Yamanaka Factors in the Budding Tunicate *Botryllus schlosseri* Show a Shared Spatio-Temporal Expression Pattern in Chordates

Virginia Vanni†, Marika Salonna‡, Fabio Gasparini†, Margherita Martini†, Chiara Anselmi§, Carmela Gissi‡, Lucia Manni†

1Department of Biology, University of Padova, Padova, Italy, 2Department of Biosciences, Biotechnologies and Biopharmaceutics, University of Bari “Aldo Moro”, Bari, Italy, 3Stanford University, Hopkins Marine Station, Pacific Grove, CA, United States, 4Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA, United States, 5IBOM, Institute of Biomembranes, Bioenergetics and Molecular Biotechnologies, Consiglio Nazionale Delle Ricerche, Bari, Italy, 6CoNISMa, Consorzio Nazionale Interuniversitario per le Scienze Del Mare, Roma, Italy

In vertebrates, the four transcription factors *Sox2*, *c-Myc*, *Pou5f1* and *Klf4* are involved in the differentiation of several tissues during vertebrate embryogenesis; moreover, they are normally co-expressed in embryonic stem cells and play roles in pluripotency, self-renewal, and maintenance of the undifferentiated state in adult cells. The *in vitro* forced co-expression of these factors, named Yamanaka factors (YFs), induces pluripotency in human or mouse fibroblasts. *Botryllus schlosseri* is a colonial tunicate undergoing continuous stem cell-mediated asexual development, providing a valuable model system for the study of pluripotency in the closest living relatives of vertebrates. In this study, we identified *B. schlosseri* orthologs of human *Sox2* and *c-Myc* genes, as well as the closest homologs of the vertebrate-specific *Pou5f1* gene, through an in-depth evolutionary analysis of the YF gene families in tunicates and other deuterostomes. Then, we studied the expression of these genes during the asexual cycle of *B. schlosseri* using *in situ* hybridization in order to investigate their possible involvement in tissue differentiation and in pluripotency maintenance. Our results show a shared spatio-temporal expression pattern consistent with the reported functions of these genes in invertebrate and vertebrate embryogenesis. Moreover, *Myc*, *SoxB1* and *Pou3* were expressed in candidate stem cells residing in their niches, while *Pou2* was found expressed exclusively in the immature previtellogenic oocytes, both in gonads and circulating in the colonial vascular system. Our data suggest that *Myc*, *SoxB1* and *Pou3* may be individually involved in the differentiation of the same territories seen in other chordates, and that, together, they may play a role in stemness even in this colonial ascidian.

**Keywords:** ascidians, asexual development, *Botryllus schlosseri*, cMyc, Pou, Klf4, Sox2, Yamanaka factors
INTRODUCTION

Stem cells (SCs) are undifferentiated cells capable of self-renewal and with the ability to differentiate into specialised cells. Given their pivotal role in animal development and regeneration, understanding the molecular mechanisms controlling their potency and evolution is of critical importance. Studies in SC biology are mainly based on mammals or other vertebrates, despite their limited regenerative capacities (Zhao et al., 2016). Tunicates are the sister group of vertebrates (Delsuc et al., 2006), and colonial tunicates are the only chordates able to regenerate entire organs from circulating SCs (Rinkevich et al., 1995; Sunanaga et al., 2006; Voskoboynik et al., 2007; Brown et al., 2009; Kassmer et al., 2019; Manni et al., 2019; Scelzo et al., 2019). As such, this group offers unique opportunities to investigate the similarities and differences in SC biology between vertebrate and invertebrate chordates.

Among colonial tunicates, the ascidian Botryllus schlosseri is recognized as a model species for the study of SCs and asexual reproduction (Manni et al., 2007; Voskoboynik and Weissman, 2015; Manni et al., 2019). This species reproduces both sexually and asexually. In a colony, each individual (zooid) is hermaphroditic and produces swimming larvae via sexual reproduction. Larvae metamorphose into sessile individuals that reproduce asexually by budding (blastogenesis), forming clonal zooids grouped in star-shaped systems and embedded in a common tunic (Figures 1A,B). In each colony, three generations coexist: adult zooids, their primary buds (1Bs), and secondary buds (2Bs). In laboratory conditions (18°C), each week all adult zooids are resorbed during a process called takeover and replaced by their 1Bs; simultaneously, 2Bs become 1Bs and produce a new generation of 2Bs. Each new 2B originates from a thickening of the peribranchial epithelium of a 1B, which further develops into a vesicle undergoing organogenesis. Organogenesis nears completion when the 2B becomes a 1B where tissue differentiation occurs. B. schlosseri SCs are able to freely circulate in the colony haemocoel where, over generations, they contribute to tissue differentiation in buds (Kowarsky et al., 2021). Candidate SCs (cSCs) were identified in B. schlosseri as hemoblast, i.e., round cells, with a diameter of about 5 μm and a high nucleus-cytoplasm ratio (Rosental et al., 2018; Kowarsky et al., 2021). Two cSC transient niches have also been proposed in adult zooids (Figures 1E–H): the haemocoel, ventral to the anterior endostyle (the gland on the pharynx floor which produces the mucus net for filtration) (Voskoboynik et al., 2008; Rosental et al., 2018), and the “cell islands”, i.e., groups of haemocytes in the ventral body wall (Rinkevich et al., 2013).

Cell islands are numerous in adult individuals belonging to the close SoxB2 group (Bowles et al., 2000; Heenan et al., 2016). While the close SoxB2 group consists of Sox14 and Sox21, invertebrates have only a single member each of the SoxB1 and SoxB2 groups (Bowles et al., 2000; Heenan et al., 2016). The role of SoxB1 in pluripotency appears to be conserved among invertebrates (Onal et al., 2012; Niwa et al., 2016), where it is fundamental for neural development and regeneration (Mizuseki et al., 1998; Buescher et al., 2002; Bylund et al., 2003; Graham et al., 2003; Miya and Nishida, 2003; Meulemans and Bronner-Fraser, 2007; Marlow et al., 2009; Ross et al., 2018) and is expressed in multiple adult self-renewing epithelia (Arnold et al., 2011).

