THE INTERNATIONAL JOURNAL OF DEVELOPMENTAL BIOLOGY
www.intjdevbiol.com

doi: 10.1387/ijdb.170205is

SUPPLEMENTARY MATERIAL

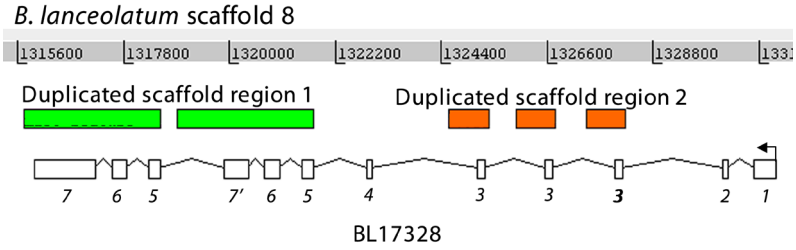
corresponding to:

Amphioxus *Sp5* is a member of a conserved Specificity Protein complement and is modulated by Wnt/β-catenin signalling

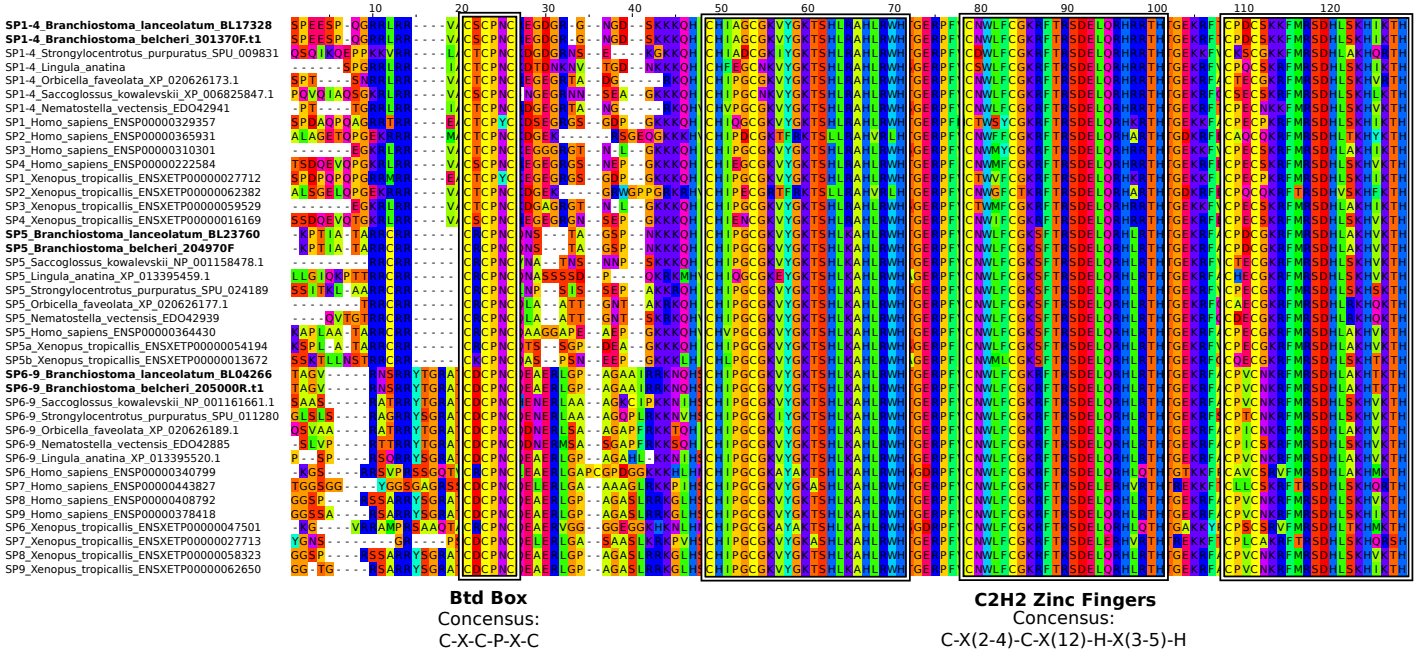
SIMON C. DAILEY, IRYNA KOZMIKOVA and ILDIKÓ M.L. SOMORJAI*

