



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

PHD PROGRAMME IN  
Innovative technologies and sustainable use of Mediterranean Sea  
fishery and biological resources (FishMed-PhD)

Cycle XXXVII

Settore Consorsuale: 05/B1 – ZOOLOGIA E ANTROPOLOGIA

Settore Scientifico Disciplinare: BIO/05 – ZOOLOGIA

**A comparative and evolutionary approach to  
study bivalve sex determination from a  
broad-phylogenetic perspective**

Candidate: Filippo Nicolini

PhD Coordinator:  
prof. Stefano Goffredo

Supervisor:  
prof. Andrea Luchetti

Final exam 2025

# Contents

<b>Table of contents . . . . .</b>	<b>i</b>
<b>List of abbreviations . . . . .</b>	<b>1</b>
<b>Chapter 1: Introduction . . . . .</b>	<b>4</b>
1.1 The diversity of sexual processes . . . . .	4
1.2 Genetic sex determination and the evolution of sex-determining genes . . . . .	5
1.3 Sex determination in bivalves: a long-standing enigma . . . . .	6
<b>Chapter 2: Bivalves as emerging model systems to study the mechanisms and evolution of sex determination: a genomic point of view. . . . .</b>	<b>9</b>
2.1 Open yet inspiring topics in bivalve sex determination . . . . .	11
2.1.1 Transitions between environmental and genetic sex determination . . . . .	11
2.1.2 Evolution of sex chromosomes . . . . .	12
2.1.3 Mito-nuclear interactions . . . . .	13
2.1.4 Evolution of sex-determination related genes . . . . .	14
2.2 The case of the Dmrt gene family in bivalves . . . . .	15
2.3 Conclusions: bivalves as new models in the study of sex determination . . . . .	17
2.4 Acknowledgments . . . . .	26
2.5 Data Availability . . . . .	26
<b>Chapter 3: Identification of putative sex-determination related genes in bivalves through comparative molecular evolutionary analyses . . . . .</b>	<b>27</b>
3.1 Introduction . . . . .	28
3.2 Materials and Methods . . . . .	31
3.2.1 Dataset of bivalve annotated genomes and transcriptomes . . . . .	31
3.2.2 Identification and classification of Dmrt, Sox and Fox genes in bivalves .	31
3.2.3 Sequence diversity of bivalve single-copy orthogroups . . . . .	33
3.2.4 Mammals and <i>Drosophila</i> spp. as test datasets . . . . .	34

3.2.5	GO-term enrichment . . . . .	34
3.3	Results . . . . .	36
3.3.1	Genomic and transcriptomic datasets . . . . .	36
3.3.2	The Dmrt, Sox, and Fox complements in bivalves . . . . .	36
3.3.3	Amino acid sequence divergence of Dmrt, Sox, and Fox genes in bivalves . . . . .	38
3.3.4	Dmrt, Sox, and Fox genes, and amino acid sequence divergence in the test datasets . . . . .	45
3.4	Discussion . . . . .	47
3.4.1	A new manually-curated and phylogenetic-based reference dataset of Dmrt, Sox, and Fox genes in bivalves . . . . .	47
3.4.2	High amino acid sequence divergence identifies putative sex-determining genes . . . . .	51
3.5	Conclusions. . . . .	56
<b>Chapter 4:</b>	<b>Expression patterns of three sex-related genes and the germline marker <i>Vasa</i> in early developmental stages of <i>Mytilus galloprovincialis</i> embryos . . . . .</b>	<b>57</b>
4.1	Introduction . . . . .	58
4.2	Materials and Methods . . . . .	58
4.2.1	Time-series gene expression . . . . .	58
4.2.2	Sample collection, MitoTracker staining and fixation . . . . .	59
4.2.3	mRNA <i>in-situ</i> Hybridization Chain Reaction (HCR) . . . . .	60
4.2.4	Immunolocalization of Vasa . . . . .	62
4.3	Results . . . . .	62
4.4	Discussion . . . . .	63
<b>References.</b>	<b>83</b>	
<b>Appendix</b>	<b>84</b>	
<b>Supplementary materials</b>	<b>90</b>	
Supplementary figures	90	
Supplementary tables	104	
<b>Data Availability.</b>	<b>173</b>	

# List of abbreviations

<b>AASD</b>	amino acid sequence divergence
<b>BSA</b>	bovine serum albumin
<b>CMS</b>	cytoplasmatic male sterility
<b>CUE</b>	coupling of ubiquitin conjugation to endoplasmic reticulum degradation [domain]
<b>DEAD/DEAH-box</b>	Asp-Glu-Ala-Asp/Asp-Glu-Ala-His box
<b>DGE</b>	differential gene expression
<b>DM</b>	<i>dsx</i> and <i>mab-3</i> [domain]
<b>DMA</b>	DM-associated [domain]
<b>Dmrt</b>	<i>doublesex</i> ( <i>dsx</i> ) and <i>mab-3</i> related transcription factor
<b>Dmrt-1L</b>	<i>Dmrt 1-like</i>
<b>Dm-W</b>	<i>dsx</i> and <i>mab-3</i> related gene <i>W</i>
<b>Dmy</b>	<i>dsx</i> and <i>mab-3</i> related gene <i>Y</i>
<b>DSFG</b>	Dmrt, Sox, and Fox gene
<b>dsx</b>	<i>doublesex</i>
<b>DUI</b>	doubly uniparental inheritance
<b>EDTA</b>	ethylenediaminetetraacetic acid
<b>ESD</b>	environmental sex determination
<b>FASW</b>	filtered artificial sea water
<b>FHA</b>	forkhead-associated [domain]
<b>Fox</b>	forkhead box

<b>GO</b>	gene ontology
<b>GRN</b>	gene regulatory network
<b>GSD</b>	genetic sex determination
<b>HCR</b>	hybridization chain reaction
<b>HeSC</b>	heteromorphic sex chromosome
<b>HMG</b>	high mobility group [box domain]
<b>HMM</b>	hidden Markov model
<b>HoSC</b>	homomorphic sex chromosome
<b>hpf</b>	hours post fertilization
<b><i>mab-3</i></b>	<i>male abnormal-3</i>
<b>MCL</b>	Markov clustering algorithm
<b>ML</b>	maximum likelihood
<b>mRNA-ISH</b>	mRNA <i>in-situ</i> hybridization
<b>Mya</b>	million years ago
<b>ORF</b>	open reading frame
<b>PBS</b>	1× phosphate-buffered saline
<b>PBS-Tw</b>	1× PBS with 0.1% Tween 20
<b>PFA</b>	paraformaldehyde
<b>PGC</b>	primordial germ cell
<b>RNAi</b>	RNA interference
<b>RT</b>	room temperature
<b>SC</b>	sex chromosome
<b>SCO</b>	single-copy orthogroup
<b>SD</b>	sex determination
<b>SDf</b>	sex differentiation
<b>SDG</b>	sex-determining gene

**SDS** sodium dodecyl sulfate  
**Sox** *Sry*-related HMG-box  
**SRG** sex-determination related gene  
***Sry*** *Sex-determining region of chromosome Y*  
**SSC-Tw** 5× saline-sodium citrate with Tween 20 buffer  
***Sxl*** *Sex-lethal*

**TBS** 1× Tris-buffered saline  
**TBS-Tx** 1× TBS with Triton X-100  
***tra*** *transformer*

# Chapter 1

## Introduction

### 1.1 The diversity of sexual processes

The process of sex determination (SD) has been traditionally associated with the very first step of gonad differentiation, where an initial trigger activates the molecular pathway that establishes organism sex. According to this view, two alternative types of SD have been recognized at first: the environmental sex determination (ESD) and the genetic sex determination (GSD), depending on whether the very first cues are of environmental or genetic origin. Conversely, all the downstream events of gonad development (i.e., after SD) have been appointed as sex differentiation (SDf), which consists of the entire set of morphogenetic, molecular, and physiological events leading to the full maturation of testes or ovaries (**Uller and Helanterä, 2011; Beukeboom and Perrin, 2014**). Lately, however, the dichotomous views of ESD/GSD and of SD/SDf have been questioned. On the one hand, a growing number of studies on non-model organisms proved that ESD and GSD represent a continuum of mixed conditions rather than two mutually exclusive phenomena. On the other, the high evolutionary dynamics and the variable expression patterns of the genes involved in the processes of gonad commitment and development make the distinction between SD and SDf of unclear utility (**Beukeboom and Perrin, 2014**).

Considering this complex scenario, **Uller and Helanterä, 2011** proposed a unified and broad-scope definition for SD, that is, “the processes within an embryo leading to the formation of differentiated gonads as either testes or ovaries”, without any actual distinction between environmental/genetic initial triggers or the downstream effectors. However, I argue that this definition should be expanded to encompass not only the embryonic stage of the animal life cycle

but also adulthood, since cases of sex reversals and sex changes (sequential hermaphroditism) legitimately express proper SD processes during post-embryonic life stages as well.

## 1.2 Genetic sex determination and the evolution of sex-determining genes

In its most intimate core, animal SD is the manifestation of complex gene regulatory networks where, in accordance with the Wilkins' theory (1995), the downstream actors appear to be nearly conserved both from functional and identity point of views, while the master top regulators (the commonly recognized sex determinants, such as the *Sex-determining region of chromosome Y (Sry)* in therians or the ratio between sex and autosome chromosomes in *Drosophila*) are often the most variable part (**Beukeboom and Perrin, 2014**). As a matter of fact, this evolutionary pattern of animal sex-determining cascades has been observed in major animal clades, including vertebrates (e.g., **Marshall Graves and Peichel, 2010**), insects (e.g., **Verhulst et al., 2010**), and nematodes (e.g., **Stothard and Pilgrim, 2003**).

Sex-determination related genes (SRGs) are of particular interest not only from a regulatory point of view but also because of their patterns of molecular evolution. In fact, transcriptionally sex-biased genes (including SRGs) often tend to evolve faster than unbiased genes at the level of protein sequences. In particular, male-biased genes generally show higher rate of sequence evolution in comparison to both female-biased and unbiased counterparts (reviewed in **Parsch and Ellegren, 2013; Grath and Parsch, 2016**), as it has been repeatedly observed in well-studied organisms such as fruit flies (e.g., **Meisel and Connallon, 2013**), nematodes (e.g., **Cutter and Ward, 2005**), mice (e.g., **Kousathanas et al., 2014**) and primates (e.g., **Khaitovich et al., 2005**), and in other emerging model systems, such as *Daphnia pulex* (**Eads et al., 2007**), aphids (**Purandare et al., 2014**), and two wasp species of the genus *Nasonia* (**Wang et al., 2015**). Growing evidence is however showing cases in which instead female-biased genes have higher rates of sequence evolution than male-biased genes, such as in mosquitoes of the genus *Anopheles* (**Papa et al., 2017**), and European and Manila clams of the genus *Ruditapes* (**Ghiselli et al., 2018**).

The pattern of molecular evolution of sex-biased genes is particularly evident in organisms with sex chromosomes (both in XY/ZW and X0 systems), such as fruit flies, birds and mammals, where the so-called fast-X (or fast-Z) effect has been extensively reported for sex-chromosome

associated genes (**Vicoso and Charlesworth, 2006; Mank et al., 2007; Meisel and Connallon, 2013**). This high rate of sequence evolution in sex-biased genes and sex chromosomes (SCs) can be the result of both adaptative and non-adaptative processes, since the observed higher ratio between non-synonymous and synonymous mutations (dN/dS) can be caused by natural selection, sexual selection or sexual antagonism, as well as genetic drift (**Vicoso and Charlesworth, 2006; Meisel and Connallon, 2013; Parsch and Ellegren, 2013; Grath and Parsch, 2016**).

### 1.3 Sex determination in bivalves: a long-standing enigma

Bivalves are the second largest clade in molluscs, counting more than 18,000 species (Catalogue of Life) distributed at all depths and in all marine environments, as well as in some freshwater habitats. Thanks to their high diversity and biological peculiarities, they have been proposed as promising model organisms for investigating a wide array of biological, ecological and evolutionary issues (**Milani and Ghiselli, 2020; Ghiselli et al., 2021**). However, despite their socio-economic and scientific importance, the knowledge concerning the molecular basis of bivalve reproduction and SD is still quite limited (**Breton et al., 2018**). Clues from various works seem to suggest that both genetic and environmental factors (e.g., temperature, food availability, and steroids) are involved in SD, and that heteromorphic sex chromosomes (HeSCs) are absent (**Breton et al., 2018; Han et al., 2022**). However, the exact process by which sex is determined and gonad commitment is established is, currently, still unknown. Actually, bivalves represent a dazzling example of how the traditional dichotomies between ESD/GSD and SD/SDf can sometimes hamper scientific research, as many bivalve species exhibit various forms of hermaphroditism and because a master environmental or genetic sex determinant inducing SDf may just not exist.

In the attempt to identify SRGs, many differential gene expression analyses have been recently performed on a variety of species covering most of the phylogenetic diversity of bivalves (e.g., **Milani et al., 2013; Zhang et al., 2014; Chen et al., 2017; Capt et al., 2018; Ghiselli et al., 2018; Shi et al., 2018**). Some of the genes that were found to be differentially expressed between gonads of different sex were systematically retrieved across species, such as those belonging to the *dsx* and *mab-3* related transcription factor (Dmrt), *Sry*-related HMG-box (Sox), and forkhead box (Fox) families, which act in concert in various

animal developmental processes including the SD cascade (**Marshall Graves and Peichel, 2010; Beukeboom and Perrin, 2014**). To this regard, **Zhang et al., 2014** proposed a working model for the sex-determining pathway of the Pacific oyster *Crassostrea gigas* in which: *CgSoxH* promotes male gonad development by activating *CgDsx*, which belong to the Dmrt family, and inhibiting *CgFoxL2*; *CgFoxL2*, when not inhibited by the pair *CgSoxH/CgDsx*, promotes female gonad development. Moreover, **Han et al., 2022** recently identified homomorphic sex chromosomes (HoSCs) in eight scallop species and appointed *FoxL2* as a putative SRG in *Patinopacten yessoensis* and *Chlamys farreri*. Though, much of the recent research effort on bivalve SRGs has been limited to their molecular cloning, differential transcription, and tissue localization (**Liang et al., 2019; Sun et al., 2022**). Furthermore, few works have directly investigated the biological functions of Dmrt, Sox, and Fox genes in bivalves so far, and most used post-transcriptional silencing of target mRNAs [RNA interference (RNAi)]. **Liang et al., 2019** studied the role of *Sox2* in the spermatogenesis of the Zhikong scallop *C. farreri* and found that it likely regulates proliferation of spermatogonia and apoptosis of spermatoocytes, since its knockdown resulted in the loss of male germ cells. **Wang et al., 2020** proposed that in the female gonads of the freshwater mussel *Hyriopsis cumingii*, *FoxL2* might be related to the *Wnt/β-catenin* signaling pathway, which takes part in ovarian differentiation also in vertebrates. **Sun et al., 2022** found instead that in *C. gigas*, *FoxL2* and *Dmrt1L* mRNA knockdown results in the size reduction of female and male mature gonads, respectively.

In this sense, bivalve molluscs represent a striking example of the difficulty to reconcile the traditional view of a single sex determinant with an apparent multifactorial model in which many genes and environmental cues act in concert to establish the sexual identity of the individual (**Breton et al., 2018**). Lately, much effort has been put in the characterisation of bivalve SD and a general framework is eventually taking shape. Functional assays with RNAi and CRISPR-Cas9 techniques (e.g., **Wang et al., 2020; Sun et al., 2022; Wang et al., 2022**), as well as with mRNA *in-situ* hybridization (mRNA-ISH) and immunohistochemistry (e.g., **Perez-Garcia et al., 2011; Milani et al., 2013**), are making their way into the study of bivalve biology and have been proved essential instruments also for the investigation of sex-related traits. However, very few works have made extensive use of the comparative and integrative approach in bivalve studies so far, which hampers the possibility to infer general patterns for such a vast class of organisms (**Milani and Ghiselli, 2020**). The high evolutionary rates and plasticity of SRGs make the situation even harder, since phylogenetic and orthology

inferences can lead to erroneous reconstructions in the presence of signal saturation and high sequence divergence (reviewed in **Natsidis et al., 2021**; **Lozano-Fernandez, 2022**).

# Chapter 2

## Bivalves as emerging model systems to study the mechanisms and evolution of sex determination: a genomic point of view

Filippo Nicolini<sup>1,2</sup>, Fabrizio Ghiselli<sup>1</sup>, Andrea Luchetti<sup>1</sup>, Liliana Milani<sup>1</sup>

<sup>1</sup>*Department of Biological, Geological and Environmental Science, University of Bologna, Bologna (BO), Italy.*

<sup>2</sup>*Fano Marine Center, Fano (PU), Italy.*

**Published in:** 2023, *Genome Biology and Evolution*, 15(10):evad181.  
10.1093/gbe/evad181

**Abstract.** Bivalves are a diverse group of molluscs that have recently attained a central role in plenty of biological research fields, thanks to their peculiar life history traits. Here we propose that bivalves should be considered as emerging model systems also in sex-determination studies, since they would allow to investigate: (i) the transition between environmental and genetic sex determination, with respect to different reproductive backgrounds and sexual systems (from species with strict gonochorism to species with various forms of hermaphroditism); (ii) the genomic evolution of sex chromosomes, considering that no heteromorphic sex chromosomes are currently known and that homomorphic sex chromosomes have been identified just in few species of scallops; (iii) the putative role of mitochondria at some level of the sex determination signaling pathway, in a mechanism that may resemble the cytoplasmatic male sterility of plants; (iv) the evolutionary history of sex-determination related gene families with respect to other

animal groups. In particular, we think that this last topic may lay the foundations for expanding our understanding of bivalve sex determination, as our current knowledge is quite fragmented and limited to few species. As a matter of fact, tracing the phylogenetic history and diversity of sex-determination related gene families (such as the Dmrt, Sox and Fox genes) would allow to perform more targeted functional experiments and genomic analyses, but also fostering the possibility of establishing a solid comparative framework.

**Significance.** In this perspective, we provide an examination of the phylogenetic diversity of Dmrt genes, a sex-determination related gene family, to address the importance of bivalves in sex determination studies. By analyzing their taxonomic distribution and sequence diversity, we show how such a comparative study may set a common ground plan to settle down targeted functional experiments and essays. This kind of approach should be applied more extensively in future studies, especially when dealing with understudied organisms.

Bivalves are the second largest clade in molluscs, counting more than 18,000 species (Catalogue of Life, accessed 16/12/2022) distributed at all depths and in all marine environments, as well as in some freshwater habitats. Thanks to their high diversity and peculiar biological features, they have been proposed as promising model organisms for investigating a wide array of biological, ecological, and evolutionary issues, from mitochondrial biology and evolution to the physiological plasticity under fluctuating environmental conditions (**Milani and Ghiselli, 2020; Ghiselli et al., 2021**). In this context, bivalves may serve as a compelling model system to investigate the evolution and characteristics of sex determination (SD) as well, thanks to the diversity of their reproductive modes and genomic features. Nonetheless, this research field has been largely overlooked and many aspects of bivalve reproductive biology remain uncharacterized. In this perspective, we address the topic by first examining the relevant questions that bivalves may help to answer regarding processes and patterns of SD, and then providing a case study in the field of comparative genomics.

## 2.1 Open yet inspiring topics in bivalve sex determination

Despite the socio-economic and scientific importance of bivalves, the knowledge concerning the genetic and molecular bases of their SD system is quite limited and its study has been mostly neglected. Yet, bivalves may constitute a novel model system in SD studies that is as intriguing and valuable as other well-established models, such as vertebrates, insects and plants (**of Sex Consortium et al., 2014**), as they may provide complementary perspectives in many aspects of SD evolutionary studies. Topics such as (i) the transition between environmental and genetic SD, (ii) the evolution of sex chromosomes, (iii) the mito-nuclear interaction, and (iv) the evolution of SD related genes, can largely benefit from the integration with bivalve studies. But many others are likely to emerge as research in the field progresses.

### 2.1.1 Transitions between environmental and genetic sex determination

Clues from several works seem to suggest that both genetic and environmental factors are involved in bivalve SD, thus implying that a mixed system may exist (reviewed in **Breton et al., 2018**). The traditional dichotomy between environmental sex determination (ESD) and genetic

sex determination (GSD) seems inapplicable in most bivalve species, where ESD and GSD rather represent the two ends of a continuum of mixed and plastic conditions. A weak distinction between ESD and GSD is also found in amphibians, reptiles and teleost fish, three clades in which environment-dependent SD has been largely studied. Here, the interaction—or even the transition—between the two sexual systems have been reported in many species, suggesting that sex-determining mechanisms can be extraordinary plastic (**Bachtrog et al., 2014; Capel, 2017**). Adding a representative and diverse group of Lophotrochozoa (Protostomia) to those vertebrate taxa, can widely expand the comparative framework of the investigation, allowing to better understand the evolution of SD as a whole. In bivalves, ESD has been studied mostly in oysters, where hermaphroditic species show an effect of temperature on SD (reviewed in **Breton et al., 2018; Fig. 2.1**). Oysters may indeed constitute a prolific model to examine how the SD pathways are shaped in the presence of different initial triggers and highly dynamic reproductive backgrounds. In fact, various sexual systems can be found in oysters, such as (i) strictly gonochoric population, (ii) the coexistence of simultaneous hermaphroditic with strictly gonochoric individuals in the same population, (iii) the possibility of sex change according to environmental conditions, and (iv) the presence of both parasitic dwarf males and free-living males in the same species (**Collin, 2013**). Consequently, oysters may be extremely useful to understand how epigenetic control is involved in sex change, how gene regulatory networks can sustain the occurrence of different hermaphroditic conditions within gonochoric populations, and whether certain SD systems are more labile than others (**Abbott, 2011**).

### 2.1.2 Evolution of sex chromosomes

So far, heteromorphic sex chromosomes (HeSCs)—i.e., sex chromosomes showing strong morphological differentiation, have never been observed in bivalves (**Breton et al., 2018**), while the first evidence of homomorphic sex chromosomes (HoSCs)—i.e., sex chromosomes showing little or no differentiation, comes from a very recent study on several scallop species, where a non-homologous origin of the SD system has been proposed for different subfamilies (**Han et al., 2022; Fig. 2.1**). Theory predicts that, once originated, sex chromosomes (SCs) will eventually turn into HeSCs, because of the recombination arrest in the sex-determining region (**Bachtrog et al., 2014; Beukeboom and Perrin, 2014; Han et al., 2022**). Nonetheless, HoSCs are much more widespread in the animal kingdom than expected, sometimes also being of ancient age (**Bachtrog et al., 2014; Han et al., 2022**).

Species from the order Pectinida may thus be useful to investigate what determines the long-term maintenance of HoSCs and which genomic architectures and molecular dynamics prevent HeSCs from evolving in bivalves. Additionally, they may be taken as model systems to investigate the origin of SCs in relation to the sexual systems and the route by which molecular pathways have been reprogrammed in the transition between different SD mechanisms (**Han et al., 2022**).

Researchers have been addressing this topic mainly in snakes, ratites and sturgeons (**Bachtrog et al., 2014; Han et al., 2022** and references therein). Though, scallops currently hold the oldest HoSC pairs, which are dated back to about 350 million years. The system is thus of great importance to investigate the role of sex-biased gene expression and selection forces in the long-term stability of SCs (**Han et al., 2022**), as well as the intertwining between SD systems.

### 2.1.3 Mito-nuclear interactions

An additional pivotal topic in bivalve biology, tentatively connected to SD, regards the doubly uniparental inheritance (DUI) of mitochondria, a process in which two highly divergent mitochondrial genomes are transmitted uniparentally through the maternal and paternal lineages, respectively through eggs and sperm. This process, which has been reported in more than a hundred bivalve species from five different orders (**Fig. 2.1; Gusman et al., 2016; Capt et al., 2020**), has been proposed to interact with the major nuclear pathways that primarily establish the sexual identity, in a way that can resemble the cytoplasmatic male sterility (CMS) of plants (**Ghiselli et al., 2013; Breton et al., 2022**). In CMS, specific mitochondrial chimeric open reading frames (ORFs) cause the pollen to be sterile, while certain nuclear loci act in counterbalance to restore male fertility when occurring in the same individual. This Red-Queen scenario, in which balancing selection shapes the evolution of both CMS and restorer-of-fertility genes and keeps the two sexes viable, has been also hypothesized to be acting on bivalve DUI species (**Ghiselli et al., 2013; Xu, Iannello, et al., 2022**), where additional and effectively-transcribed ORFs have been observed in both the male-inherited and female-inherited mitochondrial lineages (**Milani et al., 2013, 2014**).

Clearly, if a functional interplay between DUI and SD in bivalves is proven, this will provide new research questions regarding not only bivalve biology itself but also broader evolutionary topics (e.g., are there any converging trait between DUI and CMS systems? What is the degree

of plasticity of such mitochondria-related SD systems? Are mitochondria-related SD systems more widespread in eukaryotes than currently thought?).

#### 2.1.4 Evolution of sex-determination related genes

Considering this intricate scenario of SD mechanisms and the wide diversity of bivalves, in the last years many differential transcription analyses have been performed on several species with the attempt to identify the most probable sex-determination related genes (SRGs) (e.g., **Milani et al., 2013; Zhang et al., 2014; Chen et al., 2017; Capt et al., 2018; Shi et al., 2018; Fig. 2.1**). Interestingly, certain genes consistently emerged across different bivalve species as being substantially more transcribed in one sex (sex-biased) or exclusively transcribed in one sex (sex-specific), suggesting their potential involvement in the SD pathway. These genes mainly belong to the *dsx* and *mab-3* related transcription factor (Dmrt), *Sry*-related HMG-box (Sox), and forkhead box (Fox) families, which play a role in various developmental processes (including the SD cascade) in most animals (**Marshall Graves and Peichel, 2010; Bachtrog et al., 2014; Beukeboom and Perrin, 2014**). Members of these three gene families are also included in the working model for the SD regulatory network proposed for the Pacific oyster *Crassostrea gigas* by **Zhang et al., 2014**, in which: *CgSoxH* (which belong to the Sox family) promotes male gonad development by activating *CgDsx* (which belong to the Dmrt family) and inhibiting *CgFoxL2* (which belong to the Fox family); *CgFoxL2*, when not inhibited by the pair *CgSoxH/CgDsx*, promotes female gonad development. Similarly, **Han et al., 2022** appointed *FoxL2* as a putative SD gene in the two scallop species *Patinopacten yessoensis* and *Chlamys farreri*. If their pivotal role in SD of bivalves is confirmed, an evolutionary genomic analysis may help in better understanding why members of the above-mentioned gene families appear particularly prone to be recruited in the SD cascade also in distantly related species, as it is observed for *Dmrt1* and *Sox3* homologs in vertebrates (**Marshall Graves and Peichel, 2010; Bachtrog et al., 2014**; and the following section). Furthermore, considering the occurrence of mixed SD systems in bivalves, Dmrt, Sox, and Fox genes may provide new perspectives on the influence of different environmental cues on the molecular evolution of animal SRGs. However, to date, experiments have been limited to molecular cloning, differential transcription, and tissue localization of such genes (**Liang et al., 2019; Sun et al., 2022**), while only a few have directly investigated their biological functions in bivalves, for example through post-transcriptional silencing of target mRNAs [RNAi; **Fig. 2.1**; e.g., **Liang et al., 2019; Wang**

et al., 2020; Sun et al., 2022].

Overall, Dmrt, Sox, and Fox genes are highly interesting targets to be investigated in the framework of bivalve SD and have indeed obtained much more attention than the study of SCs or the role of environmental cues. However, much work is still to be done in order to understand their function in the SD signaling pathway and their evolutionary history.

## 2.2 The case of the Dmrt gene family in bivalves

Among the SRG candidates identified in bivalves, Dmrt genes (named after *doublesex* (*dsx*) from *Drosophila melanogaster* and *male abnormal-3* (*mab-3*) from *Caenorhabditis elegans*) are of particular interest. As a matter of fact, in vertebrates, besides their role in placode neurogenesis and somite patterning (reviewed in Mawaribuchi et al., 2019), Dmrt genes are also involved in the development of male gonads and the maintenance of the testicular function (Sun et al., 2022). Their role in the specification and organization of male sexual characters seems indeed to be common across Metazoa, suggesting that a similar function may have been already present in the Bilateria common ancestor (Kopp, 2012; Beukeboom and Perrin, 2014).

The first attempts to dig inside the phylogenetic history and diversity of bivalve Dmrt genes have been provided by Li et al., 2018 and Evensen et al., 2022: besides retrieving all the canonical genes (i.e., *Dmrt2*, *Dmrt3* and *Dmrt4/5*), their inferences brought to light a monophyletic Dmrt group (named *Dmrt 1-like* (*Dmrt-1L*)) which appears to be private to molluscs and present in several bivalve species. The *Dmrt-1L* monophyletic group is confirmed also when expanding the analysis by mining genomes from a wider range of bivalve taxa (Fig. 2.1; Fig. 2.2A), suggesting that *Dmrt-1L* genes are widespread in bivalves and were likely present in their common ancestor (Evensen et al., 2022). In particular, *Dmrt-1L* genes can be successfully retrieved in species of the orders Mytilida, Ostreida, Pectinida, Unionida, and from *Scapharca broughtonii* (Arcida), while the opposite holds for Venerida, *Sinonovacula constricta* (Adapedonta), and *Dreissena* spp. (Myida; Fig. 2.2B). Clearly, the absence of *Dmrt-1L* genes demands further investigations, as it may derive from errors in genome assembly and annotations.

The present analysis also supports a higher amino acid sequence divergence of the *Dmrt-1L* orthology group with respect to the other Dmrt orthology groups (Fig. 2.1C), which may

be explained by a higher rate of sequence evolution related to their sex-biased expression in certain species (**Zhang et al., 2014; Shi et al., 2015; Li et al., 2018; Evensen et al., 2022**). This is consistent with what has been already observed for the SRGs *Dmrt1* and *dsx* in vertebrates and *Drosophila*, respectively (e.g., **Bewick et al., 2011; Baral et al., 2019**). In fact, sex-biased genes (including SRGs) often tend to evolve faster than unbiased genes at the level of protein sequences, either when considering male-biased (reviewed in **Parsch and Ellegren, 2013; Grath and Parsch, 2016**) or female-biased genes (e.g., **Papa et al., 2017; Ghiselli et al., 2018**). Another possible explanation for the higher amino acid divergence of *Dmrt-1L* genes may lie on their expression breadth, that is, genes with a narrow tissue-specific expression tend to evolve faster than more ubiquitous genes (**Parsch and Ellegren, 2013; Xu, Martelossi, et al., 2022**). As a matter of fact, *Dmrt-1L* genes have been found to be significantly more transcribed in the gonadic tissue (particularly in testes) in *P. yessoensis* (**Li et al., 2018**) and *C. gigas* (**Yue et al., 2021**).

Understanding the role and molecular interactions of *Dmrt-1L* genes in bivalve SD and gonad development would greatly enhance the possibility of outlining the evolutionary causes and consequences of their high amino acid divergence (**Fig. 2.2C**), for example by linking the molecular evolution to the degree of pleiotropy. However, most of our knowledge on *Dmrt-1L* biology is currently limited to the temporal and tissue localization of transcripts in a few species of bivalves (e.g., **Li et al., 2018; Yue et al., 2021**). In fact—apart from the work by **Sun et al., 2022**, which confirmed the role of *Dmrt-1L* in the gonad development of *C. gigas* through non-invasive RNAi and found that the knocked-down phenotype results in size reduction of male gonads—no other experiments intended to elucidate the function of *Dmrt-1L* genes in bivalves have been carried out so far (**Fig. 2.1**). This clearly hinders any possible integration between molecular data with functional assays. If the role of *Dmrt-1L* as major sex determinants was confirmed, bivalves would become an intriguing clade in which investigate why, in Metazoa, certain genes (namely, the Dmrt gene family) appear particularly prone to being recruited at the top of the SD cascade. To date, this phenomenon has been widely examined in vertebrates, where *Dmrt1* genes have independently gained a primary role in male SD in fish, amphibians, and birds, and are considered candidate sex-determining genes also in monotreme mammals (**Marshall Graves and Peichel, 2010; Beukeboom and Perrin, 2014; Mawaribuchi et al., 2019**). Bivalves may provide an alternative evolutionary scenario to study the selective forces and molecular modifications that support Dmrt genes in repeatedly taking over the SD

process. In fact, since *Dmrt-1L* genes seem to be restricted to molluscs (**Fig. 2.2A**), it would be intriguing to clarify if the putative involvement in the SD cascade of extant bivalve species is the result of shared ancestry or convergent evolution, which would establish a study system for the evolution of Dmrt genes parallel to that of vertebrates (see **Capel, 2017**).

Obviously, *Dmrt-1L* should not be expected to be the sole sex-determining gene. In fact, *Fox-L2* has already been appointed as the female sex-determining gene in *P. yessoensis* and *C. farreri* (**Han et al., 2022**). Consequently, we should expect that other primary genetic determinants exist, consistently with the extremely high species diversity of the clade. Thus, bivalves may additionally serve as a valuable model system to study how genes from different families take over the SD cascade and are shaped by selection.

## 2.3 Conclusions: bivalves as new models in the study of sex determination

SD is undoubtedly a fascinating biological and evolutionary topic as much as it is challenging to investigate. Our understanding of the causes and consequences of the SD mechanism diversity strongly relies on the study of different systems and non-model model organisms (**Bachtrög et al., 2014; Milani and Ghiselli, 2020**), which provide the foundation for depicting a comprehensive evolutionary and comparative framework in which new and coherent research perspectives can be grounded.

