

## **Origin and evolution of lysyl oxidases**

Xavier Grau-Bové<sup>1</sup>, Iñaki Ruiz-Trillo<sup>1,3,4</sup>, Fernando Rodriguez-Pascual<sup>2</sup>

<sup>1</sup> Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Barcelona, Spain

<sup>2</sup> Centro de Biología Molecular “Severo Ochoa” Consejo Superior de Investigaciones Científicas (C.S.I.C.) / Universidad Autónoma de Madrid (Madrid), Madrid, Spain

<sup>3</sup> Departament de Genètica, Universitat de Barcelona, Barcelona, Spain

<sup>4</sup> Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

### **Supplementary Material**

**Tables S1 and S2.**

**Files S1 to S8.**

## Figure legends

**Table S1.** Taxon sampling used in the eukaryotic+prokaryotic phylogenies. It includes the taxonomic classification, the acronym used in the present study (in brackets) and the source of the proteome data.

**Table S2.** Extended sampling of animal genomes, used in the holozoan phylogenies. It includes the taxonomic classification, the acronym used in the present study (in brackets) and the source of the proteome data.

**File S1.** Maximum likelihood (ML) phylogenetic tree of 154 eukaryotic and prokaryotic LOX, using dataset from supplementary Table S1 and prokaryote genomes (see Materials and Methods). Trees rooted using the midpoint-rooted tree option. Bootstrap supports are indicated at each node. Sequences are color-coded according to taxonomic assignment (see legend).

**File S2.** Bayesian inference (BI) phylogenetic tree of 154 eukaryotic and prokaryotic LOX, using dataset from supplementary Table S1 and prokaryote genomes (see Materials and Methods). Trees rooted using the midpoint-rooted tree option. Bayesian Posterior Probabilities are indicated at each node. Sequences are color-coded according to taxonomic assignment (see legend).

**File S3.** Maximum likelihood (ML) phylogenetic tree of 129 metazoan and ichthyosporean LOX, using dataset from supplementary Table S2 and ichthyosporean genomes (see Materials and Methods). Trees rooted using the midpoint-rooted tree option. Bootstrap supports are indicated at each node. Sequences are color-coded according to taxonomic assignment (see legend).

**File S4.** Bayesian inference (ML) phylogenetic tree of 129 metazoan and ichthyosporean LOX, using dataset from supplementary Table S2 and ichthyosporean genomes (see Materials and Methods). Trees rooted using the midpoint-rooted tree option. Bayesian Posterior Probabilities are indicated at each node. Sequences are color-coded according to taxonomic assignment (see legend).

**File S5.** Fasta file with 154 eukaryotic and prokaryotic LOX sequences (complete proteins), extracted from dataset from supplementary Table S1 and prokaryote genomes (see Materials and Methods).

**File S6.** Fasta file with 129 metazoan and ichthyosporean LOX sequences (complete proteins), extracted from dataset from supplementary Table S2 and ichthyosporean genomes (see Materials and Methods).

**File S7.** GFF (General Feature Format) file showing molecular features in the protein domain architectures of sequences in File S5 (154 eukaryotic and prokaryotic LOX sequences).

**File S8.** GFF (General Feature Format) file showing molecular features in the protein domain architectures of sequences in File S6 (129 metazoan and ichthyosporean LOX sequences).