***Address correspondence to:** Ildikó M.L. Somorjai. University of St Andrews, Biomedical Sciences Research Complex, North Haugh, St Andrews, KY16 9ST, United Kingdom. Tfno: 01334 463628. E-mail: imls@st-andrews.ac.uk - web: https://synergy.st-andrews.ac.uk/cord http://orcid.org/0000-0001-5243-6664

Full text for this paper is available at: http://dx.doi.org/10.1387/ijdb.170205is



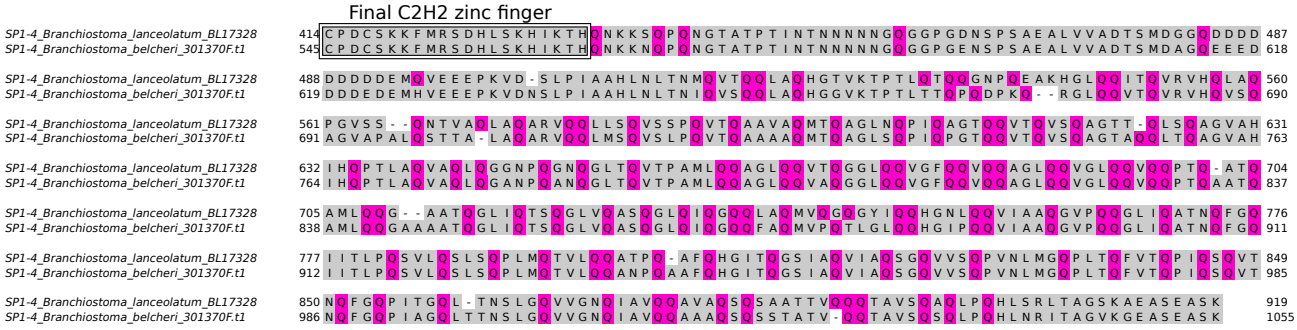
Suppl. Fig. S1. Assembly errors affecting *B. lanceolatum* scaffold 8. *The SP-like protein BL17328 is found on* B. lanceolatum *genomic scaffold 8. Whilst the predicted protein appears to be unaffected, there are two apparent assembly artefacts affecting this locus. Green and orange boxes mark two sequences which were repeated 2 and 3 times, respectively. In both cases, repeats were immediately followed by a stretch of unknown nucleotides (Ns), and were 99-100% identical across the entire sequence, including introns.* Sp5 *exons are numbered, and exon 7' represents a partial copy of this exon, which is interrupted by a patch of Ns.*



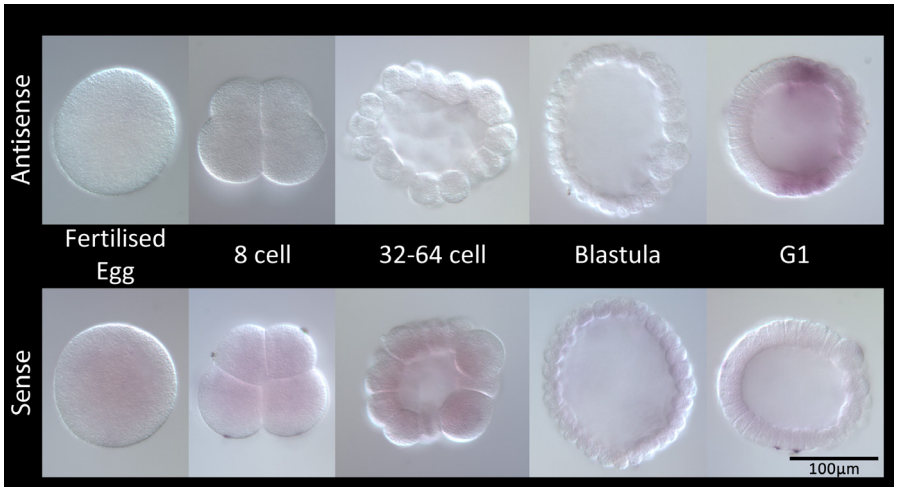
Suppl Fig. S2. Btd box and C2H2 triplet conservation. *For all six* Branchiostoma *SP-like proteins, the conserved Btd box motif and C2H2 zinc finger motifs are present. Consensus sequences taken from Presnell* et al. *(2015).* Branchiostoma *sequences are highlighted in bold.*

