

Supplementary Material

Title: *Amphioxus SYCP1*: a case of retrogene replacement and co-option of regulatory elements adjacent to the ParaHox cluster

Journal: Development, Genes and Evolution

Authors: Myles G. Garstang* and David E.K. Ferrier

Affiliation: The Scottish Oceans Institute, Gatty Marine Laboratory, University of St Andrews, East Sands, St Andrews, Fife, KY16 8LB, UK.

*** Current address:** School of Biological Sciences, University of Essex, Wivenhoe, Colchester, Essex, CO4 3SQ

Corresponding author: David E.K. Ferrier

The Scottish Oceans Institute, Gatty Marine Laboratory, University of St Andrews, East Sands, St Andrews, Fife, KY16 8LB, UK.

Tel.: +44 (0)1334 463480

e-mail: dekf@st-andrews.ac.uk

Figure S1: Full-length protein alignment of metazoan SYCP1 proteins.



Table S1. Metazoan SYCP1 intron length comparison

Species (Genome version)	No. of <i>SCP1</i> introns	Length of transcribed <i>SCP1</i> gene (bp)	Total intron + exon length (bp)	Scaffold/Chromosome
<i>Mus musculus</i> (GRCm38)	33	3437	117601	Chr3: 102,818,499-102,936,100
<i>Homo sapiens</i> (GRCh38)	32	3452	140567	Chr1: 114,854,803-114,995,370
<i>Danio rerio</i> (GRCz10)	32	3440	8345	Chr6: 49,054,088-49,062,433
<i>Callorhinchus milii</i> (ESHARK1)	24	2462	30416	scaffold_89: 538,459-573,196
<i>Ciona intestinalis</i> (Cioin2)	0	2696	2696	chr_08q:4728476-4808970
<i>Branchiostoma floridae</i>	0	3141	3141	PACS 33B4, 36D2
<i>Strongylocentrotus purpuratus</i> (Spur_v3.1)	23	2816	11135	Scaffold2806:1-29859
<i>Patiria Miniata</i> (Pmin_v1)	22	1668	12900	Scaffold2247:1-47700
<i>Aplysia californica</i> (Broad 2.0/aplCal1)	21	2196	18402	Scaffold_829:1-266,202
<i>Capitella teleta</i> (Capca1)	25	3048	5745	scaffold_46:228355-395154
<i>Lottia gigantea</i> (Lotigi1)	21	2214	15448	sca_59:503750-573256
<i>Hydra vulgaris</i> (Hydra_RP_1.0)	6	3056	26989	HYDRAScaffold_36716
<i>Amphimedon queenslandica</i> (Aqu1)	23	2331	4812	Scaffold Contig13414: 96,588-101,400