Pou5f1 belongs to the metazoan POU gene family, characterized by a modular DNA-binding domain (POU) consisting of a N-terminal POU-specific domain (POUs) and a C-terminal homeodomain (POUh) separated by a linker of variable length. Based upon similarity across the entire POU domain, Pou genes are grouped into six classes (Rosenfeld, 1991), that have evolved through multiple lineage-specific events of gene duplications and losses (Frankenberg and Renfree, 2013; Gold et al., 2014; Onichtchouk, 2016). Pou5f1 belongs to the vertebrate-specific Pou5 class (also including Pou5f3 and the primate/rodent-specific Pou5f2), hypothesised to be arisen from a Pou3-like ancestor together with the bilaterian-specific Pou2 class (Gold et al., 2014).

c-Myc belongs to the MYC gene family, characterized by a C-terminal basic-helix-loop-helix-leucine zipper (bHLHZ) domain mediating protein dimerization and DNA binding, and two N-terminal trans-activation motifs (Atchley and Fitch, 1995; James and Eisenman, 2002). Vertebrates possess multiple Myc genes (c-Myc, N-Myc, L-Myc), with c-Myc involved in cell division and growth control, and considered a central hub for pluripotent SC regulatory networks (Chappell and Dalton, 2013; Fagnocchi and Zippo, 2017). In contrast, two genes are found in a cnidarian with high regenerative ability (Harl et al., 2010), and only a single homolog is present in Drosophila and most other invertebrates (Gallant, 2006; McMerrin and Atchley, 2011). Despite differences in gene number, the expression of the single Myc gene in the undifferentiated cells of invertebrates suggests that its function is similar to that of c-Myc in vertebrates (Kawamura et al., 2008; Harl et al., 2010; Fujiwara et al., 2011).

Sox2 is a member of the ancient SOX gene family that has undergone frequent gene duplications and is therefore divided into ten subgroups (A–J) based on the similarity of the DNA-binding domain (Wegner, 1999; Bowles et al., 2000). In vertebrates, Sox2, Sox1 and Sox3 comprises the SoxB1 group, while the close SoxB2 group consists of Sox14 and Sox21. Invertebrates have only a single member each of the SoxB1 and SoxB2 groups (Bowles et al., 2000; Heenan et al., 2016). The role of SoxB1 in pluripotency appears to be conserved among invertebrates (Onal et al., 2012; Niwa et al., 2016), where it is fundamental for neural development and regeneration (Mizuseki et al., 1998; Buescher et al., 2002; Bylund et al., 2003; Graham et al., 2003; Miya and Nishida, 2003; Meulemans and Bronner-Fraser, 2007; Marlow et al., 2009; Ross et al., 2018) and is expressed in multiple adult self-renewing epithelia (Arnold et al., 2011).

Pou5f1 belongs to the metazoan POU gene family, characterized by a modular DNA-binding domain (POU) consisting of a N-terminal POU-specific domain (POUs) and a C-terminal homeodomain (POUh) separated by a linker of variable length. Based upon similarity across the entire POU domain, Pou genes are grouped into six classes (Rosenfeld, 1991), that have evolved through multiple lineage-specific events of gene duplications and losses (Frankenberg and Renfree, 2013; Gold et al., 2014; Onichtchouk, 2016). Pou5f1 belongs to the vertebrate-specific Pou5 class (also including Pou5f3 and the primate/rodent-specific Pou5f2), hypothesised to be arisen from a Pou3-like ancestor together with the bilaterian-specific Pou2 class (Gold et al., 2014).
FIGURE 1 | Botryllus schlosseri blastogenic cycle and Yamanaka Factors spatio-temporal expression pattern. (A) The weekly blastogenic cycle (dashed blue arrows) begins when a 2B emerges as a thickening of the parental peribranchial chamber epithelium accompanied by the epidermis (stage 1). The 2B then arches (stage 2) before it closes (stage 3) in a “double vesicle”: the outer vesicle is formed by the epidermis, the inner one by the original peribranchial epithelium. The inner vesicle undergoes invaginations (stage 4) to form the branchial and peribranchial chamber rudiments. In the lateral mesenchymal space between the inner and the outer vesicle germ cells reside whereas the heart rudiment forms postero-ventrally. The branchial chamber rudiment forms posteriorly the gut rudiment (stage 5), whereas on
**FIGURE 1** Its ventral floor the endostyle becomes recognizable (stage 6). The passage to stage 7 (clotted light blue arrow) marks the change in generation when the 2B becomes a 1B, producing a new generation of 2Bs at stage 1. At stage 8, when the bud is at stage 2, the heart begins beating. In the following days (days 3–6), the 2B grows and differentiates organs. At the generation change, the 1B opens its siphons and becomes an adult (stage 9; clotted blue arrow). Concurrently, the previous generation of adult zooids undergoes resorption (takeover). Every week, at 18°C, there is a generation change and the lifespan of an individual (from 2B at stage 1 to adult zooid in regression) lasts ~3 weeks. Numbers indicate the stages: numbers 1-1 refer to the 2B; numbers 7-8 to the 1B; numbers 9–11 to adult zooids. The triplet of numbers 9/7/1, 9/8/2, 9/8/3, 9/8/4, 9/8/5, and 11/8/6 refer to the colony phase; the first number referring to adult zooids, the middle to 1Bs and the last to 2Bs (Gasparini et al., 2015; Manni et al., 2007; Manni et al., 2014; Manni et al., 2019). 2Bs at stages 1–3 are in lateral view, other zooids are in ventral view. Colonies at the phases 9/7/1 and 9/8/2 are in early-cycle; colonies in 9/8/3–9/8/4 are in mid-cycle; colonies in 9/8/5–11/8/6 are in late cycle. Colour code refers to the different organs/anatomical parts. (B) Colony of *B. schlosseri* formed by seven adult zooids (az) grouped in a star-shaped system (dorsal view). 1 and 2Bs are hidden by adult individuals. Each adult zooid has its own oral siphon (os), whereas the cloacal siphons converge at the system center to open in a common atrial siphon (as). A transparent tunic (t) embeds the zooids. (C) Expression of *Myc*, *SoxB1*, *Pou2* and *Pou3* in the organs of 2Bs, 1Bs and adults based on ISH labelling data. Dark grey = strong labelling; light grey = weak labelling; white = absence of labelling. (D) Illustrations synthesizing *Myc*, *SoxB1*, *Pou2* and *Pou3* ISH labelling in 2 and 1Bs. Colour code refers to the different gene sets expressed in the different organs/anatomical parts, according to the colour of gene names. See (A) for anatomical details. The dorsal organ (in stage 7) and its derivative, i.e., the neural gland (in stage 8) are bordered by a dotted line and are found in the dorsal midline in 1Bs. Numbers below each illustration indicate the stages. (E–G) Transverse histological section of an adult zooid with its right 1B. The triplet of numbers in the upper right corner indicates the colony phase. The black square area in (E) is enlarged in (F) to show a cell island. Note some cSCs, i.e., small round cells, about 5 µm in diameter, with high nucleus/cytoplasm ratio and without cytoplasmic specializations (arrowheads). They are close to large phagocytes. The brown square area in (E) is enlarged in (G) to show the endostyle niche, connective tissue rich in cSCs (arrowheads), ventral to the endostyle (en). bc: branchial chamber; pbe: peribranchial chamber; sg: stigmata; t: tunic. D–V: dorso-ventral axes; L–R: left-right axes. Hematoxylin-Eosin. (H) A cSC of the endostyle niche. m: mitochondria; n: nucleus; nu: nucleolus; rer: rough endoplasmic reticulum cistern. Transmission Electron Microscopy.