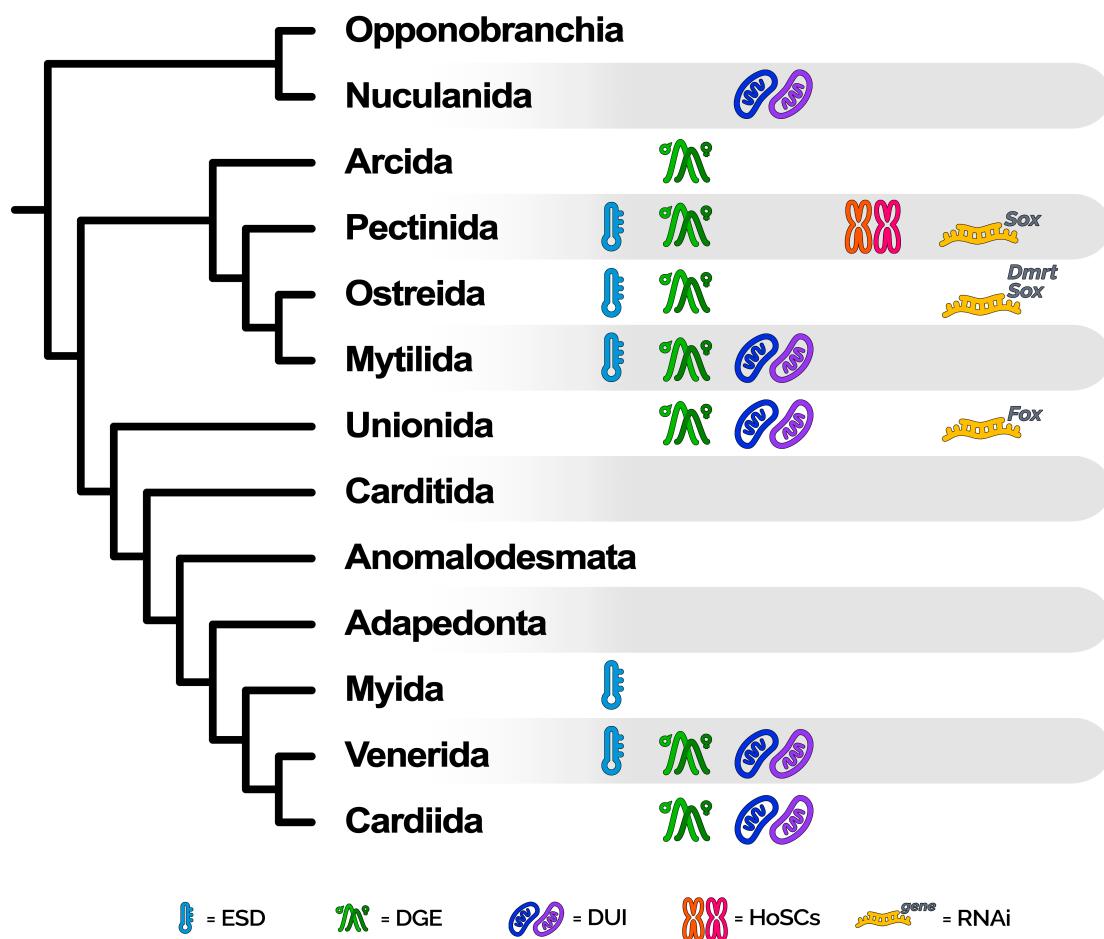
In recent years, bivalves have been achieving growing importance in many fields of biology, from ecology to genomics, and from environmental biomonitoring to mitochondrial studies (**Milani and Ghiselli, 2020; Ghiselli et al., 2021**), but they can be a valuable model to address also SD studies. The diversity of their life history traits provides indeed a challenging, yet extremely fascinating framework, to put the SD processes into an evolutionary context.

Bivalves can help us explain how ESD and GSD interplay with each other in response to the environmental conditions, as a mixed system of both has been proposed to act in the establishment of bivalve sexual identity (reviewed in **Breton et al., 2018**). Moreover, the occurrence of the many existing variants of hermaphroditism and gonochorism even in closely related species, or within the same population, strongly suggests that the basic SD pathway (whether genetic, environmental, or mixed) should be plastic enough to sustain the existence

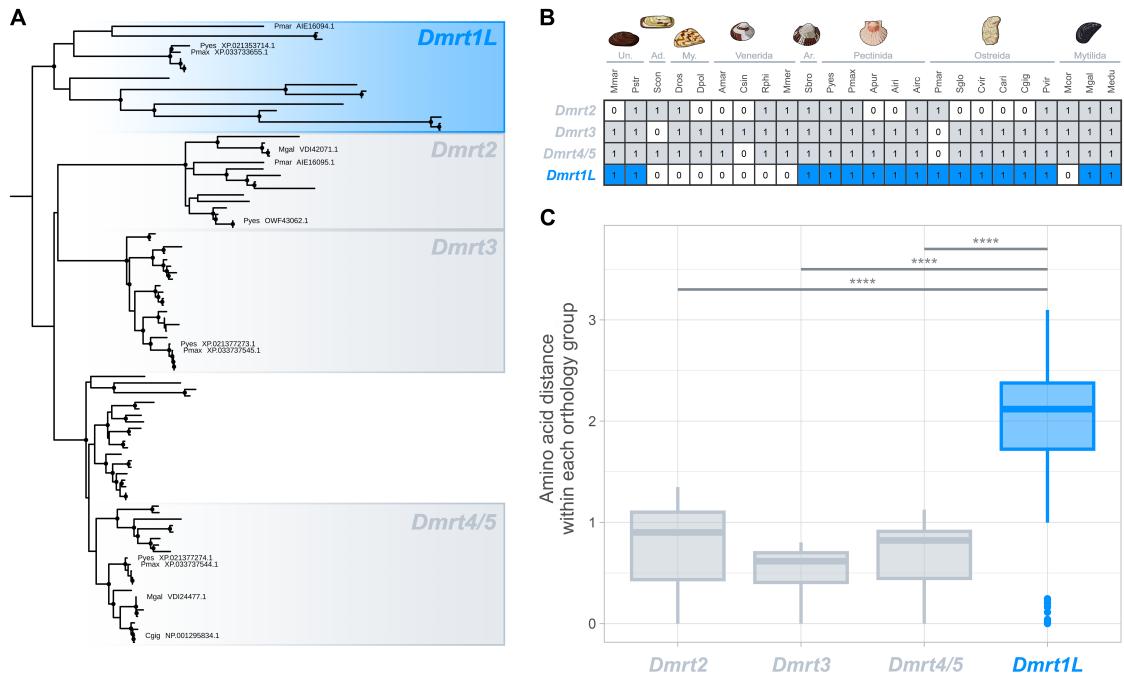
of individuals of both sexes, thus providing the opportunity to study how SD gene regulatory networks are shaped and selected throughout evolution and how epigenetic regulation may influence SD. The unique DUI system further poses an undeniable challenge in SD studies since it may represent an SD-linked mechanism which relies on the non-nuclear portion of the genome and may unfold many new research paths (**Milani and Ghiselli, 2020; Ghiselli et al., 2021**). Nonetheless, much of the research effort on bivalve SD has been devolved to specific groups of socio-economic importance, such as Mytilida, Ostreida, Pectinida, and Unionida, while the other lineages of the bivalve phylogeny have been neglected (**Fig. 2.1**). Our understanding of the SD processes of bivalves is thus restricted and is mainly lacking a broad comparative framework in which to draw comprehensive evolutionary inferences.

Genes from the Dmrt, Sox and Fox families, which are involved in SD also in other Metazoa, may be considered excellent genomic targets to study the processes and patterns of molecular evolution in sex-biased genes, as well as of the recurrent recruitment of genes in the SD cascade. Also, identifying the major genetic regulators of SD in bivalves would burst the functional study of the interaction between ESD and GSD, by providing genetic targets that can be manipulated through RNAi and/or genome editing techniques to understand the role of environmental cues in SD. In the same way, knowing the main genetic actors of SD would allow researcher to identify SCs not only on the basis of in-silico techniques (such as k-mer based or SNP methods) but also by less-expensive wet lab protocols (such as fluorescence mRNA *in-situ* hybridization (mRNA-ISH) on metaphase chromosome plates). Furthermore, it would help to understand whether and how the mitochondrial additional ORFs of DUI species interact with the SD system, by performing thorough gene expression essays.

In conclusion, we strongly urge researchers to invest more resources in the integrative study of bivalve SD to unravel the many underlying mechanisms and expand our understanding of this biological process. Given our limited knowledge in the field, one of the first routes that should be undertaken may rely on the comparative study of SRGs of bivalves from a genomic perspective, as this kind of data is nowadays growing at a rate faster than ever. Establishing such a genomic ground plan for understudied organisms will in fact allow researchers to develop evolutionary-aware experiments with better selected genetic targets.



**Figure 2.1. Graphical summary of the available knowledge and experiments concerning the genetic basis of SD in bivalves, at the level of major taxonomic orders (as reported in WoRMS; accessed before or on 14/03/2023).** For each bivalve clade it is reported: (i) the availability of records of ESD (ii) the availability of differential gene expression (DGE) experiments specifically intended to investigate sex-biased or sex-specific genes; (iii) whether the DUI of mitochondria has been reported in at least one species; (iv) whether HoSCs have been identified in at least one species; (v) the availability of RNA interference (RNAi) experiments for genes belonging to the Dmrt, Sox, and Fox gene families. The phylogenetic tree on the left has been drawn on the basis of the most widely accepted topology for bivalves, according to analyses based on nuclear markers and morphological data. The tips of the tree correspond to major bivalve orders, except for Opponobranchia and Anomalodesmata, which represent higher-level taxonomic ranks. References for the availability of data and experiments can be found throughout the main test.



**Figure 2.2. Phylogenetic tree (A) and taxonomic distribution (B) of Dmrt genes in bivalves, and comparison of amino acid pairwise distances within *Dmrt-1L* and the other Dmrts (C). (A)** Dmrt orthologs from bivalve genome assemblies were obtained with HMMsearch (HMMER toolkit; Eddy, 2011) with the Pfam HMM profile of the DM domain (PF00751). Amino acid alignment was obtained with MAFFT-DASH (Rozewicki et al., 2019), and manually inspected to remove poorly aligning sequences, and trimmed with trimAl (gap threshold of 60%; Capella-Gutiérrez et al., 2009). The phylogenetic analysis was carried out using IQ-TREE 2 (Minh et al., 2020) with default parameters. Nodes with bootstrap values greater than 84 are marked with filled black circles. The tree was rooted according to Evensen et al., 2022. Dmrt genes analysed by Evensen et al., 2022 were used as reference to annotate the various orthology groups, and accession numbers are reported in the tree. The phylogenetic tree with all annotated tips and nodes can be accessed on supplementary material online. (B) Taxonomic distribution of identified Dmrt genes in bivalve genomes. Orders as reported in WoRMS (accessed before or on 14/03/2023) and in Fig. 2.1 are specified. (C) Pairwise amino acid distances were computed for amino acid sequences within each Dmrt orthology group identified in the tree, with the R package ‘phangorn’ (Schliep, 2011) under the JTT substitution model. After checking for normality with the Shapiro-Wilk test ( $W = 0.88544$ , p-value  $\approx 2.2\text{e-}16$ ) and for group effect with the Kruskal-Wallis test (p-value  $\approx 2.2\text{e-}16$ ), the pairwise Wilcoxon rank-sum test was used to compare the distributions of pairwise amino acid distances of *Dmrt-1L* and the other Dmrts. Horizontal bars mark the significative results with  $p \leq 2.2\text{e-}16$  (\*\*\*\*) (Bonferroni correction for multiple test was applied). The list of genome assemblies used for these analyses and species identifiers can be found in Fig. 2.1. Un.: Unionida; Ad.: Adapedonta; My.: Myida; Ar.: Arcida.

**Table 2.1. List of bivalve genomes from which Dmrt genes have been extracted.** For each species, the accepted name and the most-common synonym (in parentheses) are reported. NCBI accession numbers are provided, when available, as well as BUSCO scores of the predicted proteomes against the metazoa\_odb10 dataset (Manni et al., 2021).

Species	ID	Order	Assembly level	BUSCO score	Reference	NCBI Acc. No.
<i>Anadara (Scapharca) broughtonii</i>	Sbro	Arcida	Chromosome	C:91.2% [S:85.6%,D:5.6%] F:2.6% M:6.2%	Bai et al., 2019	NA
<i>Sinonovacula consticta</i>	Scon	Adapedonta	Chromosome	C:92.5% [S:80.4%,D:12.1%] F:3.4% M:4.1%	Ran et al., 2019	GCA_007844125.1
<i>Dreissena polymorpha</i>	Dpol	Myida	Chromosome	C:86.9% [S:75.1%,D:11.8%] F:6.4% M:6.7%	McCartney et al., 2022	GCA_020536995.1
<i>Dreissena rostriformis</i>	Dros	Myida	Scaffold	C:75.2% [S:73.2%,D:2.0%] F:15.2% M:9.6%	Calcino et al., 2019	GCA_007657795.1
<i>Mytilus unguiculatus (coruscus)</i>	Mcor	Mytilida	Chromosome	C:80.0% [S:79.1%,D:0.9%] F:7.7% M:12.3%	Yang et al., 2021	GCA_017311375.1

Tab. 2.1 continued from previous page

Species	ID	Order	Assembly level	BUSCO score	Reference	NCBI Acc. No.
<i>Mytilus edulis</i>	Medu	Mytilida	Scaffold	C:83.7% [S:64.5%,D:19.2%] F:5.2% M:11.1%	Corrochano-Fraile et al., 2022	GCA_905397895.1
<i>Mytilus galloprovincialis</i>	Mgal	Mytilida	Scaffold	C:80.3% [S:47.5%,D:32.8%] F:8.8% M:10.9%	Gerdol et al., 2020	GCA_900618805.1
<i>Perna viridis</i>	Pvir	Mytilida	Scaffold	C:99.4% [S:99.0%,D:0.4%] F:0.2% M:0.4%	Inoue et al., 2021	GCA_018327765.1
<i>Magallana (Crassostrea) ariakensis</i>	Cari	Ostreida	Chromosome	C:94.6% [S:90.9%,D:3.7%] F:0.9% M:4.5%	Li et al., 2021	GCA_020567875.1
<i>Magallana (Crassostrea) gigas</i>	Cgig	Ostreida	Chromosome	C:98.5% [S:67.6%,D:30.9%] F:0.3% M:1.2%	Penaloza et al., 2021	GCF_902806645.1

Tab. 2.1 continued from previous page

Species	ID	Order	Assembly level	BUSCO score	Reference	NCBI Acc. No.
<i>Crassostrea virginica</i>	Cvir	Ostreida	Chromosome	C:98.1% [S:58.6%,D:39.5%] F:0.3% M:1.6%	Gómez-Chiarri et al., 2015	GCF_002022765.2
<i>Saccostrea glomerata</i>	Sglo	Ostreida	Scaffold	C:88.9% [S:85.3%,D:3.6%] F:5.1% M:6.0%	Powell et al., 2018	GCA_003671525.1
<i>Argopecten irradians concentricus</i>	Airc	Pectinida	Scaffold	C:94.8% [S:93.9%,D:0.9%] F:3.7% M:1.5%	Liu et al., 2020	GCA_004382765.1
<i>Argopecten irradians irradians</i>	Airi	Pectinida	Scaffold	C:94.8% [S:93.9%,D:0.9%] F:3.7% M:1.5%	Liu et al., 2020	GCA_004382745.1
<i>Argopecten purpuratus</i>	Apur	Pectinida	Scaffold	C:89.2% [S:88.5%,D:0.7%] F:5.0% M:5.8%	Liu et al., 2020	NA

Tab. 2.1 continued from previous page

Species	ID	Order	Assembly level	BUSCO score	Reference	NCBI Acc. No.
<i>Pecten maximus</i>	Pmax	Pectinida	Chromosome	C:98.5% [S:74.7%,D:23.8%] F:0.4% M:1.1%	Kenny et al., 2020	GCF_902652985.1
<i>Mizuhoppecten (Patinopecten) yessoensis</i>	Pyes	Pectinida	Scaffold	C:98.6% [S:75.2%,D:23.4%] F:0.4% M:1.0%	Wang, Zhang, et al., 2017	GCF_002113885.1
<i>Margaritifera margaritifera</i>	Mmar	Unionida	Scaffold	C:92.6% [S:82.3%,D:10.3%] F:3.2% M:4.2%	Gomes-dos-Santos et al., 2021	GCA_015947965.1
<i>Potamilius streckeroni</i>	Pstr	Unionida	Scaffold	C:74.7% [S:73.8%,D:0.9%] F:7.0% M:18.3%	Smith, 2021	GCA_016746295.1
<i>Calyptogena (Archivesica) marissinica</i>	Amar	Venerida	Chromosome	C:82.0% [S:80.0%,D:2.0%] F:6.1% M:11.9%	Ip et al., 2021	GCA_014843695.1

Tab. 2.1 continued from previous page

Species	ID	Order	Assembly level	BUSCO score	Reference	NCBI Acc. No.
<i>Cyclina sinensis</i>	Csin	Venerida	Scaffold	C:94.0% [S:83.8%,D:10.2%] F:1.9% M:4.1%	Wei et al., 2020	GCA_012932295.1
<i>Mercenaria mercenaria</i>	Mmer	Venerida	Chromosome	C:95.4% [S:70.9%,D:24.5%] F:0.5% M:4.1%	Song et al., 2021	GCF_014805675.1
<i>Ruditapes philippinarum</i>	Rphi	Venerida	Chromosome	C:83.4% [S:74.5%,D:8.9%] F:8.8% M:7.8%	Xu, Martelossi, et al., 2022	GCA_026571515.1

## **2.4 Acknowledgments**

The authors are extremely thankful to Sofía Blanco González from the University of Vigo for her willingness to engage in discussions and for genuinely sharing her opinion on this work.

## **2.5 Data Availability**

Analyzed data and R scripts used to generate plots can be accessed in supplementary material online deposited at the following GitHub repository: [filonico/bivalve\\_sex\\_perspective](https://github.com/filonico/bivalve_sex_perspective).

# Chapter 3

## Identification of putative sex-determination related genes in bivalves through comparative molecular evolutionary analyses

Filippo Nicolini<sup>1,2</sup>, Mariangela Iannello<sup>1</sup>, Giovanni Piccinini<sup>1</sup>, Sergey Nuzhdin<sup>3</sup>, Fabrizio Ghiselli<sup>1</sup>, Andrea Luchetti<sup>1</sup>, Liliana Milani<sup>1</sup>

<sup>1</sup>*Department of Biological, Geological and Environmental Science, University of Bologna, Bologna (BO), Italy.*

<sup>2</sup>*Fano Marine Center, Fano (PU), Italy.*

<sup>3</sup>*Department of Molecular and Computational Biology, University of Southern California, Los Angeles, CA, USA.*

***In preparation.***

### 3.1 Introduction

In sexually reproducing organisms, the modes of sex determination (SD), i.e., the process by which the male or female identity of an organism (or of the gonadic tissue) is established, is highly diverse, ranging from strictly genetic systems to environmentally-dependent processes (**Haag and Doty, 2005; Uller and Helanterä, 2011; Bachtrog et al., 2014; Beukeboom and Perrin, 2014**). Characterising the molecular basis of SD is crucial for understanding not only reproductive biology but also the evolutionary pressures shaping these systems (**Wilkins, 1995; Ellegren and Parsch, 2007; Grath and Parsch, 2016; Nicolini, Ghiselli, et al., 2023**), as sex-determination related genes (SRGs), including primary sex-determining genes (SDGs), are those responsible for the phenotypic differences of males and females, thanks to their sex-biased expression and interactions (**Ellegren and Parsch, 2007; Beukeboom and Perrin, 2014; Grath and Parsch, 2016**). One key aspect of SRGs is that they often exhibit accelerated rates of sequence evolution, due to their involvement in sex-related traits and reproduction. This represents the effects of sexual and/or adaptive selection, which act in sex-biased genes and produce high-divergent proteins at the interspecific level (**Civetta and Singh, 1998; Ellegren and Parsch, 2007; Meisel, 2011; Grath and Parsch, 2016**). Rapid sequence evolution is known for *Sex-determining region of chromosome Y (Sry)* of therians (**Pamilo and O'Neill, 1997; Mawaribuchi et al., 2012**), *dsx* and *mab-3* related gene *W (Dm-W)* of the African clawed frog *Xenopus laevis*, and *dsx* and *mab-3* related gene *Y (Dmy)* of the medaka fish *Oryzias latipes* (**Mawaribuchi et al., 2012**), all of which are master SDGs, that is, genes whose expression is primarily responsible for the establishment of the sexual fate of the organism. Evolution under episodic diversifying selection has been detected also in *Drosophila* for genes involved in the SD cascade [e.g., *Sex-lethal (Sxl)*, *transformer (tra)*, and *doublesex (dsx)*], in correspondence with its establishment in the genus common ancestor (**Mullon et al., 2012; Baral et al., 2019**); though, rapid sequence evolution seems to not be concerning extant amino acid sequences (**Haerty et al., 2007; Baral et al., 2019**), as they are globally evolving under purifying selection, especially in their catalytic domain (**Mullon et al., 2012; Baral et al., 2019**). Concerning the *dsx* genes, higher rates of nucleotide and amino acid sequence evolution can be however observed for male-specific regions, if compared to female-specific and oligomerization regions (**Baral et al., 2019**).

While SD has been extensively studied in model organisms, like mammals, insects, and nematodes, comparatively little is known about the molecular ground plans in non-model or-

ganisms. A remarkable example of this is represented by bivalve molluscs, which exhibit a wide variety of reproductive strategies and sexual systems (**Breton et al., 2018**). Notwithstanding the considerable importance in the human socio-economic landscape (reviewed in **Haszprunar and Wanninger, 2012; Gomes-dos-Santos et al., 2020**), the study of SD mechanisms in bivalves has been hampered by the striking divergence among species (**Li et al., 2022**), and thus largely overlooked and limited to few case studies (**Breton et al., 2018; Nicolini, Ghiselli, et al., 2023**). So far, no master SDG has been unambiguously identified, and the only working hypothesis on the functioning of the SD gene regulatory network is available for the Pacific oyster *Crassostrea gigas* (now *Magallana gigas*; **Zhang et al., 2014**). Nonetheless, the field still lacks both a robust functional investigation and an evolutionary framework in which to place the current knowledge (**Nicolini, Ghiselli, et al., 2023**). As a matter of fact, major efforts have been dedicated to identify sex-biased genes through differential gene expression (DGE) analyses (e.g., **Milani et al., 2013; Teaniniuraitemoana et al., 2014; Zhang et al., 2014; Capt et al., 2018; Afonso et al., 2019**), but very few have leveraged cutting-edge techniques to investigate their actual role in SD and/or gonad differentiation and development (e.g., **Liang et al., 2019; Sun et al., 2022**).

Components of the Dmrt, Sox, and Fox gene (DSFG) families are notoriously known as key actors in several developmental processes across Metazoa (**Benayoun et al., 2011; Matson and Zarkower, 2012; Sarkar and Hochedlinger, 2013; Mawaribuchi et al., 2019**), including SD in certain clades: the aforementioned *Dm-W*, *Dmy*, and *dsx* all belong to the *dsx* and *mab-3* related transcription factor (Dmrt) gene family, while *Sry* belongs to the *Sry*-related HMG-box (Sox) gene family; *Fox-L2*, which takes part in most of the vertebrate SD processes as a downstream effector of the female pathway, belongs to the forkhead box (Fox) gene families. Members of the DSFGs have been identified as putative SRGs also in bivalves, thanks to both DGE analyses and mRNA *in-situ* hybridization (mRNA-ISH) (e.g., **Naimi et al., 2009; Li et al., 2018; Liang et al., 2019; Yue et al., 2021**), suggesting that their role in morphological and sexual development is maintained also in the clade. However, the clear role of DSFGs has yet to be elucidated, probably as a consequence to the lack of (i) a systematic classification of the families and (ii) a comprehensive understanding of their evolutionary history.

In order to overcome such limitations, this study aims to perform a thorough investigation of the DSFG families in bivalves, with the attempt to provide a high-quality resource to be used as a reference for future studies. Through the analysis of more than 40 annotated bivalve genomes

and transcriptomes, we aim (i) to describe the complete set and evolutionary history of DSFGs in bivalves by means of phylogenetic inferences, manual curation, and orthology prediction; furthermore, we aim (ii) to identify DSFGs potentially involved in bivalve SD by investigating their sequence evolution in a genome-wide context. As a matter of fact, our hypothesis is that, if any of the DSFGs is directly involved in SD (i.e., is a SDG), then we should expect it to be experiencing a higher rate of sequence evolution, as already found in previous studies (**Pamilo and O'Neill, 1997; Mawaribuchi et al., 2012**) and discussed earlier; this characteristic, in turn, would be reflected in a high diversity of the extant amino acid sequences across the bivalve clade. To assess the robustness and reliability of our approach, we additionally applied our pipeline to two non-bivalve datasets, composed of mammal and *Drosophila* species, respectively (hereon referred to as the ‘mammal dataset’ and the ‘fruit fly dataset’). By choosing two clades for which SD is well characterised, we wanted to compare our results with those obtained on taxa for which a deeper and detailed knowledge is available. Particularly, mammals and *Drosophila* provide two different frameworks to study the patterns of molecular evolution in SDGs: the former is a system where SD is completely genetic (i.e., the development into a male or into a female is triggered by the up- or downregulation of *Sry* in undifferentiated gonads, respectively), while the latter is a system where SD is chromosomal, thus lacks a master SDG (the sexual fate of the individual is determined by the ratio between autosomal and X chromosomes). Hence, they represent opposing control datasets to be compared to bivalves, as it is expected that a higher rate of sequence evolution concerns only master SDGs (as *Sry* in therians; i.e., the top regulatory part of the SD cascade), but not also the downstream genes (i.e., the bottom effectors). If our method is robust, we should thus expect that, (i) in the mammalian dataset *Sry* is detected as rapidly-evolving, while (ii) in the fruit fly dataset no gene among those working within the sex-determining cascade is evolving at a higher pace. By testing the performance of the pipeline in mammals and fruit flies, we were able to assess the reliability of results in bivalves.

This work offers novel insights into the evolutionary dynamics of SRGs and contributes a valuable genomic resource for understanding SD in bivalves, one of the most ecologically and economically important groups of marine organisms. Particularly, here we provide the first extensive phylogenetic-based classification of DSFGs in bivalves, covering many species from the major bivalve orders, along with a comprehensive investigation of their sequence evolution.

## 3.2 Materials and Methods

### 3.2.1 Dataset of bivalve annotated genomes and transcriptomes

Annotated genome assemblies of bivalves were obtained from various publicly available resources, while reference genome assemblies for gastropods and cephalopods were downloaded from NCBI (**Supp. Tab. S1**). Isoforms were removed from genome annotations using a perl script from the AGAT toolkit (v0.8.0; **Dainat et al., 2022**). Concerning *Sinonovacula constricta* (Adapedonta), the nucleotide coding sequence fasta file was not available for download. To avoid excluding the species from our analyses, the file was generated in-house by mapping the annotated protein sequences on the reference genome using miniprot (v0.13-0; **Li, 2023**). Then, the corresponding nucleotide sequences were extracted using AGAT on the resulting gff annotation file.

In order to provide an extensive identification of SRGs also for underrepresented bivalve orders (mainly belonging to the Heterodonta clade), 14 additional species represented by sequenced transcriptomes were included in the analyses. Assembled and annotated transcriptomes were obtained from **Piccinini et al., 2021** and **Iannello et al., 2023**. Briefly, raw reads were trimmed using Trimmomatic (**Bolger et al., 2014**) and assembled using Trinity (**Grabherr et al., 2011**) with default parameters. Isoforms were removed using the dedicated perl script from the Trinity utilities. Open reading frames were predicted through TransDecoder (**Haas, n.d.**), by also including diamond (**Buchfink et al., 2015**) and HMMER (v3.3.2; <http://hmmer.org/>) annotation of hits.

The resulting set of annotated genomes and transcriptomes (hereafter referred to as the “comprehensive set”) was checked for completeness using BUSCO with the Metazoa reference dataset (v5.2.2; **Manni et al., 2021**).

### 3.2.2 Identification and classification of Dmrt, Sox and Fox genes in bivalves

Members of DSFG families were retrieved in the comprehensive set with hmmsearch from the HMMER package (v3.3.2; <http://hmmer.org/>). The signature catalytic domains of each family were used as queries. Specifically, hidden Markov model (HMM) profiles were built after the Pfam databases for the *dsx* and *mab-3* (DM) domain (PF00751), the high mobility

group (HMG) box (PF00505) and the forkhead domain (PF00250) to retrieve members of the DSFG families, respectively. The e-value for both the per-target and the per-domain inclusion threshold was set to 1.0e–5.

Obtained hits were then annotated using (i) the PANTHER HMM standalone sequence scoring against the PANTHER library v18.0 and (ii) RPS-BLAST (v2.5.0+) against the Conserved Domain Database (CDD; pre-compiled version, downloaded from ftp.ncbi.nih.gov on 09/11/23). In both cases, hits with an e-value of 1.0e–5 were retained. Genes which were correctly annotated by both systems (on the basis of the PANTHER gene family and CDD domain identifiers; **Supp. Tab. S2**) were kept for subsequent analyses.

DSFGs from *Homo sapiens*, *Drosophila melanogaster*, and *Caenorhabditis elegans* (**Supp. Tab. S3**; hereafter referred to as ‘reference species’) were retrieved from NCBI and were used as reference genes for annotation (see below). Classification and nomenclature of each family was retrieved from: **Mawaribuchi et al., 2019** for Dmrt genes; **Phochanukul and Russell, 2010** and **Sarkar and Hochedlinger, 2013** for Sox genes; **Mazet et al., 2003** for Fox genes.

The alignments of mollusc and reference DSFGs were guided by the aforementioned Pfam HMM profiles and performed with Clustal Omega (v1.2.3; **Sievers et al., 2011**), then trimmed with trimAl (v1.4.rev15; **Capella-Gutiérrez et al., 2009**) with a gap threshold of 40%. Resulting alignments were manually inspected to remove sequences with incomplete catalytic domains, then aligned and trimmed again as before. Phylogenetic trees were inferred using IQ-TREE (v2.1.4-beta COVID-edition; **Minh et al., 2020**) with automatic model selection (**Kalyaanamoorthy et al., 2017**), 1000 bootstrap replicates and 5 independent runs. The phylogenetic tree of Dmrt genes was midpoint rooted, as no clear homology relationship has been found with other gene families or zinc-finger proteins so far (**Wexler et al., 2014**). Phylogenetic trees of Sox and Fox gene families were rooted using two fungi mating protein A (Mat-A) sequences (XP\_62685912.1, CCD57795.1) and two Amoebozoa forkhead-like domains (XP\_004368148.1, XP\_004333268.1), respectively (**Nakagawa et al., 2013; Heenan et al., 2016**). The rooting was performed with Gotree (v0.4.5; **Lemoine and Gascuel, 2021**). To identify and annotate bivalve homology groups within each gene family, we employed a species overlap algorithm followed by a Markov clustering algorithm (MCL) weighted by node supports as implemented in Possvm (v1.2; **Grau-Bové and Sebé-Pedrós, 2021**). DSFGs from *H. sapiens*, *D. melanogaster*, and *C. elegans* were used as reference annotation.

In order to better establish the orthology relationships among ambiguous groups of Dmrt and Fox genes, we run a series of other phylogenetic reconstructions (see 3.4), by using the same pipeline as before. In the case of *Fox-Y* genes, we also employed Fox gene sequences from the sea urchin *Strongylocentrotus purpuratus*, as given by **Tu et al., 2006**. All the phylogenetic trees were plotted using the R package ‘ggtree’ (**Yu et al., 2017**).

### 3.2.3 Sequence diversity of bivalve single-copy orthogroups

As a metrics to measure the sequence diversity of bivalve DSFGs, and test whether those putatively involved in SD show higher values than other genes, we employed the amino acid sequence divergence. As a matter of fact, this metric is fast and straightforward to obtain, as it only requires the amino acid alignment and the corresponding best-fit substitution mode.

To this purpose, we produced amino acid alignments of bivalve single-copy orthogroups (SCOs) groups and built the distribution of their median amino acid sequence divergence (AASD). Specifically, we assembled a second dataset (hereafter referred to as the ‘reduced bivalve dataset’) which includes, for each bivalve genus, only the best genomes and transcriptomes in terms of either BUSCO scores (on the metazoan\_odb10 dataset; **Manni et al., 2021**) or assembly statistics (**Supp. Tab. S1**), in order to reduce computational time. *Archivesica marissinica* (now *Calyptogena marissinica*) and *Saccostrea glomerata* were also removed, as their annotated coding sequences contain many stop codons, which prevent accurate amino acid guided alignments. Genes were clustered in orthologous groups using OrthoFinder (v2.5.5; **Emms and Kelly, 2019**) with DIAMOND ultra-sensitive and default parameters. Resulting orthogroups were splitted into SCOs using DISCO (v1.3.1; **Willson et al., 2022**), and orthogroups with at least 17 species (50% of the species included in the bivalve reduced dataset) were retained. Amino acid and nucleotide sequences of SCOs were then aligned using Clustal Omega as implemented in TranslatorX (v1.1; **Abascal et al., 2010**), and jointly trimmed using trimAl with a gap threshold of 40% and the removal of spurious sequences (`-resoverlap 50 -seqoverlap 50`). Eventually, orthogroups containing (i) internal stop codons, (ii) with less than 17 species left (50% of the species included in the bivalve reduced dataset), or (iii) containing DSFGs were removed from downstream analyses. The best amino acid substitution model was inferred for each trimmed alignment using ModelFinder as implemented in IQTREE2 (model search was restricted to matrices accepted by the ‘phangorn’ R library; i.e., Blosum62, cpREV, Dayhoff, DCMut, FLU, HIVb, HIVw, JTT, JTTDCMut,

LG, mtART, mtMAM, mtREV, mtZOA, rtREV, VT, WAG) and the corresponding pairwise amino acid distances were computed with the function ‘dist.ml’ from the ‘phangorn’ R package (**Schliep, 2011**). We decided to employ the pairwise amino acid distance instead of the tip-to-tip phylogenetic distance (which accounts for a more comprehensive evolutionary signal) in order to save computational time. However, to check whether the two metrics were comparable to each other, we randomly selected 200 decomposed orthogroups (including orthogroups from the DSFGs) and computed the maximum likelihood (ML) trees using IQTREE2, with ModelSelection restricted as before. Then, the tip-to-tip pairwise distances were obtained with the R package ‘adephylo’ (**Jombart and Dray, 2010**). The same pipeline was also employed to obtain pairwise amino acid distances for each DSFG single-copy orthologous group.