Table S1

Taxonomic classification	Species	Source	
<b>METAZOA</b>	Deuterostomia	<i>Homo sapiens (Hsap)</i>	Ensembl
		<i>Ciona intestinalis (Cint)</i>	Ensembl
		<i>Saccoglossus kowalevskii (Skow)</i>	NCBI
		<i>Strongylocentrotus purpuratus (Spur)</i>	NCBI
	Ecdysozoa	<i>Drosophila melanogaster (Dmel)</i>	Flybase <a href="http://flybase.org/">http://flybase.org/</a>
		<i>Caenorhabditis elegans (Cele)</i>	Ensembl
	Lophotrochozoa	<i>Capitella teleta (Ctel)</i>	JGI
		<i>Echinococcus multilocularis (Emul)</i>	Wellcome Trust Sanger Institute <a href="http://www.sanger.ac.uk/resources/downloads/helminths/echinococcus-multilocularis.html">http://www.sanger.ac.uk/resources/downloads/helminths/echinococcus-multilocularis.html</a>
		<i>Lottia gigantea (Lgig)</i>	JGI <a href="http://genome.jgi.doe.gov/">http://genome.jgi.doe.gov/</a>
		<i>Crassostrea gigas (Cgig)</i>	Oyster Genome Project webpage <a href="http://www.oysterdb.cn">http://www.oysterdb.cn</a>
	Cnidaria	<i>Nematostella vectensis (Nvec)</i>	NCBI
		<i>Acropora digitifera (Adig)</i>	<i>Acropora digitifera</i> Genome Project webpage <a href="http://marinegenomics.oist.jp/">http://marinegenomics.oist.jp/</a>
		<i>Hydra magnipapillata (Hmag)</i>	NCBI
	Placozoa	<i>Trichoplax adhaerens (Tadh)</i>	NCBI
	Ctenophora	<i>Mnemiopsis leidyi (Mlei)</i>	<i>M. leidyi</i> Genome Project webpage <a href="http://research.nhgri.nih.gov/mnemiopsis/">http://research.nhgri.nih.gov/mnemiopsis/</a>
	Porifera	<i>Oscarella carmela (Ocar)</i>	Courtesy of Scott Nichols
		<i>Amphimedon queenslandica (Aque)</i>	NCBI
		<i>Sycon ciliatum (Scil)</i>	Courtesy of Maja Adamska
		<i>Leucosolenia complicate (Lcom)</i>	Courtesy of Maja Adamska
		<i>Monosiga brevicollis (Mbre)</i>	NCBI
<b>CHOANOFLAGELLATA</b>	<i>Salpingoeca rosetta (Sros)</i>	Broad Institute <a href="http://www.broadinstitute.org/annotation/genome/multicellularity_project/MultiHome.html">http://www.broadinstitute.org/annotation/genome/multicellularity_project/MultiHome.html</a>	
	<i>Capsaspora owczarzaki (Cowc)</i>	Broad Institute	
<b>FILASTEREA</b>	<i>Ministeria vibrans (Mvib)</i>	RNASeq data	
	<i>Sphaeroforma arctica (Sarc)</i>	Broad Institute	
<b>ICHTHYOSPOREA</b>	<i>Creolimax fragrantissima (Cfra)</i>	Beijing Genome Institute (Beijing, PRC)	
	<i>Abeoforma whisleri (Awhi)</i>	RNASeq data	
	<i>Pirum gemmata (Pgem)</i>	RNASeq data	
	<i>Amoebidium parasiticum (Apar)</i>	RNASeq data	
	<i>Incertae sedis</i>	<i>Corallochytrium limacisporum (Clim)</i>	The Genome Analysis Centre (Norwich, UK)
<b>FUNGI</b>	Ascomycota	<i>Schizosaccharomyces pombe (Spom)</i>	NCBI
		<i>Tuber melanosporum (Tmel)</i>	NCBI
		<i>Neurospora crassa (Ncra)</i>	Broad Institute <a href="http://www.broadinstitute.org/scientific-community/science/projects/fungal-genome-initiative">http://www.broadinstitute.org/scientific-community/science/projects/fungal-genome-initiative</a>
	Basidiomycota	<i>Saccharomyces cerevisiae (Scer)</i>	Broad Institute
		<i>Cryptococcus neoformans (Cneo)</i>	NCBI
		<i>Coprinopsis cinerea (Ccin)</i>	Broad Institute
	<i>Incertae sedis</i>	<i>Ustilago maydis (Umay)</i>	NCBI
		<i>Mortierella verticillata (Mver)</i>	Broad Institute
	Mucoromycotina	<i>Phyomyces blakesleeanus (Pbla)</i>	JGI
		<i>Mucor circinelloides (Mcir)</i>	JGI
		<i>Rhizopus oryzae (Rory)</i>	Broad Institute
		<i>Umbelopsis ramanniana (Umar)</i>	JGI Fungi Portal <i>Fueling the future with fungal genomics</i> , I Grigoriev et al., Mycology 2011
	Blastocladiomycota	<i>Allomyces macrogynus (Amac)</i>	Broad Institute
		<i>Catenaria anguillulae (Cang)</i>	JGI Fungi Portal <i>Fueling the future with fungal genomics</i> , I Grigoriev et al., Mycology 2011
	Chytridiomycota	<i>Batrachochytrium dendrobatidis (Bden)</i>	JGI
		<i>Spizellomyces punctatus (Spun)</i>	Broad Institute
	Glomeromycota	<i>Rhizophagus irregularis (Rirr)</i>	JGI Fungi Portal <i>Fueling the future with fungal genomics</i> , I Grigoriev et al., Mycology 2011
	Kickxellomycotina	<i>Coemansia reversa (Crev)</i>	JGI Fungi Portal <i>Fueling the future with fungal genomics</i> , I Grigoriev et al., Mycology 2011
	Entomophthoromycota	<i>Conidiobolus coronatus (Ccor)</i>	JGI Fungi Portal <i>Fueling the future with fungal genomics</i> , I Grigoriev et al., Mycology 2011