*Klf4* is a member of the Kruppel like factors (*Klf*) family, characterized by the presence of three C-terminal highly conserved zinc finger motifs and a high diversification of the associated transactivation/repression domains (Vangapandu and Ai, 2009; Pressnell et al., 2015). Depending on their structure and/or function, the *Klf* members are partitioned into different groups whose number varies across studies (McConnell and Yang, 2010; Pei and Grishin, 2015; Pressnell et al., 2015; Pollak et al., 2018). Pei and Grishin (2015) found a minimum of 18 *Klf* genes in vertebrates and classified them into seven groups (A–G), with invertebrate deuterostomes possessing a single member of each group and the loss of *KlfE* in the ascidian *Ciona intestinalis*. Together with *Klf1/2/17*, the human *Klf4* belongs to the group *KlfA*.

In this study we investigate the presence and the evolution of the *YF* genes in ascidians, and the loss of *KlfE* in the ascidian *Ciona intestinalis*. Together with *Klf1/2/17*, the human *Klf4* belongs to the group *KlfA*.

### METHODS

**Identification and Evolution of Yamanaka Factors in Chordates**

The *in silico* identification of ascidian *YF* genes and the evolutionary analyses of the *YF* gene families in chordates were performed on the species listed in Supplementary Table 1, as detailed in Supplementary Data Sheet 1.

**B. schlosseri Colonies Rearing**

*B. schlosseri* colonies were collected in the Venice Lagoon and reared at the Department of Biology (University of Padova) in standard laboratory condition (Sabbadin, 1955; Sabbadin, 1960). Colonies were observed daily and fixed at 6 developmental phases as defined by the staging method developed by Sabbadin (1955) (Figure 1A).

**RNA Extraction, Yamanaka Factors cDNA Sequencing, and *in Situ* Hybridization**

In order to validate the *B. schlosseri* predictions of *SoxB1*, *Myc*, *Pou2* and *Pou3*, and to develop probes for ISH experiments, total RNA was extracted from *B. schlosseri* single colonies according to the protocol of Campagna et al. (2016), followed by cDNA synthesis, cloning and sequencing. Whole-mount ISHs were carried out as described in Franchi and Ballarin (2014) using at least three samples per stage. Detailed protocols are in Supplementary Data Sheet 1.

**Histology and Transmission Electron Microscopy**

Samples have been treated according to the classical protocols for histology (Hematoxylin–Eosin) and Transmission Electron Microscopy (TEM); details on the protocols are reported in Supplementary Data Sheet 1.

**RESULTS**

**Annotation of the *YF* Gene Families in Ascidians**

A total of 47 coding sequences (CDS) of the *YF* families were re-annotated or annotated *ex novo* in the reference genomes of the 11 analysed ascidians (see genes with "_p" suffix in Supplementary Tables 2–5, Fasta sequences in Supplementary Data Sheet 2–5, and details in Supplementary
FIGURE 2 | Myc evolution and expression pattern in B. schlosseri. (A) ML evolutionary tree of the MYC family (978 amino acid sites; best-fit substitution model JTT+G+I+F; bootstrap on 100 replicates; Myc of the echinoderm Strongylocentrotus purpuratus used as outgroup). Only bootstrap values ≥70% are shown near each node. Red: B. schlosseri sequence used in ISH; yellow: ascidians; light blue: vertebrates; grey: other deuterostomes. Sequence data are in Supplementary Table 2. Sequences with the "_p" suffix were annotated ex novo or re-annotated in this study. Species classification is reported in Supplementary Table 1, and species abbreviations are: Ac: Anolis carolinensis; Am: Alligator mississippiensis; Bb: Branchiostoma belcheri; Bf: Branchiostoma floridae; Bl: Botrylloides leachi; (Continued)

(B) Continued

(C) Continued

(D) Continued

(E) Continued

(F) Continued

(G) Continued

(H) Continued

(I) Continued

(J) Continued
analyzed vertebrates and invertebrates genes. Deviation: 118 ± 31 amino acids), which are absent in all other ascidians the second exon encodes for both the bHLHZ domain and bootstrap values in Supplementary Table 1), although not all nodes have statistically significant bootstrap values (Figure 3C). Both SoxB1 and SoxB2 have conserved gene structure in ascidians (with 2 and 4 coding exons, respectively), and were re-annotated in B. schlosseri and in several other ascidians (Supplementary Table 3 and Supplementary Data Sheet 1).

SoxB1 expression was absent in early stages of blastogenesis until the late 2B stages (stages 4–5–6). Here, SoxB1 transcript was found in the developing branchial epithelium, oesophagus, heart and neural complex (Figures 1C,D, 3). In 1Bs, the gene was additionally expressed in stigmata and endostyle. In adult zooids in early-, mid-, and late-cycle, the expression was restricted to the stigmata and to a few cSCs (round cells, with a diameter of 5.3 ± 1.06 µm and a high nucleus-cytoplasm ratio; Figure 3A: Supplementary Data Sheet 6) aggregated in the region ventral to the endostyle and in cell islands. Among hemocytes, these cells were the only positive ones during the blastogenetic cycle (Figure 3L). However, cSCs expressing SoxB1 were less numerous than those expressing the other YFs, probably in relationship with their functional state (Figure 3B, Supplementary Data Sheet 6).