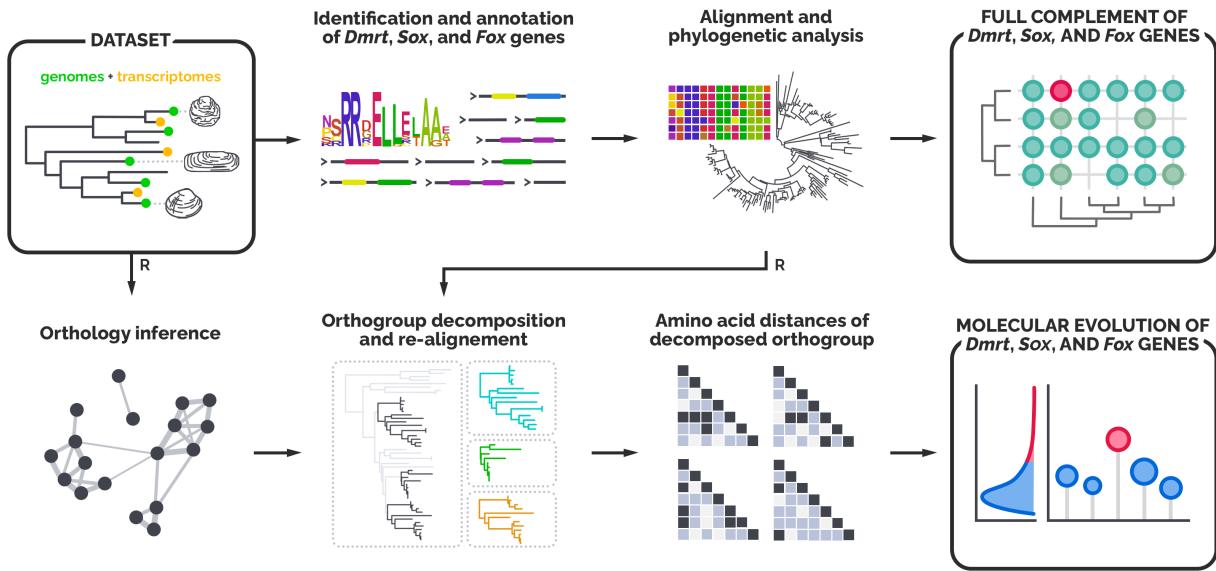
The distribution of amino acid distances was then built after the median values of pairwise distances of each SCO, and genes were categorised accordingly into three groups: Group 1, consisting of genes from the 1% upper quantile of the distribution; Group 2, consisting of genes between the 1% and 5% upper quantiles; and Group 3, consisting of all the remaining genes. Group 1 and Group 2 genes will be referred to as ‘highly divergent genes’.

### 3.2.4 Mammals and *Drosophila* spp. as test datasets

To validate our approach for the study of bivalve SRG molecular evolution, we run the same analysis on two additional datasets, consisting of reference genomes of mammals and *Drosophila* species (**Supp. Tab. S4–S5**, respectively), whose sex-determining mechanisms are well studied and characterised. As a matter of fact, despite it is well known that SDGs tend to evolve faster than genes not involved in SD, the hypothesis has never been tested extensively across the entire phylogenetic diversity of a group: molecular evolution of SDGs and SRGs has mainly been tested on single species or inside the boundaries of taxonomic genera (REFERENCE REFERENCE). For both mammals and fruit flies, annotated genomes were downloaded from NCBI using the command-line tool ‘datasets’, then processed using the same pipeline and scripts as before (**Figure 3.1**).

### 3.2.5 GO-term enrichment

After having obtained the distributions of AASD in the three datasets (Bivalvia, Mammalia, and *Drosophila*) and having sorted SCOs genes up into 3 groups (Group 1, Group 2, and Group 3), we performed a gene ontology (GO) enrichment analysis of genes from Group 1 and genes



**Figure 3.1. Workflow of the analyses for the bivalve dataset.** Starting from a set of both genomes and transcriptomes covering a great portion of bivalve taxonomic diversity, we first characterized the entire complement of glsdsfg genes (upper row). In particular, we used sequence annotation and phylogenetic tools to obtain reliable sequences and filter out any putative mis-assembled or mis-annotated sequence. Afterwards, we built a reduced set of transcriptomes and genomes (the reduced bivalve dataset, where we minimized the redundancy of congeneric species) from which to draw the molecular evolution patterns of orthologous genes (bottom row). In particular, after having obtained gene single-copy orthologous groups, we calculated the amino acid distances within each orthogroup and then we built the distribution of median values. The same pipeline was also employed for the mammal and the fruit fly datasets, with just two minor differences: the starting dataset was composed of only genomes, and that the reduction step (R) was not necessary.

from Group 1 + Group 2. To do so, we firstly selected one gene per SCO, giving priority to few chosen species: (i) for bivalves, we selected genes from *Pecten maximus*, or alternatively from *C. gigas*, *Hyriopsis bialata* (now *Unio delphinus*), *Tridacna squamosa*, and *Solen grandis*; (ii) for mammals, we selected genes from *H. sapiens*, or alternatively from *Bubalus bubalis*, *Panthera tigris*, *Camelus dromedarius*, and *Monodelphis domestica*; (iii) for fruit flies, we selected genes from *D. melanogaster*, or alternatively from *Drosophila hydei*, *Drosophila pseudoobscura*, and *Drosophila suzukii*. By doing so, we ensured that each SCO was represented by one gene. Afterwards, we annotated the obtained datasets with the corresponding GO terms using the OMA browser (accessed 18/09/2024; **Altenhoff et al., 2024**). The GO-term enrichment of Group 1 genes and Group 1 + Group 2 genes was performed with the R package ‘topGO’ with the Fisher exact test (**Alexa and Rahnenführer, 2009**).

## 3.3 Results

### 3.3.1 Genomic and transcriptomic datasets

The complete bivalve dataset consists of 29 bivalve genomes, 14 bivalve transcriptomes, and 7 outgroup genomes (5 gastropods and 2 *Octopus* spp.; **Supp. Tab. S1**). BUSCO statistics for complete single-copy genes spanned from the 64.9% in *Modiolus modiolus* to the 99.4% of *Perna viridis*, with a median value of 94.7%. We were able to get at least one representative species for 11 different bivalve orders, covering a good proportion of the phylogenetic diversity of the clades Pteriomorpha, Palaeoheterodonta, and Imparidentia, and thus building the most extensive genomic and transcriptomic dataset for bivalve comparative analyses so far (**Supp. Tab. S1**). Unfortunately, no genomes or transcriptomes for Protobranchia, Archiheterodonta, and Anomalodesmata were available at the time of the project, thus we were not able to include any of those clades in our analysis. The reduced bivalve dataset (used for the orthology inference and the molecular evolution analysis; **Fig. 3.1**) consists instead of 36 genomes and transcriptomes (**Supp. Tab. S1**), and was built to retain just one species for each taxonomic genera.

The mammal dataset consists of 32 species and 1 outgroup (*Gallus gallus*, Aves; **Supp. Tab. S4**), and covers 12 major orders, while the fruit fly dataset consists of 17 species and 1 outgroup (*Anopheles gambiae*, Culicidae; **Supp. Tab. S5**), and covers 2 *Drosophila* subgenera (i.e., *Drosophila* and *Sophophora*). BUSCO statistics for complete single-copy genes were generally higher than those of bivalves, with a median of 98.3% for mammals and of 99.8% for fruit flies (**Supp. Tab. S4–S5**).

### 3.3.2 The Dmrt, Sox, and Fox complements in bivalves

Our annotation pipeline managed to successfully identify and annotate DSFGs in bivalves, as proved by the same analysis in mammals and fruit flies (see the paragraph **The Dmrt, Sox, and Fox complements and their amino acid divergence in the testing datasets COMPLETARE** COMPLETARE).

We retrieved four main orthology groups of Dmrt genes in bivalves (**Fig. 3.2**; **Supp. Fig. S1**; **Supp. Tab. S6**), three corresponding to the groups present in the Bilateria common ancestor (*Dmrt-2*, *Dmrt-3*, and *Dmrt-4/5*; **Mawaribuchi et al., 2019**), and one additional group with no unambiguous ortholog among reference genes, and thus putatively specific to

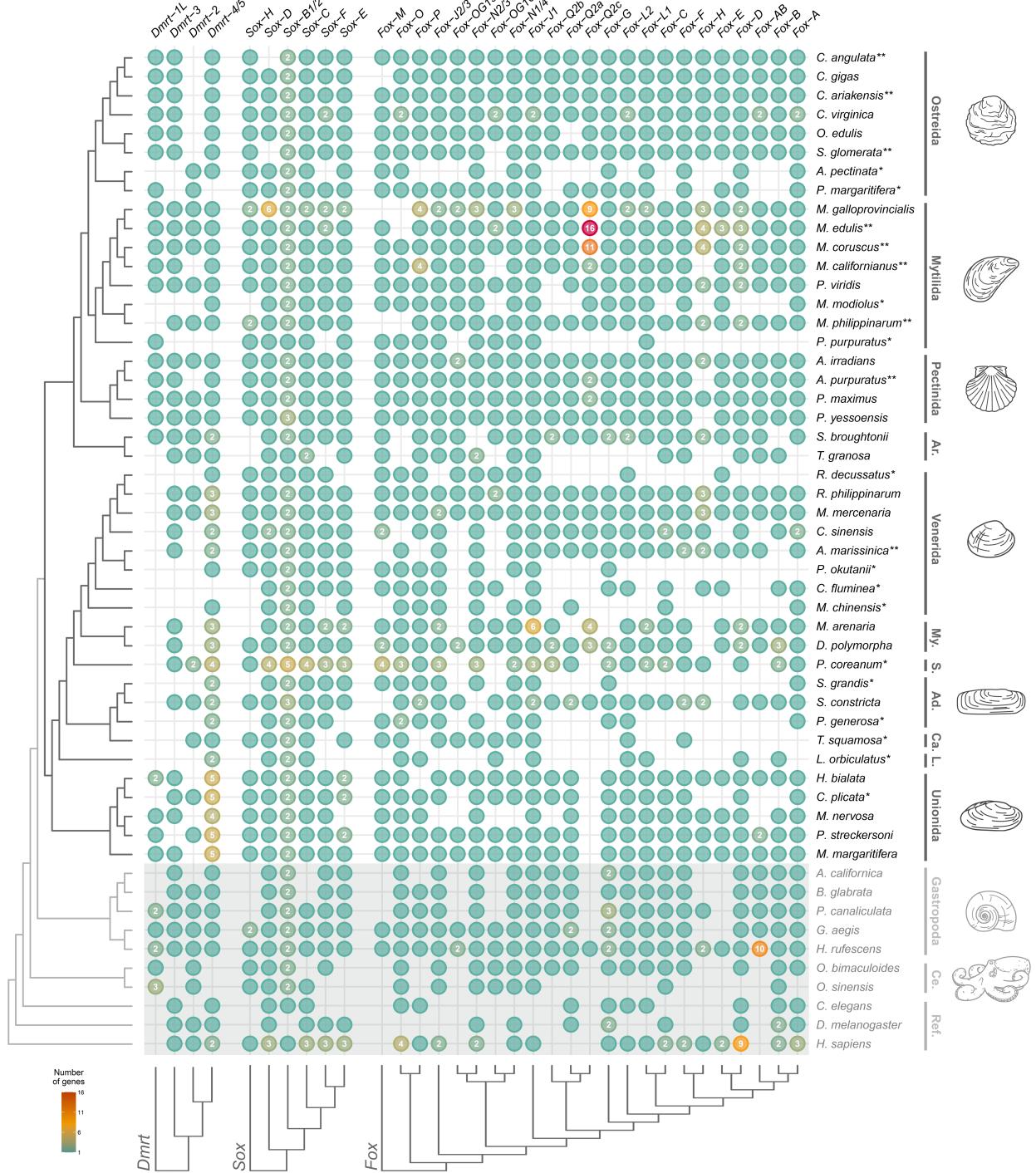
molluscs (named *Dmrt 1-like* (*Dmrt-1L*), as per **Li et al., 2018; Evensen et al., 2022**). The majority of identified *Dmrt* genes are present in single-copy in each species, but *Dmrt-4/5s* show a group-specific expansion in Palaeoheterodonta and Heterodonta, while *Dmrt-1L* is completely absent from Heterodonta. The degree of missing data for *Dmrt* genes in bivalves is about 35%, with *Dmrt-2* having the highest (about 56%) and *Dmrt-4/5* the lowest (about 7%; **Supp. Tab. S7**). The coupling of ubiquitin conjugation to endoplasmic reticulum degradation (CUE)-like DM-associated (DMA) domain has been annotated in most of the *Dmrt-3* and *Dmrt-4/5* genes, while an additional DM domain has been annotated in *Dmrt-1L* genes in Mytilida and the gastropod *Pomacea canaliculata* (**Supp. Tab. S6**). Additionally, we retrieved six main orthology groups of *Sox* genes, none of which is restricted to molluscs or bivalves (**Fig. 3.2; Supp. Fig. S2; Supp. Tab. S6**). Five *Sox* groups (*Sox-B1/2*, *Sox-C*, *Sox-D*, *Sox-E*, and *Sox-F*) are those traditionally considered to be present in the Bilateria common ancestor (**Phochanukul and Russell, 2010**), while one has been identified outside mammals only recently (*Sox-H*, or *Sox-30*; **Han et al., 2010**). *Sox-B2* and *Sox-B1* have been grouped in the same clade, as in our phylogenetic reconstruction the former results in a paraphyletic group with the latter (**Supp. Fig. S2**), despite being traditionally recognised as a separate paralogy group in humans, fruit flies, and nematodes. The degree of missing data for *Sox* genes in bivalves is about 8%, with *Sox-H* having the highest (about 21%) and *Sox-B1/2* and *Sox-C* both having no missing genes (**Supp. Tab. S7**). The *Sox* N-terminal signature domain was annotated for *Sox-E* genes (**Supp. Tab. S6**). Concerning *Fox* genes, we retrieved 27 main orthology groups (**Fig. 3.2; Supp. Fig. S3; Supp. Tab. S6**), two of which are specific to molluscs (*Fox-OG13/NA*, *Fox-OG16/NA*). Additionally, other potential mollusc-specific *Fox* groups have been identified, but these have been excluded from the final orthology analysis as they are present in less than half of bivalve species (see **Materials and Methods** REFERENCE REFERENCE; **Supp. Tab. S6**). The two major *Fox* gene subgroups, Group I (monophyletic, specific to Metazoa; includes *Fox-A*, *Fox-B*, *Fox-C*, *Fox-D*, *Fox-E*, *Fox-F*, *Fox-G*, *Fox-H*, *Fox-L1*, *Fox-L2*, *Fox-Q2*) and Group II (paraphyletic, specific to Opisthokonta; includes *Fox-O*, *Fox-P*, *Fox-J2*, *Fox-J1*, *Fox-K*, *Fox-N2/3*, *Fox-N1/4*; **Larroux et al., 2008**), have been recovered, including the four *Fox* genes that were present in the Bilateria common ancestor (*Fox-C*, *Fox-F*, *Fox-L1*, and *Fox-Q1*; **Shimeld et al., 2010**). Two putative lineage-specific expansions have been recovered for *Fox-OG28/NA*, one regarding *Mytilus* spp. and one regarding the two Myida species (**Fig. 3.2; Supp. Fig. S3**). The degree of missing data for *Fox* genes in bivalves is about 22%, with *Fox-H* having the highest (about 42%) and

*Fox-J1* having no missing genes (**Supp. Tab. S7**). The forkhead-associated (FHA) domain was annotated for *Fox-K* genes, the *Fox-P* coiled-coil signature domain was annotated for *Fox-P* genes, while both the forkhead N- and C-terminal signature domains were annotated for *Fox-A* genes (**Supp. Tab. S6**). Regarding bivalve species, the amount of missing data greatly differs between genomes and transcriptomes, with a mean of about 9% and about 45%, respectively. *Argopecten irradians concentricus*, *Mytilus unguiculatus* (formerly *coruscus*), and *Pecten maximus* have no missing data, while *Loripes orbiculatus* has the highest proportion (about 64%; **Supp. Tab. S7**).

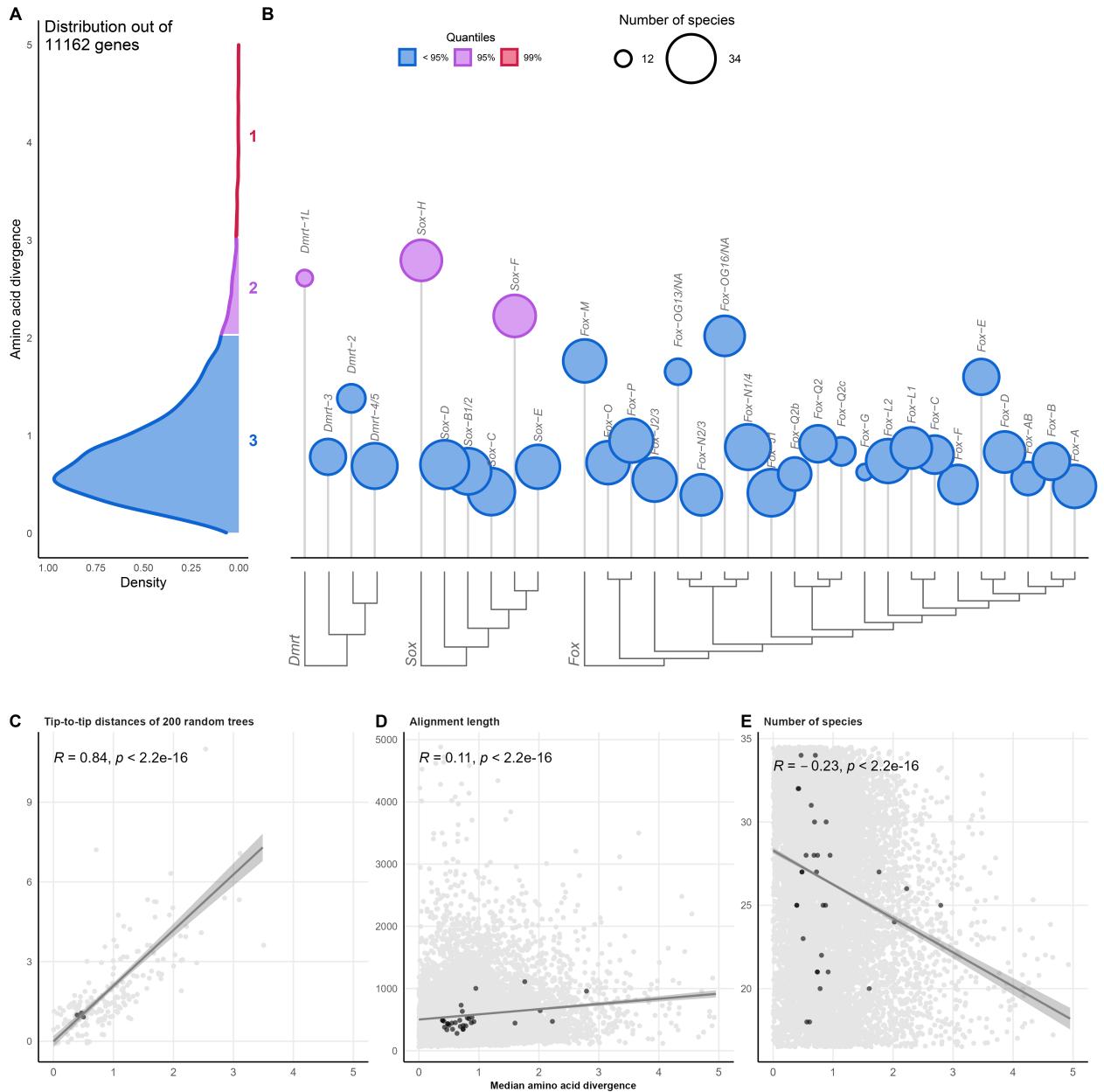
### 3.3.3 Amino acid sequence divergence of Dmrt, Sox, and Fox genes in bivalves

In the reduced bivalve dataset, OrthoFinder collectively analysed >1.2G genes distributed in 34 species. 89.4% of these genes were placed in orthogroups, while 10.6% were not. The number of retrieved SCOs is 5, which is drastically low but can be explained considering the mixed nature of the dataset, that is, it includes both genomes and transcriptomes with highly different BUSCO scores (**Supp. Tab. S1**). In order to be able to analyse a greater number of genes, we decomposed OrthoFinder orthogroups using DISCO and eventually obtained 11k SCOs with at least 50% of the species. By running the same pipeline on DSFGs, we included in the AASD analysis 32 SCOs (**Fig. 3.2**) out of 33 initial Possvm-identified groups (*Fox-H* didn't meet the species occupancy threshold; **Fig. 3.3**).

From the distribution of median AASD, 112 genes were assigned to Group 1 (1% upper quantile), 447 to Group 2 (5% upper quantile), and 10.603 to Group 3. Most of the DSFGs (29/32) fell in Group 3 (**Fig. 3.3**), which means they have a median AASD comparable to the vast majority of other genes in bivalves (median level of the genomes). Just *Dmrt-1L*, *Sox-H*, and *Sox-F* showed higher divergences, and have been accordingly placed in Group 2. Overall, pairwise AASD proved to be a good approximation of the tip-to-tip distances ( $R = 0.84, p < 2.2\text{e-}16$ , calculated on 200 randomly-selected trees; **Fig. 3.3C**), while it showed no influence from the alignment length ( $R = 0.11$ ) or the number of represented species ( $R = -0.23$ ; **Fig. 3.3D-E**). Genes from Group 1 and Group 2 are strongly involved in cellular regulatory processes (such as those related to the metabolism of nucleic acids, proteins, and other macromolecules), but also in development and response to external stimuli, as shown by the GO-term enrichment analysis (**Tab. 3.1**; **Supp. Tab. S10**).



**Figure 3.2. DSFG complement in bivalves and their outgroups.** Presence/absence of genes in various species are indicated by filled circles. Numbers inside each circle specify genes with 2 or more copies. The shaded area highlights non-bivalve species, belonging either to other molluscs or to the references. The phylogenetic tree of analyzed species, as inferred from literature, is shown on the left, while major taxonomic groups are reported on the right. Species represented by transcriptomic data are marked with an asterisk (\*), and species not present in the reduced bivalve dataset are marked with two asterisks (\*\*; see main text and Fig. 3.1); note that the two categories do not overlap. DSFG trees are shown on the bottom (full trees can be found in Supp. Fig. S1–S3). Full species names, along with all assembly and taxonomic information, can be found in Supp. Tab. S1. Ad.: Adapedonta; Ar.: Arcida; Ca.: Cardiida; Ce.: Cephalopoda; L.: Lucinida; My.: Myida; Ref.: reference genes; S.: Sphaeriida.



**Figure 3.3. Distribution of AASD of single-copy orthogroups in bivalves (A), including DSFGs (B), and their correlations with tip-to-tip distances (C), alignment lengths (D), and number of species (E).** The distribution of AASD has been computed on the median values of pairwise distances of >11k SCOs from the reduced bivalve dataset (see main text and Fig. 3.1). Genes have been divided according to their median AASD value into three different groups, which are indicated by different colors and increasing numbers (Groups 1, 2, and 3). Circle heights of DSFGs show the median value of their AASD, while the size indicates the number of represented species. DSFG trees are shown on the bottom (full trees can be found in Supp. Fig. S1–S3). Darker points in C–E indicate DSFG SCOs. The correlation between the amino acid distance and the tip-to-tip distance has been computed on 200 randomly-selected orthogroups.

**Table 3.1. Top enriched GO terms for Group 1 and Group 2 genes of bivalves, mammals, and *Drosophila*.** The extended version of the table, which includes also the expected number of annotated genes per GO term and all the other enriched GO terms, can be accessed in Supp. Tab. S10.

Dataset	GO.ID	Term	Annotated genes		Significant genes	Corrected p-value
			Annotated genes	Significant genes		
Bivalvia	GO:0060255	regulation of macromolecule metabolic process	737	59	0.04525	
	GO:0080090	regulation of primary metabolic process	673	53	0.01818	
	GO:0019219	regulation of nucleobase-containing compound metabolic process	541	41	0.02388	
	GO:0006351	DNA-templated transcription	571	39	0.03767	
	GO:0032774	RNA biosynthetic process	579	39	0.04490	
	GO:0051252	regulation of RNA metabolic process	517	37	0.02719	
	GO:0006355	regulation of DNA-templated transcription	490	35	0.03751	
	GO:2001141	regulation of RNA biosynthetic process	491	35	0.03844	
	GO:0006950	response to stress	370	33	0.01949	
	GO:0032502	developmental process	261	27	0.04445	
<i>Drosophila</i>	GO:0006468	protein phosphorylation	345	23	0.02483	
	GO:0031325	positive regulation of cellular metabolic process	125	17	0.00801	
	GO:0010604	positive regulation of macromolecule metabolic process	151	17	0.04047	
	GO:0051172	negative regulation of nitrogen compound metabolic process	117	16	0.00814	
	GO:0051173	positive regulation of nitrogen compound metabolic process	137	15	0.02454	
	GO:0006310	DNA recombination	66	14	0.00087	
	GO:0048513	animal organ development	83	12	0.04088	
	GO:0010629	negative regulation of gene expression	78	11	0.00048	
	GO:0023051	regulation of signaling	133	11	0.02872	
	GO:0045934	negative regulation of nucleobase-containing compound metabolic process	64	11	0.03637	
Mammals	GO:0009605	response to external stimulus	90	11	0.04544	
	GO:0006355	regulation of macromolecule metabolic process	737	59	0.04525	
	GO:0080090	regulation of primary metabolic process	673	53	0.01818	
	GO:0019219	regulation of nucleobase-containing compound metabolic process	541	41	0.02388	
	GO:0006351	DNA-templated transcription	571	39	0.03767	
	GO:0032774	RNA biosynthetic process	579	39	0.04490	
	GO:0051252	regulation of RNA metabolic process	517	37	0.02719	
	GO:0006355	regulation of DNA-templated transcription	490	35	0.03751	
	GO:2001141	regulation of RNA biosynthetic process	491	35	0.03844	
	GO:0006950	response to stress	370	33	0.01949	
<i>Drosophila</i>	GO:0031325	positive regulation of cellular metabolic process	125	17	0.00801	
	GO:0010604	positive regulation of macromolecule metabolic process	151	17	0.04047	
	GO:0051172	negative regulation of nitrogen compound metabolic process	117	16	0.00814	
	GO:0051173	positive regulation of nitrogen compound metabolic process	137	15	0.02454	
	GO:0006310	DNA recombination	66	14	0.00087	
	GO:0048513	animal organ development	83	12	0.04088	
	GO:0010629	negative regulation of gene expression	78	11	0.00048	
	GO:0023051	regulation of signaling	133	11	0.02872	
	GO:0045934	negative regulation of nucleobase-containing compound metabolic process	64	11	0.03637	
	GO:0009605	response to external stimulus	90	11	0.04544	

Tab. 3.1 continued from previous page

Dataset	GO.ID	Term	Annotated genes	Significant genes	Corrected p-value
Bivalvia	GO:0044419	biological process involved in interspecies interaction between organisms	63	11	0.04761
	GO:0006955	immune response	1297	145	0.00061
	GO:0098542	defense response to other organism	853	112	0.02066
	GO:0045087	innate immune response	647	82	8.5e-10
	GO:0001817	regulation of cytokine production	630	51	0.04660
	GO:0042742	defense response to bacterium	233	45	1.7e-07
	GO:0006954	inflammatory response	642	45	0.01735
	GO:0019221	cytokine-mediated signaling pathway	382	44	3.9e-07
	GO:0002250	adaptive immune response	342	44	1.3e-05
	GO:0001819	positive regulation of cytokine production	402	41	0.02723
	GO:0002697	regulation of immune effector process	308	37	0.04426
	GO:0042110	T cell activation	432	35	0.02564
	GO:0051607	defense response to virus	257	34	1.9e-07
	GO:0048232	male gamete generation	491	32	0.02255
	GO:0007283	spermatogenesis	478	31	0.02801
	GO:0070661	leukocyte proliferation	273	29	0.01285
	GO:0002449	lymphocyte mediated immunity	221	29	0.04833
	GO:0070663	regulation of leukocyte proliferation	212	25	0.01870
	GO:0050727	regulation of inflammatory response	300	24	0.00235
	GO:0031349	positive regulation of defense response	240	24	0.01239
	GO:0002768	immune response-regulating cell surface receptor signaling pathway	177	22	0.00336
	GO:0050829	defense response to Gram-negative bacterium	66	17	1.7e-10
	GO:0071222	cellular response to lipopolysaccharide	164	17	0.00012

Tab. 3.1 continued from previous page

Dataset	GO.ID	Term	Annotated genes		Significant genes	Corrected p-value
			Annotated genes	Significant genes		
Mammalia	GO:0010466	negative regulation of peptidase activity	163	16	0.00036	
	GO:0002429	immune response-activating cell surface receptor signaling pathway	164	16	0.00243	
	GO:1903555	regulation of tumor necrosis factor superfamily cytokine production	137	16	0.01244	
	GO:0071706	tumor necrosis factor superfamily cytokine production	137	16	0.01244	
	GO:0070665	positive regulation of leukocyte proliferation	132	16	0.02765	
	GO:0045089	positive regulation of innate immune response	113	16	0.03224	
	GO:0071356	cellular response to tumor necrosis factor	175	15	0.00219	
	GO:0002695	negative regulation of leukocyte activation	148	15	0.01151	
	GO:0002456	T cell mediated immunity	82	15	0.01605	
	GO:0002705	positive regulation of leukocyte mediated immunity	113	15	0.01837	
<i>Drosophila</i>	GO:0032680	regulation of tumor necrosis factor production	133	15	0.03262	
	GO:0032640	tumor necrosis factor production	133	15	0.03262	
	GO:0050866	negative regulation of cell activation	165	15	0.04048	
	GO:0000819	sister chromatid segregation	140	11	0.02927	
	GO:0070192	chromosome organization involved in meiotic cell cycle	54	9	0.00849	
	GO:0007131	reciprocal meiotic recombination	37	7	0.00066	
	GO:0007143	female meiotic nuclear division	54	6	0.02270	
	GO:0035967	cellular response to topologically incorrect protein	44	5	0.03334	
	GO:0035966	response to topologically incorrect protein	47	5	0.04266	
	GO:0007141	male meiosis I	13	4	0.00150	
	GO:0140543	positive regulation of piRNA transcription	3	3	6.9e-05	
	GO:0010526	retrotransposon silencing	8	3	0.00331	
	GO:0007130	synaptonemal complex assembly	10	3	0.00666	

Tab. 3.1 continued from previous page

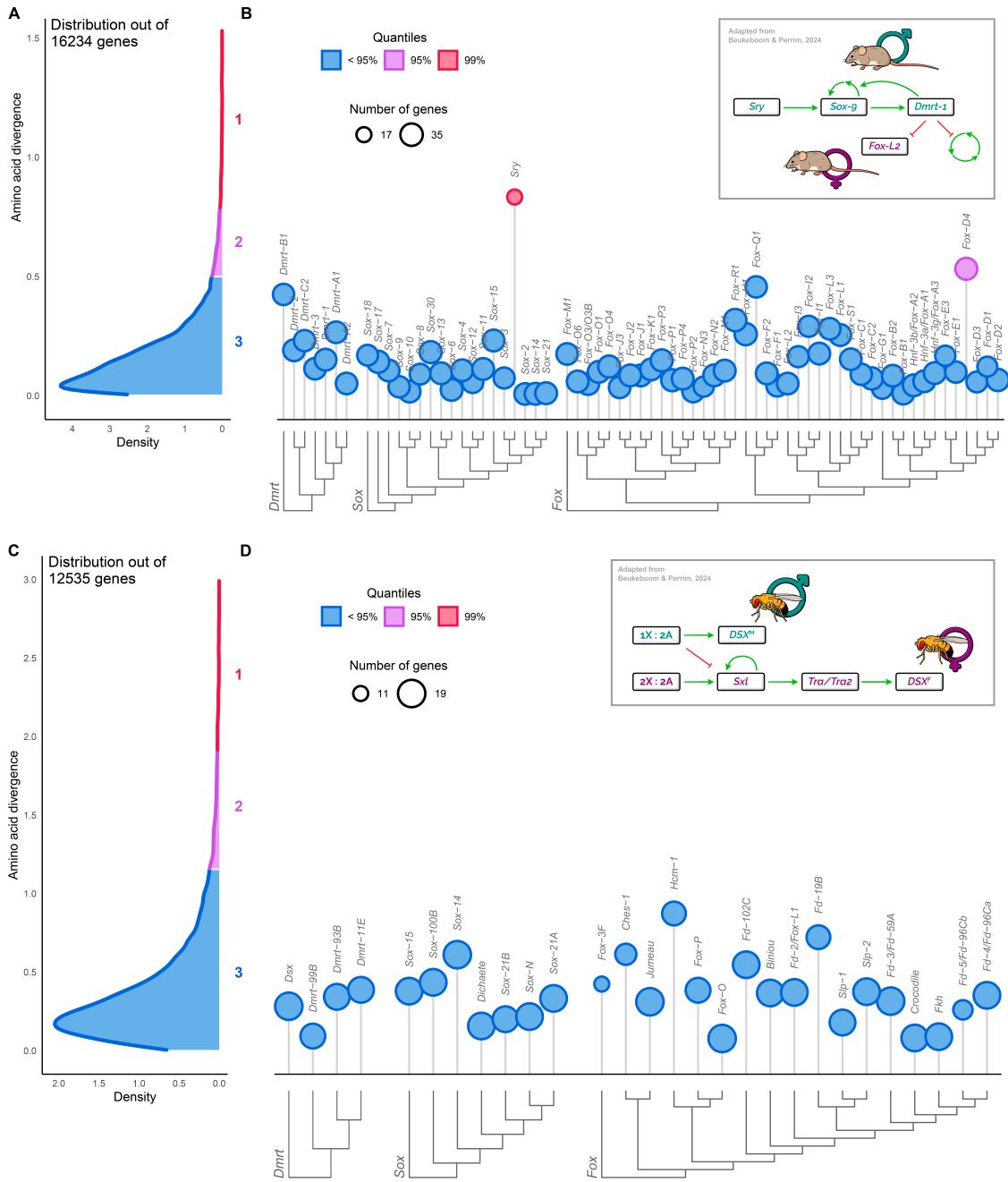
Dataset	GO.ID	Term	Annotated genes		Significant genes	Corrected p-value
			Annotated genes	Significant genes		
<i>Drosophila</i>	GO:0030719	P granule organization	11	3	0.00888	
	GO:0071218	cellular response to misfolded protein	12	3	0.01149	
	GO:0051788	response to misfolded protein	12	3	0.01149	
	GO:0007135	meiosis II	15	3	0.02169	
	GO:0034508	centromere complex assembly	19	3	0.04094	

### 3.3.4 Dmrt, Sox, and Fox genes, and amino acid sequence divergence in the test datasets

The DSFG datasets retrieved in mammals and fruit flies are far more complete than those in bivalves, and most of the already-recognised orthology groups have been identified.