	Monoblepharidomycota	<i>Gonapodya prolifera</i> (Gpro)	JGI Fungi Portal <i>Fueling the future with fungal genomics</i> , I Grigoriev et al., Mycology 2011
	Neocallimastigomycota	<i>Piromyces</i> sp. E2 (Pisp)	JGI Fungi Portal <i>Fueling the future with fungal genomics</i> , I Grigoriev et al., Mycology 2011
	Microsporidia	<i>Encephalitozoon cuniculi</i> (Ecu)	NCBI
		<i>Nematocida parisii</i> (Npar)	NCBI
	Cryptomycota	<i>Rozella allomycis</i> (Rall)	JGI
<b>NUCLEARIIDAE AND FONTICULA GROUP</b>		<i>Nuclearia</i> spp. (Nspp)	RNAseq data
		<i>Fonticula alba</i> (Falb)	Broad Institute <a href="http://www.broadinstitute.org/annotation/genome/multicellularity_project/Downloads.html">http://www.broadinstitute.org/annotation/genome/multicellularity_project/Downloads.html</a>
<b>APUSOZOA</b>		<i>Thecamonas trahens</i> (Ttra)	Broad Institute <a href="http://www.broadinstitute.org/annotation/genome/multicellularity_project/Downloads.html">http://www.broadinstitute.org/annotation/genome/multicellularity_project/Downloads.html</a>
<b>AMOEBOZOA</b>	Mycetozoa	<i>Dictyostelium discoideum</i> (Ddis)	NCBI
		<i>Polysphondylium pallidum</i> (Ppal)	NCBI
	Archamoeba	<i>Entamoeba histolytica</i> (Ehis)	Wellcome Trust Sanger Institute <a href="http://www.sanger.ac.uk/resources/downloads/protozoa/entamoeba.html">http://www.sanger.ac.uk/resources/downloads/protozoa/entamoeba.html</a>
	Centramoebida	<i>Acanthamoeba castellanii</i> (Acas)	<i>Ab initio</i> protein prediction
<b>VIRIDIPLANTAE</b>	Embryophyta	<i>Arabidopsis thaliana</i> (Atha)	NCBI
		<i>Mimulus guttatus</i> (Mgut)	Phytozome (JGI) <a href="http://www.phytozome.net/">http://www.phytozome.net/</a>
		<i>Aquilegia coerulea</i> (Acoe)	Phytozome (JGI)
		<i>Brachypodium distachyon</i> (Bdis)	Phytozome (JGI)
		<i>Sorghum bicolor</i> (Sbic)	Phytozome (JGI)
		<i>Selaginella moellendorffii</i> (Smoe)	NCBI
		<i>Physcomitrella patens</i> (Ppat)	NCBI
	Chlorophyta	<i>Chlamydomonas reinhardtii</i> (Crei)	NCBI
		<i>Volvox cartieri</i> (Vcar)	NCBI
		<i>Chlorella variabilis</i> (Cvar)	NCBI
		<i>Ostreococcus tauri</i> (Otau)	NCBI
		<i>Micromonas pusilla</i> (Mpus)	JGI
	Rhodophyta	<i>Cyanidioschyzon merolae</i> (Cmer)	<i>C. merolae</i> Genome Project webpage <a href="http://merolae.biol.s.u-tokyo.ac.jp/">http://merolae.biol.s.u-tokyo.ac.jp/</a>