Table S2. Amphioxus SYCP1 predicted promoter sequences

<i>B.floridae</i>	
Identifier	Sequence (Bold italic font indicates predicted TSS)
NNPP 1	TAGTACAGTGTATAAGAGTCTATTATAACGAAAAAAATC T GACTTTCTG
NNPP 2	TACAAGTAGGAAAAAAGCTGCAGAGAAGTCAGAAATCAG A AAGTCAGAT
NNPP 3	TCCCTTTTCCAAAAATGGCGCCTCCAAGTCAGAATGTCG T CTGCAAGTG
NNPP 4	TATGTGTAGATATATGAATTGCGATCACGCCTAGCGTAGC A AAGTATTGT
NNPP 5	ATCTATTGTTCTATAAAAGGAAACGAACCCCTTGACACTC A CCAAAGGCA
TSSW Promoter TSS	TCACTTGCAGACGACATTCTGACTT G GAGGCGCCATTTTGGAAAAAGGG
ProScan 1	CCATAAGCCGCTCATCTTCGTCATCTTCCTCGTAAATCGCATCGAAGTCGCCATGATGGAAACATCTGCG CACGGTATGCTGGGAAATATCTCCGGGGTGCAAGTTCGAGGTGACGTCACTTGCAGACGACATTCTGACT TGGAGGCGCCATTTTGGAAAAAGGGAGTTTCTTCCAACAAGAACTGTAAGTGACCTCCTGTTAATCCAGT TCTCTATGTGTATGTAGTTCTGTTAATATCGTAATACA
Proscan 2	ATATCTCCGGGGTGCAAGTTCGAGGTGACGTCACTTGCAGACGACATTCTGACTTGGAGGCGCCATTTT GGAAAAAGGGAGTTTCTTCCAACAAGAACTGTAAGTGACCTCCTGTTAATCCAGTTCTCTATGTGTATGTA GTTCTGTTAATATCGTAATACAACATCTTAACGCGTTAAATATTATTTTCGGTGGGAGGGCGAA
<i>B.lanceolatum</i>	
Identifier	Sequence (Bold italic font indicates predicted TSS)
NNPP1	ATTCTACATGTATTAATAACCTACTACACGATAGGGAC A CTTTGCAGA
NNPP2	CCCTACATTTTATATAGACATGCCTGCTTGAATTGTTGG T TCAGACGCA
NNPP3	CAGGCATGTCTATATAAAATGTAGGGTGAAAAGGTCTTT T TTATGTCCT
NNPP4	GCATGGAGGACATAAAAAAGACCTTTTCACTTACATTT T ATATAGACA
NNPP5	CTCCTCTCTGCAAAAATGGCGCCTCCGAGTCAAAATGTCG T CTGCAAGTG
NNPP6	CACTGAATCTAAATATAACAAGTTACCGCAACTCGTATT C A GACACATAT
NNPP7	TCTTTTGTTCTTAAAAAGGACATGAACCTCTGCTGACACT C TTATAAGG
NNPP8	CTGACACTCCTTATAAGGCAGCGGATTGAGGTGTGCGGT A TAGATAAGG
NNPP9	ACACTTACTTTATCTATACCGCACACCTCAAATCCGCTG C TTATAAGGA
TSSW Promoter 1 TSS	GGCATACCGTGCGCAG A TGTTTCCATCATG
TSSW Promoter 2 TSS	GAGTCAAAATGTCGT C T GCAAGTGACGTCA
Proscan1	ACGCGTTGATATGTTTTATTCCGATGCTAGTCTGATACATTAAGTCATACATAGCGAACTGGAATAACAG TAGATCACTTACTTCTGGGAGAAAAATCCCTGAAGAAAAATCCCTCTGTTCTCTCTCTGCAAAAATGG CGCCTCCGAGTCAAAATGTCGTCTGCAAGTGACGTCACTTCGAAGTGACCCCGGGAGATATTTCCCGGC ATACCGTGCGCAGATGTTTCCATCATGGCGGACTTCGA