In mammals, we retrieved 7 Dmrt orthology groups with about 3.1% of missing data, 20 Sox orthology groups with about 8.1% of missing data, and 42 Fox orthology groups with about 4.6% of missing data (**Supp. Fig. S4A, S5–S7; Supp. Tab. S8**). Of these, just *Sox-5* was not included in the subsequent AASD analysis, as it did not meet the 50%-species occupancy threshold. OrthoFinder analysed about 650M genes, and the number of SCOs used in the AASD analysis (thus resulting from the DISCO-based orthogroup decomposition pipeline) is >16k (**Fig. 3.4A**). From the distribution of median AASD, 163 genes were assigned to Group 1, 649 to Group 2, and 15.355 to Group 3. Most of the DSFGs (66/68) fell in Group 3 (**Fig. 3.4B**), while *Sry* and *Fox-D4* showed higher divergences, and have been accordingly placed in Group 1 and 2, respectively. Genes from Group 1 and Group 2 show a strong enrichment in immune-related functions (such as innate and adaptive immune response, defence response to bacteria and viruses, lymphocyte methabolism, etc.), but also in reproductive processes (such as spermatogenesis; **Tab. 3.1; Supp. Tab. S10**).

Concerning *Drosophila*, we retrieved 4 Dmrt orthology groups with about 1.7% of missing data, 7 Sox orthology groups with about 3.9% of missing data, and 17 Fox genes with about 8.3% of missing data (**Supp. Fig. S4B, S8–S10; Supp. Tab. S9**). OrthoFinder analysed about 240M, and the distribution of median AASD was built after >12k SCOS (**Fig. 3.4C**). 126 genes were assigned to Group 1, 501 to Group 2, and 11.880 to Group 3. All of the DSFGs have been used in the AASD analysis, but none of them have been placed in Group 1 or 2, that is, all the DSFGs in *Drosophila* have an AASD comparable to the median level of the genome (**Fig. 3.4D**). Genes of Group 1 and Group 2 show a GO-term enrichment in meiotic processes, such as chromosome/chromatid organisation, and retrotransposon silencing (**Tab. 3.1; Supp. Tab. S10**).



**Figure 3.4. Distribution of amino acid divergence (AASD) of single-copy orthogroups in Mammalia (A) and *Drosophila* (C), including Dmrt, Sox, and Fox genes (DSFGs; B-D).** The distributions of AASD in mammals and fruit flies have been computed on the median values of pairwise distances of over 16k and 12k SCOs, respectively. Genes have been divided according to their median AASD value into three different groups, which are indicated by different colors and increasing numbers (Groups 1, 2, and 3). Circle heights of DSFGs show the median value of their AASD, while the size indicates the number of represented species. DSFG trees are shown on the bottom (full trees can be found in **Supp. Fig. S5–S7** for mammals and in **Supp. Fig. S8–S10** for fruit flies). Insets: scheme of the sex-determination molecular pathways in *Mus musculus* and in *Drosophila melanogaster*, with shown the main genes involved (adapted from **Beukeboom and Perrin, 2014**). Green arrows indicate transcription activations, red arrows indicate transcription suppressions. X: sex chromosomes; A: autosomal chromosomes; *DSX<sup>M/F</sup>*: *DSX* splicing variants present in males or females, respectively.

## 3.4 Discussion

### 3.4.1 A new manually-curated and phylogenetic-based reference dataset of Dmrt, Sox, and Fox genes in bivalves

The annotation and characterisation process of a gene family in a certain clade of organisms may harbour many overlooked challenges (**Vizueta Moraga et al., 2020**). For example, the presence of highly-conserved catalytic domains may hamper the correct identification of the components of a gene family because of insufficient phylogenetic signal, as it is the case for Hox and ParaHox genes and their homeobox motif (**Baldwin-Brown et al., 2018; Nicolini, Martelossi, et al., 2023**). Conversely, the components of dynamic gene families characterised by abrupt and sequential duplication events may be difficult to sort into separate groups. As a matter of fact, varying levels of sequence heterogeneity and gene copy numbers makes the inference of orthologous groups hard, as for certain clans of the P450 gene family (**Dermauw et al., 2020**). Regardless of the causes, having a solid and wide phylogenetic context in which to study gene duplications and losses, and orthology relationships, is crucial to overcome these difficulties. In the same way, manual curation and visual inspection of multiple sequence alignments, phylogenetic trees, and gene structures (in terms of domain annotation, start and stop codons, and other feature representations) is helpful, despite being time-demanding and possibly low reproducible. In this study, we characterised the full complement of DSFGs in the vast class of bivalves, by leveraging sequence domain annotation, phylogenetics, and manual curation of the dataset. Our aim was to obtain the most reliable gene complements as possible, combined with a vast taxonomic dataset, a solid phylogenetic inference, an openly-available dataset of gene sequences, and a reproducible pipeline for the annotation of gene identity. By doing so, we want to provide a reliable resource for future studies of DSFGs, either focused on bivalves or generally in Metazoa.

Concerning the Dmrt gene family, we identified orthologs of the vertebrate *Dmrt-2*, *Dmrt-3*, and *Dmrt-4/5* (or *A1/A2*; **3.2; Supp. Fig. S1; Supp. Tab. S6**), which are also expected to have been present in the Bilateria common ancestor (**Mawaribuchi et al., 2019**). **Wang et al., 2023** found that *Dmrt-4/5* is duplicated in *Mercenaria mercenaria* and *Cyclina sinensis* (Venerida), and in *Dreissena polymorpha* (Myida), and we confirm this result by tracing back the duplication event to the split between Palaeoheterodonta (here represented by Unionida) and Heterodonta (here represented by Venerida, Myida, Sphaeriida, Adapedonta,

Cardiida, and Lucinida; **Fig. 3.2**). Furthermore, we confirm *Dmrt-1L* to be present in many bivalve species (mainly belonging to the Ostreida, Pectinida, Mytilida, and Unionida orders; **Fig. 3.2**), as well as in gastropods and *Octopus*. Though, our phylogenetic analysis did not retrieve any unambiguous orthology relationship among *Dmrt-1L* and either vertebrate *Dmrt-1* or *Drosophila dsx* genes, as instead it was proposed in previous works (**Li et al., 2018; Evensen et al., 2022**). As a matter of fact, the amino acid sequence of the *Dmrt-1L* DM domain does not recall that of any other *Dmrt* gene. Furthermore, it must be considered that various phylogenetic analyses have recovered both *Dmrt-1* and *dsx* genes to be restricted to vertebrates and arthropods, respectively (**Wexler et al., 2014; Mawaribuchi et al., 2019; Panara et al., 2019**), that is, they do not have any direct ortholog outside their relative clades. Thus, if *Dmrt-1L*, *dsx*, and *Dmrt-1* are true orthologs, their origin would need to be placed at least in the Bilateria common ancestor, which seems however to be not the case. All considered, we thus confirm that *Dmrt-1L* is not orthologous to *Dmrt-1* and *dsx* and is rather a mollusc-specific gene (**Evensen et al., 2022**). The monophyly of the group is not supported by the phylogenetic tree inferred with *Dmrt* genes from molluscs and the reference species (**Supp. Fig. S1**); though, it is recovered when analysing just genes from mollusc species (**Supp Fig. S11**). To this regard, we speculate that in our analysis, the difficulty in obtaining the monophyly of *Dmrt-1L* genes may have arisen primarily because of the many *C. elegans*-restricted genes (**Supp. Tab. S3**), which are placed among the other bivalve genes (**Supp. Fig. S1**), but also because of the high AASD of *Dmrt-1L* genes (see the following section), which hampers a straight-forward phylogenetic reconstruction. Furthermore, our broad-context analysis allowed us to identify some cases of incorrect gene identification in bivalves, which have arisen because of erroneous or ambiguous annotations in previous works, as a result of limited datasets or analyses. For example, (i) the scallop-specific cluster of *Dmrt* genes retrieved by **Wang et al., 2023** rather belongs to the *Dmrt-1L* group, and (ii) the classification of *Dmrt* genes in *Crassostrea* species provided by **Zeng et al., 2024** needs to be revised following the one of this work: *Dmrt-1* genes are *Dmrt-4/5*; *Dmrt-2* genes are *Dmrt-3*; *Dmrt-3* genes are *Dmrt-1L*; hence, *Crassostrea* species do not have *Dmrt-2* genes.

For what concerns the *Sox* gene family, bivalves (or molluscs) do not show any major clade-restricted gene, as only the five Bilateria-specific *Sox* groups (*Sox-B1/2*, *Sox-C*, *Sox-D*, *Sox-E*, and *Sox-F*) and *Sox-H* have been identified (**Fig. 3.2; Supp. Fig. S2; Supp. Tab. S6**), in accordance with previous findings (**Evensen et al., 2022; Wang and Nie, 2024; Yu**

et al., 2017). *Sox-B1/2* is clearly made up of two subgroups (i.e., *Sox-B1* and *Sox-B2*), as expected, but their respective identity could not be unambiguously established, as *Sox-B1/2* genes of reference species do not form separate clusters (Supp. Fig. S2). Even when inferring the phylogenetic tree only of components of the *Sox-B1/2* group from molluscs and reference species, the identity can not be properly established (Supp. Fig. S12).

Compared to Dmrt and Sox genes, the Fox gene family appears as the most dynamic in terms of gene presence/absence, as already shown by other works (Wu et al., 2020; Schomburg et al., 2022; Seudre et al., 2022). Our phylogenetic analysis successfully recovered Group I and Group II of Fox genes (Larroux et al., 2008), which include the four Fox genes that were present in the Bilateria common ancestor (*Fox-C*, *Fox-F*, *Fox-L1*, and *Fox-Q1*; Fig. 3.2; Supp. Fig. S3; Supp. Tab. S6; Shimeld et al., 2010). To our knowledge, this is the first broad-taxonomic identification and classification of Fox genes in bivalves, as up to now they have been systematically characterised only in *C. gigas* (Yang et al., 2014), *Patinopacten yessoensis* (now *Mizuhopecten yessoensis*; Wu et al., 2020), and *Ruditapes philippinarum* (Liu et al., 2024). Firstly, our analysis confirms the absence in molluscs of *Fox-I*, *Fox-Q1*, *Fox-R*, *Fox-S* (Supp. Fig. S3), which are in fact thought to have emerged with the diversification of deuterostomes or vertebrates (Yang et al., 2014; Wu et al., 2020; Schomburg et al., 2022; Seudre et al., 2022). Furthermore, we have found many Fox groups that appeared as mollusc-specific and/or still-unnamed at a first analysis. However, a more in-depth investigation revealed a different scenario. *Fox-OG2/NA* appears close to the human *Fox-M* gene in the phylogenetic tree, but they do not form a monophyletic group (Supp. Fig. S3). However, by comparing *Fox-OG2/NA* sequences and phylogenetic tree with those analysed by Yang et al., 2014, Wu et al., 2020, Schomburg et al., 2022, and Seudre et al., 2022, it appears clear that this group of Fox genes is indeed **Fox-M**. However, our analysis has failed to retrieve a monophyletic relationship among bivalve and human *Fox-M* genes, even when inferring a tree with just *Fox-J2*, *Fox-M*, *Fox-O*, and *Fox-P* complements (Supp. Fig. S13), which belong to the same Fox group. Regarding the *Fox-OG39/NA* group, it does not have any homolog in reference species (Supp. Fig. S3) but is found to belong to the *Fox-AB* group by sequence comparison with previous works (Yang et al., 2014; Wu et al., 2020; Seudre et al., 2022). *Fox-AB* was formerly described only in the sea urchin *S. purpuratus* and the lancelet *Branchiostoma floridae* (Tu et al., 2006; Yu et al., 2008), but was later identified also in several Spiralia lineages, including molluscs (e.g., Yang et al., 2014; Wu et al., 2020;

**Seudre et al., 2022**). A similar situation concerns *Fox-OG15/NA* and *Fox-OG28/NA*, which again could not be named based on orthology relationships with the reference species genes (**Supp. Fig. S3**), but actually represent two lineage-specific expansions of the *Fox-Q2* group (named *Fox-Q2b* and *Fox-Q2c*), as already appointed in previous studies (**Yang et al., 2014; Wu et al., 2020**). This observation fits within the wider context of the *Fox-Q2* group expansion in Bilateria and, particularly, in Spiralia, that led to remarkable differences in their gene copy numbers across various clades (**Seudre et al., 2022**). Two additional Fox genes have been previously identified in bivalves, and were named *Sox-Y* and *Sox-Z* (**Yang et al., 2014; Wu et al., 2020**). In our analysis, these Fox groups were identified as *Fox-OG13/NA* and *Fox-OG16/NA*, after sequence comparison of Fox genes from *C. gigas* and *P. yessoensis*. On one hand, *Fox-Y* was firstly identified in *S. purpuratus* (**Tu et al., 2006**) and only recently in a few bivalve species (**Yang et al., 2014; Wu et al., 2020**). However, when analysing bivalve and *S. purpuratus* Fox genes, we failed in retrieving such a clear orthology relationship, as *S. purpuratus* *Fox-Y* does not fall within the phylogenetic range of bivalve *Fox-OG13/NA*, which contains the supposed *Fox-Y* orthologs (**Supp. Fig. S14**). Also, the forkhead domains of *Fox-OG13/NA* genes were annotated as ‘forkhead domain P’ (**Supp. Tab S6**). On the other hand, *Fox-Z* was firstly identified in bivalves and in several other protostomes, thanks to a phylogenetic work including the brachiopod *Lingula unguis*, the annelid *Capitella teleta*, the scorpion *Centruroides sculpturatus*, and the centipede *Strigamia maritima* (**Wu et al., 2020**). However, later works have not recovered this Fox gene, even when analysing annelids (**Seudre et al., 2022**) and panarthropods (**Schomburg et al., 2022**) in a more focused effort. In this case, the forkhead domains were annotated as either a generic ‘forkhead domain’ or a ‘forkhead domain Q2’ (**Supp. Tab. S6**). All considered, we argue that bivalves possess two additional Fox groups (here *Fox-OG13/NA* and *Fox-OG16/NA*; **Fig. 3.2**; **Supp. Fig. S3**; **Supp. Tab. S6**) which are shared with other mollusc species, as revealed also by other authors. However, given the discordant results of the phylogenetic hypothesis and domain annotation, we think that a more thorough investigation on their orthology relationships with Fox genes from other Metazoa is needed, and thus we chose to not employ their former names *Fox-Y* and *Fox-Z*.

Besides the DSFG groups discussed so far, it must be also considered that many orphan genes have been identified (**Supp. Fig. S1–S3**; **Supp. Tab. S6**). For example, **Wu et al., 2020** identified a duplication event of *Fox-H* genes in *C. gigas*, which has been recovered also in our analysis for the entire Ostreida clade (*Fox-OG36/NA*; **Supp. Fig. S3**). Similarly, a

gene orthology group putatively specific to Pteriomorphia has been identified among Sox genes (*Sox-OG1/NA*). Of course, these genes deserve as much attention as their widely-distributed paralogs, as they may constitute true group-specific expansions and may play fundamental roles in some biological processes. However, they have not been discussed here or included in **Fig. 3.2** for clarity purposes, but they are freely available in supplementary materials.

Overall, our analysis clearly shows the importance of adopting a wide-angle approach when characterising the members of a gene family, especially for large ones such as the Fox genes (**Schomburg et al., 2022**). As a matter of fact, the presence of duplication events and orphan genes needs to be addressed with a broad taxonomic dataset, in order to account for possible mis-annotations, gene phylogenetic mis-placements, and sequence heterogeneity. Additionally, many reference species need to be included for the gene identification process, in order to consider distantly-related genes and obtain a solid annotation. Our gene annotation pipeline also resulted to be very solid, even with non-model organisms and sub-optimal genomic and transcriptomic resources as they are those of bivalves. As a matter of fact, by running the same pipeline on two additional datasets composed of mammal and fruit fly genomes, we were able to obtain high-quality orthology groups in accordance with previous knowledge on the clades (**Supp. Fig. S5–S10; Supp. Tab. S8–S9**), with little or no manual curation. Furthermore, this represents also the first broad analysis of DSFGs in both mammals and fruit flies, as so far attention has been mainly dedicated to single well-studied organisms or little clades (e.g., **Jackson et al., 2010**).

### 3.4.2 High amino acid sequence divergence identifies putative sex-determining genes

Sex-biased genes tend to evolve more rapidly than unbiased genes at the level of their protein sequences. Accelerated rates have been observed in both male-biased genes (reviewed in **Parsch and Ellegren, 2013; Grath and Parsch, 2016**) and female-biased genes (e.g., **Papa et al., 2017; Ghiselli et al., 2018**), but also in SRGs and primary SDGs (**O’Neil and Belote, 1992; Whitfield et al., 1993; de Bono and Hodgkin, 1996**). For example, it has been shown that *Dm-W*, *Dmy*, and *Sry* (which are SDGs in the African clawed frog *X. laevis*, in the medaka fish *O. latipes*, and in eutherians, respectively) all have higher substitution rates than their paralogues (*Dmrt-1* for *Dm-W* and *Dmy*, *Sox-3* for *Sry*), particularly when considering their DNA-binding domains (**Mawaribuchi et al., 2012**). Similarly, both a burst

of positive selection and a relaxation of purifying selection has been detected in *Drosophila Sxl* in correspondence with its recruitment at the top of the sex-determining cascade. The same signs of relaxed purifying selection have been found in the downstream targets of *Sxl*, that is, *tra* and *dsx*, despite no evidence of positive selection has been detected (Mullon et al., 2012).

Considering these shared features of SRGs and SDGs, we decided to look for signs of accelerated sequence evolution in DSFGs of bivalves, in order to evaluate if any of them could be *a-priori* associated with SD by employing the tools of molecular evolution. However, we wanted to analyse patterns of sequence evolution not only among putative SRGs and their close paralogs, but also considering the genomic context in which these genes evolve. In fact, our aim was to check whether higher rates of sequence evolution of SRGs hold true also when compared to other genes not involved in SD and not belonging to the same gene family. To do so, we obtained the AASD median values of more than 11k SCOs from bivalve genomes (Fig. 3.3A), in order to build a statistical distribution to be used as a reference: if SRGs/SDGs (in this case, DSFGs) truly evolve faster than other genes, we may expect them to fall within the 5% (or even 1%) upper quantile of the distribution (Fig. 3.3B), i.e., within highly divergent genes (Group 1 and Group 2 genes of the distribution; see Section 3.2). We chose to use the AASD as a metric of sequence evolution (instead of the tip-to-tip distances of phylogenetic trees, which account for more comprehensive evolutionary models) in order to save computational time. As a matter of fact, the AASD median values proved to be a good approximation of the tip-to-tip median distances in 200 randomly-selected genes (Fig. 3.3C;  $R = 0.84, p < 2.2\text{e-}6$ ).

Among DSFGs, three fell within the 5% upper quantile, namely *Dmrt-1L*, *Sox-H*, and *Sox-F*. Interestingly, *Dmrt-1L* and *Sox-H* have been already proposed to be involved in the male SD pathway of *C. gigas* (*inset* in Fig. 3.3B; Zhang et al., 2014), on the basis of DGE analyses. Specifically, *Sox-H* would play a major role in *C. gigas* SD, by interacting with *Dmrt-1L* and determining the onset of the male phenotype development; at the same time, both *Sox-H* and *Dmrt-1L* would inhibit *Fox-L2*, which instead is necessary to start the female phenotype development. *Dmrt-1L* and *Sox-H* have been appointed several other times to be involved in male-gonad development and differentiation, through DGE (e.g., Teaniniuraitemoana et al., 2014; Capt et al., 2018; Afonso et al., 2019), mRNA-ISH (e.g., Naimi et al., 2009; Li et al., 2018; Liang et al., 2019; Yue et al., 2021) and RNA interference (RNAi) (Liang et al., 2019; Sun et al., 2022). Therefore, the high AASD of *Dmrt-1L* and *Sox-H* is coherent with previous works, strengthening their role as putative SRGs.

The relationship between high gene AASD and the involvement in SD is particularly enforced when looking at the patterns of AASD in the test datasets, which corroborates the solidity of our analysis: (i) from one side, in the mammal dataset—which represents a strictly genetic SD system, thus with a master and rapidly-evolving SDG, one of the genes from the 5% upper quantile of the distribution is *Sry* (**Fig. 3.4A–B**), the male sex-determining gene in eutherians (*inset* in **Fig. 3.4A–B**); (ii) from the other side, in the fruit fly dataset—which represents a chromosomal SD system, thus without any expected difference in the rates of sequence evolution among SRGs, none of the DSFG exhibit significantly high AASD (**Fig. 3.4C–D**), including the downstream effector *dsx* (*inset* in **Fig. 3.4D**). Also *Sxl* and *tra*, both involved in the SD pathway of *Drosophila* (*inset* in **Fig. 3.4D**) do not belong to the group of highly-divergent genes, as they have a mean amino acid divergence of about 0.09 and 0.9, respectively (**Fig. 3.4D**). Therefore, it can be argued that both *Dmrt-1L* and *Sox-H* may not only be SRGs, but may participate in bivalve SD as primary SDGs, which is reflected in their high AASD, as it is observed for *Sry* in mammals. As a matter of fact, if they were involved in SD just as intermediate actors of the signalling cascade, then we should have not observed a high AASD, as *Drosophila Sxl*, *tra*, and *dsx* seem to suggest. Overall, these patterns of molecular evolution concerning SRGs and SDG are also supported by the way SD regulatory networks evolve. As a matter of fact, it has been proposed that the sex-determining cascades tend to arise and be established with a bottom-up mechanism (**Wilkins, 1995; Mullon et al., 2012; Beukeboom and Perrin, 2014; Capel, 2017**). This means that the regulative relationships among genes at the bottom of the cascade are settled up prior to the regulative relationships among genes at the top and, consequently, upstream regulators are progressively recruited to fine-tune diverse SD signals. These evolutionary patterns eventually produce gene-regulatory networks in which the divergence of the upstream triggers is higher than that of downstream effectors, in terms of both identity and sequence composition (**Beukeboom and Perrin, 2014**). This mechanism has been proposed for *Drosophila* species (**Mullon et al., 2012**), *C. elegans* (**Stothard and Pilgrim, 2003**), and vertebrates, despite in the latter case it has been questioned several times (reviewed in **Capel, 2017**).

At this point, two main objections can be moved against our approach: (1) the distribution of AASD is not appropriate for this kind of inference, as it does not represent the true gene evolutionary (or substitution) rates (which instead are those usually employed when dealing with SRGs and SDGs); (2) the three datasets are not comparable one to each other, as they take

into consideration very different animal groups, with different taxonomic rankings and different divergence times (thus, the patterns of AASD are the products of other confounding factors not directly related to SD). Concerning the first objection, we are aware that the AASD does not represent the evolutionary rate itself, but rather its product. However, the two features are tightly linked, as on the long term highly-divergent proteins tend to be produced by genes with high evolutionary (or substitution) rates (**Echave et al., 2016**). By performing a GO-term enrichment, it emerged that highly-divergent genes of the mammal dataset are mainly involved in the immune response and male spermatogenesis (**Tab. 3.1; Supp. Tab. S10**), which are two processes notoriously connected with rapid sequence evolution (i.e., higher evolutionary rates; **Swanson and Vacquier, 2002; Murat et al., 2023; Vinkler et al., 2023**). Similarly, highly-divergent genes from the fruit fly dataset show an enrichment for GO-terms associated with meiotic-related functions (such as the formation of the synaptonemal complex by the *c(2)M*, *c(3)G*, *corona*, and *corolla* proteins; **Tab. 3.1; Supp. Tab. S10**), which again are known to be rapidly evolving (**Hemmer and Blumenstiel, 2016**). In other words, the test datasets allow us to directly link the high AASD (as found in this work) with high rates of sequence evolution (as found in previous works), as they represent well-studied and characterised model systems. This consideration can thus be extended also to the bivalve dataset: highly-divergent genes in terms of AASD, which include some DSFGs and show an enrichment for GO-terms associated to macromolecule metabolism and morphological development (**Tab. 3.1; Supp. Tab. S10**), are also genes with accelerated substitution rates. Concerning the second objection, we chose two test datasets with different characteristics as we wanted to check the extent of our hypothesis (i.e., molecular evolution can be used to look for putative primary SDGs in taxonomic-wide analyses). As a matter of fact, the difference in divergence times and taxonomy ranks for bivalves and therians [Late Cambrian, about 498 million years ago (Mya), **Song et al., 2023**; and Early Mesozoic, 166–123 Mya, **Álvarez-Carretero et al., 2022**, respectively] seems to not influence the sequence diversity of SRGs, as both *Dmrt-1L/Sox-H* for bivalves and *Sry* for mammals exhibit high AASD with respect to their own distributions, regardless of their age. *Dmrt-1L* and *Sox-H* (which are mollusc- and Bilateria-specific, respectively) are undoubtedly older than *Sry* (which, instead, emerged in the Theria common ancestor; **Foster et al., 1992**), but each of them can be considered a highly-divergent gene in bivalves and mammals, respectively (i.e., genes that are included in the 5% upper quantile of bivalve and mammal AASD distributions). Conversely, the difference in divergence times and taxonomic ranks for *Drosophila* (Paleocene/Eocene boundary, about 56 Mya; **Russo et al.,**

2013) may seem to be influencing the results for the dataset, resulting in a false negative. In other words, it can be argued that: (i) the genes included in the SD cascade of *Drosophila* (such as *Sxl*, *tra*, and *dsx*; *inset* in Fig. 3.4D) have indeed a high AASD, which however has not been detected by our methodological approach (for example, this may be traced back to the young diversification age of *Drosophila* species if compared to bivalves); (ii) the species included in the analysis are all congeneric, thus the sequence differentiation of SRGs may exist not at the amino acid level but at the nucleotide one. To better disentangle this issue and further discuss the fruit fly dataset, we repeated the analysis of the AASD only on species of the *Crassostrea* genus (*C. gigas*, *Crassostrea angulata*, *Crassostrea ariakensis*, and *Crassostrea virginica*), which are all congeneric and much younger (Middle Cretaceous, less than 100 Mya; Qi et al., 2023), thus comparable to *Drosophila*. Results showed that, even when analysing a smaller bivalve dataset, encompassing only 4 species of recent origin, the high AASD of *Dmrt-1L* persists, that is, *Dmrt-1L* is still grouped together with highly-divergent genes (Supp. Fig. S15). The same has not been recovered for *Sox-H*, which fell in genes from Group 3 (the group corresponding to the 95% interval of the AASD distribution) but still have the second highest AASD median value among DSFGs (Supp. Fig. S15).

Of course we should not expect that highly-divergent genes are only those involved in SD, but may participate also in other processes (as discussed earlier and shown by GO-term enrichments; Tab. 3.1; Supp. Tab. S10). Besides the genes of interest for SD (*Dmrt-1L/Sox-H* for bivalves, and *Sry* for mammals), also other components of the DSFG families have been retrieved with a high AASD, despite they have never been linked directly to SD so far: *Sox-F* in bivalves (Fig. 3.3B) and *Fox-D4* in mammals (Fig. 3.4B). This implies that our approach can't be used to unambiguously identify SDGs alone, as high AASD is exhibited also by many other genes. Instead, the analysis is meant to be used to detect highly-divergent genes and, subsequently, by comparison with literature and a more thorough and focused functional investigation, putative SDGs among them. In this sense, the mammal dataset exemplify the importance of putting the results of our pipeline (as those of any other comparative genomics analysis) into the correct evolutionary and genomic context: among DSFGs of mammals, two genes exhibit high AASD, one of which is directly related to SD (*Sry*), while the other has a function connected with neural development (*Fox-D4*; Klein et al., 2013). Thus, the high AASD may arise either because of the involvement in the upper SD pathway or because of other life-history traits connected with the gene, respectively. Regarding bivalves, *Dmrt-1L* and *Sox-*

*H* show a sharp connection with SD as a putative primary SDG, either when considering their molecular evolutionary features or when looking at their gene expression and possible function in gonad development (Naimi et al., 2009; Teaniniuraitemoana et al., 2014; Zhang et al., 2014; Capt et al., 2018; Li et al., 2018; Afonso et al., 2019; Liang et al., 2019; Yue et al., 2021). It is difficult to further speculate on the actual involvement in SD of *Dmrt-1L* and *Sox-H* without any additional information on their biology. Nonetheless, molecular evolution proves to be a valuable tool to investigate genes putatively involved in SD, and to identify major targets onto which dedicate future research effort.

### 3.5 Conclusions.

*In preparation.*

# Chapter 4

## Expression patterns of three sex-related genes and the germline marker *Vasa* in early developmental stages of *Mytilus galloprovincialis* embryos

Filippo Nicolini<sup>1,2</sup>, Sergey Nuzhdin<sup>3</sup>, Fabrizio Ghiselli<sup>1</sup>, Andrea Luchetti<sup>1</sup>, Liliana Milani<sup>1</sup>

<sup>1</sup>*Department of Biological, Geological and Environmental Science, University of Bologna, Bologna (BO), Italy.*

<sup>2</sup>*Fano Marine Center, Fano (PU), Italy.*

<sup>3</sup>*Department of Molecular and Computational Biology, University of Southern California, Los Angeles, CA, USA.*

*In preparation.*

## 4.1 Introduction

*In preparation.*

## 4.2 Materials and Methods

### 4.2.1 Time-series gene expression

Miglioli et al., 2024 recently produced one of the very first detailed developmental transcriptome of *M. galloprovincialis*, spanning from the unfertilized oocyte to the larval stage at 72 hpf, with time points sampled every 4 hpf. A total of 30 different mRNA libraries was sequenced, consisting of fifteen developmental time points per two technical replicates. These data are very useful to thoroughly investigate the transcription patterns of genes throughout the first three days of development in *M. galloprovincialis* and to obtain hints on the expected outcomes of mRNA-ISH experiments.

Raw reads were downloaded from the Sequence Read Archive (SRA) in NCBI (BioProject: PRJNA996031) and trimmed using Trimmomatic v0.39 (Bolger et al., 2014; LEADING:5 TRAILING:5 SLIDINGWINDOW:4:15 MINLEN:65). Read quality was checked using FastQC v0.12.1 (Andrews et al., 2010). Trimmed reads were mapped against the *M. galloprovincialis* annotated genome (GCA\_900618805.1; Gerdol et al., 2020) using STAR v2.7.10b (Dobin et al., 2013) in alignReads mode with default parameters. The resulting gene count matrix was extracted with StringTie v2.2.1 (Pertea et al., 2015, 2016) in expression estimation mode followed by the python script prepDE.py (-1 99).

The resulting matrix was processed in R. Raw gene counts were normalized using the built-in function `vst` of the package DESeq2 (Love et al., 2014). The function `plotPCA` was then used to run a principal component analysis (PCA) on read mapping counts and visualize the corresponding results. Normalized gene counts were also used to plot expression values of target genes (i.e., *Vasa*, *Dmrt1L*, *SoxH* and *FoxL2*), as well as in maSigPro (Conesa et al., 2006) to run a differential gene expression analysis in a time course experiment.

The entire pipeline was automated through custom python and bash scripts, which are available in a private repository on GitHub.

#### 4.2.2 Sample collection, MitoTracker staining and fixation

Adult Mediterranean mussels (*M. galloprovincialis*) were hand collected from various locations surrounding the AltaSea institute at the port of Los Angeles (CA, USA). Sampling took place during the late spawning season of the species in California, i.e., from October 2023 to early January 2024. Specimens were checked for species and sexual maturity before usage.

Selected mussels were thoroughly cleaned from epibionts and placed in ice for approximately 30-60 minutes, then transferred in filtered artificial sea water (FASW) at 16°C and acclimatized for 30 minutes. All the individuals were then placed in a common tank and spawning was induced by cyclical thermal shock, that is, by exposing mussels alternatively to FASW at 24-26°C and 14-16°C for 30-40 minutes. As soon as individual mussels started spawning, they were promptly removed from the common tank, carefully washed and then allowed to continue spawning in isolated containers of about 250 ml 16°C FASW.

Sperm from six males and oocytes from six females were separately mixed to increase the number of crosses. An hour after the spawning started, oocytes were filtered through a 75 over a 30 µm mesh and aged in 1 L of FASW for 40-60 minutes to let them assume a proper circular shape. Oocyte abundance was estimated under a stereo microscope, by counting the number of gametes in five aliquotes of 1 mL and then calculating the mean value. Sperm mitochondria were labeled with MitoTracker™ Red CMXRos (Thermo Fisher Scientific) at a working concentration of 500 nM for 30 minutes. MitoTracker is a vital and fixation-resistant mitochondrial dye and was used to be able to detect the sex of developing embryos (as early as the two-blastomere stage) according to the distribution pattern of sperm mitochondria (**Cao et al., 2004; Obata and Komaru, 2005**). From this step onward, samples were always kept in the dark.

Fertilization was performed by mixing oocytes and sperm at a ratio of 1:10. Fertilization success was checked after 20-30 minutes by the formation of polar bodies. The suspension was then carefully washed to remove excess sperm and brought to a concentration of 250 zygotes/mL. The resulting suspension was transferred into cell-culture flasks of 40 mL and embryos/larvae were reared at 16°C in the dark. Water was changed every 24 hours. After 48 hpf, larvae were fed with *Isochrysis galbana* at a final concentration of circa 100,000 cells/mL following **Helm et al., 2004**.

Embryos/larvae were sampled at 1, 2, 3 and 4 hpf, and then every 12 hours until 72 hpf,

Target	Amplifier	Fluorophore	No. of probe pairs
<i>Vasa</i>	B1	ALEXA-488	33
<i>Dmrt1L</i>	B2	ALEXA-647	18
<i>SoxH</i>	B3	ALEXA-546	22
<i>FoxL2</i>	B4	ALEXA-700	28

**Table 4.1.** List of genes targeted through HCR, with the corresponding amplifiers, fluorophores and number of generated probe pairs.

every time after checking for proper development and vitality. After concentration in a mesh of proper size, embryos/larvae were fixed in 3.2% paraformaldehyde (PFA) in 1× PBS at 4°C overnight under constant and gentle shaking. Fixed samples were washed 3 × 20 minutes in 1× PBS 0.1% Tween 20 (PBST) and then dehydrated 3 × 30 minutes in absolute methanol at room temperature (RT). Dehydrated samples were stored at -20°C until usage.