		<i>Chondrus crispus</i> (Ccri)	<i>C. crispus</i> Genome Project webpage Courtesy of Mark Cock
		<i>Pyropia yezoensis</i> (Pyez)	<i>P. yezoensis</i> Genome Project webpage <a href="http://nrifs.fra.affrc.go.jp/ResearchCenter/5_AG/genome/s/nori/">http://nrifs.fra.affrc.go.jp/ResearchCenter/5_AG/genome/s/nori/</a>
	Glaucophyta	<i>Cyanophora paradoxa</i> (Cpar)	<i>C. paradoxa</i> Genome Project webpage <a href="http://cyanophora.rutgers.edu/cyanophora/">http://cyanophora.rutgers.edu/cyanophora/</a>
<b>HETEROKONTA/ STRAMENOPILES</b>	Brown algae/ Phaeophyceae	<i>Ectocarpus siliculosus</i> (Esil)	EMBL
		<i>Nannochloropsis gaditana</i> (Ngad)	<i>N. gaditana</i> Genome Project webpage <a href="http://nannochloropsis.genomeprojectsolutions-databases.com/">http://nannochloropsis.genomeprojectsolutions-databases.com/</a>
		<i>Aureococcus anophagefferens</i> (Aano)	Uniprot
		<i>Phaeodactylum tricorutum</i> (Ptri)	NCBI
		<i>Thalassiosira pseudonana</i> (Tpse)	NCBI
	Oomycota	<i>Phytophthora infestans</i> (Pinf)	NCBI
	Labyrinthulomycetes	<i>Aplanochytrium kerguelense</i> (Aker)	JGI
		<i>Schizochytrium aggregatum</i> (Sagg)	JGI
<b>ALVEOLATA</b>	Apicomplexa	<i>Toxoplasma gondii</i> (Tgon)	NCBI
		<i>Plasmodium falciparum</i> (Pfal)	Uniprot
	Ciliophora	<i>Paramecium tetraurelia</i> (Ptet)	Uniprot
		<i>Tetrahymena thermophila</i> (Tthe)	NCBI
	Perkinsidae	<i>Perkinsus marinus</i> (Pmar)	NCBI
	Dinoflagellata	<i>Symbiodinium minutum</i> (Smin)	<i>Symbiodinium minutum</i> Clade B1 Genome Project webpage <a href="http://marinegenomics.oist.jp/">http://marinegenomics.oist.jp/</a>
<b>RHIZARIA</b>		<i>Bigelowiella natans</i> (Bnat)	JGI
<b>HAPTOPHYTA</b>		<i>Emiliania huxleyi</i> (Ehux)	JGI
<b>CRYPTOPHYTA</b>		<i>Guillardia theta</i> (Gthe)	JGI
<b>EXCAVATA</b>	Heterolobosea	<i>Naegleria gruberi</i> (Ngru)	NCBI
	Kinetoplastida	<i>Bodo saltans</i> (Bsal)	Wellcome Trust Sanger Institute <a href="http://www.sanger.ac.uk/resources/downloads/protozoa/bodo-saltans.html">http://www.sanger.ac.uk/resources/downloads/protozoa/bodo-saltans.html</a>
		<i>Trypanosoma cruzi</i> (Tcru)	NCBI
		<i>Leishmania major</i> (Lmaj)	NCBI
	Metamonada	<i>Trichomonas vaginalis</i> (Tvag)	NCBI
		<i>Giardia lamblia</i> (Glam)	NCBI