Pou2 Expression Is Female Gonad-Restricted Whereas Pou3 has a Wider Expression Pattern

Due to its complexity, the evolutionary history of the POU family was reconstructed including the POU2, POU3, POU4 and POU5 classes,
FIGURE 3 | Morphometrics analyses of labelled candidate SCs in niches, SoxB evolution and SoxB1 expression pattern in B. schlosseri. (A) Histogram showing the diameter of cSCs found labelled in niches. Statistical tests show no significant difference among the diameter of cells expressing Myc, SoxB1 and Pou3. (B) Histogram showing the percentage of cSCs found labelled in niches. Candidate SCs found positive for the expression of SoxB1 are significantly less numerous than cSCs expressing Myc and Pou3 (** *= p-value < 0.0001, ANOVA test), but the high percentage of cells expressing these last two genes suggest a coexpression in at least a part of cSCs. (C) ML evolutionary tree of the SoxB subgroup (524 amino acid sites; best-fit substitution model JTT+G+I+F; bootstrap on 100 replicates; SoxB2 cluster (Continued))
with vertebretate representatives of the POU6 class as outgroups (Figure 4 and Supplementary Image 1). In the reconstructed tree all basal nodes remain unresolved and a polytomous POU3 “busb” can be observed (Figure 4A). Increasing the number of replicates does not improve the support of these nodes, indeed bootstrap tests did not converge even after 1,000 trees. However, five highly supported monophyletic clades are identified (see nodes with dots in Figure 4A), corresponding to: 1) the POU2 class; 2) the POU4 class; 3) the vertebrate-specific POU5 class; 4) all analysed vertebrete Pou3f1 genes; and 5) four fast-evolving ascidian sequences (POU-X clade in Figure 4A). Thus our tree confirms the absence of POU5 in ascidians and allows the unambiguous identification of the ascidian orthologs of POU2 and POU4 classes, remarkably present in all 11 analysed ascidian genomes (Figure 4A). Moreover, the topology of both the POU2 and POU4 ascidian clades corresponds to the expected species tree based on the ascidian classification, and a nearly conserved intron/exon structure is observed within each of these clades (see column “Ascidian Classification” and “N° of coding exons” in Supplementary Table 4).

Since a published phylogenetic reconstruction performed on all metazoa and all POU classes hypothesized a Pou3-like gene as the last common ancestor of the POUS5/3/2 polytomous clade (Gold et al., 2014), we focused our expression pattern analyses on Pou2 and Pou3 (see Discussion).

Two Pou2 genes, named Pou2a and Pou2b, have been identified in the B. schlosseri genome (Bsp12092_Pou2a and Bsp13218_Pou2b in Figure 4A). They differ only for the first and last exon while the remaining intron/exon structure is identical (10 exons, with 16 over a total of 18 intron/exon boundaries falling in exactly the same protein position). Thus, the encoded proteins are nearly identical (97% amino acid identity over 255 alignable sites) but have short and completely different N- and C-terminal ends, and different total lengths (359 for Pou2a vs. 293 amino acids for Pou2b). Since the C-terminal region of Pou2a is more similar than Pou2b to the homologous regions of the closest species B. leachii, only Pou2a was experimentally validated (Supplementary Data Sheet 1) and used in ISH experiments. Indeed, we hypothesize that Pou2b is the result of errors in genome assembly as the entire mapping regions and even the intron sequences of Pou2a and Pou2b are almost identical (94% nucleotide identity with 2% gap over 4.6 kb). However, a very recent gene duplication cannot be ruled out.

The POU-X clade only includes representatives of the families Styelidae and Pyuridae and was putatively assigned to the POU3 class based on two considerations: the single-exon gene structure, typical of all other Pou3 of deuterostomes here analysed (Supplementary Table 4), and the presence of elements known to impair the tree resolution (i.e., the very long branches of POU-X sequences associated with the intrinsic polytomy of the POU3 class). Therefore, the Bs_g52290 sequence of POU-X clade (Figure 4A) was considered a valid B. schlosseri Pou3 and used in subsequent ISH experiments.

Pou2 was found expressed in the female germline and, more specifically, in oocytes in previtellogenesis in 1Bs, 2Bs, adult zooids, and circulating in the haemocoe (Figures 1C,D, 4A–D). Moreover, its transcript was retrieved in cells of early developing 2B. Pou3 was found expressed in early-cycle of blastogenic development (Figures 1C,D, 4E–I). Subsequently its expression was restricted to the germline and the rudiments of the neural complex, heart, and digestive system. In the following phases of bud development, the expression persisted at high levels only in the female and male germlines, while lower levels were found in the nervous system. In adult zooids, the presence of Pou3 transcript persisted in germline and in cSCs (round cells, with a diameter of 5.2 ± 1.08 µm and a high nucleus-cytoplasm ratio; Figure 3A, Supplementary Data Sheet 6) found near the endostyle and in cell islands in all the phases. Moreover, it could be retrieved in the endostyle, with an antero-posterior pattern of respectively high and low expression levels. As reported for Myc and SoxB1, cSCs were the only hemocytes expressing Pou3 during the blastogenetic cycle (Figure 4H).