#### 4.2.3 mRNA *in-situ* Hybridization Chain Reaction (HCR)

##### HCR probe design

*Vasa*, *Dmrt1L*, *SoxH*, and *FoxL2* spliced-transcript nucleotide sequences of *M. galloprovincialis* were obtained from previous analyses with OrthoFinder v2.5.5 (**Emms and Kelly, 2019**) and 30 annotated bivalve genomes (see **Chapter 3**). Accession numbers of spliced transcripts are 10B017427, 10B093608, 10B014180, and 10B094018, respectively. The `insitu_probe_generator` script from Ozpolat Lab (**Kuehn et al., 2022**) was used to generate pairs of probes specifically designed for third-generation HCR (**Choi et al., 2018**). The built-in BLASTN search against the annotated *M. galloprovincialis* transcriptome was employed to check for putative off-target bindings of probe pairs. B1-488, B2-647, B3-546, and B4-700 pairs of HCR amplifiers and fluorophores were chosen as in **Tab. 4.1**. Resulting probes were synthetized by Integrated DNA Technologies (IDT™) in different oligo pools.

##### Fluorescent *in-situ* hybridization through hybridization chain reaction and microscope imaging

HCR mRNA-FISH in *M. galloprovincialis* embryos was performed following **Miglioli et al., 2024**. All the steps were carried out in the dark to prevent MitoTracker from fading. Probe hybridization buffer, probe wash buffer and amplification buffer were manufactured by Molecular Instruments, Inc.

Target	Dye	Excitation (nm)	Emission (nm)
dsDNA (nuclei)	DAPI	360	460
Sperm mitochondria	MitoTracker™ Red CMXRos	575	600
<i>Vasa</i>	ALEXA-488	499	520
<i>Dmrt1L</i>	ALEXA-647	653	670
<i>SoxH</i>	ALEXA-546	557	575
<i>FoxL2</i>	ALEXA-700	685	700

**Table 4.2.** List of dyes used for every target, together with the excitation and emission peaks as returned by the Las X software.

Dehydrated samples stored in methanol were washed 4 times per 5 minutes and 1 time per 10 minutes in a phosphate-buffered saline solution (PBS; 128 mM NaCl, 2 mM KCl, 8 mM Na<sub>2</sub>HPO<sub>4</sub> · 2H<sub>2</sub>O, 2 mM KH<sub>2</sub>PO<sub>4</sub>) with 0.1% Tween 20 (PBST). Samples were then permeabilized for 30 minutes in a detergent solution (1.0% SDS, 0.5% Tween 20, 50 mM Tris-HCl, 1.0 mM ethylenediaminetetraacetic acid (EDTA), 150.0 mM NaCl) and washed again 2 times per 5 minutes in PBST. Samples were prepared for the HCR detection stage by incubation in probe hybridization buffer for 30 minutes at 37 °C. Detection stage was then performed with 4 nM of each probe set in hybridization solution overnight (>12 h) at 37 °C.

Excess probes was removed by washing 4 times per 20 minutes with probe wash buffer at 37 °C and 3 times per 5 minutes with 5× saline-sodium citrate Tween 20 buffer (SSCT; 5× SSC, 0.1% Tween 20) at room temperature. Samples were incubated for 30 minutes in amplification buffer at room temperature. Hairpins were heated at 95 °C for 90 seconds and then snap-cooled at room temperature for 30 minutes. The amplification step of HCR was performed with 6 pmol of each hairpin in amplification buffer overnight (>12 h) at room temperature.

Excess hairpins was removed by washing 2 times per 5 minutes, 2 times per 30 minutes, and 1 time per 5 minutes with SSCT. If not immediately mounted on slides, samples were stored in SSCT at +4 °C. Otherwise, samples were immersed in 50% and 75% glycerol for 30-60 minutes each, and then mounted with VECTASHIELD® PLUS Antifade Mounting Medium with DAPI (H-2000). Slides were imaged on a Stellaris 5 Confocal Package system with the software Las X (Leica Microsystems). Each dye was imaged sequentially in a separate channel, to enhance the yield and avoid any crosstalks. **Tab. 4.2** summarises the excitation and emission peaks for each dye. Images were then manipulated and post-produced using Fiji v2.14.0.

#### 4.2.4 Immunolocalization of Vasa

Vasa immunolocalization in *M. galloprovincialis* embryos was performed following **Milani et al., 2011** with modifications. All the steps were carried out in the dark to prevent MitoTracker from fading.

Dehydrated samples stored in methanol were rinsed 3 times per 10 minutes and 1 time for 2 hours in Tris-buffered saline (TBS; 10 mM Tris-HCl, 155 mM NaCl), following an additional wash for 10 minutes with PBS. Samples were then digested for 6 minutes and 30 seconds with 0.01% pronase E (Merck) in PBS, and washed again 2 times for 5 minutes in PBS. Permeabilization was then performed in TBS-Triton (TBST) 0.1% for 5 minutes at RT and in TBST 1% overnight at 4°C.

After an additional rinse for 5 minutes in TBST 0.1%, non-specific protein-binding sites were blocked with a TBST 0.1% solution containing 3% bovine serum albumin (BSA). Samples were then incubated at 4°C for 32-48 hours with primary anti-VASA/VAS antibody (Abcam ab209710; polyclonal anti-Vasa developed in rabbit), diluted 1:100.

Excess primary antibody was rinsed from samples with 4 washes of 30 minutes in TBST 0.1%, while non-specific protein-binding sites were blocked again with an incubation of 1 hour in TBST 0.1% containing 3% BSA. Samples were then incubated at 4°C for 24-32 hours with secondary antibody HRP anti-rabbit in goat (Santa Cruz Biotechnology Inc.) dilutied 1:400. Excess secondary antibody was rinsed with 4 washes of 30 minutes in TBST 0.1% and 1 wash of 1 hour in TBST 1%.

Samples were immersed in 50% and 75% glycerol for 30-60 minutes each, and then mounted with VECTASHIELD® PLUS Antifade Mounting Medium with DAPI (H-2000). Slides were imaged COMPLETECOMPLETECOMPLETECOMPLETE. Each dye was imaged sequentially in a separate channel, to enhance the yield and avoid any crosstalks. **Tab. 4.2** summarises the excitation and emission peaks for each dye. Images were then manipulated and post-produced using Fiji v2.14.0.

### 4.3 Results

*In preparation.*

## 4.4 Discussion

*In preparation.*

# References

- Abascal, F., Zardoya, R., & Telford, M. J. (2010). Translatorx: Multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic acids research*, 38(suppl\_2), W7–W13.
- Abbott, J. K. (2011). Intra-locus sexual conflict and sexually antagonistic genetic variation in hermaphroditic animals. *Proceedings of the Royal Society B: Biological Sciences*, 278(1703), 161–169.
- Afonso, L. F., Americo, J. A., Soares-Souza, G. B., Torres, A. L. Q., Wajsenzon, I. J. R., & Rebelo, M. d. F. (2019). Gonad transcriptome of golden mussel *Limnoperna fortunei* reveals potential sex differentiation genes. *biorxiv*, 818757.
- Albertin, C. B., Simakov, O., Mitros, T., Wang, Z. Y., Pungor, J. R., Edsinger-Gonzales, E., Brenner, S., Ragsdale, C. W., & Rokhsar, D. S. (2015). The octopus genome and the evolution of cephalopod neural and morphological novelties. *Nature*, 524(7564), 220–224.
- Alexa, A., & Rahnenführer, J. (2009). Gene set enrichment analysis with topgo. *Bioconductor Improv*, 27, 1–26.
- Allen, S. L., Delaney, E. K., Kopp, A., & Chenoweth, S. F. (2017). Single-molecule sequencing of the *Drosophila serrata* genome. *G3: Genes, Genomes, Genetics*, 7(3), 781–788.
- Altenhoff, A. M., Warwick Vesztrocy, A., Bernard, C., Train, C.-M., Nicheperovich, A., Prieto Baños, S., Julca, I., Moi, D., Nevers, Y., Majidian, S., et al. (2024). Oma orthology in 2024: Improved prokaryote coverage, ancestral and extant go enrichment, a revamped synteny viewer and more in the oma ecosystem. *Nucleic Acids Research*, 52(D1), D513–D521.
- Álvarez-Carretero, S., Tamuri, A. U., Battini, M., Nascimento, F. F., Carlisle, E., Asher, R. J., Yang, Z., Donoghue, P. C., & Dos Reis, M. (2022). A species-level timeline of mammal evolution integrating phylogenomic data. *Nature*, 602(7896), 263–267.

- Andrews, S., et al. (2010). Fastqc: A quality control tool for high throughput sequence data.
- Bachtrog, D., Mank, J. E., Peichel, C. L., Kirkpatrick, M., Otto, S. P., Ashman, T.-L., Hahn, M. W., Kitano, J., Mayrose, I., Ming, R., et al. (2014). Sex determination: Why so many ways of doing it? *PLoS biology*, 12(7), e1001899.
- Bai, C.-M., Xin, L.-S., Rosani, U., Wu, B., Wang, Q.-C., Duan, X.-K., Liu, Z.-H., & Wang, C.-M. (2019). Chromosomal-level assembly of the blood clam, *Scapharca (Anadara) broughtonii*, using long sequence reads and hi-c. *GigaScience*, 8(7), giz067.
- Baldwin-Brown, J. G., Weeks, S. C., & Long, A. D. (2018). A new standard for crustacean genomes: The highly contiguous, annotated genome assembly of the clam shrimp *Eulimnadia texana* reveals hox gene order and identifies the sex chromosome. *Genome biology and evolution*, 10(1), 143–156.
- Baral, S., Arumugam, G., Deshmukh, R., & Kunte, K. (2019). Genetic architecture and sex-specific selection govern modular, male-biased evolution of *doublesex*. *Science advances*, 5(5), eaau3753.
- Benayoun, B. A., Caburet, S., & Veitia, R. A. (2011). Forkhead transcription factors: Key players in health and disease. *Trends in Genetics*, 27(6), 224–232.
- Beukeboom, L. W., & Perrin, N. (2014). *The evolution of sex determination*. Oxford University Press.
- Bewick, A. J., Anderson, D. W., & Evans, B. J. (2011). Evolution of the closely related, sex-related genes *DM-W* and *DMRT1* in african clawed frogs (*Xenopus*). *Evolution*, 65(3), 698–712.
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.
- Bredemeyer, K. R., Hillier, L., Harris, A. J., Hughes, G. M., Foley, N. M., Lawless, C., Carroll, R. A., Storer, J. M., Batzer, M. A., Rice, E. S., et al. (2023). Single-haplotype comparative genomics provides insights into lineage-specific structural variation during cat evolution. *Nature genetics*, 55(11), 1953–1963.
- Breton, S., Capt, C., Guerra, D., & Stewart, D. (2018). Sex-determining mechanisms in bivalves. In J. L. Leonard (Ed.), *Transitions between sexual systems: Understanding the mechanisms of, and pathways between, dioecy, hermaphroditism and other sexual systems* (pp. 165–192). Springer International Publishing.

- Breton, S., Stewart, D. T., Brémaud, J., Havird, J. C., Smith, C. H., & Hoeh, W. R. (2022). Did doubly uniparental inheritance (dui) of mtDNA originate as a cytoplasmic male sterility (cms) system? *BioEssays*, 44(4), 2100283.
- Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using diamond. *Nature methods*, 12(1), 59–60.
- Calcino, A. D., de Oliveira, A. L., Simakov, O., Schwaha, T., Zieger, E., Wollesen, T., & Wanninger, A. (2019). The quagga mussel genome and the evolution of freshwater tolerance. *DNA Research*, 26(5), 411–422.
- Cao, L., Kenchington, E., & Zouros, E. (2004). Differential segregation patterns of sperm mitochondria in embryos of the blue mussel (*Mytilus edulis*). *Genetics*, 166(2), 883–894.
- Capel, B. (2017). Vertebrate sex determination: Evolutionary plasticity of a fundamental switch. *Nature Reviews Genetics*, 18(11), 675–689.
- Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). Trimal: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25(15), 1972–1973.
- Capt, C., Bouvet, K., Guerra, D., Robicheau, B. M., Stewart, D. T., Pante, E., & Breton, S. (2020). Unorthodox features in two venerid bivalves with doubly uniparental inheritance of mitochondria. *Scientific reports*, 10(1), 1087.
- Capt, C., Renaut, S., Ghiselli, F., Milani, L., Johnson, N. A., Sietman, B. E., Stewart, D. T., & Breton, S. (2018). Deciphering the link between doubly uniparental inheritance of mtDNA and sex determination in bivalves: Clues from comparative transcriptomics. *Genome biology and evolution*, 10(2), 577–590.
- Chakraborty, M., Chang, C.-H., Khost, D. E., Vedanayagam, J., Adrion, J. R., Liao, Y., Montooth, K. L., Meiklejohn, C. D., Larracuente, A. M., & Emerson, J. (2021). Evolution of genome structure in the *Drosophila simulans* species complex. *Genome research*, 31(3), 380–396.
- Chen, H., Xiao, G., Chai, X., Lin, X., Fang, J., & Teng, S. (2017). Transcriptome analysis of sex-related genes in the blood clam *Tegillarca granosa*. *PLoS One*, 12(9), e0184584.
- Choi, H. M., Schwarzkopf, M., Fornace, M. E., Acharya, A., Artavanis, G., Stegmaier, J., Cunha, A., & Pierce, N. A. (2018). Third-generation *in situ* hybridization chain reaction: Multiplexed, quantitative, sensitive, versatile, robust. *Development*, 145(12), dev165753.
- Civetta, A., & Singh, R. S. (1998). Sex-related genes, directional sexual selection, and speciation. *Molecular biology and evolution*, 15(7), 901–909.

- Collin, R. (2013). Phylogenetic patterns and phenotypic plasticity of molluscan sexual systems. *Integrative and Comparative Biology*, 53(4), 723–735.
- Conesa, A., Nueda, M. J., Ferrer, A., & Talón, M. (2006). Masigpro: A method to identify significantly differential expression profiles in time-course microarray experiments. *Bioinformatics*, 22(9), 1096–1102.
- Corrochano-Fraile, A., Davie, A., Carboni, S., & Bekaert, M. (2022). Evidence of multiple genome duplication events in *Mytilus* evolution. *BMC genomics*, 23(1), 340.
- Cutter, A. D., & Ward, S. (2005). Sexual and temporal dynamics of molecular evolution in *C. elegans* development. *Molecular Biology and Evolution*, 22(1), 178–188.
- Dainat, J., Hereñú, D., Davis, E., Crouch, K., Sol, L., & Agostinho, N. (2022). Another gff analysis toolkit to handle annotations in any gtf/gff format (version v1. 0). *Zenodo. doi, 10.*
- de Bono, M., & Hodgkin, J. (1996). Evolution of sex determination in *Caenorhabditis*: Unusually high divergence of *tra-1* and its functional consequences. *Genetics*, 144(2), 587–595.
- Deng, T., Pang, C., Lu, X., Zhu, P., Duan, A., Tan, Z., Huang, J., Li, H., Chen, M., & Liang, X. (2016). *De novo* transcriptome assembly of the chinese swamp buffalo by rna sequencing and ssr marker discovery. *PLoS One*, 11(1), e0147132.
- Dermauw, W., Van Leeuwen, T., & Feyereisen, R. (2020). Diversity and evolution of the p450 family in arthropods. *Insect biochemistry and molecular biology*, 127, 103490.
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). Star: Ultrafast universal rna-seq aligner. *Bioinformatics*, 29(1), 15–21.
- Dong, C., Zhang, L., Xia, S., Sosa, D., Arsala, D., & Long, M. (2022). New gene evolution with subcellular expression patterns detected in pacbio-sequenced genomes of *Drosophila* genus. *bioRxiv*, 2022–11.
- Eads, B. D., Colbourne, J. K., Bohuski, E., & Andrews, J. (2007). Profiling sex-biased gene expression during parthenogenetic reproduction in *Daphnia pulex*. *BMC genomics*, 8, 1–14.
- Echave, J., Spielman, S. J., & Wilke, C. O. (2016). Causes of evolutionary rate variation among protein sites. *Nature Reviews Genetics*, 17(2), 109–121.
- Eddy, S. R. (2011). Accelerated profile hmm searches. *PLoS computational biology*, 7(10), e1002195.

- Elbers, J. P., Rogers, M. F., Perelman, P. L., Proskuryakova, A. A., Serdyukova, N. A., Johnson, W. E., Horin, P., Corander, J., Murphy, D., & Burger, P. A. (2019). Improving illumina assemblies with hi-c and long reads: An example with the north african dromedary. *Molecular Ecology Resources*, 19(4), 1015–1026.
- Ellegren, H., & Parsch, J. (2007). The evolution of sex-biased genes and sex-biased gene expression. *Nature Reviews Genetics*, 8(9), 689–698.
- Emms, D. M., & Kelly, S. (2019). Orthofinder: Phylogenetic orthology inference for comparative genomics. *Genome biology*, 20, 1–14.
- Evensen, K. G., Robinson, W. E., Krick, K., Murray, H. M., & Poynton, H. C. (2022). Comparative phylotranscriptomics reveals putative sex differentiating genes across eight diverse bivalve species. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 41, 100952.
- Fan, H., Wu, Q., Wei, F., Yang, F., Ng, B. L., & Hu, Y. (2019). Chromosome-level genome assembly for giant panda provides novel insights into carnivora chromosome evolution. *Genome biology*, 20, 1–12.
- Farhat, S., Bonnivard, E., Pales Espinosa, E., Tanguy, A., Boutet, I., Guiglielmoni, N., Flot, J.-F., & Allam, B. (2022). Comparative analysis of the *Mercenaria mercenaria* genome provides insights into the diversity of transposable elements and immune molecules in bivalve mollusks. *BMC genomics*, 23(1), 192.
- Foote, A. D., Liu, Y., Thomas, G. W., Vinař, T., Alföldi, J., Deng, J., Dugan, S., van Elk, C. E., Hunter, M. E., Joshi, V., et al. (2015). Convergent evolution of the genomes of marine mammals. *Nature genetics*, 47(3), 272–275.
- Foster, J. W., Brennan, F. E., Hampikian, G. K., Goodfellow, P. N., Sinclair, A. H., Lovell-Badge, R., Selwood, L., Renfree, M. B., Cooper, D. W., & Marshall Graves, J. A. (1992). Evolution of sex determination and the y chromosome: *SRY*-related sequences in marsupials. *Nature*, 359(6395), 531–533.
- Fouret, J., Brunet, F. G., Binet, M., Aurine, N., Enchéry, F., Croze, S., Guinier, M., Goumaidi, A., Preininger, D., Volff, J.-N., et al. (2020). Sequencing the genome of indian flying fox, natural reservoir of nipah virus, using hybrid assembly and conservative secondary scaffolding. *Frontiers in Microbiology*, 11, 1807.
- Gerdol, M., Moreira, R., Cruz, F., Gómez-Garrido, J., Vlasova, A., Rosani, U., Venier, P., Naranjo-Ortiz, M. A., Murgarella, M., Greco, S., et al. (2020). Massive gene presence-

- absence variation shapes an open pan-genome in the mediterranean mussel. *Genome biology*, 21, 1–21.
- Ghiselli, F., Iannello, M., Piccinini, G., & Milani, L. (2021). Bivalve molluscs as model systems for studying mitochondrial biology. *Integrative and Comparative Biology*, 61(5), 1699–1714.
- Ghiselli, F., Iannello, M., Puccio, G., Chang, P. L., Plazzi, F., Nuzhdin, S. V., & Passamonti, M. (2018). Comparative transcriptomics in two bivalve species offers different perspectives on the evolution of sex-biased genes. *Genome biology and evolution*, 10(6), 1389–1402.
- Ghiselli, F., Milani, L., Guerra, D., Chang, P. L., Breton, S., Nuzhdin, S. V., & Passamonti, M. (2013). Structure, transcription, and variability of metazoan mitochondrial genome: Perspectives from an unusual mitochondrial inheritance system. *Genome biology and evolution*, 5(8), 1535–1554.
- Gomes-dos-Santos, A., Lopes-Lima, M., Castro, L. F. C., & Froufe, E. (2020). Molluscan genomics: The road so far and the way forward. *Hydrobiologia*, 847(7), 1705–1726.
- Gomes-dos-Santos, A., Lopes-Lima, M., Machado, A. M., Marcos Ramos, A., Usié, A., Bolotov, I. N., Vikhrev, I. V., Breton, S., Castro, L. F. C., da Fonseca, R. R., et al. (2021). The crown pearl: A draft genome assembly of the european freshwater pearl mussel *Margaritifera margaritifera* (linnaeus, 1758). *DNA research*, 28(2), dsab002.
- Gomes-dos-Santos, A., Lopes-Lima, M., Machado, M. A., Teixeira, A., C. Castro, L. F., & Froufe, E. (2023). Pacbio hi-fi genome assembly of the iberian dolphin freshwater mussel *Unio delphinus* spengler, 1793. *Scientific Data*, 10(1), 340.
- Gómez-Chiarri, M., Warren, W. C., Guo, X., & Proestou, D. (2015). Developing tools for the study of molluscan immunity: The sequencing of the genome of the eastern oyster, *Crassostrea virginica*. *Fish & shellfish immunology*, 46(1), 2–4.
- González, V. L., Andrade, S. C., Bieler, R., Collins, T. M., Dunn, C. W., Mikkelsen, P. M., Taylor, J. D., & Giribet, G. (2015). A phylogenetic backbone for bivalvia: An rna-seq approach. *Proceedings of the Royal Society B: Biological Sciences*, 282(1801), 20142332.
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Trinity: Reconstructing a full-length transcriptome without a genome from rna-seq data. *Nature biotechnology*, 29(7), 644.
- Grath, S., & Parsch, J. (2016). Sex-biased gene expression. *Annual review of genetics*, 50, 29–44.

- Grau-Bové, X., & Sebé-Pedrós, A. (2021). Orthology clusters from gene trees with possvm. *Molecular Biology and Evolution*, 38(11), 5204–5208.
- Gusman, A., Lecomte, S., Stewart, D. T., Passamonti, M., & Breton, S. (2016). Pursuing the quest for better understanding the taxonomic distribution of the system of doubly uniparental inheritance of mtDNA. *PeerJ*, 4, e2760.
- Haag, E. S., & Doty, A. V. (2005). Sex determination across evolution: Connecting the dots. *PLoS biology*, 3(1), e21.
- Haas, B. J. (n.d.). *Transdecoder*. <https://github.com/TransDecoder>
- Habtewold, T., Wagah, M., Tambwe, M. M., Moore, S., Windbichler, N., Christophides, G., Johnson, H., Heaton, H., Collins, J., Krasheninnikova, K., et al. (2023). A chromosomal reference genome sequence for the malaria mosquito, *Anopheles gambiae*, Giles, 1902, ifakara strain. *Wellcome Open Research*, 8.
- Haerty, W., Jagadeeshan, S., Kulathinal, R. J., Wong, A., Ravi Ram, K., Sirot, L. K., Levesque, L., Artieri, C. G., Wolfner, M. F., Civetta, A., et al. (2007). Evolution in the fast lane: Rapidly evolving sex-related genes in *Drosophila*. *Genetics*, 177(3), 1321–1335.
- Han, F., Wang, Z., Wu, F., Liu, Z., Huang, B., & Wang, D. (2010). Characterization, phylogeny, alternative splicing and expression of *Sox30* gene. *BMC Molecular Biology*, 11, 1–11.
- Han, W., Liu, L., Wang, J., Wei, H., Li, Y., Zhang, L., Guo, Z., Li, Y., Liu, T., Zeng, Q., et al. (2022). Ancient homomorphy of molluscan sex chromosomes sustained by reversible sex-biased genes and sex determiner translocation. *Nature Ecology & Evolution*, 1–16.
- Hart, S. F., Yonemitsu, M. A., Giersch, R. M., Garrett, F. E., Beal, B. F., Arriagada, G., Davis, B. W., Ostrander, E. A., Goff, S. P., & Metzger, M. J. (2023). Centuries of genome instability and evolution in soft-shell clam, *Mya arenaria*, bivalve transmissible neoplasia. *Nature Cancer*, 4(11), 1561–1574.
- Haszprunar, G., & Wanninger, A. (2012). Molluscs. *Current Biology*, 22(13), R510–R514.
- Heenan, P., Zondag, L., & Wilson, M. J. (2016). Evolution of the sox gene family within the chordate phylum. *Gene*, 575(2), 385–392.
- Helm, M. M., Bourne, N., & Lovatelli, A. (2004). *Hatchery culture of bivalves: A practical manual*.
- Hemmer, L. W., & Blumenstiel, J. P. (2016). Holding it together: Rapid evolution and positive selection in the synaptonemal complex of *Drosophila*. *BMC evolutionary biology*, 16, 1–17.

- Hoskins, R. A., Carlson, J. W., Wan, K. H., Park, S., Mendez, I., Galle, S. E., Booth, B. W., Pfeiffer, B. D., George, R. A., Svirskas, R., et al. (2015). The release 6 reference sequence of the *Drosophila melanogaster* genome. *Genome research*, 25(3), 445–458.
- Iannello, M., Forni, G., Piccinini, G., Xu, R., Martelossi, J., Ghiselli, F., & Milani, L. (2023). Signatures of extreme longevity: A perspective from bivalve molecular evolution. *Genome Biology and Evolution*, 15(11), evad159.
- Inoue, K., Yoshioka, Y., Tanaka, H., Kinjo, A., Sassa, M., Ueda, I., Shinzato, C., Toyoda, A., & Itoh, T. (2021). Genomics and transcriptomics of the green mussel explain the durability of its byssus. *Scientific Reports*, 11(1), 5992.
- Ip, J. C.-H., Xu, T., Sun, J., Li, R., Chen, C., Lan, Y., Han, Z., Zhang, H., Wei, J., Wang, H., et al. (2021). Host–endosymbiont genome integration in a deep-sea chemosymbiotic clam. *Molecular Biology and Evolution*, 38(2), 502–518.
- Jackson, B. C., Carpenter, C., Nebert, D. W., & Vasiliou, V. (2010). Update of human and mouse forkhead box (fox) gene families. *Human genomics*, 4, 1–8.
- Jombart, T., & Dray, S. (2010). Adephylo: Exploratory analyses for the phylogenetic comparative method. *Bioinformatics*, 26(15), 1–21.
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Von Haeseler, A., & Jermiin, L. S. (2017). Modelfinder: Fast model selection for accurate phylogenetic estimates. *Nature methods*, 14(6), 587–589.
- Kenny, N. J., McCarthy, S. A., Dudchenko, O., James, K., Betteridge, E., Corton, C., Dolucan, J., Mead, D., Oliver, K., Omer, A. D., et al. (2020). The gene-rich genome of the scallop *Pecten maximus*. *Gigascience*, 9(5), giaa037.
- Khaitovich, P., Hellmann, I., Enard, W., Nowick, K., Leinweber, M., Franz, H., Weiss, G., Lachmann, M., & Paabo, S. (2005). Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science*, 309(5742), 1850–1854.
- Kim, B. Y., Wang, J. R., Miller, D. E., Barmina, O., Delaney, E., Thompson, A., Comeault, A. A., Peede, D., D'Agostino, E. R., Pelaez, J., et al. (2021). Highly contiguous assemblies of 101 drosophilid genomes. *Elife*, 10, e66405.
- Klein, S. L., Neilson, K. M., Orban, J., Yaklichkin, S., Hoffbauer, J., Mood, K., Daar, I. O., & Moody, S. A. (2013). Conserved structural domains in foxd4l1, a neural forkhead box transcription factor, are required to repress or activate target genes. *PloS one*, 8(4), e61845.

- Knudsen, B., Kohn, A. B., Nahir, B., McFadden, C. S., & Moroz, L. L. (2006). Complete dna sequence of the mitochondrial genome of the sea-slug, *Aplysia californica*: Conservation of the gene order in euthyneura. *Molecular Phylogenetics and Evolution*, 38(2), 459–469.
- Kopp, A. (2012). Dmrt genes in the development and evolution of sexual dimorphism. *Trends in Genetics*, 28(4), 175–184.
- Kousathanas, A., Halligan, D. L., & Keightley, P. D. (2014). Faster-x adaptive protein evolution in house mice. *Genetics*, 196(4), 1131–1143.
- Kuehn, E., Clausen, D. S., Null, R. W., Metzger, B. M., Willis, A. D., & Özpolat, B. D. (2022). Segment number threshold determines juvenile onset of germline cluster expansion in *Platynereis dumerilii*. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 338(4), 225–240.
- Lan, Y., Sun, J., Chen, C., Sun, Y., Zhou, Y., Yang, Y., Zhang, W., Li, R., Zhou, K., Wong, W. C., et al. (2021). Hologenome analysis reveals dual symbiosis in the deep-sea hydrothermal vent snail *Gigantopelta aegis*. *Nature communications*, 12(1), 1165.
- Larroux, C., Luke, G. N., Koopman, P., Rokhsar, D. S., Shimeld, S. M., & Degnan, B. M. (2008). Genesis and expansion of metazoan transcription factor gene classes. *Molecular biology and evolution*, 25(5), 980–996.
- Lemoine, F., & Gascuel, O. (2021). Gotree/goalign: Toolkit and go api to facilitate the development of phylogenetic workflows. *NAR Genomics and Bioinformatics*, 3(3), lqab075.
- Li, A., Dai, H., Guo, X., Zhang, Z., Zhang, K., Wang, C., Wang, X., Wang, W., Chen, H., Li, X., et al. (2021). Genome of the estuarine oyster provides insights into climate impact and adaptive plasticity. *Communications Biology*, 4(1), 1287.
- Li, F., Bian, L., Ge, J., Han, F., Liu, Z., Li, X., Liu, Y., Lin, Z., Shi, H., Liu, C., et al. (2020). Chromosome-level genome assembly of the east asian common octopus (*Octopus sinensis*) using pacbio sequencing and hi-c technology. *Molecular ecology resources*, 20(6), 1572–1582.
- Li, H. (2023). Protein-to-genome alignment with miniprot. *Bioinformatics*, 39(1), btad014.
- Li, R., Zhang, L., Li, W., Zhang, Y., Li, Y., Zhang, M., Zhao, L., Hu, X., Wang, S., & Bao, Z. (2018). *FOXL2* and *DMRT1L* are yin and yang genes for determining timing of sex differentiation in the bivalve mollusk *Patinopecten yessoensis*. *Frontiers in Physiology*, 9, 1166.

- Li, X.-Y., Mei, J., Ge, C.-T., Liu, X.-L., & Gui, J.-F. (2022). Sex determination mechanisms and sex control approaches in aquaculture animals. *Science China Life Sciences*, 65(6), 1091–1122.
- Liang, S., Liu, D., Li, X., Wei, M., Yu, X., Li, Q., Ma, H., Zhang, Z., & Qin, Z. (2019). SOX2 participates in spermatogenesis of zhikong scallop *Chlamys farreri*. *Scientific Reports*, 9(1), 76.
- Liao, Y., Zhang, X., Chakraborty, M., & Emerson, J. (2021). Topologically associating domains and their role in the evolution of genome structure and function in *Drosophila*. *Genome research*, 31(3), 397–410.
- Liu, C., Zhang, Y., Ren, Y., Wang, H., Li, S., Jiang, F., Yin, L., Qiao, X., Zhang, G., Qian, W., et al. (2018). The genome of the golden apple snail *Pomacea canaliculata* provides insight into stress tolerance and invasive adaptation. *Gigascience*, 7(9), giy101.
- Liu, T., Zhang, Y., Nie, H., Sun, J., & Yan, X. (2024). Characterization and expression patterns of the fox gene family under heat and cold stress in manila clam *Ruditapes philippinarum* based on genome-wide identification. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 52, 101313.
- Liu, X., Li, C., Chen, M., Liu, B., Yan, X., Ning, J., Ma, B., Liu, G., Zhong, Z., Jia, Y., et al. (2020). Draft genomes of two atlantic bay scallop subspecies *Argopecten irradians irradians* and *A. i. concentricus*. *Scientific Data*, 7(1), 99.
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome biology*, 15(12), 1–21.
- Lozano-Fernandez, J. (2022). A practical guide to design and assess a phylogenomic study. *Genome Biology and Evolution*, 14(9), evac129.
- Mahajan, S., Wei, K. H.-C., Nalley, M. J., Gibilisco, L., & Bachtrog, D. (2018). De novo assembly of a young *Drosophila* y chromosome using single-molecule sequencing and chromatin conformation capture. *PLoS biology*, 16(7), e2006348.
- Mai, D., Nalley, M. J., & Bachtrog, D. (2020). Patterns of genomic differentiation in the *Drosophila nasuta* species complex. *Molecular biology and evolution*, 37(1), 208–220.
- Mank, J. E., Axelsson, E., & Ellegren, H. (2007). Fast-x on the z: Rapid evolution of sex-linked genes in birds. *Genome Research*, 17(5), 618–624.
- Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., & Zdobnov, E. M. (2021). Busco update: Novel and streamlined workflows along with broader and deeper phylogenetic

- coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Molecular biology and evolution*, 38(10), 4647–4654.
- Marshall Graves, J. A., & Peichel, C. L. (2010). Are homologies in vertebrate sex determination due to shared ancestry or to limited options? *Genome biology*, 11, 1–12.
- Matson, C. K., & Zarkower, D. (2012). Sex and the singular dm domain: Insights into sexual regulation, evolution and plasticity. *Nature Reviews Genetics*, 13(3), 163–174.
- Mawaribuchi, S., Ito, Y., & Ito, M. (2019). Independent evolution for sex determination and differentiation in the *DMRT* family in animals. *Biology Open*, 8(8), bio041962.
- Mawaribuchi, S., Yoshimoto, S., Ohashi, S., Takamatsu, N., & Ito, M. (2012). Molecular evolution of vertebrate sex-determining genes. *Chromosome Research*, 20, 139–151.
- Mazet, F., Yu, J.-K., Liberles, D. A., Holland, L. Z., & Shimeld, S. M. (2003). Phylogenetic relationships of the fox (forkhead) gene family in the bilateria. *Gene*, 316, 79–89.
- McCartney, M. A., Auch, B., Kono, T., Mallez, S., Zhang, Y., Obille, A., Becker, A., Abrahante, J. E., Garbe, J., Badalamenti, J. P., et al. (2022). The genome of the zebra mussel, *Dreissena polymorpha*: A resource for comparative genomics, invasion genetics, and biocontrol. *G3*, 12(2), jkab423.
- Mead, D., Fingland, K., Cripps, R., Miguez, R. P., Smith, M., Corton, C., Oliver, K., Skelton, J., Betteridge, E., Doulcan, J., et al. (2020). The genome sequence of the eastern grey squirrel, *Sciurus carolinensis* gmelin, 1788. *Wellcome Open Research*, 5.
- Meisel, R. P. (2011). Towards a more nuanced understanding of the relationship between sex-biased gene expression and rates of protein-coding sequence evolution. *Molecular Biology and Evolution*, 28(6), 1893–1900.
- Meisel, R. P., & Connallon, T. (2013). The faster-x effect: Integrating theory and data. *Trends in genetics*, 29(9), 537–544.
- Miglioli, A., Tredez, M., Boosten, M., Sant, C., Carvalho, J. E., Dru, P., Canesi, L., Schubert, M., & Dumollard, R. (2024). The mediterranean mussel *Mytilus galloprovincialis*: A novel model for developmental studies in mollusks. *Development*, 151(4), dev202256.
- Milani, L., & Ghiselli, F. (2020). Faraway, so close. the comparative method and the potential of non-model animals in mitochondrial research. *Philosophical Transactions of the Royal Society B*, 375(1790), 20190186.
- Milani, L., Ghiselli, F., Maurizii, M. G., Nuzhdin, S. V., & Passamonti, M. (2014). Paternally transmitted mitochondria express a new gene of potential viral origin. *Genome biology and evolution*, 6(2), 391–405.