Suppl Fig. S3 (opposite). C2H2 domain phylogeny of SP and KLF proteins. *The C2H2 triplets for the SP proteins used in this study were compared to those of KLF proteins from the same species, where previously identified KLF proteins were readily available. The same domain from the human Wilms tumor1 (Wt1) protein was used to root the tree. Percentage RAxML bootstraps (above node) and MrBayes posterior probabilities (below node) are indicated only for key nodes of the SP proteins only. Both models strongly support an SP-Specific clade, confirming the identification of this gene family. The MrBayes analysis supported the existence of SP1-4, SP5 and SP6-9 clades; RAxML had weak support values for the internal topology of SP proteins. Trees used the LF model of protein evolution, with 4 gamma rate categories. Tree branches are coloured according to their RAxML bootstrap support (green low, blue high).*

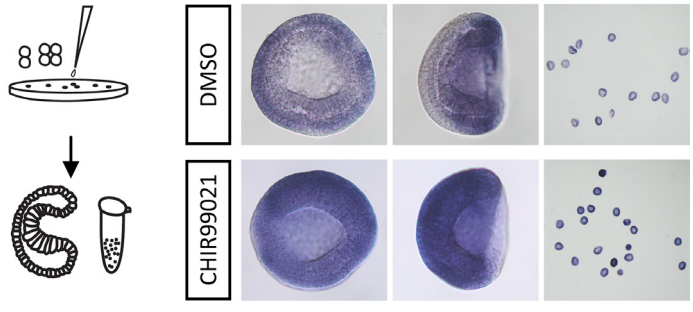

KLF9/13b Nematostella vectensis
KLF10 Homo sapiens
KLF10 Xenopus tropicalis
KLF9/13b Strongylocentrotus purpuratus
KLF11 Homo sapiens
KLF11 Xenopus tropicalis
KLF9 Xenopus tropicalis
KLF9 Homo sapeins
KLF13 Xenopus tropicalis
KLF13 Homo sapiens
KLF9/13a Strongylocentrotus purpuratus
SP1-4 Orbicella faveolata
SP1-4 Strongylocentrotus purpuratus
SP4 Homo sapiens
SP4 Xenopus tropicallis
56
99
SP1-4 Saccoglossus kowalevskii
SP2 Homo sapiens
SP2 Xenopus tropicallis
SP1-4 Nematostella vectensis
SP3 Homo sapiens
SP1 Xenopus tropicallis
SP1 Homo sapiens
44
97
SP3 Xenopus tropicallis
SP1-4 Lingula ana na
SP1-4 Branchiostoma belcheri
SP1-4 Branchiostoma lanceolatum
SP5 Branchiostoma belcheri
SP5 Branchiostoma lanceolatum
SP5 Homo sapiens
57
97
SP5 Saccoglossus kowalevskii
SP5a Xenopus tropicallis
SP5 Orbicella faveolata
SP5 Nematostella vectensis
SP5 Strongylocentrotus purpuratus
SP5 Lingula ana na
SP5b Xenopus tropicallis
SP6-9 Nematostella vectensis
SP6-9 Orbicella faveolata
SP6-9 Strongylocentrotus purpuratus
SP7 Xenopus tropicallis
SP7 Homo sapiens
SP8 Xenopus tropicallis
SP6-9 Saccoglossus kowalevskii
SP8 Homo sapiens
SP9 Homo sapiens
SP9 Xenopus tropicallis
SP6-9 Lingula ana na
88
100
SP6 Homo sapiens
SP6 Xenopus tropicallis
SP6-9 Branchiostoma belcheri
SP6-9 Branchiostoma lanceolatum
KLF9/13a Nematostella vectensis
KLF14 Homo sapiens
KLF16 Homo sapeins
KLF5b Nematostella vectensis
KLF6/7 Strongylocentrotus purpuratus
KLF6 Homo sapiens
KLF7 Homo sapiens
KLF7 Xenopus tropicalis
KLF6/7 Nematostella vectensis
KLF5a Nematostella vectensis
KLF3 Xenopus tropicalis
KLF3 Homo sapeins
KLF8 Homo sapiens
KLF8 Xenopus tropicalis
KLF12 Homo sapiens
KLF12 Xeonpus tropicalis
KLF1/2/4a Nematostella vectensis
KLF1/2/4b Nematostella vectensis
KLF5 Strongylocentrotus purpuratus
KLF5b Xenopus tropicalis
KLF5 Homo sapiens
KLF5a Xenopus tropicalis
KLF1 Homo sapiens
KLF4 Xenopus tropicalis
KLF17_Homo_sapiens
KLF4 Homo sapiens
KLF2 Homo sapiens
KLF2 Xenopus tropicalis
KLF1 Xenopus tropicalis
KLF1/2/4 Strongylocentrotus purpuratus
KLF15 Homo sapiens
KLF15a Xenopus tropicalis
KLF5b Xenopus tropicalis
KLF15 Strongylocentrotus purpuratus
KLF15 Nematostella vectensis
WT1 Homo sapiens
0.3
Bootstraps: 0 100
SP1-4
SP5
SP6-9