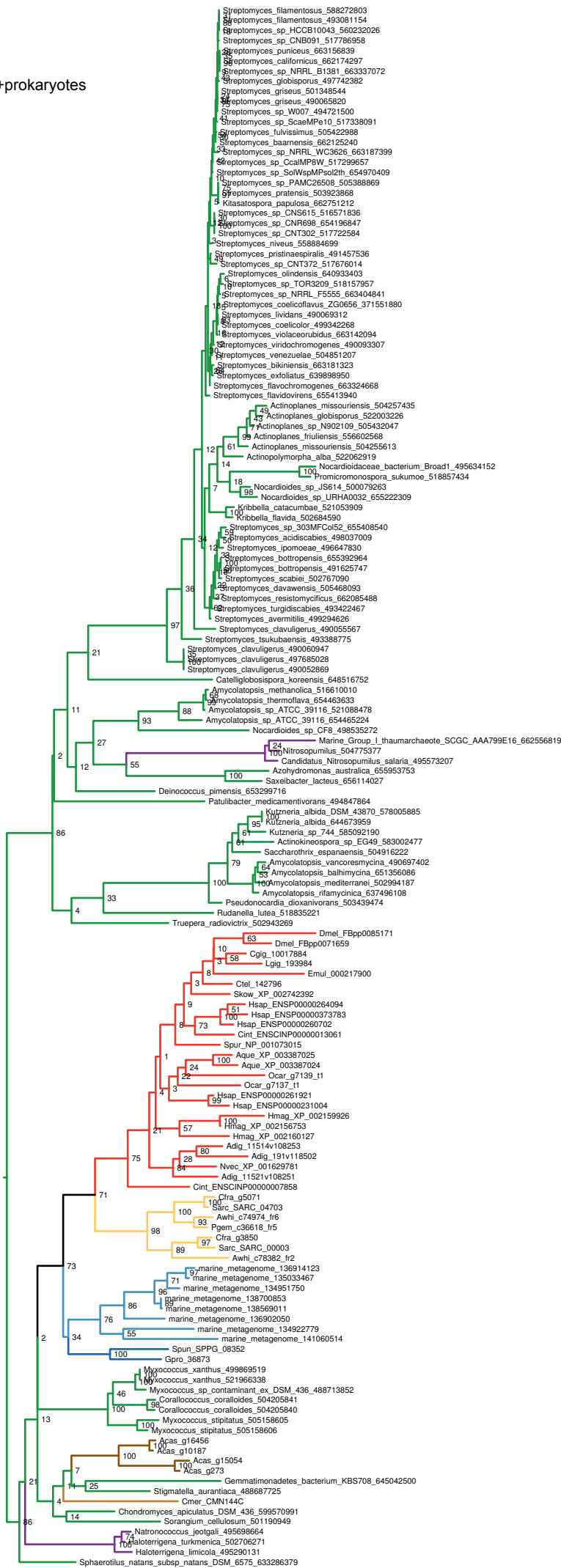
Table S2

Taxonomic classification		Species	Source	
DEUTEROSTOMIA	Chordata	Vertebrata	<i>Homo sapiens (Hsap)</i>	Ensembl
			<i>Mus musculus (Mmus)</i>	Ensembl
			<i>Gallus gallus (Ggal)</i>	Ensembl
			<i>Anolis carolinensis (Acar)</i>	Ensembl
			<i>Xenopus tropicalis (Xtro)</i>	Ensembl
			<i>Danio rerio (Drer)</i>	Ensembl
			<i>Latimeria chalumnae (Lcha)</i>	Ensembl
			<i>Callorhynchus milii (Cmil)</i>	Elephant Shark Genome Project <a href="http://esharkgenome.imcb.a-star.edu.sg/">http://esharkgenome.imcb.a-star.edu.sg/</a>
	<i>Petromyzon marinus (Pema)</i>	Ensembl		
	Tunicata	<i>Ciona intestinalis (Cint)</i>	Ensembl	
		<i>Ciona savignyi (Csav)</i>	Ensembl	
		<i>Oikopleura dioica (Oodio)</i>	Genoscope <a href="http://www.genoscope.cns.fr/externe/GenomeBrowser/Oikopleura/">http://www.genoscope.cns.fr/externe/GenomeBrowser/Oikopleura/</a>	
	Cephalochordata	<i>Branchiostoma floridae (Bflo)</i>	JGI	
	Ambulacraria	Hemichordata	<i>Saccoglossus kowalevskii (Skow)</i>	NCBI
Echinodermata		<i>Strongylocentrotus purpuratus (Spur)</i>	NCBI	
ECDYSOZOA	Arthropoda	Hexapoda	<i>Drosophila melanogaster (Dmel)</i>	Flybase <a href="http://flybase.org/">http://flybase.org/</a>
			<i>Tribolium castaneum (Tcas)</i>	Beetlebase <a href="http://www.beetlebase.org/">http://www.beetlebase.org/</a>
		Crustacea	<i>Daphnia pulex (Dpul)</i>	JGI
		Arachnida	<i>Ixodes scapularis (Isca)</i>	NCBI
	Nematoda	<i>Caenorhabditis elegans (Cele)</i>	Ensembl	
		<i>Trichinella spiralis (Tspi)</i>	NCBI	
		<i>Brugia malayi (Bmal)</i>	NCBI	