The Klf Gene Family
The KLF proteins are characterized by a high N-terminal sequence variability that makes their alignment unreliable except in the three zinc finger motifs. The reconstructed evolutionary tree, based on the 93 reliably alignable sites of the zinc finger domain and including the KlfA/B/C/D/F/G groups, is almost completely unresolved (Supplementary Image 2) and shows a mildly supported KlfA clade (75% bootstrap) inside a wide polytomic KlfA/B/C/D clade. Only a partial KlfA sequence, corresponding only to the Zn finger domain, was found and annotated by us in an earlier assembly version of the B. schlosseri genome (see “Notes” in Supplementary Table 5), suggesting that the KlfA absence in the
FIGURE 4 | Pou evolution and expression pattern of Pou2 and Pou3 in B. schlosseri. (A) ML evolutionary tree of the POU family based on the C-terminal region (199 amino acid sites; best-fit substitution model JTT+I+G; bootstrap on 100 replicates; Pou6 used as outgroup). The cladogram of the full ML tree, containing all bootstrap values and all Pou6 outgroups, is shown in Supplementary Image 1. Black dots: bootstrap values ≥70%; red: B. schlosseri sequences used in ISH; yellow: ascidians; light blue: vertebrates; grey: other deuterostomes. Species abbreviations are as in Figure 2. Species classification and sequence data are in Supplementary Table 1 and in Supplementary Table 4, respectively. Sequences with the "_p" suffix were annotated ex novo or re-annotated in this study; the sequence with "_h" suffix (Continued)
current genome version was due to assembly problems. The very high conservation among groups of the available Klf/A region (i.e., the Zn finger domain) prevented us from designing target-specific ISH probes able to discriminate between the different Klf groups, therefore no gene expression analyses were performed on this gene.

**Yamanaka Factors Are Individually Expressed in Several Organ Rudiments of the Early Bud and, Except Pou2, in Candidate Adult Stem Cells in Proposed Niches**

An overview of the spatio-temporal expression pattern of Myc, SoxB1, Pou2 and Pou3 during the blastogenic cycle is summarised in Figures 1C,D. The endostyle, neural complex, heart, gut, undifferentiated germline, and female germline all expressed three factors during the same developmental stages: Myc, SoxB1 and Pou3. The same factors were expressed during the entire cycle in some cSCs, showing comparable dimensions among samples (Figure 3A), in both the SC niches of adult individuals and in hemocel. Almost 70% of adult cSCs in niches expressed Myc and Pou3 (Figures 3A,B, Supplementary Data Sheet 6), suggesting that these two factors should be coexpressed in at least 20% of cSCs. However, SoxB1 is expressed in a lower percentage of cells (Figures 3A,B, Supplementary Data Sheet 6). Of note is that only the female gonad additionally expressed Pou2. Myc and Pou (both Pou2 and Pou3) genes were expressed in early 2Bs. Myc and SoxB1 were recognised during the development of the branchial and peribranchial epithelia, whereas SoxB1 and Pou3 were present during brain formation. Finally, the testis expressed only Myc, and only during development, whereas the adult stigmata maintain the expression of SoxB1.

**DISCUSSION**

All YF gene families have a complex evolutionary history characterized by several gene duplications, some of which also occurred in the ancestor of vertebrates. Our evolutionary analyses show that a one-to-many relationship exists between ascidian and human orthologs of some YF genes, which means that the ascidian gene has more than one ortholog in humans. Indeed, the ascidian SoxB1 is the ortholog not only of the human/vertebrate YF Sox2, but also of Sox1 and Sox3. Likewise, the ascidian Myc gene is the ortholog not only of the human/vertebrate YF c-Myc, but also of N-Myc and L-Myc. This implies that the functional role and the expression pattern of these ascidian genes could have been inherited by one or many human/vertebrate orthologs. With that in mind, we have analysed the expression pattern of the YF genes in *B. schlosseri*, a highly regenerative model species belonging to the taxon evolutionarily closest to vertebrates.

Our results show that Myc is expressed in all developing tissues of buds and in cSCs in niches (Figures 1C,D), suggesting a crucial role in undifferentiated territories. Noteworthy, Myc expression in cSCs in niches is a new finding, not observed in previous studies on Myc expression in two other colonial ascidians. Indeed, in *Botryllus primigenius*, Myc was proposed as a marker of undifferentiated, but not proliferative, somatic and germine SCs, as it was expressed by multipotent epithelial cells in buds, undifferentiated coelomic cells, and developing gonads (Kawamura et al., 2008). In *Polyandrocarpa misakiensis*, Myc expression was observed in the atrial epithelium in the organ-forming region of developing bud, and in fibroblast-like cells, which participate in the organogenesis together with the epithelial cells (Fujiiwara et al., 2011). Moreover, the noteworthy identification of an ascidian-specific C-terminal tail, located just downstream the bHLHZ domain, incentivises further investigation on the role and the possible additional functions performed by this protein region.

As summarized in Figure 1C, SoxB1 is widely expressed during *B. schlosseri* asexual development. It is found in the developing nervous system, as has also been reported for Sox2 and other SoxB1 group genes in many metazoans (Wood and Episkopou, 1999; Penzel et al., 1997; Rex et al., 1997; Meulemans and Bronner-Fraser, 2007; Ferrero et al., 2014; Ross et al., 2018). In the developing nervous system we also detected the expression of Pou3 (Figure 1C), known to be involved in neural system development in vertebrates and other invertebrates, (Candidi et al., 2002; Cole and Arnone, 2009; Wollesen et al., 2014; Garner et al., 2016; Wei et al., 2016; Prünster et al., 2019). The expression, in samples at the same developmental stages, of SoxB1 and Pou3 in the developing nervous system, observed here and in previous studies (Iwafuchi-Doi et al., 2011; Hui et al., 2015), suggests a co-regulation of these two factors during neurogenesis. We also retrieved SoxB1 expression in the developing oesophagus, stomach and proximal intestine, suggesting a role in the differentiation and maintenance of these epithelia. This result corresponds with the Sox2 expression in the adult stomach tissues of mammals (Arnold et al., 2011). Moreover, SoxB1 was found expressed in the endostyle of 1B, and in cSCs of the endostyle niche and cell islands, indicating...
potential involvement in stemness. Supporting this hypothesis, Niwa et al. (2016) demonstrated that C. robusta and other invertebrate SoxB1 proteins are able to replace the pluripotency-associated function of mouse Sox2. Thus the role of SoxB1 in the control of pluripotency in B. schlosseri is plausible.