Suppl Fig. S4. Glutamine-rich trailing region of *Branchiostoma* SP1-4 proteins. *Alignment of the region following the final C2H2 domain of* B. lanceolatum *and* B. belcheri *SP1-4 proteins. Glutamine (Q) residues, which comprise approximately 25% of these sequences, are coloured magenta.*



Suppl Fig. S5. Expression of *Sp5* in early *B. lanceolatum* development. *The early developmental stages of* B. lanceolatum *were assayed by WMISH for the expression of* Sp5 using an antisense probe. *A sense probe, which would be expected to produce no signal, was included as a control. Neither probe produced a signal between fertilised egg to early blastula stages* **(A-D)**, *however by late blastula stage* **(E)**, *a hemispherical domain of expression is observed for the* Sp5 *antisense probe, whilst the sense control remains unstained. Sense control embryos were exposed to staining solution for an excessively long duration, in order to be confident that no specific staining was observed. This over-exposure results in the diffuse purple colouration and aggregation of dirt seen here in some embryos. All images are at the same scale.*

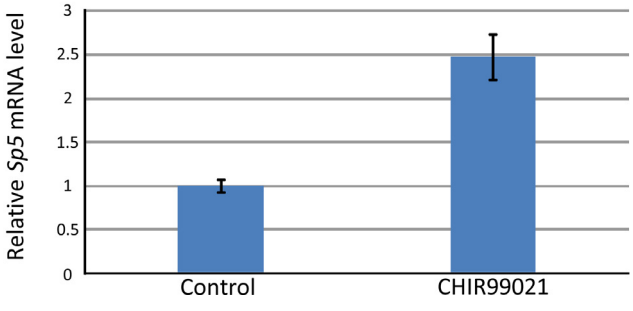


SUPPL. TABLE S1

SUPPORTING TRANSCRIPTS FOR EXPRESSION OF *B. LANCEOLATUM SP* GENES

Gene	Supporting transcripts
Sp1-4	JT854271.1 JT904194.1
Sp5	JT904194.1 JT873481.1 JT860870.1
Sp6-9	JT904753.1 JT848050.1

Accession codes for transcripts that originated from the combined developmental and adult *B. lanceolatum* transcriptome (Oulion *et al.*, 2012) are indicated.



Suppl Fig. S6. CHIR99021 treatment increases *Sp5* expression. *Expression of* Sp5 *is upregulated by CHIR99021 pharmacological treatment.* **(A)** *Scheme of the treatment schedule. CHIR99021 was applied at the 8-cell stage and embryos harvested for WMISH or qRT-PCR at the gastrula stage.* **(B)** *WMISH of control (DMSO-treated) and CHIR99021-treated (10 µM) embryos. Right-most panels show low resolution image of many embryos. The fragment used to synthesise the* Sp5 *probe was generated via PCR using primers 5' TGACCGCATCTTGCTGAGAGT 3' (forward) and 5' CTAGTCTTCTTGTACGTCGAT 3' (reverse), ligated into pCR-BluntII-TOPO.* **(C)** *Quantitation of* Sp5 *mRNA changes determined by qRT-PCR in control (DMSO-treated) and CHIR99021-treated embryos; p<0.001.*