- Milani, L., Ghiselli, F., Maurizii, M. G., & Passamonti, M. (2011). Doubly uniparental inheritance of mitochondria as a model system for studying germ line formation. *PLoS One*, 6(11), e28194.
- Milani, L., Ghiselli, F., Nuzhdin, S. V., & Passamonti, M. (2013). Nuclear genes with sex bias in *Ruditapes philippinarum* (bivalvia, veneridae): Mitochondrial inheritance and sex determination in dui species. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 320(7), 442–454.
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., & Lanfear, R. (2020). Iq-tree 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular biology and evolution*, 37(5), 1530–1534.
- Moreno, J. A., Dudchenko, O., Feigin, C. Y., Mereby, S. A., Chen, Z., Ramos, R., Almet, A. A., Sen, H., Brack, B. J., Johnson, M. R., et al. (2024). *Emx2* underlies the development and evolution of marsupial gliding membranes. *Nature*, 1–9.
- Mullon, C., Pomiankowski, A., & Reuter, M. (2012). Molecular evolution of *Drosophila Sex-lethal* and related sex determining genes. *BMC evolutionary biology*, 12, 1–11.
- Murat, F., Mbengue, N., Winge, S. B., Trefzer, T., Leushkin, E., Sepp, M., Cardoso-Moreira, M., Schmidt, J., Schneider, C., Mößinger, K., et al. (2023). The molecular evolution of spermatogenesis across mammals. *Nature*, 613(7943), 308–316.
- Murata, Y., Nikaido, M., Sasaki, T., Cao, Y., Fukumoto, Y., Hasegawa, M., & Okada, N. (2003). Afrotherian phylogeny as inferred from complete mitochondrial genomes. *Molecular phylogenetics and evolution*, 28(2), 253–260.
- Naimi, A., Martinez, A.-S., Specq, M.-L., Diss, B., Mathieu, M., & Sourdaine, P. (2009). Molecular cloning and gene expression of *Cg-Foxl2* during the development and the adult gametogenetic cycle in the oyster *Crassostrea gigas*. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 154(1), 134–142.
- Nakagawa, S., Gisselbrecht, S. S., Rogers, J. M., Hartl, D. L., & Bulyk, M. L. (2013). Dna-binding specificity changes in the evolution of forkhead transcription factors. *Proceedings of the National Academy of Sciences*, 110(30), 12349–12354.
- Natsidis, P., Kapli, P., Schiffer, P. H., & Telford, M. J. (2021). Systematic errors in orthology inference and their effects on evolutionary analyses. *Iscience*, 24(2).
- Nicolini, F., Ghiselli, F., Luchetti, A., & Milani, L. (2023). Bivalves as emerging model systems to study the mechanisms and evolution of sex determination: A genomic point of view. *Genome Biology and Evolution*, 15(10), evad181.

- Nicolini, F., Martelossi, J., Forni, G., Savojardo, C., Mantovani, B., & Luchetti, A. (2023). Comparative genomics of hox and parahox genes among major lineages of brachiopoda with emphasis on tadpole shrimps. *Frontiers in Ecology and Evolution*, 11, 1046960.
- Obata, M., & Komaru, A. (2005). Specific location of sperm mitochondria in mussel *Mytilus galloprovincialis* zygotes stained by mitotracker. *Development, growth & differentiation*, 47(4), 255–263.
- of Sex Consortium, T., et al. (2014). Tree of sex: A database of sexual systems. *Scientific Data*, 1.
- O'Neil, M. T., & Belote, J. M. (1992). Interspecific comparison of the *transformer* gene of *Drosophila* reveals an unusually high degree of evolutionary divergence. *Genetics*, 131(1), 113–128.
- Orkin, J. D., Montague, M. J., Tejada-Martinez, D., De Manuel, M., Del Campo, J., Cheves Hernandez, S., Di Fiore, A., Fontseré, C., Hodgson, J. A., Janiak, M. C., et al. (2021). The genomics of ecological flexibility, large brains, and long lives in capuchin monkeys revealed with fecalfacs. *Proceedings of the National Academy of Sciences*, 118(7), e2010632118.
- Paggeot, L. X., DeBiase, M. B., Escalona, M., Fairbairn, C., Marimuthu, M. P., Nguyen, O., Sahasrabudhe, R., & Dawson, M. N. (2022). Reference genome for the californian ribbed mussel, *Mytilus californianus*, an ecosystem engineer. *Journal of Heredity*, 113(6), 681–688.
- Pamilo, P., & O'Neill, R. (1997). Evolution of the *Sry* genes. *Molecular biology and evolution*, 14(1), 49–55.
- Panara, V., Budd, G. E., & Janssen, R. (2019). Phylogenetic analysis and embryonic expression of panarthropod dmrt genes. *Frontiers in zoology*, 16, 1–18.
- Papa, F., Windbichler, N., Waterhouse, R. M., Cagnetti, A., D'Amato, R., Persampieri, T., Lawniczak, M. K., Nolan, T., & Papathanos, P. A. (2017). Rapid evolution of female-biased genes among four species of *Anopheles* malaria mosquitoes. *Genome research*, 27(9), 1536–1548.
- Paris, M., Boyer, R., Jaenichen, R., Wolf, J., Karageorgi, M., Green, J., Cagnon, M., Parinello, H., Estoup, A., Gautier, M., et al. (2020). Near-chromosome level genome assembly of the fruit pest *Drosophila suzukii* using long-read sequencing. *Scientific reports*, 10(1), 11227.

- Parsch, J., & Ellegren, H. (2013). The evolutionary causes and consequences of sex-biased gene expression. *Nature Reviews Genetics*, 14(2), 83–87.
- Patnaik, B. B., Wang, T. H., Kang, S. W., Hwang, H.-J., Park, S. Y., Park, E. B., Chung, J. M., Song, D. K., Kim, C., Kim, S., et al. (2016). Sequencing, de novo assembly, and annotation of the transcriptome of the endangered freshwater pearl bivalve, *Cristaria plicata*, provides novel insights into functional genes and marker discovery. *PLoS One*, 11(2), e0148622.
- Penaloza, C., Gutierrez, A. P., Eöry, L., Wang, S., Guo, X., Archibald, A. L., Bean, T. P., & Houston, R. D. (2021). A chromosome-level genome assembly for the pacific oyster *Crassostrea gigas*. *GigaScience*, 10(3), giab020.
- Perez-Garcia, C., Moran, P., & Pasantes, J. J. (2011). Cytogenetic characterization of the invasive mussel species *Xenostrobus securis* lmk. (bivalvia: Mytilidae). *Genome*, 54(09), 771–778.
- Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., & Salzberg, S. L. (2016). Transcript-level expression analysis of rna-seq experiments with hisat, stringtie and ballgown. *Nature protocols*, 11(9), 1650–1667.
- Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., & Salzberg, S. L. (2015). Stringtie enables improved reconstruction of a transcriptome from rna-seq reads. *Nature biotechnology*, 33(3), 290–295.
- Phochanukul, N., & Russell, S. (2010). No backbone but lots of sox: Invertebrate sox genes. *The international journal of biochemistry & cell biology*, 42(3), 453–464.
- Piccinini, G., Iannello, M., Puccio, G., Plazzi, F., Havird, J. C., & Ghiselli, F. (2021). Mitonuclear coevolution, but not nuclear compensation, drives evolution of oxphos complexes in bivalves. *Molecular Biology and Evolution*, 38(6), 2597–2614.
- Powell, D., Subramanian, S., Suwansa-Ard, S., Zhao, M., O'Connor, W., Raftos, D., & Elizur, A. (2018). The genome of the oyster *Saccostrea* offers insight into the environmental resilience of bivalves. *DNA Research*, 25(6), 655–665.
- Purandare, S. R., Bickel, R. D., Jaquière, J., Rispe, C., & Brisson, J. A. (2014). Accelerated evolution of morph-biased genes in pea aphids. *Molecular biology and evolution*, 31(8), 2073–2083.
- Qi, H., Cong, R., Wang, Y., Li, L., & Zhang, G. (2023). Construction and analysis of the chromosome-level haplotype-resolved genomes of two *Crassostrea* oyster congeners: *Crassostrea angulata* and *Crassostrea gigas*. *GigaScience*, 12, giad077.

- Ran, Z., Li, Z., Yan, X., Liao, K., Kong, F., Zhang, L., Cao, J., Zhou, C., Zhu, P., He, S., et al. (2019). Chromosome-level genome assembly of the razor clam *Sinonovacula constricta* (lamarck, 1818). *Molecular ecology resources*, 19(6), 1647–1658.
- Ranz, J. M., Go, A. C., González, P. M., Clifton, B. D., Gomes, S., Jaberyzadeh, A., Woodbury, A., Chan, C., Gandasetiawan, K. A., Jayasekera, S., et al. (2023). Gene expression differentiation in the reproductive tissues of *Drosophila willistoni* subspecies and their hybrids. *Molecular Ecology*, 32(13), 3605–3623.
- Renschler, G., Richard, G., Valsecchi, C. I. K., Toscano, S., Arrigoni, L., Ramírez, F., & Akhtar, A. (2019). Hi-c guided assemblies reveal conserved regulatory topologies on x and autosomes despite extensive genome shuffling. *Genes & development*, 33(21-22), 1591–1612.
- Rogers, R. L., Grizzard, S. L., Titus-McQuillan, J. E., Bockrath, K., Patel, S., Wares, J. P., Garner, J. T., & Moore, C. C. (2021). Gene family amplification facilitates adaptation in freshwater unionid bivalve *Megalonaia nervosa*. *Molecular Ecology*, 30(5), 1155–1173.
- Rozewicki, J., Li, S., Amada, K. M., Standley, D. M., & Katoh, K. (2019). Mafft-dash: Integrated protein sequence and structural alignment. *Nucleic acids research*, 47(W1), W5–W10.
- Russo, C. A., Mello, B., Frazão, A., & Voloch, C. M. (2013). Phylogenetic analysis and a time tree for a large drosophilid data set (diptera: Drosophilidae). *Zoological Journal of the Linnean Society*, 169(4), 765–775.
- Sanchez-Flores, A., Peñaloza, F., Carpinteyro-Ponce, J., Nazario-Yepiz, N., Abreu-Goodger, C., Machado, C. A., & Markow, T. A. (2016). Genome evolution in three species of cactophilic *Drosophila*. *G3: Genes, Genomes, Genetics*, 6(10), 3097–3105.
- Sarkar, A., & Hochedlinger, K. (2013). The sox family of transcription factors: Versatile regulators of stem and progenitor cell fate. *Cell stem cell*, 12(1), 15–30.
- Schliep, K. P. (2011). Phangorn: Phylogenetic analysis in r. *Bioinformatics*, 27(4), 592–593.
- Schomburg, C., Janssen, R., & Prpic, N.-M. (2022). Phylogenetic analysis of forkhead transcription factors in the panarthropoda. *Development genes and evolution*, 232(1), 39–48.
- Seudre, O., Martín-Zamora, F. M., Rapisarda, V., Luqman, I., Carrillo-Baltodano, A. M., & Martín-Durán, J. M. (2022). The Fox gene repertoire in the annelid *Owenia fusiformis* reveals multiple expansions of the foxQ2 class in spiralia. *Genome Biology and Evolution*, 14(10), evac139.

- Shi, J., Hong, Y., Sheng, J., Peng, K., & Wang, J. (2015). De novo transcriptome sequencing to identify the sex-determination genes in *Hyriopsis schlegelii*. *Bioscience, Biotechnology, and Biochemistry*, 79(8), 1257–1265.
- Shi, Y., Liu, W., & He, M. (2018). Proteome and transcriptome analysis of ovary, intersex gonads, and testis reveals potential key sex reversal/differentiation genes and mechanism in scallop *Chlamys nobilis*. *Marine Biotechnology*, 20, 220–245.
- Shimeld, S. M., Boyle, M. J., Brunet, T., Luke, G. N., & Seaver, E. C. (2010). Clustered fox genes in lophotrochozoans and the evolution of the bilaterian fox gene cluster. *Developmental biology*, 340(2), 234–248.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Molecular systems biology*, 7(1), 539.
- Smith, C. H. (2021). A high-quality reference genome for a parasitic bivalve with doubly uniparental inheritance (bivalvia: Unionida). *Genome Biology and Evolution*, 13(3), evab029.
- Song, H., Guo, X., Sun, L., Wang, Q., Han, F., Wang, H., Wray, G. A., Davidson, P., Wang, Q., Hu, Z., et al. (2021). The hard clam genome reveals massive expansion and diversification of inhibitors of apoptosis in bivalvia. *BMC biology*, 19, 1–20.
- Song, H., Wang, Y., Shao, H., Li, Z., Hu, P., Yap-Chiongco, M. K., Shi, P., Zhang, T., Li, C., Wang, Y., et al. (2023). Scaphopoda is the sister taxon to bivalvia: Evidence of ancient incomplete lineage sorting. *Proceedings of the National Academy of Sciences*, 120(40), e2302361120.
- Stammnitz, M. R., Gori, K., Kwon, Y. M., Harry, E., Martin, F. J., Billis, K., Cheng, Y., Baez-Ortega, A., Chow, W., Comte, S., et al. (2023). The evolution of two transmissible cancers in tasmanian devils. *Science*, 380(6642), 283–293.
- Stothard, P., & Pilgrim, D. (2003). Sex-determination gene and pathway evolution in nematodes. *Bioessays*, 25(3), 221–231.
- Sun, D., Yu, H., & Li, Q. (2022). Examination of the roles of *Foxl2* and *Dmrt1* in sex differentiation and gonadal development of oysters by using rna interference. *Aquaculture*, 548, 737732.
- Sun, J., Zhang, Y., Xu, T., Zhang, Y., Mu, H., Zhang, Y., Lan, Y., Fields, C. J., Hui, J. H. L., Zhang, W., et al. (2017). Adaptation to deep-sea chemosynthetic environments as revealed by mussel genomes. *Nature ecology & evolution*, 1(5), 0121.

- Swanson, W. J., & Vacquier, V. D. (2002). The rapid evolution of reproductive proteins. *Nature reviews genetics*, 3(2), 137–144.
- Teaniniuraitemoana, V., Huvet, A., Levy, P., Klopp, C., Lhuillier, E., Gaertner-Mazouni, N., Gueguen, Y., & Le Moullac, G. (2014). Gonad transcriptome analysis of pearl oyster *Pinctada margaritifera*: Identification of potential sex differentiation and sex determining genes. *BMC genomics*, 15, 1–20.
- Teng, W., Fu, H., Li, Z., Zhang, Q., Xu, C., Yu, H., Kong, L., Liu, S., & Li, Q. (2023). Parallel evolution in *Crassostrea* oysters along the latitudinal gradient is associated with variation in multiple genes involved in adipogenesis. *Molecular Ecology*, 32(19), 5276–5287.
- Tu, Q., Brown, C. T., Davidson, E. H., & Oliveri, P. (2006). Sea urchin forkhead gene family: Phylogeny and embryonic expression. *Developmental biology*, 300(1), 49–62.
- Tvedte, E. S., Gasser, M., Sparklin, B. C., Michalski, J., Hjelmen, C. E., Johnston, J. S., Zhao, X., Bromley, R., Tallon, L. J., Sadzewicz, L., et al. (2021). Comparison of long-read sequencing technologies in interrogating bacteria and fly genomes. *G3*, 11(6), jkab083.
- Uller, T., & Helanterä, H. (2011). From the origin of sex-determining factors to the evolution of sex-determining systems. *The Quarterly review of biology*, 86(3), 163–180.
- Verhulst, E. C., van de Zande, L., & Beukeboom, L. W. (2010). Insect sex determination: It all evolves around *transformer*. *Current opinion in genetics & development*, 20(4), 376–383.
- Vicoso, B., & Charlesworth, B. (2006). Evolution on the x chromosome: Unusual patterns and processes. *Nature Reviews Genetics*, 7(8), 645–653.
- Vilstrup, J. T., Seguin-Orlando, A., Stiller, M., Ginolhac, A., Raghavan, M., Nielsen, S. C., Weinstock, J., Froese, D., Vasiliev, S. K., Ovodov, N. D., et al. (2013). Mitochondrial phylogenomics of modern and ancient equids. *PloS one*, 8(2), e55950.
- Vinkler, M., Fiddaman, S. R., Těšický, M., O'Connor, E. A., Savage, A. E., Lenz, T. L., Smith, A. L., Kaufman, J., Bolnick, D. I., Davies, C. S., et al. (2023). Understanding the evolution of immune genes in jawed vertebrates. *Journal of evolutionary biology*, 36(6), 847–873.
- Vizueta Moraga, J., Sánchez-Gracia, A., & Rozas Liras, J. A. (2020). Bitacora: A comprehensive tool for the identification and annotation of gene families in genome assemblies. *Molecular Ecology Resources*, 2020, vol. 20, num. 5, p. 1445-1452.
- Wang, C., Wallerman, O., Arendt, M.-L., Sundström, E., Karlsson, Å., Nordin, J., Mäkeläinen, S., Pielberg, G. R., Hanson, J., Ohlsson, Å., et al. (2021). A novel canine reference

- genome resolves genomic architecture and uncovers transcript complexity. *Communications biology*, 4(1), 185.
- Wang, G., Dong, S., Guo, P., Cui, X., Duan, S., & Li, J. (2020). Identification of *Foxl2* in freshwater mussel *Hyriopsis cumingii* and its involvement in sex differentiation. *Gene*, 754, 144853.
- Wang, J., & Nie, H. (2024). Genome-wide identification and expression analysis of sox gene family in the manila clam (*Ruditapes philippinarum*). *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 50, 101244.
- Wang, M., Xia, J., Jawad, M., Wei, W., Gui, L., Liang, X., Yang, J.-L., & Li, M. (2022). Transcriptome sequencing analysis of sex-related genes and miRNAs in the gonads of *Mytilus coruscus*. *Frontiers in Marine Science*, 9, 1013857.
- Wang, Q., Cao, T., & Wang, C. (2023). Genome-wide identification and expression analysis of dmrt genes in bivalves. *BMC genomics*, 24(1), 457.
- Wang, S., Zhang, J., Jiao, W., Li, J., Xun, X., Sun, Y., Guo, X., Huan, P., Dong, B., Zhang, L., et al. (2017). Scallop genome provides insights into evolution of bilaterian karyotype and development. *Nature ecology & evolution*, 1(5), 0120.
- Wang, X., Werren, J. H., & Clark, A. G. (2015). Genetic and epigenetic architecture of sex-biased expression in the jewel wasps *Nasonia vitripennis* and *giraulti*. *Proceedings of the National Academy of Sciences*, 112(27), E3545–E3554.
- Wang, X., Liu, Z., & Wu, W. (2017). Transcriptome analysis of the freshwater pearl mussel (*Cristaria plicata*) mantle unravels genes involved in the formation of shell and pearl. *Molecular genetics and genomics*, 292, 343–352.
- Wei, M., Ge, H., Shao, C., Yan, X., Nie, H., Duan, H., Liao, X., Zhang, M., Chen, Y., Zhang, D., et al. (2020). Chromosome-level clam genome helps elucidate the molecular basis of adaptation to a buried lifestyle. *IScience*, 23(6).
- Wexler, J. R., Plachetzki, D. C., & Kopp, A. (2014). Pan-metazoan phylogeny of the dmrt gene family: A framework for functional studies. *Development Genes and Evolution*, 224, 175–181.
- Whitfield, L. S., Lovell-Badge, R., & Goodfellow, P. N. (1993). Rapid sequence evolution of the mammalian sex-determining gene sry. *Nature*, 364 (6439), 713–715.
- Wilkins, A. S. (1995). Moving up the hierarchy: A hypothesis on the evolution of a genetic sex determination pathway. *Bioessays*, 17(1), 71–77.

- Willson, J., Roddur, M. S., Liu, B., Zaharias, P., & Warnow, T. (2022). Disco: Species tree inference using multicopy gene family tree decomposition. *Systematic biology*, 71(3), 610–629.
- Wu, S., Zhang, Y., Li, Y., Wei, H., Guo, Z., Wang, S., Zhang, L., & Bao, Z. (2020). Identification and expression profiles of fox transcription factors in the yesso scallop (*patinopecten yessoensis*). *Gene*, 733, 144387.
- Xiong, Y., Brandley, M. C., Xu, S., Zhou, K., & Yang, G. (2009). Seven new dolphin mitochondrial genomes and a time-calibrated phylogeny of whales. *BMC Evolutionary Biology*, 9, 1–13.
- Xu, R., Iannello, M., Havird, J. C., Milani, L., & Ghiselli, F. (2022). Lack of transcriptional coordination between mitochondrial and nuclear oxidative phosphorylation genes in the presence of two divergent mitochondrial genomes. *Zoological Research*, 43(1), 111.
- Xu, R., Martelossi, J., Smits, M., Iannello, M., Peruzza, L., Babbucci, M., Milan, M., Dunham, J. P., Breton, S., Milani, L., et al. (2022). Multi-tissue rna-seq analysis and long-read-based genome assembly reveal complex sex-specific gene regulation and molecular evolution in the manila clam. *Genome Biology and Evolution*, 14(12), evac171.
- Yang, J.-L., Feng, D.-D., Liu, J., Xu, J.-K., Chen, K., Li, Y.-F., Zhu, Y.-T., Liang, X., & Lu, Y. (2021). Chromosome-level genome assembly of the hard-shelled mussel *Mytilus coruscus*, a widely distributed species from the temperate areas of east asia. *Gigascience*, 10(4), giab024.
- Yang, M., Xu, F., Liu, J., Que, H., Li, L., & Zhang, G. (2014). Phylogeny of forkhead genes in three spiralians and their expression in pacific oyster *Crassostrea gigas*. *Chinese journal of oceanology and limnology*, 32(6), 1207–1223.
- Yu, J., Zhang, L., Li, Y., Li, R., Zhang, M., Li, W., Xie, X., Wang, S., Hu, X., & Bao, Z. (2017). Genome-wide identification and expression profiling of the sox gene family in a bivalve mollusc *Patinopecten yessoensis*. *Gene*, 627, 530–537.
- Yu, J.-K., Mazet, F., Chen, Y.-T., Huang, S.-W., Jung, K.-C., & Shimeld, S. M. (2008). The fox genes of *Branchiostoma floridae*. *Development genes and evolution*, 218, 629–638.
- Yue, C., Li, Q., & Yu, H. (2021). Variance in expression and localization of sex-related genes *CgDsx*, *CgBHMG1* and *CgFoxl2* during diploid and triploid pacific oyster *Crassostrea gigas* gonad differentiation. *Gene*, 790, 145692.
- Zeng, Y., Zheng, H., He, C., Zhang, C., Zhang, H., & Zheng, H. (2024). Genome-wide identification and expression analysis of dmrt gene family and their role in gonad development

- of pacific oyster (*Crassostrea gigas*). *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 269, 110904.
- Zhang, N., Xu, F., & Guo, X. (2014). Genomic analysis of the pacific oyster (*Crassostrea gigas*) reveals possible conservation of vertebrate sex determination in a mollusc. *G3: Genes, Genomes, Genetics*, 4(11), 2207–2217.
- Zhu, C., Zhang, L., Ding, H., & Pan, Z. (2019). Transcriptome-wide identification and characterization of the *Sox* gene family and microsatellites for *Corbicula fluminea*. *PeerJ*, 7, e7770.

# **Appendix**

The appendix includes the titles and abstracts of the papers published during my PhD that are not part of this thesis.

# Taxonomic revision of the Australian stick insect genus *Candovia* (Phasmida: Necrosciinae): insight from molecular systematics and species-delimitation approaches.

Giobbe Forni<sup>1,2</sup>, Alex Cussigh<sup>1,2</sup>, Paul D. Brock<sup>3</sup>, Braxton R. Jones<sup>4</sup>, Filippo Nicolini<sup>1</sup>, Jacopo Martelossi<sup>1</sup>, Andrea Luchetti<sup>1</sup>, Barbara Mantovani<sup>1</sup>

<sup>1</sup>*Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, Italy.*

<sup>2</sup>*Department of Agricultural and Environmental Sciences, University of Milan, Milano, Italy.*

<sup>3</sup>*The Natural History Museum, Cromwell Road, London, UK.*

<sup>4</sup>*School of Life and Environmental Sciences, The University of Sydney, Sydney NSW 2006, Australia.*

**Published in:** 2023, *Zoological Journal of the Linnean Society*, 197:189–210.  
10.1093/zoolinnean/zlac074

**Abstract.** The Phasmida genus *Candovia* comprises nine traditionally recognized species, all endemic to Australia. In this study, *Candovia* diversity is explored through molecular species-delimitation analyses using the *COI<sub>Fol</sub>* gene fragment and phylogenetic inferences leveraging seven additional mitochondrial and nuclear loci. Molecular results were integrated with morphological observations, leading us to confirm the already described species and to the delimitation of several new taxa and of the new genus *Paracandovia*. New *Candovia* species from various parts of Queensland and New South Wales are described and illustrated (*C. alata* sp. nov., *C. byfieldensis* sp. nov., *C. dalgleishae* sp. nov., *C. eungellensis* sp. nov., *C. karasi* sp. nov., *C. koensi* sp. nov. and *C. wollumbinensis* sp. nov.). New combinations are proposed and species removed from synonymy with the erection of the new genus *Paracandovia* (*P. cercata* stat. rev., comb. nov., *P. longipes* stat. rev., comb. nov., *P. pallida* comb. nov., *P. peridromes* comb. nov., *P. tenera* stat. rev., comb. nov.). Phylogenetic analyses suggest that the egg capitulum may have independently evolved multiple times throughout the evolutionary history of these insects. Furthermore, two newly described species represent the first taxa with fully developed wings in this previously considered apterous clade.

# Comparative genomics of *Hox* and *ParaHox* genes among major lineages of Branchiopoda with emphasis on tadpole shrimps.

Filippo Nicolini<sup>1,2</sup>, Jacopo Martelossi<sup>1</sup>, Giobbe Forni<sup>3</sup>,  
Castrense Savojardo<sup>4</sup>, Barbara Mantovani<sup>1</sup>, Andrea Luchetti<sup>1</sup>

<sup>1</sup>*Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, Italy.*

<sup>2</sup>*Fano Marine Center, Fano (PU), Italy.*

<sup>3</sup>*Department of Agricultural and Environmental Sciences, University of Milan, Milan, Italy.*

<sup>4</sup>*Department of Pharmacy and Biotechnology, University of Bologna, Bologna, Italy.*

**Published in:** 2023, *Frontiers in Ecology and Evolution*, 11:1046960.  
10.3389/fevo.2023.1046960

**Abstract.** *Hox* and *ParaHox* genes (HPHGs) are key developmental genes that pattern regional identity along the anterior–posterior body axis of most animals. Here, we identified HPHGs in tadpole shrimps (Pancrustacea, Branchiopoda, Notostraca), an iconic example of the so-called “living fossils” and performed a comparative genomics analysis of HPHGs and the *Hox* cluster among major branchiopod lineages. Notostraca possess the entire *Hox* complement, and the *Hox* cluster seems to be split into two different subclusters, although we were not able to support this finding with chromosome-level assemblies. However, the genomic structure of *Hox* genes in Notostraca appears more derived than that of *Daphnia* spp., which instead retains the plesiomorphic condition of a single compact cluster. Spinicaudata and *Artemia franciscana* show instead a *Hox* cluster subdivided across two or more genomic scaffolds with some orthologs either duplicated or missing. Yet, branchiopod HPHGs are similar among the various clades in terms of both intron length and number, as well as in their pattern of molecular evolution. Sequence substitution rates are in fact generally similar for most of the branchiopod *Hox* genes and the few differences we found cannot be traced back to natural selection, as they are not associated with any signals of diversifying selection or substantial switches in selective modes. Altogether, these findings do not support a significant stasis in the Notostraca *Hox* cluster and further confirm how morphological evolution is not tightly associated with genome dynamics.

# Multiple and diversified transposon lineages contribute to early and recent bivalve genome evolution.

Jacopo Martelossi<sup>1</sup>, Filippo Nicolini<sup>1,2</sup>, Simone Subacchi<sup>1</sup>, Daniela Pasquale<sup>1</sup>, Fabrizio Ghiselli<sup>1</sup>, Andrea Luchetti<sup>1</sup>

<sup>1</sup>*Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, Italy.*

<sup>2</sup>*Fano Marine Center, Fano (PU), Italy.*

**Published in:** 2023, *BMC Biology*, 21:145.

10.1186/s12915-023-01632-z

**Abstract.** **Background.** Transposable elements (TEs) can represent one of the major sources of genomic variation across eukaryotes, providing novel raw materials for species diversification and innovation. While considerable effort has been made to study their evolutionary dynamics across multiple animal clades, molluscs represent a substantially understudied phylum. Here, we take advantage of the recent increase in mollusc genomic resources and adopt an automated TE annotation pipeline combined with a phylogenetic tree-based classification, as well as extensive manual curation efforts, to characterize TE repertoires across 27 bivalve genomes with a particular emphasis on DDE/D class II elements, long interspersed nuclear elements (LINEs), and their evolutionary dynamics. **Results.** We found class I elements as highly dominant in bivalve genomes, with LINE elements, despite less represented in terms of copy number per genome, being the most common retroposon group covering up to 10% of their genome. We mined 86,488 reverse transcriptases (RVT) containing LINE coming from 12 clades distributed across all known superfamilies and 14,275 class II DDE/D-containing transposons coming from 16 distinct superfamilies. We uncovered a previously underestimated rich and diverse bivalve ancestral transposon complement that could be traced back to their most recent common ancestor that lived about 500 Mya. Moreover, we identified multiple instances of lineage-specific emergence and loss of different LINEs and DDE/D lineages with the interesting cases of CR1-Zenon, Proto2, RTE-X, and Academ elements that underwent a bivalve-specific amplification likely associated with their diversification. Finally, we found that

this LINE diversity is maintained in extant species by an equally diverse set of long-living and potentially active elements, as suggested by their evolutionary history and transcription profiles in both male and female gonads. **Conclusions.** We found that bivalves host an exceptional diversity of transposons compared to other molluscs. Their LINE complement could mainly follow a “stealth drivers” model of evolution where multiple and diversified families are able to survive and co-exist for a long period of time in the host genome, potentially shaping both recent and early phases of bivalve genome evolution and diversification. Overall, we provide not only the first comparative study of TE evolutionary dynamics in a large but understudied phylum such as Mollusca, but also a reference library for ORF-containing class II DDE/D and LINE elements, which represents an important genomic resource for their identification and characterization in novel genomes.

# Towards a time-tree solution for Branchiopoda diversification: a jackknife assessment of fossil age priors.

Niccolò Righetti<sup>1\*</sup>, Filippo Nicolini<sup>2\*</sup>, Giobbe Forni<sup>2</sup>, Andrea Luchetti<sup>2</sup>

<sup>1</sup>*Laboratoire de Biologie Computationnelle et Quantitative (LCQB), Sorbonne Université, CNRS, IBPS, UMR7238, Paris, France.*

<sup>2</sup>*Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, Italy.*

\* the authors equally contributed to this work.