Our evolutionary analyses demonstrate the absence of the POU5 class in ascidians, indicating that POU5 arose after the tunicate-vertebrate divergence. Since there is no true ortholog of Pou5f1 to be investigated in B. schlosseri, the ancestor of the POU5 class could instead be investigated in its place. Indeed, it remains unclear whether the role of vertebrate Pou5f1 in SC pluripotency was inherited from the ancestral paralog (see discussion in Gold et al., 2014). Our POU evolutionary tree shows a basal polytomy that prevents identifying the POU5 ancestor (Figure 4A). However, our tree is in agreement with a previous wide evolutionary analysis of the entire POU family (including POU1) in all Metazoa (Gold et al., 2014). In that study, a polytomous clade including POU5, POU3 and POU2 classes was identified, and the POU4 class was found as the paralog that diverged after the POU+POU1 basal clade. Relying on the ancestral state reconstruction and on the taxonomic distribution of POU2 and POU5, Gold et al. (2014) hypothesized that a Pou3-like gene was the last common ancestor of the POU5/3/2 polytomy and vertebrate-specific and vertebrate-specific duplications, respectively. On this basis, and given the goal of our study, we investigated the expression pattern in B. schlosseri of both Pou3, as representative of the Pou3-like ancestor of Pou5f1, and Pou2, as the closest paralog of Pou5f1.

Remarkably, while POU2 and POU4 are present in all analysed ascidians, Pou3 was found only in ascidians of two Stolidobranchia families (i.e., Styeliidae and Pyuridae; Supplementary Table 1 and Figure 4A) and seems to be absent even in the two widely studied Ciona species with highly curated genomic sequence annotations. This leads us to hypothesize a Pou3 loss in most ascidian groups, although the long branch/fast evolution of the ascidian Pou3 clade (Figure 4A) could also explain the failure to identify these genes. As summarized in Figure 1D, we found that B. schlosseri Pou3 is expressed in the endostyle, even in the adult phase, especially in the anterior-most part, close to the previously identified SC niche (Rinkevich et al., 2013). Moreover, Pou3 is expressed in the few pluripotent cells of the early 2B and cSCs in their niches. This expression pattern highlights a likely role of B. schlosseri Pou3 in undifferentiated cells, as also recently evidenced in Botrylloides diegensis/leachii (Viard et al., 2019; Kassmer et al., 2020). Overall, our and previous results support a role in SC pluripotency for the Pou3-like ancestor that gave rise to the vertebrate Pou5f1, here represented by the Pou3 of ascidians. Finally, Pou3 was found expressed in the differentiating heart, gut and gonads, and in the developing nervous system (see above).

As for Pou2, it is expressed exclusively in the immature previtellogenic oocytes, both in gonads and circulating in the colonial vascular system. Remarkably, this expression pattern significantly differs from those of vertebrates and invertebrates, where Pou2 is expressed in many more territories. For example, Pou2 is involved in neurogenesis in cephalopods (together with Pou3) (Wollesen et al., 2014), in the establishment of the dorsal/ventral axis in echinoderms (Range and Lepage, 2011) and in immunity in vertebrates (Strubin et al., 1995; Karnowski et al., 2012). Notably, this is the first report of Pou2 expression in tunicates.

Overall, these results point to an expression pattern of the vertebrate Pou5f1 more similar to that of B. schlosseri Pou3 rather than Pou2, thus supporting the Gold’s and coll. (2014) hypothesis of a Pou3-like gene as ancestor of Pou5f1.

To summarize, we identified in the model species B. schlosseri the orthologs of vertebrate YFs through deep evolutionary analyses. Then, we found that Myc, SoxB1 and Pou3 are individually expressed in several differentiating territories, consistently with the reported function of these genes in chordate embryogenesis. Moreover, we found their expression in cells with features of cSCs exhibiting the same morphological and morphometric features. These cells are located close to the endostyle and in cell islands, and according to our analyses, should coexpress at least Pou3 and Myc. Further studies, such as a deep analysis of niche dynamic and double in situ hybridization experiments, are needed to understand the actual co-expression of these factors in cSCs, which could in turn suggest a shared ancestral pluripotency program in chordates.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository and accession numbers can be found below: GenBank, OLS828248-OL828252.

AUTHOR CONTRIBUTIONS

Conception and design: LM, CG; Mariculture, observation and sample collection: FG, MM, VV; Gene annotation and evolutionary analyses: CG, MS, CA; Cloning and sequencing: MS, MM, FG; ISH: MM, VV, MS, FG, LM; Manuscript writing: CG, MS, LM, VV. All authors participated in the data discussion and manuscript finalization.

FUNDING

This study was supported by MIUR, Italy (PRIN-2015, grant number: 2015NSFHXF) to LM and CG; by the Molecular Biodiversity Laboratory of the Italian node of Lifewatch (CNR, Consiglio Nazionale delle Ricerche) to CG; by the University of Padova, grant BIRD213252 to LM.

ACKNOWLEDGMENTS

The authors thank Ayelet Voskoboynik and Mark Kowarsky for providing contigs of an earlier assembly of the B. schlosseri genome and for discussion on gene annotation; Timothy Frawley for critical review and edits; and a reviewer for valuable suggestions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2022.782722/full#supplementary-material
REFERENCES

Arnold, K., Sarkar, A., Yram, M. A., Polo, J. M., Bronson, R., Sengupta, S., et al. (2011). Sox2+ Adult Stem and Progenitor Cells Are Important for Tissue Regeneration and Survival of Mice. Cell Stem Cell 9, 317–329. doi:10.1016/j.stem.2011.09.001

Atchley, W. R., and Fitch, W. M. (1995). Myc and Max: Molecular Evolution of a Proto-OrcoGene Products and Their Dimerization Partner. Proc. Natl. Acad. Sci. 92, 10217–10221. doi:10.1073/pnas.92.22.10217

Avilion, A. A., Nicolas, S. K., Peny, L. H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent Cell Lineages in Early Mouse Development Depend on Sox2 Function. Genes Dev. 17, 126–140. doi:10.1101/gad.224503

Ballarin, L., and Cima, F. (2005). Cytochemical Properties of Botryllus Schlosseri Haemoocytes: Indications for Morpho-Functional Characterisation. Eur. J. Haematol. 49, 255–264.