LOPHOTROCHOZOA	Spiralia	Mollusca	<i>Lottia gigantea (Lgig)</i>	JGI <a href="http://genome.jgi.doe.gov/">http://genome.jgi.doe.gov/</a>
			<i>Aplysia californica (Acal)</i>	NCBI
			<i>Crassostrea gigas (Cgig)</i>	Oyster Genome Project <a href="http://oysterdb.cn/">http://oysterdb.cn/</a>
		Annelida	<i>Capitella teleta (Ctel)</i>	JGI
			<i>Helobdella robusta (Hrob)</i>	JGI
	Platyhelminthes	<i>Echinococcus multilocularis (Emul)</i>	Wellcome Trust Sanger Institute <a href="http://www.sanger.ac.uk/resources/downloads/helminths/echinococcus-multilocularis.html">http://www.sanger.ac.uk/resources/downloads/helminths/echinococcus-multilocularis.html</a>	
		<i>Schistosoma mansoni (Sman)</i>	NCBI	
CNIDARIA	<i>Nematostella vectensis (Nvec)</i>	NCBI		
	<i>Acropora digitifera (Adig)</i>	<i>Acropora digitifera</i> Genome Project <a href="http://marinegenomics.oist.jp/">http://marinegenomics.oist.jp/</a>		
PLACOZOA	<i>Trichoplax adhaerens (Tadh)</i>	NCBI		
CTENOPHORA	<i>Mnemiopsis leidyi (Mlei)</i>	<i>M. leidyi</i> Genome Project <a href="http://research.nhgri.nih.gov/mnemiopsis/">http://research.nhgri.nih.gov/mnemiopsis/</a>		
PORIFERA	<i>Oscarella carmela (Ocar)</i>	Courtesy of Scott Nichols		
	<i>Amphimedon queenslandica (Aque)</i>	NCBI		
	<i>Sycon ciliatum (Scil)</i>	Courtesy of Maja Adamska		
	<i>Leucosolenia complicata (Lcom)</i>	Courtesy of Maja Adamska		

# Figure S1

ML 154 eukaryotes+prokaryotes

- Metazoa
- Ichthyosporia
- Fungi
- Amoebozoa
- Rhodophyta
- Bacteria
- Archaea
- Metagenomic



# Figure S2

BI 154 eukaryotes+prokaryotes

- Metazoa
- Ichthyospora
- Fungi
- Amoebozoa
- Rhodophyta
- Bacteria
- Archaea
- Metagenomic

