All these putative molecular markers co-localise in the sub-endostylar haemolimph lacunae, a region identified as stem cell niche (Voskoboynik et al, 2008). Moreover, they showed a common over-expression in the mid-cycle phase, characterised by secondary bud tissues formation that require new undifferentiated cells (Kawamura et al, 2008). Collectively these data strongly suggests that the molecular vertebrate orthologues can be used as haematopoietic molecular markers even in the colonial ascidian *B. schlosseri*, that haemocyte turnover occurs during the mid-cycle phase and that haemopoiesis takes place in endostylar niche.
Vertebrate neural markers orthologues, involved in the recognition of neural stem cells, were even found. I identified five putative transcripts, BsSox2, BsMsi2, BsPax2/5/8, BsHox1 and BsHox3 with similarity to vertebrate gene. The *Botryllus* orthologous, co-localise in the endostyle niche and in the primordia of the branchial basket nerves region. The former is considered a stem cell niche, whereas the latter is a region in which, when the bud complete its growth and became adult. The putative neural markers, like haematopoietic ones, had an high expression levels in the mid-cycle phase, in which the new bud neural complex take place. The co-localisation of the considered neural transcript in the stem cell niche, in addition to the over-expression in the mid-cycle phase, where new cells are needed for the secondary bud tissues formation, allow us to conclude that *bssox2, bsmsi2, bspax2/5/8, bshox1* and *bshox3* may be good neural stem cells markers even in *B. schlosseri* and hypothesise that cells labelled by both haematopoietic and neural markers, in the endostyle stem niche, are a pool of long- living pluripotent stem cells.

For every haematopoietic and neural considered genes, *ISH* assay reveal labelling of a small fraction of morula cells. This scenario could mean that there are morula cells in different differentiation stage in the *Botryllus* haemolymph, with a population not fully differentiated, or

that, such as vertebrates, that molecules used as stem cells markers play even alternative roles in other tissues. A detailed study of the function of these genes in circulating haemocytes will be the subject of future investigations.

References

Brand AH, Livesey FJ. 2011. Neural stem cell biology in vertebrates and invertebrates: more alike than different? Neuron 70: 719-729

Kawamura K, Tachibana M, Sunanaga T. 2008. Cell proliferation dynamics of somatic and germline tissues during zooidal life span in the colonial tunicate *Botryllus primigenus*. Dev Dyn 237: 1812-1825

Rinkevich B. 2002. The colonial urochordate *Botryllus schlosseri*: from stem cells and natural tissue transplantation to issues in evolutionary ecology. Bioessays 24: 730-740

Voskoboynik A, Soen Y, Rinkevich Y, Rosner A, Ueno H, et al. 2008. Identification of the endostyle as a stem cell niche in a colonial chordate. Cell Stem Cell 3: 456-464