Bowles, J., Schepers, G., and Koopman, P. (2000). Phylogeny of the SOX Family of Developmental Transcription Factors Based on Sequence and Structural Indicators. Dev. Biol. 227, 239–255. doi:10.1006/dbio.2000.9883

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., et al. (2005). Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. Cell 122, 947–956. doi:10.1016/j.cell.2005.08.020

Brown, F. D., Tiozzo, S., Roux, M. M., Ishizuka, K., Swalla, B. J., and De Tomaso, A. W. (2009). Early Lineage Specification of Long-Lived Germline Precursors in the Colonial Ascidian Botryllus Schlosseri. Development 136, 3485–3494. doi:10.1242/dev.037754

Brozovic, M., Dantec, C., Dardaillon, J., Dauga, D., Faure, E., Gineste, M., et al. (2018). ANISEED 2017: Extending the Integrated Ascidian Database to the Exploration and Evolutionary Comparison of Genome-Scale Datasets. Nucleic Acids Res. 46, D718–D725. doi:10.1093/nar/gkx1108

Buescher, M., Hing, F. S., and Chia, W. (2002). Formation of Neuroblasts in the Embryonic central Nervous System of Drosophila melanogaster Is Controlled by SoxNeuro. Development 129, 4193–4203. doi:10.1242/dev.129.18.4193

Bylund, M., Anderson, E., Novitch, B. G., and Muhr, J. (2003). Vertebrate Stem-cell-specific Activation of an Ancestral Myc Protooncogene with Conserved Basic Functions in the Early Metazoan Hydra. Proc. Natl. Acad. Sci. USA 107, 4051–4056. doi:10.1073/pnas.0911060107

Heenan, P., Zondag, L., and Wilson, M. J. (2016). Evolution of the Sox Gene Family within the Chordate Phylog. Phym 255, 385–392. doi:10.1016/j.gene.2015.09.013

Hui, S. P., Nog, T. C., and Gauthier, S. (2015). Characterization of Proliferating Neural Progenitors after Spinal Cord Injury in Adult Zebrafish. PLoS ONE 10, e0143595. doi:10.1371/journal.pone.0143595

Iwafuchi-Doi, M., Yoshida, Y., Onichtchouk, D., Leichsenring, M., Driever, W., Takemoto, T., et al. (2011). The Pouf1/Pouf3-dependent but Sox8-independent Regulation of Conserved Enhancer N2 Initiates Sox2 Expression during Epiblast to Neural Plate Stages in Vertebrates. Dev. Biol. 352, 354–366. doi:10.1016/j.ydbio.2010.12.027

Kassmer, S. H., Langenbacher, A. D., and De Tomaso, A. W. (2020). Integrin-alpha+ Candidate Stem Cells Are Responsible for Whole Body Regeneration in the Invertebrate Chordate Botryllusidiies Diegenis. Nat. Commun. 11, 4435. doi:10.1038/s41467-020-18288-w

Kassmer, S. H., Langenbacher, A., and De Tomaso, A. W. (2019). Primordial Blasts, a Population of Blood Borne Stem Cells Responsible for Whole Body Regeneration in a Basal Chordate. BioRxiv. doi:10.1101/675785

Kawamura, K., Tachibana, M., and Sunanaga, T. (2008). Cell Proliferation Dynamics of Somatic and Germline Tissues during Zooidal Life Span in the Colonial Tunicate Botryllus Primigenius. Dev. Dyn. 237, 1812–1825. doi:10.1002/dvdy.21992

Kowarsky, M., Anselmi, C., Hotta, K., Burighel, P., Zaniolo, G., Caicci, F., et al. (2019). Myc Stimulates Nuclearly Encoded Mitochondrial Genes and Mitochondrial Maintenance of Long-Lived Germline Precursors in the Colonial Chordate Botryllusidiies Diegenis. Nat. Commun. 11, 4435. doi:10.1038/s41467-020-18288-w

Kowarsky, M., Anselmi, C., Hotta, K., Burighel, P., Zaniolo, G., Caicci, F., et al. (2021). Sexual and Asexual Development: Two Distinct Programs Producing the Same Tunicate. Cell Rep 34, 108681. doi:10.1016/j.celrep.2020.108681

Kuwarsky, M., Anselmi, C., Hotta, K., Burighel, P., Zaniolo, G., Caicci, F., et al. (2021). Sexual and Asexual Development: Two Distinct Programs Producing the Same Tunicate. Cell Rep 34, 108681. doi:10.1016/j.celrep.2020.108681

Li, F., Wang, Y., Zeller, K. I., Potter, J. I., Wonsy, D. R., O’Donnell, K. A., et al. (2005). Myc Stimulates Nuclearly Encoded Mitochondrial Genes and Mitochondrial Biogenesis. Mol. Cell. Biol. 25, 6225–6234. doi:10.1128/MCB.25.14.6225-6234.2005

Manni, I., Gasparini, F., Hotta, K., Ishizuka, K. J., Ricci, L., Tiozzo, S., et al. (2014). Ontology for the Asexual Development and Anatomy of the Colonial Chordate Botryllusidiies Diegenis. Mol. Biol. Evol. 297. doi:10.1080/08945630.2014.903360

Manni, I., Anselmi, C., Cima, F., Gasparini, F., Vokoboyan, A., Martini, M., et al. (2019). Sixty Years of Experimental Studies on the Blastogenesis of the Colonial Tunicate Botryllusidiies Diegenis. Dev. Biol. 448, 293–308. doi:10.1016/j.ydbio.2018.09.009

Manni, I., Zaniolo, G., Cima, F., Burighel, P., and Ballarin, L. (2007). Botryllus Schlosseri: A Model Ascidian for the Study of Asexual Reproduction. Dev. Dyn. 236, 335–352. doi:10.1002/dvdy.21037

Marlow, H. Q., Sriaravata, M., Matus, D. Q., Rokhsar, D., and Martindale, M. Q. (2009). Anatomy and Development of the Nervous System of Nemastostella vectensis, an Anchostran Cnidarian. Dev. Neurobiol. 69, 235–254. doi:10.1002/dnreb.20698

McConnell, R. B., and Yang, V. W. (2010). Mammalian Kupffer-like Factors in Health and Disease. Physiol. Rev. 90, 1337–1381. doi:10.1152/physrev.00058.2009

McFerrin, L. G., and Atchley, W. R. (2011). Evolution of the Max and Mlx Networks in Animals. Genome Biol. Evol. 3, 915–937. doi:10.1093/gbe/evr082