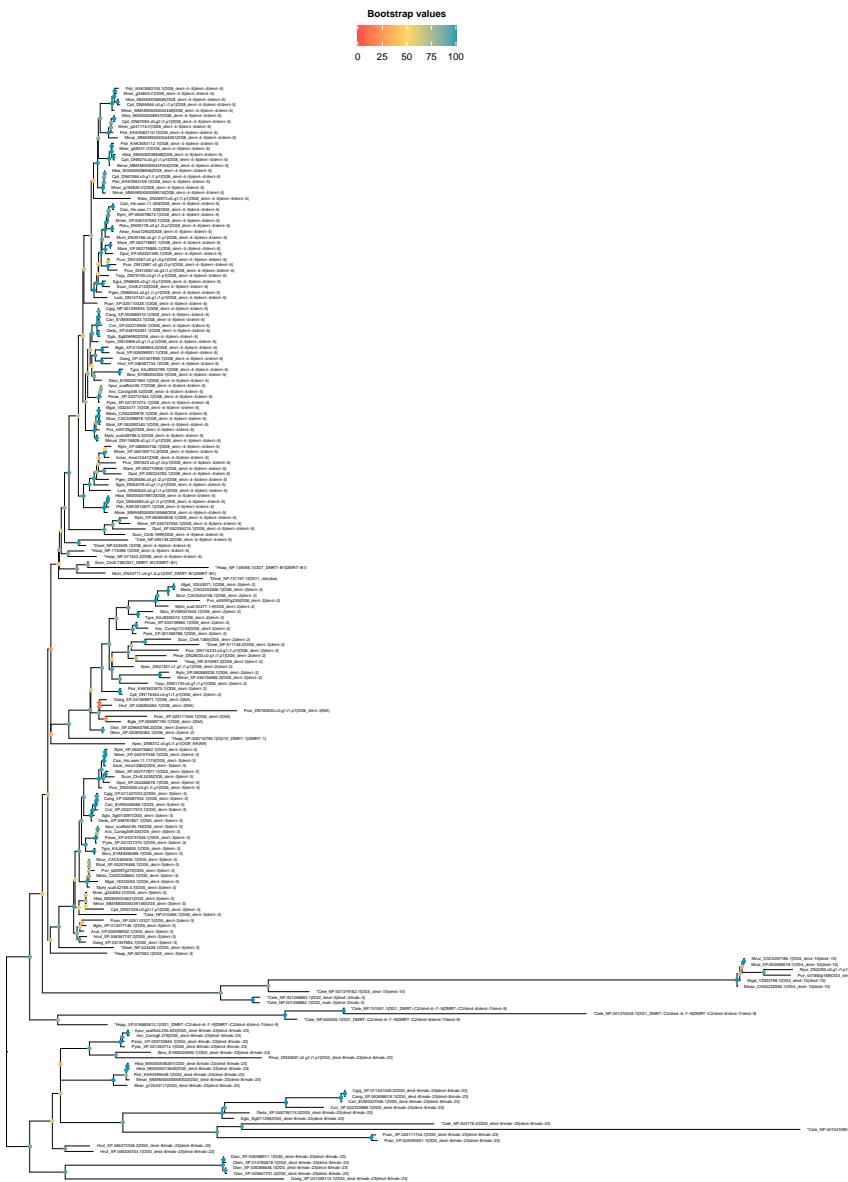
***Submitted for peer-review.***

**Abstract.** An understanding of Branchiopoda's evolutionary history is crucial for a comprehensive knowledge of the Pancrustacea tree of life, given their close evolutionary relationship with Hexapoda. Despite significant advances in molecular and morphological phylogenetics that have resolved much of the branchiopod backbone topology, a reliable temporal framework remains elusive. Key challenges include a sparse fossil record, long-term morphological stasis, and past topological inconsistencies. Leveraging a Bayesian Inference approach and the most extensive phylogenomic dataset for branchiopod to date, encompassing 46 species and over 130 genes, we inferred a time-calibrated phylogenetic tree. Furthermore, to strengthen the confidence in our divergence times estimation, we assessed the impact of age priors, topological uncertainties, and gene trees which are discordant from the species trees. Our results are largely consistent with the fossil record and with previous studies, indicating that Branchiopoda originated between 400 and 500 million years ago, and the orders of large branchiopods diversified during the Mesozoic. Concerning Cladocera, results remain problematic, with a sharper uncertainty in the diversification time with respect to the fossil record. Though, the jackknife resampling of fossils and the other sensitivity analyses proved our calibration method to be robust, suggesting that the difficulties in obtaining a paleontological-consistent time tree may be hindered by the variability in branchiopod substitution rates and topological instability within certain clades.

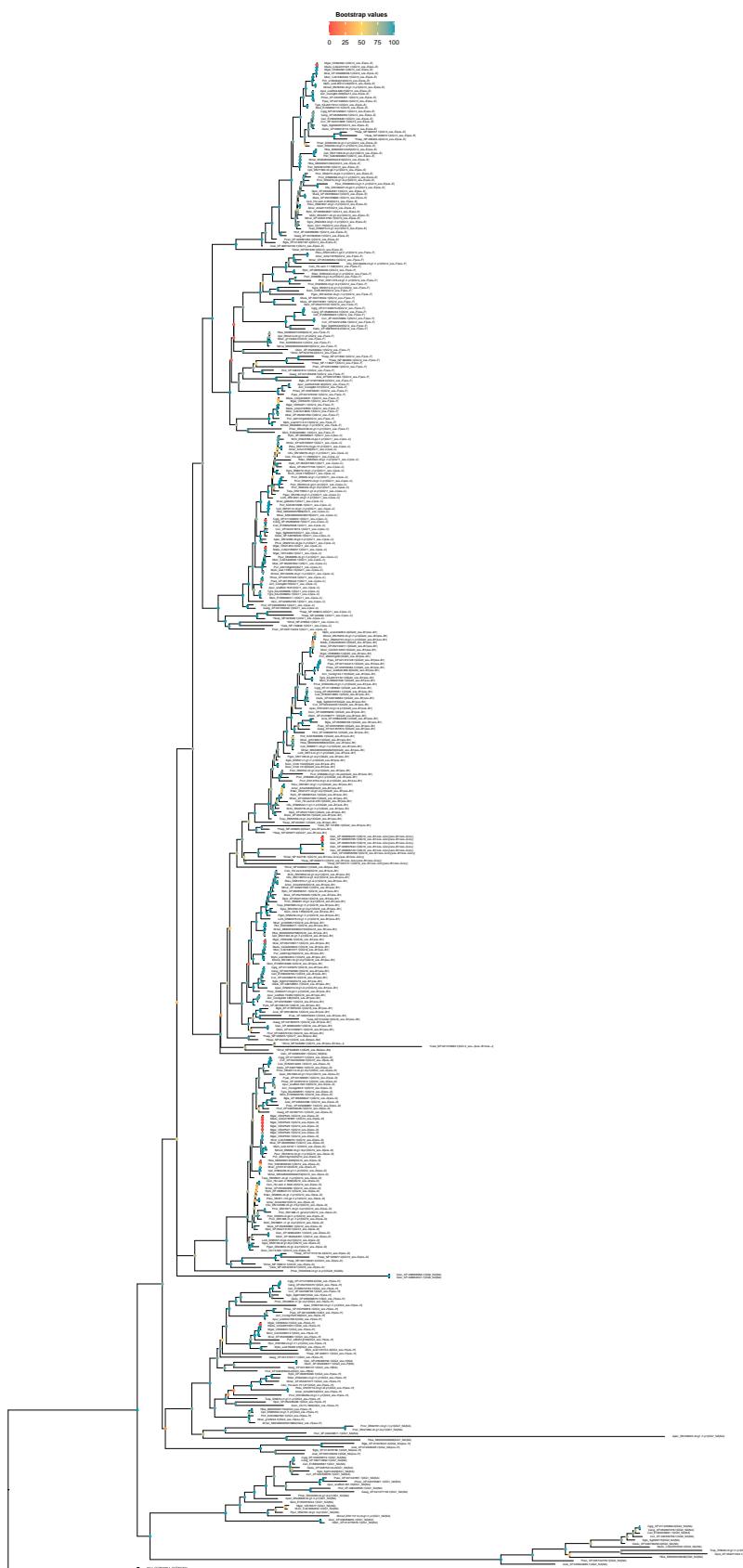
# Supplementary materials

## Supplementary figures

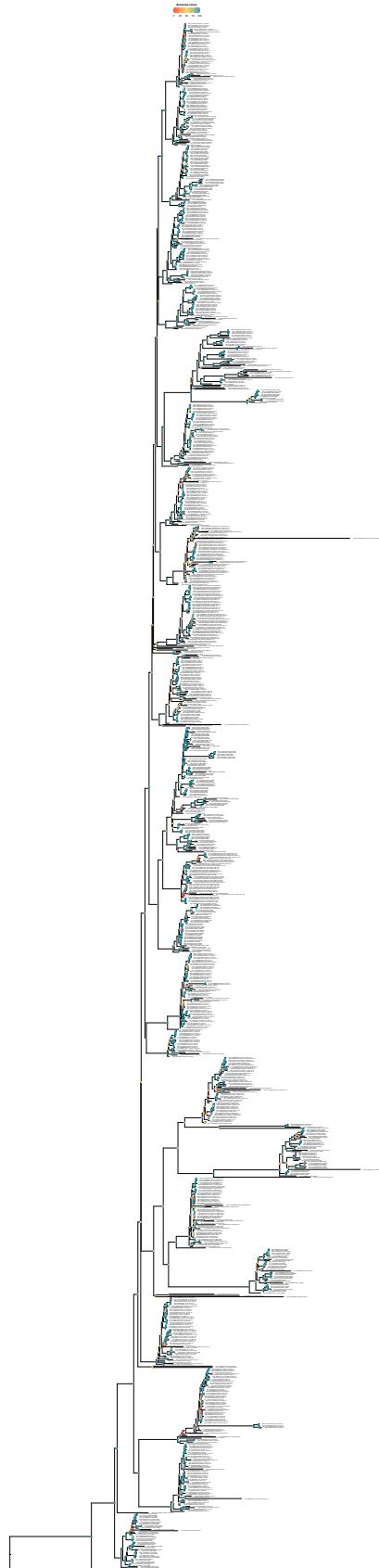
High-quality supplementary figures are available at the following GitHub repository: [LINK](#) [LINK](#) [LINK](#).



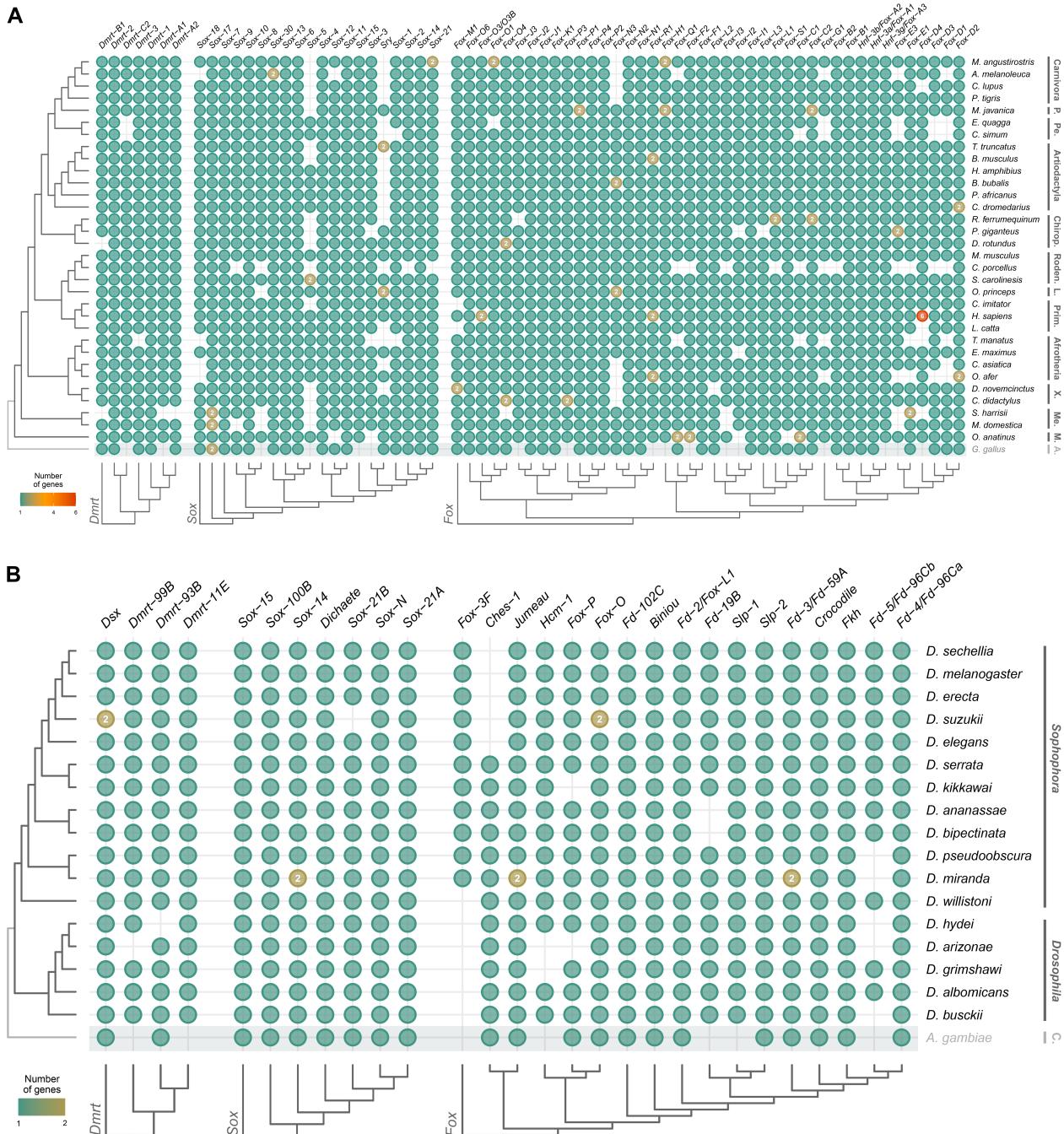
**Supplementary Figure S1. maximum likelihood (ML) phylogenetic tree of the Dmrt gene family in molluscs, including the possvm orthology inference.** Reference genes from *Homo sapiens*, *Caenorhabditis elegans*, and *Drosophila melanogaster* are marked with an asterisk at the beginning of the tip names. Species ID can be found in **Supp. Tab. S1**. The tree has been midpoint rooted. Bootstrap values are shown for each node.



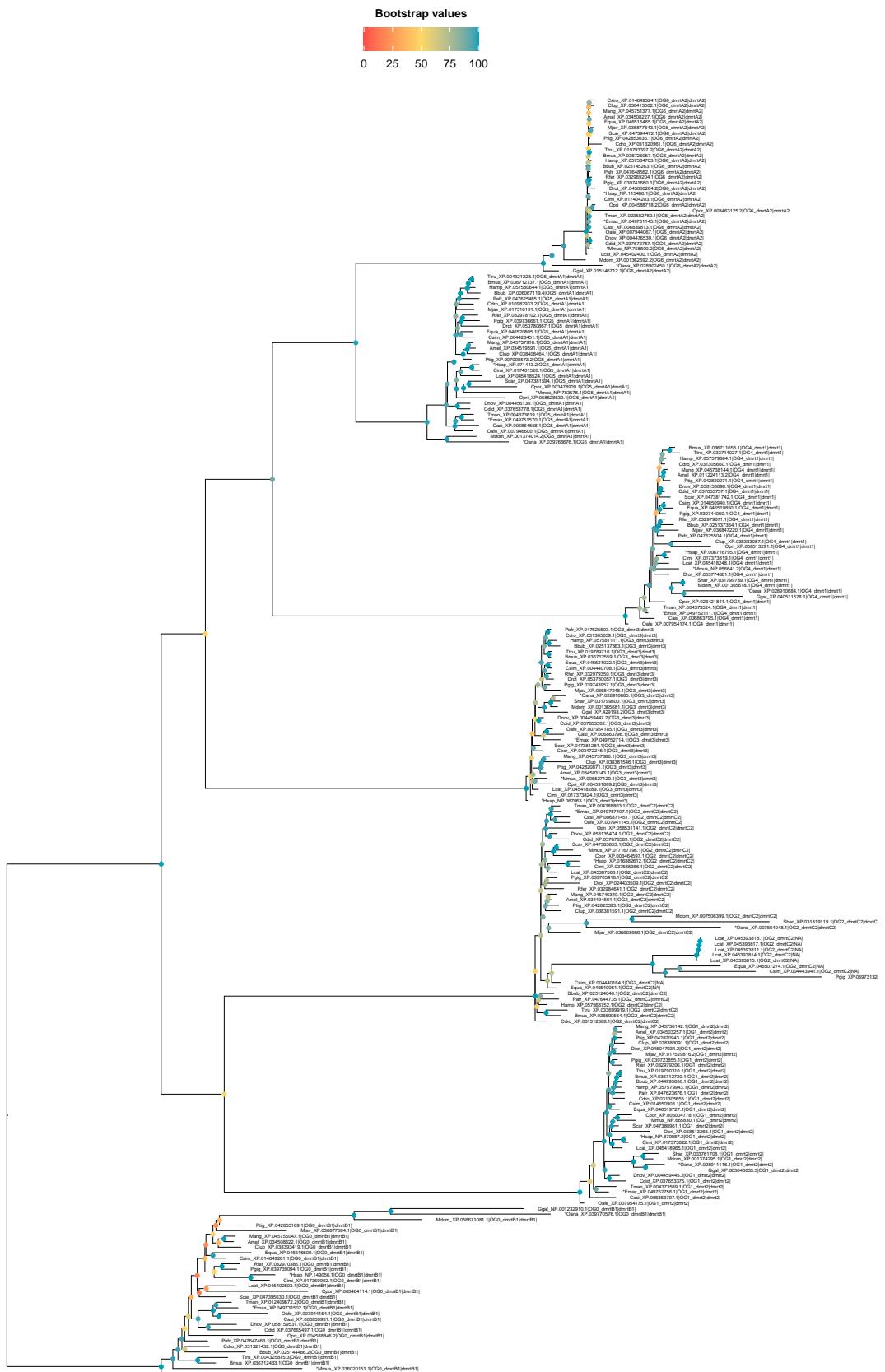
**Supplementary Figure S2.** ML phylogenetic tree of the Sox gene family in molluscs, including the possvm orthology inference. Reference genes from *H. sapiens*, *C. elegans*, and *D. melanogaster* are marked with an asterisk at the beginning of the tip names. Species ID can be found in Supp. Tab. S1. Bootstrap values are shown for each node.



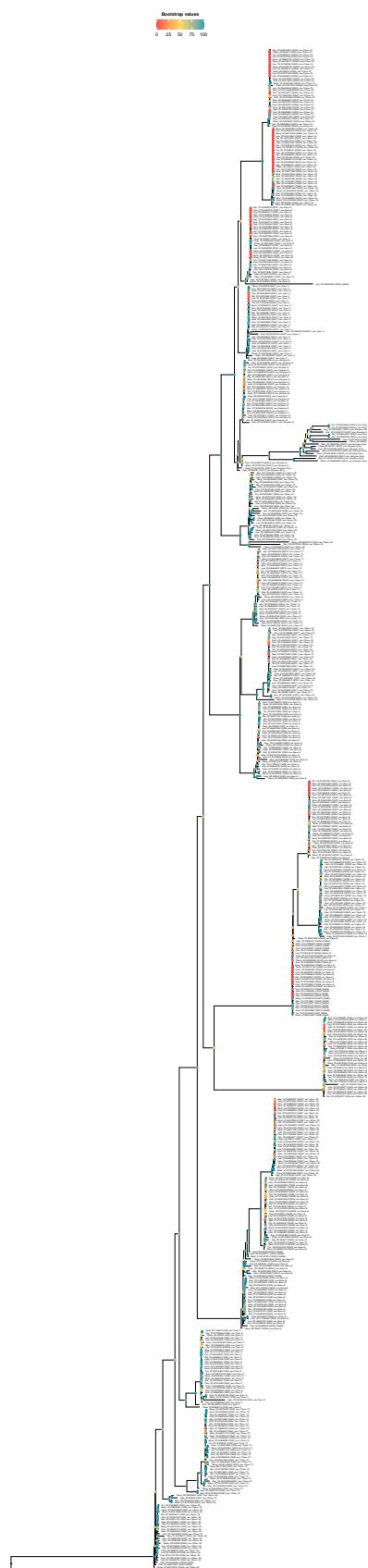
**Supplementary Figure S3.** ML phylogenetic tree of the Fox gene family in molluscs, including the possvm orthology inference. Reference genes from *H. sapiens*, *C. elegans*, and *D. melanogaster* are marked with an asterisk at the beginning of the tip names. Species ID can be found in **Supp. Tab. S1**. Bootstrap values are shown for each node.



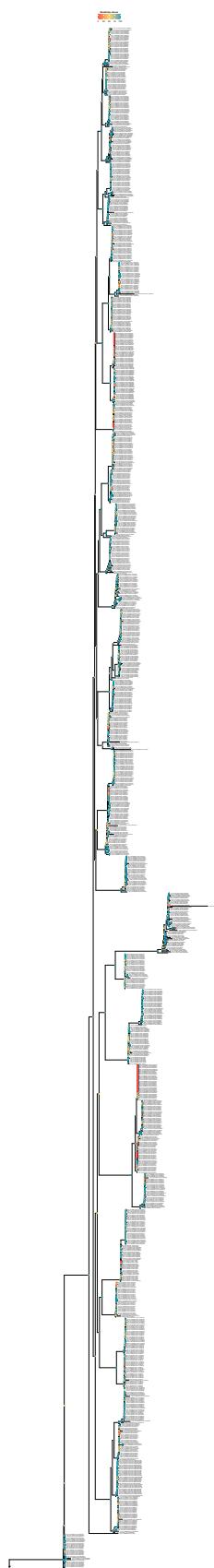
**Supplementary Figure S4. The DSFG complement in Mammalia and *Drosophila* spp.**  
 Presence/absence of genes in various species are indicated by filled circles. Numbers inside each circle specify genes with 2 or more copies. The shaded area highlights outgroup species, *Gallus gallus* (Aves) for mammals and *Anopheles gambiae* (Culicidae) for fruit flies. The phylogenetic tree of analysed species, as inferred from literature, is shown on the left, while major taxonomic groups are reported on the right. All species are represented by genomic data. Dmrt, Sox, and Fox gene (DSFG) trees are shown on the bottom (full trees can be found in **Supp. Fig. S5–S7**). Full species names for both mammals and fruit flies, along with all assembly and taxonomic information, can be found in **Supp. Tab. S4** and **Supp. Tab. S5**, respectively. A.: Aves; Chiroptera; L.: Lagomorpha; M.: Monotremata; Me.: Metatheria; P.: Pholidota; Pe.: Perissodactyla; Prim.: Primates; Roden.: Rodentia; X.: Xenarthra; C.: Culicidae.



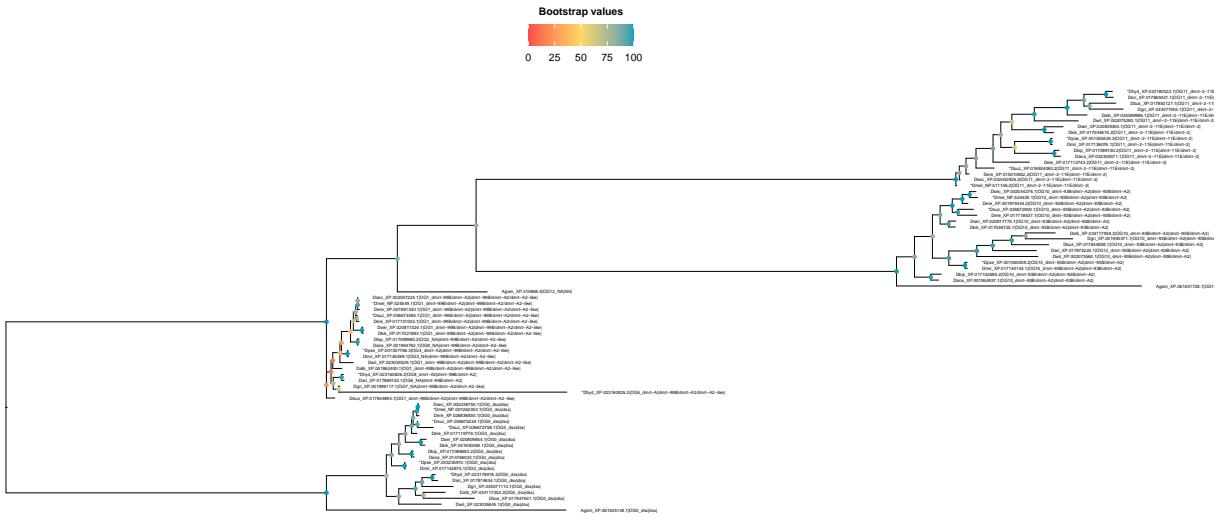
**Supplementary Figure S5. ML phylogenetic tree of the *dsx* and *mab-3* related transcription factor (Dmrt) gene family in mammals, including the Possvm orthology inference.** Reference genes from *H. sapiens*, *Mus musculus*, *Elephas maximus indicus*, and *Ornithorhynchus anatinus* are marked with an asterisk at the beginning of the tip names. Species ID can be found in Supp. Tab. S4. The tree has been midpoint rooted. Bootstrap values are shown for each node.



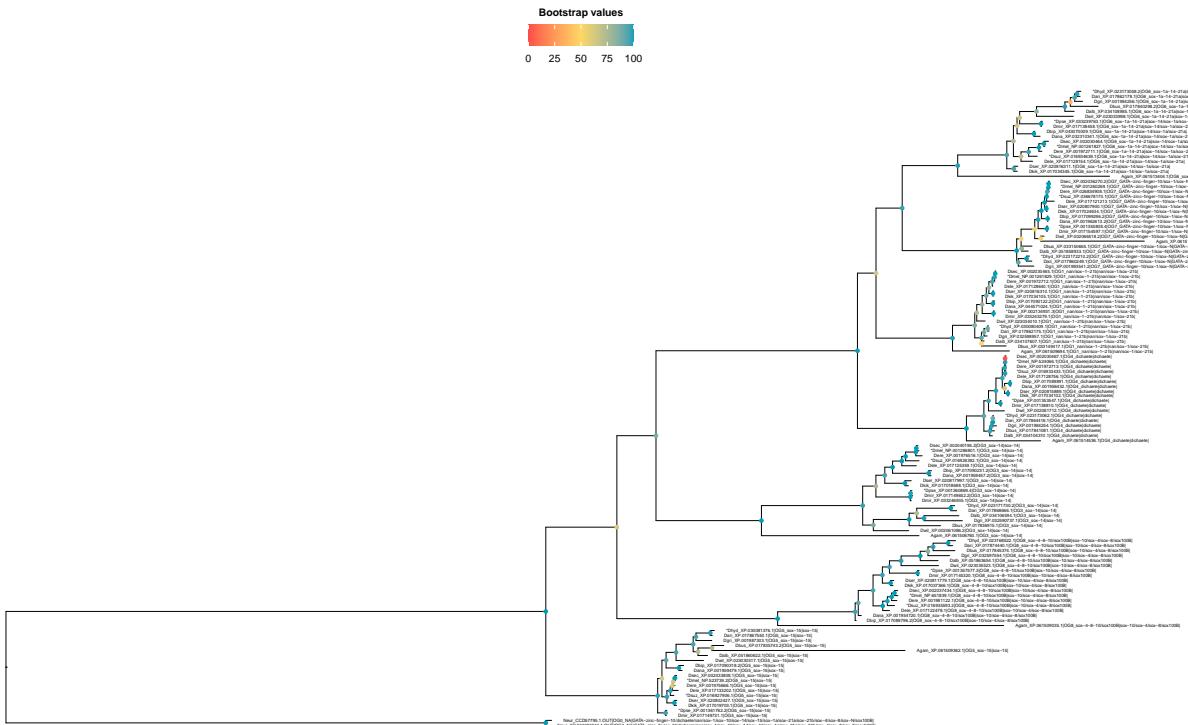
**Supplementary Figure S6. ML phylogenetic tree of the *Sry*-related HMG-box (Sox) gene family in mammals, including the Possvm orthology inference.** Reference genes from *H. sapiens*, *M. musculus*, *E. maximus indicus*, and *O. anatinus* are marked with an asterisk at the beginning of the tip names. Species ID can be found in **Supp. Tab. S4**. Bootstrap values are shown for each node.



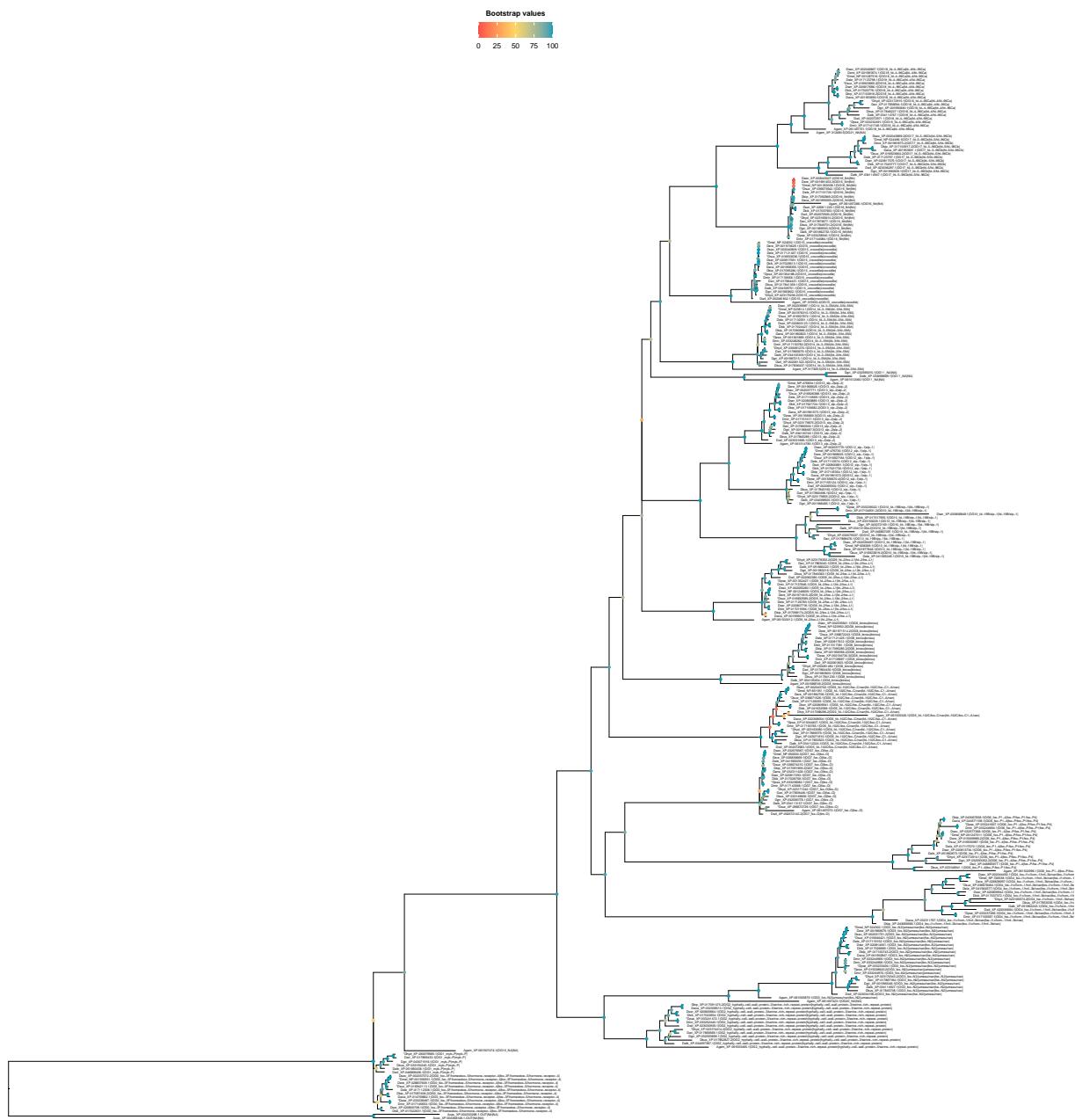
**Supplementary Figure S7. ML phylogenetic tree of the forkhead box (Fox) gene family in mammals, including the Possvm orthology inference.** Reference genes from *H. sapiens*, *M. musculus*, *E. maximus indicus*, and *O. anatinus* are marked with an asterisk at the beginning of the tip names. Species ID can be found in **Supp. Tab. S4**. Bootstrap values are shown for each node.



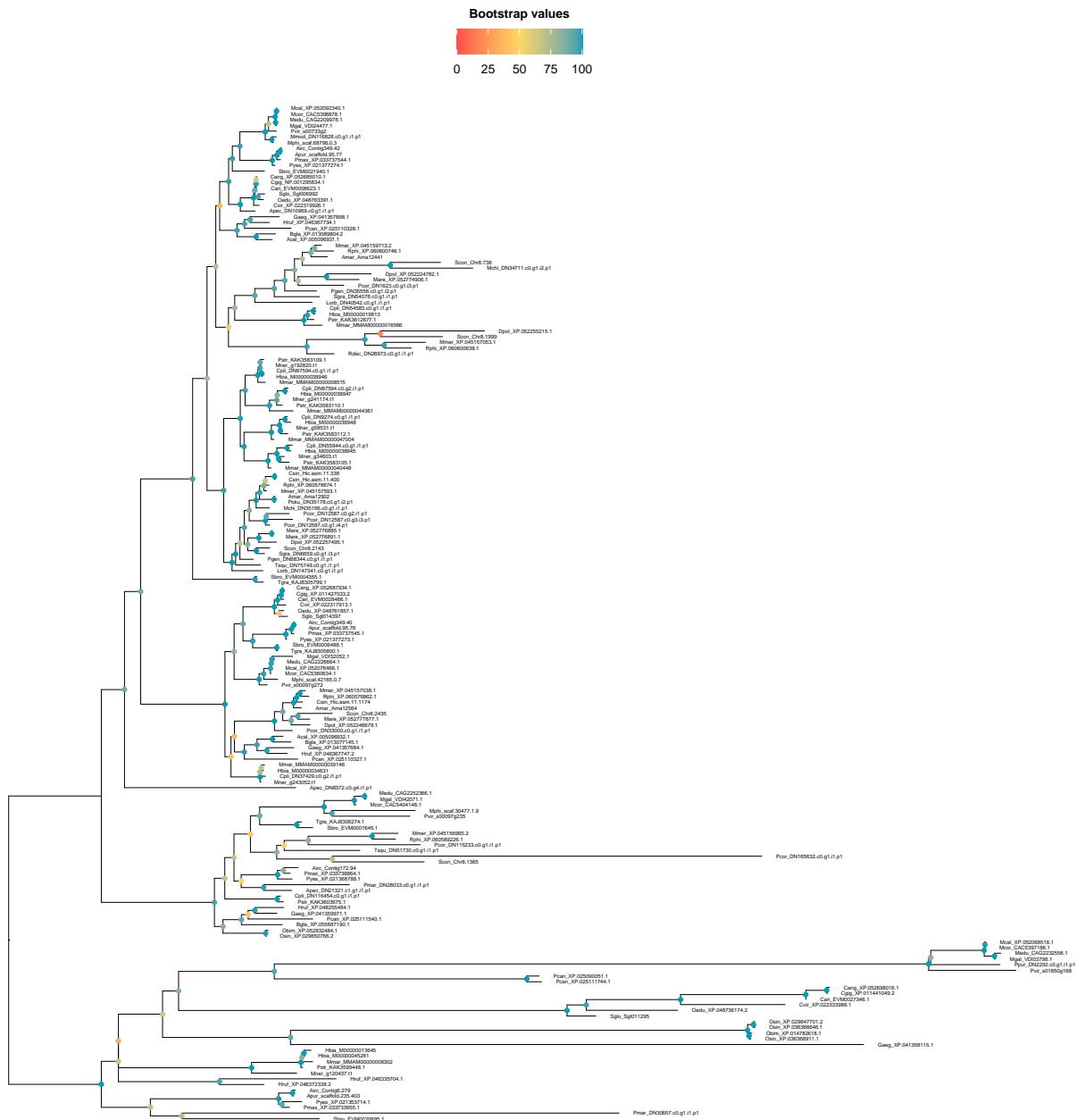
**Supplementary Figure S8. ML phylogenetic tree of the Dmrt gene family in fruit flies, including the Possvm orthology inference.** Reference genes from *D. melanogaster*, *Drosophila hydei*, *Drosophila pseudoobscura*, and *Drosophila suzukii* are marked with an asterisk at the beginning of the tip names. Species ID can be found in **Supp. Tab. S5**. The tree has been midpoint rooted. Bootstrap values are shown for each node.