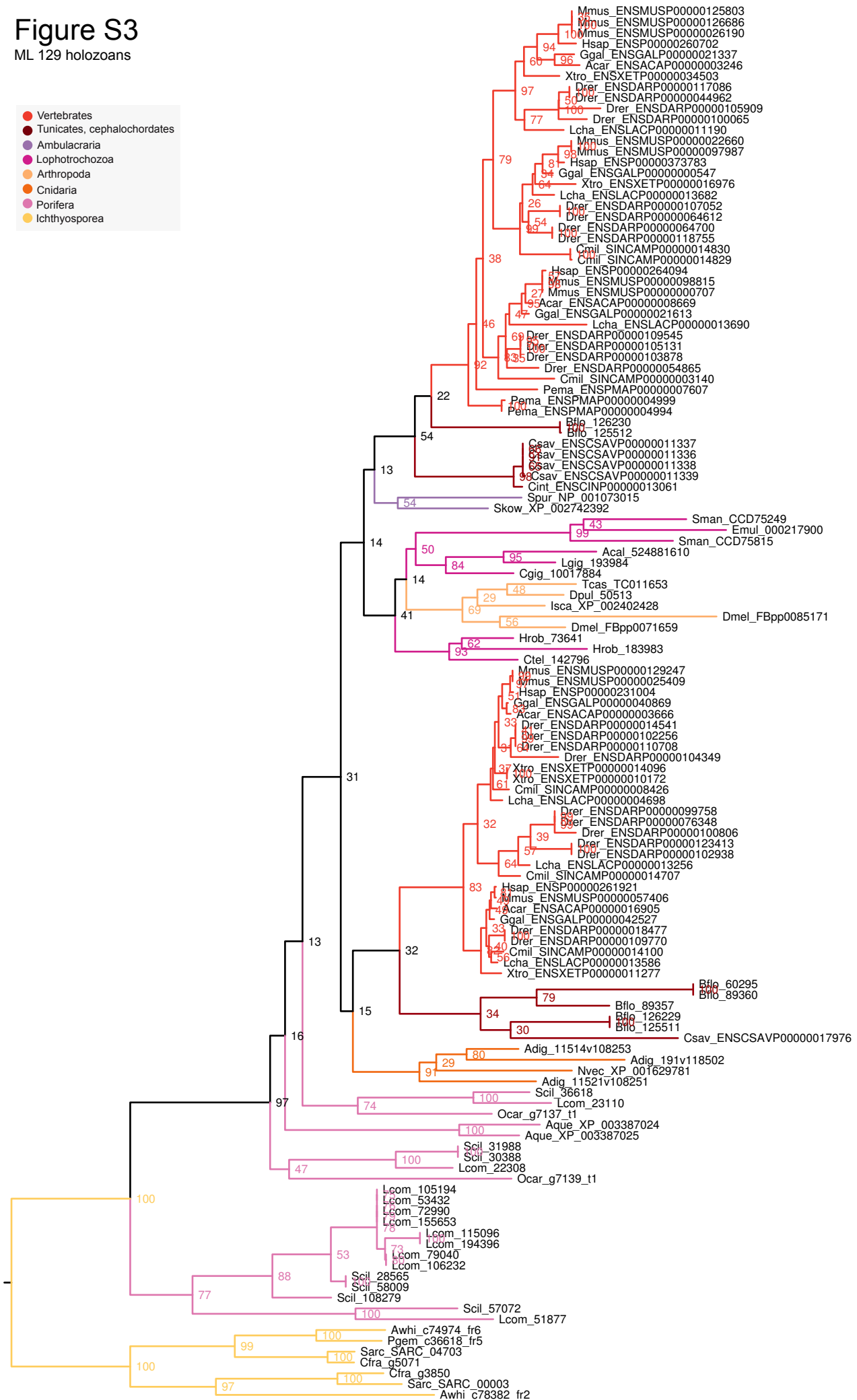




# Figure S3

ML 129 holozoans

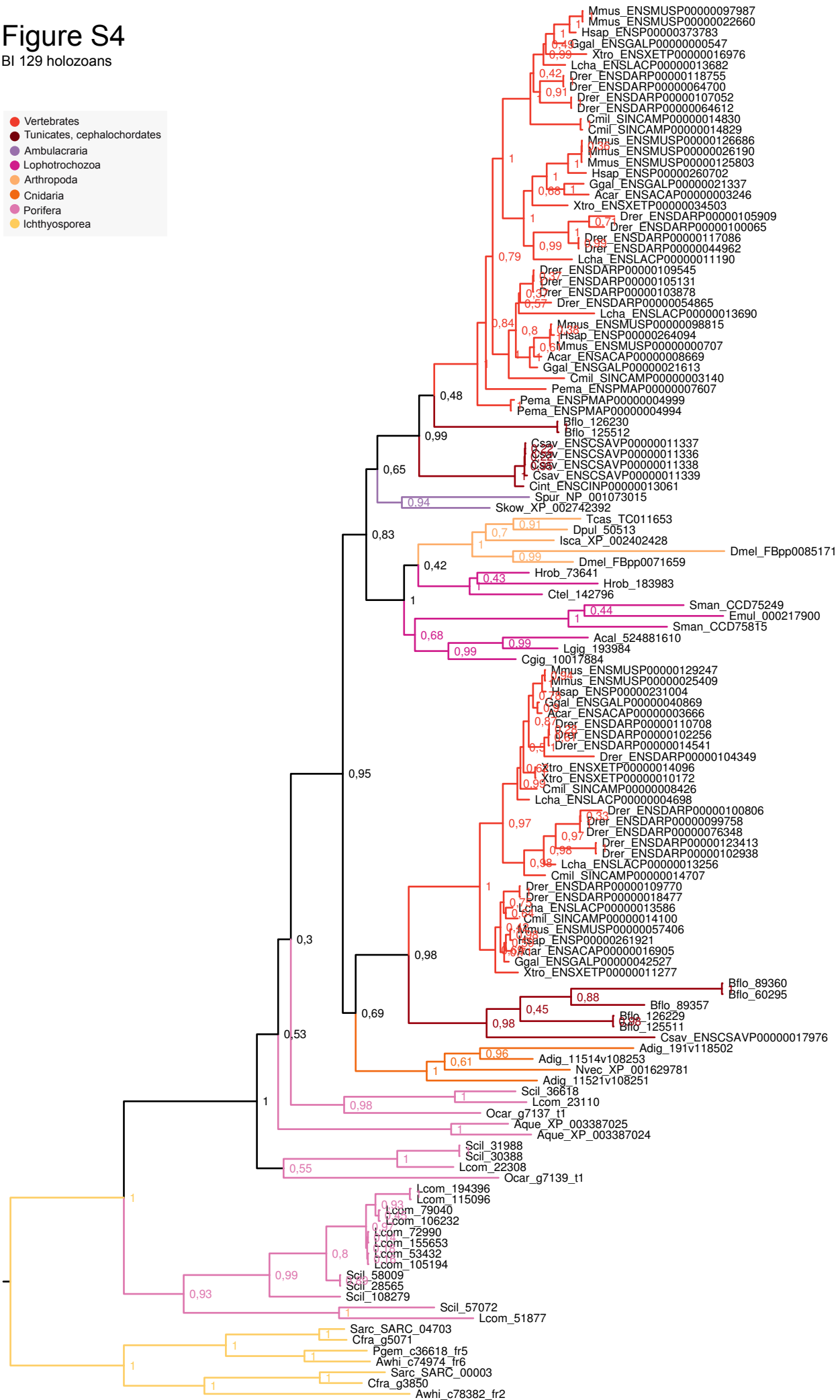
- Vertebrates
- Tunicates, cephalochordates
- Ambulacraria
- Lophotrochozoa
- Arthropoda
- Cnidaria
- Porifera
- Ichthyosporea



# Figure S4

BI 129 holozoans

- Vertebrates
- Tunicates, cephalochordates
- Ambulacraria
- Lophotrochozoa
- Arthropoda
- Cnidaria
- Porifera
- Ichthyosporia



0.3