Table S3. Metazoan SYCP1 sequences

Species	Group	Accession number/ ID
<i>Alvinella pompejana</i> (Pompeii worm)	Lophotrochozoa-Polychaeta	GO222799.1
<i>Amphimedon queenslandica</i> (demosponge)	Porifera	ACUQ01003074.1
<i>Anolis carolinensis</i> (lizard)	Tetrapoda	XP_008108235.1
<i>Aplysia californica</i> (Sea hare)	Lophotrochozoa-Gastropoda	XP_005108708.1
<i>Asterias amurensis</i> (northern pacific sea star)	Echinodermata-Asteroidia	GAVL01043045.1
<i>Branchiostoma floridae</i> (amphioxus)	Chordata-Cephalochordata	AC129948, AC129947
<i>Branchiostoma lanceolatum</i> (amphioxus)	Chordata-Cephalochordata	BI71nemh Sc0000037:73,417-80,768, Unpublished cDNA
<i>Branchiostoma belcheri</i> (amphioxus)	Chordata-Cephalochordata	XM_019769777.1
<i>Callorhinchus milii</i> (elephant shark)	Chordata-Chimaera	XP_007899851.1
<i>Canis lupis familiaris</i> (dog)	Chordata-Tetrapoda	XP_857086.1
<i>Ciona intestinalis</i> (Sea Squirt)	Chordata-Tunicata	KH2012:KH.C8.516.v1.A.ND1-1
<i>Ciona savignyi</i> (Sea Squirt)	Chordata-Tunicata	AACT01000684.1
<i>Crassostrea gigas</i> (pacific oyster)	Lophotrochozoa--Bivalvia	XP_011438578.1
<i>Danio rerio</i> (zebrafish)	Chordata-Actinopterygii	NP_001112366.1
<i>Equus caballus</i> (horse)	Chordata-Tetrapoda	XP_001496166.2
<i>Gallus gallus</i> (chicken)	Chordata-Tetrapoda	XP_004935063.1
<i>Hapalochlaena maculosa</i> (southern blue-ringed octopus)	Lophotrochozoa-Cephalopoda	GEXH01115743.1
<i>Homo sapiens</i> (human)	Chordata-Tetrapoda	EAW56621.1
<i>Hydra vulgaris</i> (freshwater hydroid)	Cnidaria-Hydrozoa	JQ906934.1
<i>Hydractinia symbiolongicarpus</i> (marine hydroid)	Cnidaria-Hydrozoa	GCHW01005973.1
<i>Lamellibrachia satsuma</i> (Giant tube worm)	Lophotrochozoa-Polychaeta	GEHO01210685.1
<i>Lottia Gigantea</i> (owl limpet)	Lophotrochozoa-Bivalva	FC693207.1
<i>Lytechinus variegatus</i> (green sea urchin)	Echinodermata-Echinoidea	GAUR01060760.1
<i>Macaca mulatta</i> (rhesus macaque)	Chordata-Tetrapoda	XP_001111808.1
<i>Mnemiopsis leidyi</i> (sea Walnut)	Ctenophora	GFAT01124492
<i>Mus musculus</i> (mouse)	Chordata-Tetrapoda	CAA86262.1
<i>Mytilus galloprovincialis</i> (mediterranean mussel)	Lophotrochozoa--Bivalvia	GAEN01007290.1
<i>Nematostella vectensis</i> (starlet sea anemone)	Cnidaria-Anthozoa	HADP01191031.1
Species	Group	Accession number/Scaffold ID

<i>Octopus bimaculoides</i> (california two-spot octopus)	Lophotrochozoa-Cephalopoda	XM_014935107.1
<i>Olavius algarvensis</i> (gutless oligochaete worm)	Lophotrochozoa-Oligochaeta	HACZ01084717.1
<i>Orbicella faveolata</i> (mountainous star coral)	Cnidaria-Anthozoa	XM_020759056.1
<i>Orzias latipes</i> (medaka)	Chordata-Actinopterygii	JQ906936.1
<i>Patiria pectinifera</i> (blue bat star)	Echinodermata-Asteroidia	GAVT01095647.1
<i>Pleurobrachia pileus</i> (sea gooseberry)	Ctenophora	FP999277
<i>Pomacea canaliculate</i> (channelled applesnail)	Lophotrochozoa-Gastropoda	GBZZ01052649.1
<i>Rattus norvegicus</i> (rat)	Chordata-Tetrapoda	NM_012810.1
<i>Ruditapes philippinarum</i> (littleneck clam)	Lophotrochozoa--Bivalvia	JO101748.1
<i>Saccoglossus kowalevskii</i> (acorn worm)	hemichordata	XM_006821456.1
<i>Sepiella maindroni</i> (cuttlefish)	Lophotrochozoa-Cephalopoda	GFLT01025932.1
<i>Strongylocentrotus purpuratus</i> (purple sea urchin)	Echinodermata-Echinoidea	XM_011679322.1
<i>Turritopsis</i> sp. (immortal jellyfish)	Cnidaria-Hydrozoa	IAAF01017580.1
<i>Xenopus tropicalis</i> (western clawed frog)	Chordata-Tetrapoda	XP_012811338.1
CCDC39-Drosophila Melanogaster	Ecdysozoa-Hexapoda	ACD81657.1
CCDC39-Homo Sapiens	Chordata-Tetrapoda	NP_852091.1
CCDC39-Strongylocentrotus purpuratus	Echinodermata-Echinoidea	XP_781717.3