Meulemans, D., and Bronner-Fraser, M. (2007). Insights from Amphioxus into the Evolution of Vertebrate Cartilage. PLoS ONE 2, e877. doi:10.1371/journal.pone.0000787

Mol. Biol. Evol. 34, 108681. doi:10.1016/j.celrep.2020.108681

Myo/Max/Mad in Invertebrates: the Evolution of the Max Network. Dev. Growth Differ. 53, 1004–1014. doi:10.1111/j.1440-169X.2011.01312.x

Polyandrocarpa Misakensis. Dev. Growth Differ. 53, 1004–1014. doi:10.1111/j.1440-169X.2011.01312.x

Yamana et al. Factors in a Colonial Tunicate

Yamana, Factors in a Colonial Tunicate
Mizuseki, K., Kishi, M., Matsui, M., Nakanishi, S., and Sasai, Y. (1998). Xenopus Zic-Related-1 and Sox-2. Two Factors Induced by Chordin, Have Distinct Activities in the Initiation of Neural Induction. Development 125, 579–587. doi:10.1242/dev.125.4.579

Niwa, H., Miyazaki, J., and Smith, A. G. (2000). Quantitative Expression of Oct-3/4 Defines Differentiation, Dedifferentiation or Self-Renewal of ES Cells. Nat. Genet. 24, 372–376. doi:10.1038/74199

Niwa, H., Nakamura, A., Urata, M., Shirae-Kurabayashi, M., Kuraku, S., Russell, S., et al. (2016). The Evolutionally-Conserved Function of Group B1 Sox Family Members Confers the Unique Role of Sox2 in Mouse ES Cells. BMC Evol. Biol. 16, 173. doi:10.1186/s12862-016-0755-4

Omal, P., Grün, D., Adamidi, C., Rybak, A., Solana, J., Mastrobuoni, G., et al. (2012). Gene Expression of Pluripotency Determinants Is Conserved between Mammalian and Planarian Stem Cells. EMBO J. 31, 2755–2769. doi:10.1038/embob.2012.110

Onichtchouk, D. (2016). Evolution and Functions of Oct4 Homologs in Non-mammalian Vertebrae. Biochim. Biophys. Acta 1859, 770–779. doi:10.1016/j.bbamcr.2016.03.013

Pattengale, N. D., Alipour, M., Bininda-Emonds, O. R. P., Moret, B. M. E., and Pattengale, R., Oschwald, R., Chen, Y., Tacke, L., and Grunz, H. (1997). Rosental, B., Kowarsky, M., Seita, J., Corey, D. M., Ishizuka, J. K., Palmeri, K. J., et al. (2007). Striving for Normality: Whole Body Regeneration through a Series of Abnormal Generations. FASEB J. 21, 1335–1344. doi:10.1096/fj.06-7337com

Penzel, R., Osswald, R., Chen, Y., Tacke, L., and Grunz, H. (1997). Characterization and Early Embryonic Expression of a Neural Specific Transcription Factor sX03X in Xenopus laevis. Int. J. Dev. Biol. 41, 667–677.

Pollak, N. M., Hoffman, M., Goldberg, I. J., and Drosatos, K. (2018). Krüppel-like Factors: Crippling and Un-crippling Metabolic Pathways. JACC Basic Transl. Sci. 3, 132–156. doi:10.1016/j.jactbs.2017.09.001

Prensell, J. S., Schnitzler, C. E., and Browne, W. E. (2015). KLF/SP Transcription Factor Family Evolution: Expansion, Diversification, and Innovation in Eukaryotes. Genome Biol. Evol. 7, 2289–2309. doi:10.1093/gbe/evv141

Prünster, M. M., Ricci, L., Brown, F. D., and Tiozzo, S. (2019). De Novo neurogenesis in a Budding Chordate: Co-option of Larval Anteroposterior Patterning Genes in a Transitory Neurogenic Organ. Dev. Biol. 438, 342–352. doi:10.1016/j.ydbio.2018.10.009

Range, R., and Lepage, T. (2011). Maternal Oct1/2 Is Required for Nodal and Vg1/Univin Expression during Dorsal-Ventral axis Specification in the Sea Urchin Embryo. Dev. Biol. 357, 440–449. doi:10.1016/j.ydbio.2011.07.005

Rex, M., Orme, A., Uwanoaho, D., Tointon, K., Wigmore, P. M., Sharpe, P. T., et al. (1997). Dynamic Expression of Chicken Sox2 and Sox3 Genes in Ectoderm Induced to Form Neural Tissue. Dev. Dyn. 209, 323–332. doi:10.1002/(SICI)1097-0177(199707)203:3<323::AID-AJA7>3.0.CO;2-K

Rinkevich, B., Shlemborg, Z., and Fishelson, L. (1995). Whole-body Protocadherase Retegiation from Totipotent Blood Cells. Proc. Natl. Acad. Sci. USA 92, 7695–7699. doi:10.1073/pnas.92.17.7695

Rinkevich, Y., Voskoboyinik, A., Rosner, A., Rabinowicz, C., Paz, G., Oren, M., et al. (2013). Repeated, Long-Term Cycling of Putative Stem Cells between Niches in a Basal Chordate. Dev. Cell 24, 76–88. doi:10.1016/j.devcel.2012.11.010

Rosenfeld, M. G. (1991). POU-domain Transcription Factors: Pou-er Full Developmental Regulators. Genes Dev. 5, 897–907. doi:10.1101/gad.5.6.897

Rosenthal, B., Kowarsky, M., Seita, J., Corey, D. M., Ishizuka, J. K., Palmeri, K. J., et al. (2018). Complex Mammalian-Like Haematopoietic System Found in a Colonial Chordate. Nature 564, 425–429. doi:10.1038/s41586-018-0783-x

Ross, K. G., Molinario, A. M., Romero, C., Dockter, B., Cable, K. L., Gonzaletz, K., et al. (2018). Soxh1 Activity Regulates Sensory Neuron Regeneration, Maintenance, and Function in Planarians. Dev. Cel. 47, 331–347. doi:10.1016/j.devcel.2018.10.014

Sabbadin, A. (1960). Nuove Ricerche Sull’Inversione Sperimentale Situs Viscerum Botryllus Schlosseri Sabbadin. Estratto dall’Archivio di Oceanografia e Limnologia 11.