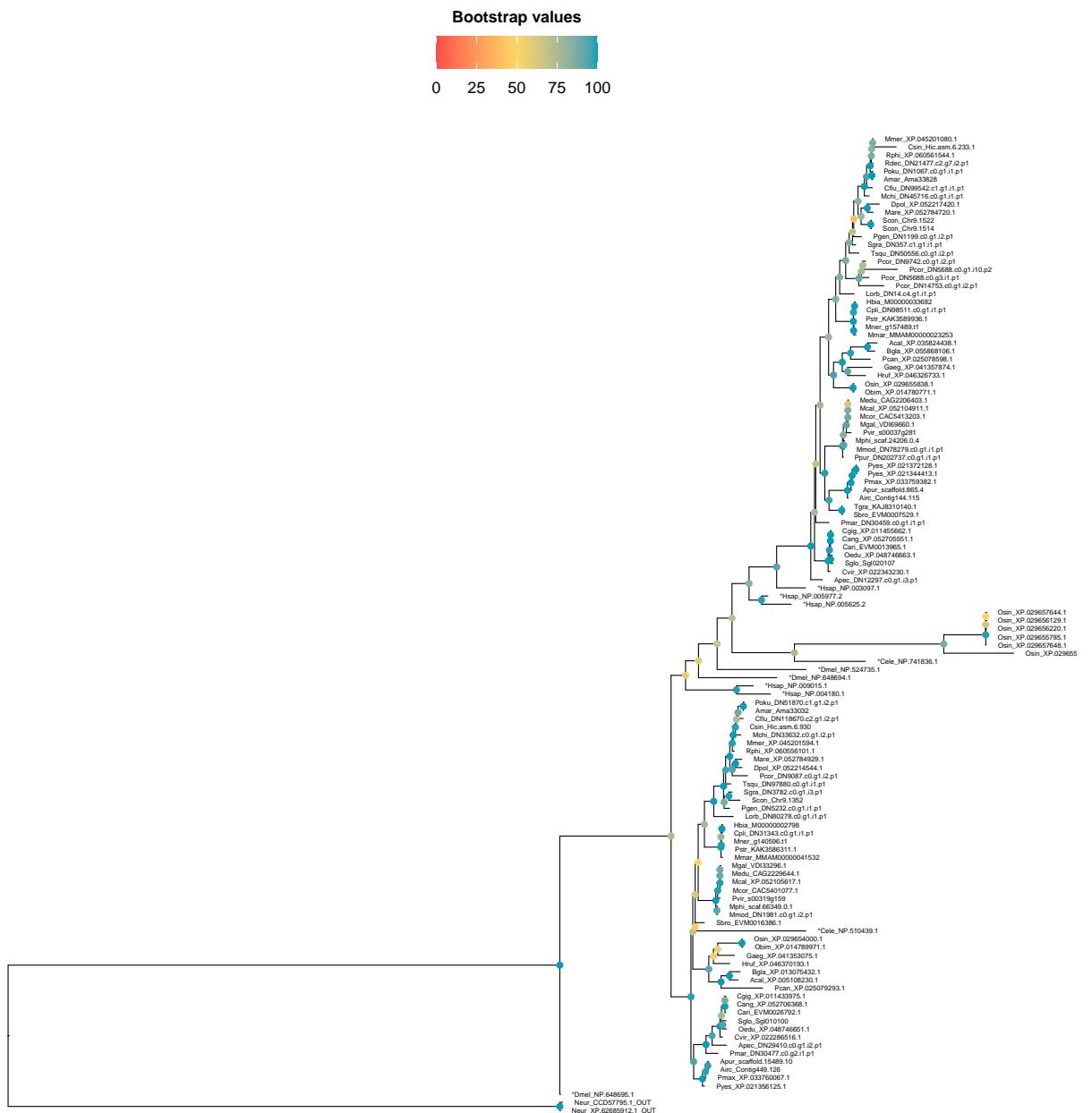
**Supplementary Figure S9. ML phylogenetic tree of the Sox gene family in fruit flies, including the Possvm orthology inference.** Reference genes from *D. melanogaster*, *D. hydei*, *D. pseudoobscura*, and *D. suzukii* are marked with an asterisk at the beginning of the tip names. Species ID can be found in **Supp. Tab. S5**. Bootstrap values are shown for each node.



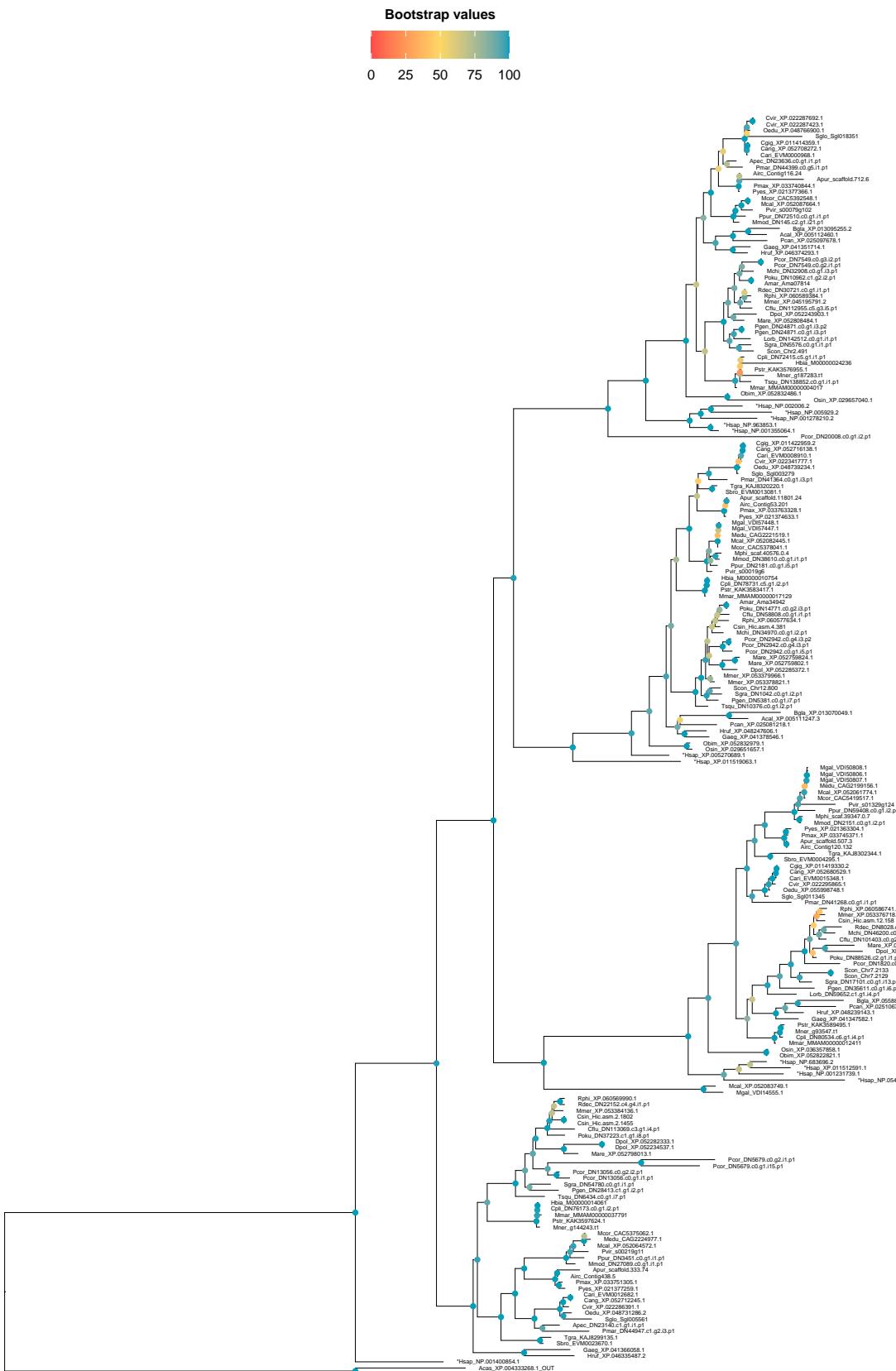
**Supplementary Figure S10. ML phylogenetic tree of the Fox gene family in fruit flies, including the Possvm orthology inference.** Reference genes from *D. melanogaster*, *D. hydei*, *D. pseudoobscura*, and *D. suzukii* are marked with an asterisk at the beginning of the tip names. Species ID can be found in **Supp. Tab. S5**. Bootstrap values are shown for each node.



**Supplementary Figure S11.** ML phylogenetic tree of the Dmrt gene family in mollusc species. Species ID can be found in Supp. Tab. S1. The tree has been midpoint rooted. Bootstrap values are shown for each node.



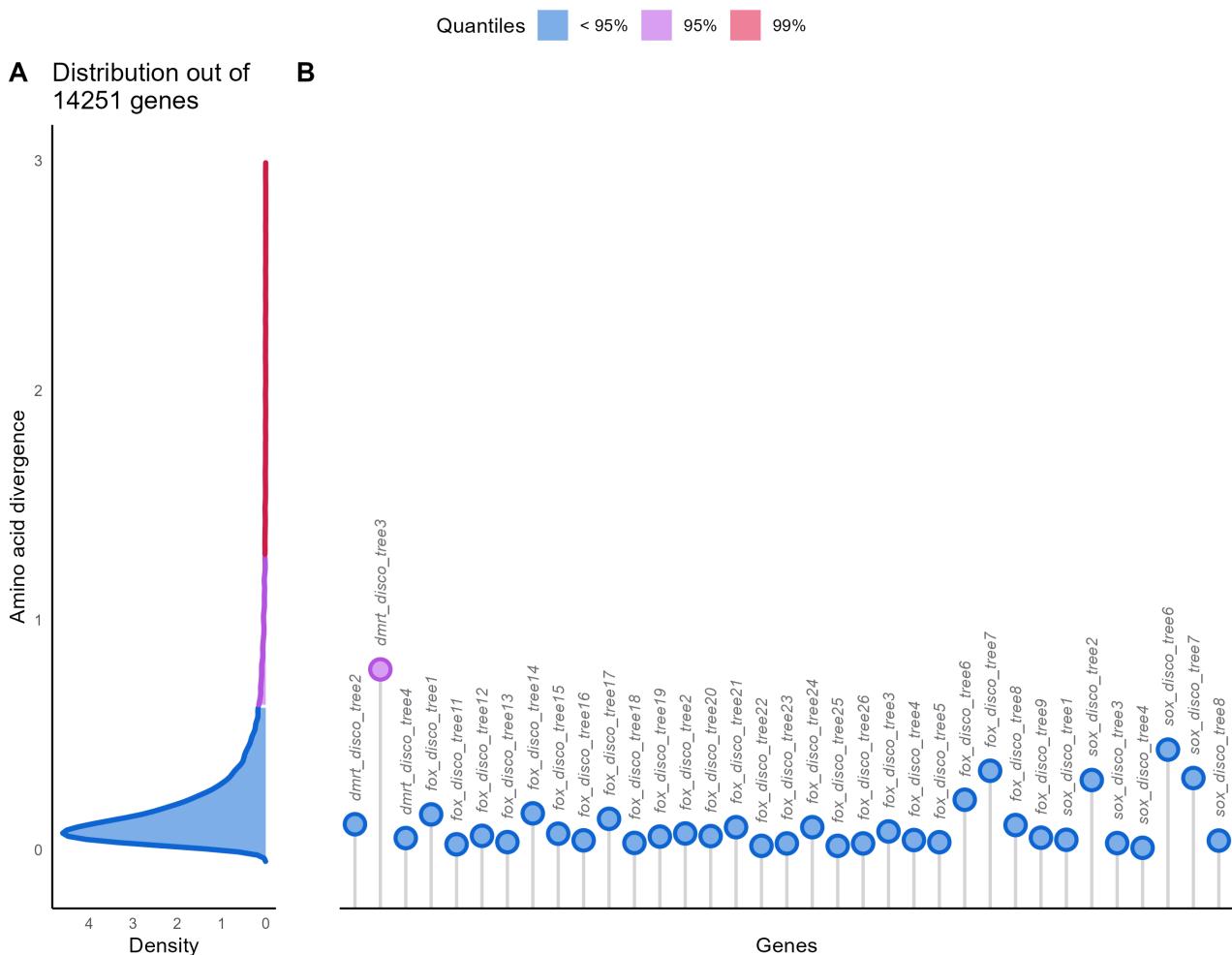
**Supplementary Figure S12. ML phylogenetic tree of *Sox-B1* and *Sox-B2* genes in mollusc and reference species.** Reference genes from *H. sapiens*, *C. elegans*, and *D. melanogaster* are marked with an asterisk at the beginning of the tip names. Species ID can be found in **Supp. Tab. S1**. Bootstrap values are shown for each node.



**Supplementary Figure S13. ML phylogenetic tree of *Fox-J2*, *Fox-M*, *Fox-O*, and *Fox-P* genes in mollusc and reference species.** Reference genes from *H. sapiens*, *C. elegans*, and *D. melanogaster* are marked with an asterisk at the beginning of the tip names. Species ID can be found in Supp. Tab. S1. Bootstrap values are shown for each node.



**Supplementary Figure S14. ML phylogenetic tree of the Fox gene family in bivalves and the sea urchin *Strongylocentrotus purpuratus* (Spur).** Reference genes from *S. purpuratus* are marked with an asterisk at the beginning of the tip names. Species ID can be found in **Supp. Tab. S1**. *S. purpuratus* genes are those given by **Tu et al., 2006**. Bootstrap values are shown for each node.



**Supplementary Figure S15.** Distribution of amino acid sequence divergence (AASD) of single-copy orthogroups in *Crassostrea gigas*, *Crassostrea angulata*, *Crassostrea ariakensis*, and *Crassostrea virginica* (A), including DSFG (B). The distribution of AASD in *Crassostrea* has been computed on the median values of pairwise distances of over 14k single-copy orthogroups (SCOs). Circle heights of DSFGs show the median value of their AASD. *Dmrt-1L* genes are indicated as ‘dmrt<sub>-</sub>disco<sub>-</sub>tree3’.

## Supplementary tables

*All the supplementary tables are available in a parsable version at the following GitHub repository: [LINK](#) [LINK](#) [LINK](#).*

**Supplementary Table S1. Genomic and transcriptomic data of bivalves and other molluscs.** For each species, the relative ID, taxonomic information, BUSCO statistics, NCBI accession number, and source publication are reported. Biv: Bivalvia; Ca: Caenogastropoda; Cep: Cephalopoda; Co: Coleoidea; Gas: Gastrooda; Gen: Genome; He: Heterobranchia; Im: Imparidaria; In: Neomphallones; Ne: Neomphallones; Pa: Palaeoheterodonta; Pt: Pteriomorpha; Tra: Transcriptome; Ve: Vetigastropoda

Species	ID	Class	Group	Order	Type	Reduced dataset	BUSCO statistics (metazoa.ncbi10)	NCBI acc. no.	Reference	Annotation source
<i>Magallana (Crassostrea) angulata</i>	Cang	Biv	Pt	Ostreida	Gen	No	C:99.1%[S:97.1%,D:2.0%], F:0.3%,M:0.6%	GCF_022612915.1	Teng et al., 2023	NCBI
<i>Magallana (Crassostrea) gigas</i>	Cgig	Biv	Pt	Ostreida	Gen	Yes	C:98.2%[S:93.1%,D:5.1%], F:0.4%,M:1.4%	GCF_902806645.1	Penaloza et al., 2021	NCBI
<i>Magallana (Crassostrea) ariakensis</i>	Cari	Biv	Pt	Ostreida	Gen	No	C:94.8%[S:91.2%,D:3.6%], F:0.7%,M:4.5%	GCA_020567875.1	Li et al., 2021	FigShare
<i>Crassostrea virginica</i>	Cvir	Biv	Pt	Ostreida	Gen	Yes	C:98.2%[S:73.1%,D:25.1%], F:0.3%,M:1.5%	GCF_002202765.2	Gómez-Chiarri et al., 2015	NCBI
<i>Ostrea edulis</i>	Oedu	Biv	Pt	Ostreida	Gen	Yes	C:98.7%[S:97.8%,D:0.9%], F:0.5%,M:0.8%	GCF_947568905.1	Darwin Tree of Life	N/A
<i>Saccostrea glomerata</i>	Sglo	Biv	Pt	Ostreida	Gen	No	C:89.1%[S:85.5%,D:3.6%], F:4.9%,M:6.0%	GCA_003671525.1	Powell et al., 2018	dbSROG
<i>Atrina pectinata</i>	Apec	Biv	Pt	Ostreida	Tra	Yes	C:95.6%[S:93.1%,D:2.5%], F:1.9%,M:2.5%	N/A	-	N/A
<i>Pinctada margaritifera</i>	Pmar	Biv	Pt	Ostreida	Tra	Yes	C:94.3%[S:93.9%,D:0.4%], F:1.7%,M:4.0%	SRR1030667	Teaniniuraitemoana et al., 2014	N/A
<i>Mytilus galloprovincialis</i>	Mgal	Biv	Pt	Mytilida	Gen	Yes	C:80.5%[S:50.4%,D:30.1%], F:8.6%,M:10.9%	GCA_900618805.1	Gerdol et al., 2020	NCBI
<i>Mytilus edulis</i>	Medu	Biv	Pt	Mytilida	Gen	No	C:83.8%[S:70.9%,D:12.9%], F:5.1%,M:11.1%	GCA_905397895.1	Corrochano-Fraile et al., 2022	NCBI
<i>Mytilus unguiculatus (coruscus)</i>	Mcor	Biv	Pt	Mytilida	Gen	No	C:80.8%[S:78.8%,D:2.0%], F:4.3%,M:14.9%	GCA_011752425.2	Yang et al., 2021	NCBI
<i>Mytilus californianus</i>	Mcal	Biv	Pt	Mytilida	Gen	No	C:96.2%[S:95.0%,D:1.2%], F:0.4%,M:3.4%	GCF_0241869535.1	Paggeot et al., 2022	NCBI
<i>Perna viridis</i>	Pvir	Biv	Pt	Mytilida	Gen	Yes	C:99.4%[S:99.0%,D:0.4%], F:0.2%,M:0.4%	GCA_0183227765.1	Inoue et al., 2021	Google Drive
<i>Modiolus modiolus</i>	Mmod	Biv	Pt	Mytilida	Tra	Yes	C:95.7%[S:92.3%,D:3.4%], F:2.1%,M:2.2%	N/A	-	N/A
<i>Modiolus philippinarum</i>	Mphi	Biv	Pt	Mytilida	Gen	No	C:64.9%[S:63.0%,D:1.9%], F:18.8%,M:16.3%	GCA_002080025.1	Sun et al., 2017	Dryad
<i>Perumytilus purpuratus</i>	Ppur	Biv	Pt	Mytilida	Tra	Yes	C:84.2%[S:83.3%,D:3.9%], F:11.8%,M:4.0%	N/A	-	N/A
<i>Argopecten irradians concentricus</i>	Airc	Biv	Pt	Pectinida	Gen	Yes	C:94.9%[S:94.0%,D:0.9%], F:3.6%,M:1.5%	GCA_004382765.1	Liu et al., 2020	Dryad
<i>Argopecten purpuratus</i>	Apur	Biv	Pt	Pectinida	Gen	No	C:89.2%[S:88.6%,D:0.6%], F:5.0%,M:5.8%	N/A	-	GigaDB
<i>Pecten maximus</i>	Pmax	Biv	Pt	Pectinida	Gen	Yes	C:98.5%[S:94.5%,D:4.0%], F:0.4%,M:1.1%	GCF_902652985.1	Kenny et al., 2020	NCBI
<i>Mizuhpecten (Patinopecten) yessoensis</i>	Pyes	Biv	Pt	Pectinida	Gen	Yes	C:98.3%[S:96.1%,D:2.2%], F:0.5%,M:1.2%	GCF_002113885.1	Wang, Zhang, et al., 2017	NCBI
<i>Anadara (Scapharca) broughtonii</i>	Sbro	Biv	Pt	Arcida	Gen	Yes	C:91.2%[S:85.6%,D:5.6%], F:2.6%,M:6.2%	N/A	Bai et al., 2019	GigaDB
<i>Tegillarca granosa</i>	Tgra	Biv	Pt	Arcida	Gen	Yes	C:70.6%[S:61.3%,D:9.3%], F:11.7%,M:17.7%	GCA_029721355.1	N/A	NCBI
<i>Ruditapes decussatus</i>	Rdec	Biv	Im	Venerida	Tra	Yes	C:84.8%[S:84.1%,D:0.7%], F:7.3%,M:7.9%	SRR527740, 41, 43, 44, 47, 51, 52, 57	Ghiselli et al., 2018	N/A
<i>Ruditapes philippinarum</i>	Rphi	Biv	Im	Venerida	Gen	Yes	C:97.8%[S:85.5%,D:12.3%], F:0.7%,M:1.5%	GCF_026571515.1	Xu, Martelossi, et al., 2022	NCBI
<i>Mercenaria mercenaria</i>	Mmer	Biv	Im	Venerida	Gen	Yes	C:96.0%[S:89.8%,D:6.2%], F:1.0%,M:3.0%	GCF_0241730395.1	Farhat et al., 2022	NCBI

Table S1 continued from previous page

Species	ID	Class	Group	Order	Type	Reduced dataset	BUSCO statistics (metazoa.ncbi10)	NCBI acc. no.	Reference	Annotation source
<i>Cyclina sinensis</i>	Csin	Biv	Im	Venerida	Gen	Yes	C:94.1%[S:83.9%,D:10.2%], F:1.8%,M:4.1%	GCA_012932295.1	Wei et al., 2020	Dryad
<i>Calyptogena (Archaeistica) marissinica</i>	Amar	Biv	Im	Venerida	Gen	No	C:82.1%[S:80.1%,D:2.0%], F:6.0%,M:11.9%	GCA_014843695.1	Ip et al., 2021	FigShare
<i>Phreagena okutanii</i>	Poku	Biv	Im	Venerida	Tra	Yes	C:92.9%[S:85.8%,D:7.1%], F:3.0%,M:4.1%	N/A	-	N/A
<i>Corbicula fluminea</i>	Cflu	Biv	Im	Venerida	Tra	Yes	C:83.7%[S:79.9%,D:3.8%], F:10.3%,M:6.0%	SRR1559272 SRR512046	González et al., 2015 Zhu et al., 2019	N/A
<i>Mactra chinensis</i>	Mchi	Biv	Im	Venerida	Tra	Yes	C:81.5%[S:80.8%,D:0.7%], F:10.2%,M:8.3%	SRR1263980	N/A	N/A
<i>Mya arenaria</i>	Mare	Biv	Im	Myida	Gen	Yes	C:98.5%[S:80.4%,D:18.1%], F:0.4%,M:1.1%	GCF_026914265.1	Hart et al., 2023	NCBI
<i>Dreissena polymorpha</i>	Dpol	Biv	Im	Myida	Gen	Yes	C:97.2%[S:80.1%,D:17.1%], F:0.4%,M:2.4%	GCF_020536995.1	McCartney et al., 2022	NCBI
<i>Pisidium corneum</i>	Pcor	Biv	Im	Sphaeriida	Tra	Yes	C:94.5%[S:81.6%,D:12.9%], F:3.6%,M:1.9%	N/A	-	N/A
<i>Solen grandis</i>	Sgra	Biv	Im	Adapedonta	Tra	Yes	C:92.7%[S:90.0%,D:2.7%], F:2.5%,M:4.8%	N/A	-	N/A
<i>Simonovacula constricta</i>	Scon	Biv	Im	Adapedonta	Gen	Yes	C:90.8%[S:79.2%,D:11.6%], F:3.5%,M:5.7%	GCA_007844125.1	Ran et al., 2019	Dryad
<i>Panopea generosa</i>	Pgen	Biv	Im	Adapedonta	Tra	Yes	C:84.1%[S:81.9%,D:2.2%], F:9.7%,M:6.2%	N/A	-	N/A
<i>Tridacna squamosa</i>	Tsqu	Biv	Im	Cardiida	Tra	Yes	C:89.7%[S:86.9%,D:2.8%], F:3.5%,M:6.8%	N/A	-	N/A
<i>Loripes orbiculatus</i>	Lorb	Biv	Im	Lucinida	Tra	Yes	C:76.1%[S:74.9%,D:1.2%], F:14.3%,M:9.6%	N/A	-	N/A
<i>Hyriopsis biatalata</i> ( <i>Unio delphinus</i> )	Hbia	Biv	Pa	Unionida	Gen	Yes	C:97.5%[S:94.9%,D:2.6%], F:12.0%,M:0.5%	GCA_029339505.1	Gomes-dos-Santos et al., 2023	FigShare
<i>Cristaria plicata</i>	Cpli	Biv	Pa	Unionida	Tra	Yes	C:93.6%[S:92.8%,D:0.8%], F:12.1%,M:4.3%	SRR12175868 SRR3095781	Patraik et al., 2016 Wang, Liu, and Wu, 2017	N/A
<i>Meganotus nervosa</i>	Mner	Biv	Pa	Unionida	Gen	Yes	C:65.0%[S:63.3%,D:1.7%], F:14.0%,M:21.0%	GCA_016617755.1	Rogers et al., 2021	Dryad
<i>Potamius streckeroni</i>	Pstr	Biv	Pa	Unionida	Gen	Yes	C:94.9%[S:93.4%,D:1.5%], F:1.2%,M:3.9%	GCA_016746295.1	Smith, 2021	NCBI
<i>Margaritifera margaritifera</i>	Mmar	Biv	Pa	Unionida	Gen	Yes	C:92.6%[S:92.1%,D:0.5%], F:3.0%,M:4.4%	GCA_015947965.1	Gomes-dos-Santos et al., 2021	FigShare
<i>Aplysia californica</i>	Acal	Gas	He	Aplysiida	Gen	No	C:97.8%[S:97.0%,D:0.8%], F:0.7%,M:1.5%	GCF_000002075.1	Knudsen et al., 2006	NCBI
<i>Biomphalaria glabrata</i>	Bgla	Gas	He	N/A	Gen	No	C:98.9%[S:98.2%,D:0.7%], F:0.1%,M:1.0%	GCF_947242115.1	N/A	NCBI
<i>Pomacea canaliculata</i>	Pcan	Gas	C <sub>a</sub>	Architaenioglossa	Gen	No	C:98.2%[S:97.0%,D:1.2%], F:0.4%,M:1.4%	GCF_003073045.1	Liu et al., 2018	NCBI
<i>Giantopelta aegia</i>	Gaeg	Gas	Ne	Neomphalida	Gen	No	C:98.4%[S:94.2%,D:4.2%], F:0.8%,M:0.8%	GCF_01609755.1	Lan et al., 2021	NCBI
<i>Haliotis rufescens</i>	Hruf	Gas	Ve	Lepetellida	Gen	No	C:99.0%[S:98.3%,D:0.7%], F:0.0%,M:1.0%	GCF_023055435.1	N/A	NCBI
<i>Octopus bimaculoides</i>	Obim	Cep	Co	Octopoda	Gen	No	C:94.9%[S:94.4%,D:0.5%], F:2.3%,M:2.8%	GCF_001941352	Albertin et al., 2015	NCBI
<i>Octopus sinensis</i>	Osin	Cep	Co	Octopoda	Gen	No	C:98.1%[S:96.9%,D:1.2%], F:0.9%,M:1.0%	GCF_006345805.1	Li et al., 2020	NCBI

**Supplementary Table S2.** DSFG family and domain identifiers (IDs) in PANTHER and CDD, respectively. After having retrieved putative DSFGs on the basis of hidden Markov model (HMM) profiles, IDs have been used to retain only reliable hits.

Gene family	PANTHER/CDD	ID	Description
Dmrt	CDD	gnl—CDD—214606	Doublesex DNA-binding motif
	CDD	gnl—CDD—425850	DM DNA binding domain
	PANTHER	PTHR12322	DOUBLESEX AND MAB-3 RELATED TRANSCRIPTION FACTOR DMRT PROTEIN CBR-MAB-23
	PANTHER	PTHR12322-SF115	DOUBLESEX- AND MAB-3-RELATED TRANSCRIPTION FACTOR 1
	PANTHER	PTHR12322-SF116	DOUBLESEX- AND MAB-3-RELATED TRANSCRIPTION FACTOR DMD-4
	PANTHER	PTHR12322-SF123	DOUBLESEX-MAB RELATED 99B
	PANTHER	PTHR12322-SF53	DOUBLESEX- AND MAB-3 RELATED TRANSCRIPTION FACTOR 2
	PANTHER	PTHR12322-SF71	DOUBLESEX- AND MAB-3-RELATED TRANSCRIPTION FACTOR A1
	PANTHER	PTHR16897-SF2	STRESS RESPONSE PROTEIN NST1
	PANTHER	PTHR46888-SF11	RIBONUCLEASE H
Sox	CDD	gnl—CDD—432488	SOX transcription factor
	CDD	gnl—CDD—432558	Sox developmental protein N terminal
	CDD	gnl—CDD—438790	high mobility group (HMG)-box found in group B SRY-related high-mobility group (HMG) box (Sox) transcription factors
	CDD	gnl—CDD—438820	high mobility group (HMG)-box found in sex-determining region Y (SRY)-box (SOX) family transcription factors
	CDD	gnl—CDD—438837	high mobility group (HMG)-box found in group A, group B and group G of SRY-related high-mobility group (HMG) box (Sox) transcription factors
	CDD	gnl—CDD—438838	high mobility group (HMG)-box found in group C SRY-related high-mobility group (HMG) box (Sox) transcription factors
	CDD	gnl—CDD—438839	high mobility group (HMG)-box found in group D SRY-related high-mobility group (HMG) box (Sox) transcription factors
	CDD	gnl—CDD—438840	high mobility group (HMG)-box found in group E SRY-related high-mobility group (HMG) box (Sox) transcription factors
	CDD	gnl—CDD—438841	high mobility group (HMG)-box found in group F SRY-related high-mobility group (HMG) box (Sox) transcription factors
	CDD	gnl—CDD—438842	high mobility group (HMG)-box found in sex determining region Y (SRY)-box 30 (SOX30) and similar proteins
	CDD	gnl—CDD—438843	high mobility group (HMG)-box found in sex-determining region Y protein (SRY) and similar proteins
	CDD	gnl—CDD—438844	high mobility group (HMG)-box found in sex determining region Y (SRY)-box 15 (SOX15) and similar proteins
	CDD	gnl—CDD—438845	high mobility group (HMG)-box found in sex determining region Y (SRY)-box 4 (SOXA) and similar proteins
	CDD	gnl—CDD—438846	high mobility group (HMG)-box found in sex determining region Y (SRY)-box 11 (SOX11) and similar proteins
	CDD	gnl—CDD—438847	high mobility group (HMG)-box found in sex determining region Y (SRY)-box 12 (SOX12) and similar proteins
	CDD	gnl—CDD—438849	high mobility group (HMG)-box found in sex determining region Y (SRY)-box 7 (SOXT) and similar proteins
	CDD	gnl—CDD—438850	high mobility group (HMG)-box found in sex determining region Y (SRY)-box 17 (SOX17) and similar proteins
	CDD	gnl—CDD—438851	high mobility group (HMG)-box found in sex determining region Y (SRY)-box 18 (SOX18) and similar proteins
	PANTHER	PTHR10270-SF107	"TRANSCRIPTION FACTOR SOX-14"
	PANTHER	PTHR10270-SF161	"SOX DOMAIN-CONTAINING PROTEIN DICHAETE-RELATED PROTEIN"
	PANTHER	PTHR10270-SF199	"SEX-DETERMINING REGION Y PROTEIN SOX-2"
	PANTHER	PTHR10270-SF231	"TRANSCRIPTION FACTOR SOX-4"
	PANTHER	PTHR10270-SF27	"TRANSCRIPTION FACTOR SOX-21"
	PANTHER	PTHR10270-SF313	"TRANSCRIPTION FACTOR SOX-1A-RELATED PROTEIN"
	PANTHER	PTHR10270-SF315	"TRANSCRIPTION FACTOR SOX-15-RELATED PROTEIN"
	PANTHER	PTHR10270-SF317	"TRANSCRIPTION FACTOR SOX-3"
	PANTHER	PTHR10270-SF322	"TRANSCRIPTION FACTOR SOX-3"
	PANTHER	PTHR10270-SF324	"TRANSCRIPTION FACTOR SOX-3"
	PANTHER	PTHR10270-SF326	"TRANSCRIPTION FACTOR SOX-3"
	PANTHER	PTHR10270-SF327	"TRANSCRIPTION FACTOR SOX-3"
	PANTHER	PTHR45789	F118025P1
	PANTHER	PTHR45789-SF2	F118025P1
	PANTHER	PTHR45803-SF1	TRANSCRIPTION FACTOR SOX-9
	PANTHER	PTHR45803-SF2	TRANSCRIPTION FACTOR SOX-8
	PANTHER	PTHR45803-SF5	SOX100B
	PANTHER	PTHR45803-SF1	SOX100B
	PANTHER	PTHR47279-SF1	TRANSCRIPTION FACTOR SOX-30
	PANTHER	PTHR47279-SF19	TRANSCRIPTION FACTOR SOX-30
Fox	CDD	gnl—CDD—410788	Forkhead (FH) domain found in Forkhead box protein A (FOXA) subfamily
	CDD	gnl—CDD—410789	Forkhead (FH) domain found in Forkhead box protein B (FOXB) subfamily
	CDD	gnl—CDD—410790	Forkhead (FH) domain found in Forkhead box protein C (FOXC) subfamily
	CDD	gnl—CDD—410791	Forkhead (FH) domain found in Forkhead box protein D (FOXD) subfamily
	CDD	gnl—CDD—410792	Forkhead (FH) domain found in Forkhead box protein E (FOXE) subfamily
	CDD	gnl—CDD—410793	Forkhead (FH) domain found in Forkhead box protein F (FOXF) subfamily
	CDD	gnl—CDD—410794	Forkhead (FH) domain found in Forkhead box protein G (FOXG) subfamily
	CDD	gnl—CDD—410795	Forkhead (FH) domain found in Forkhead box protein H (FOXH) subfamily
	CDD	gnl—CDD—410796	Forkhead (FH) domain found in Forkhead box protein J1 (FOXJ1) and similar proteins
	CDD	gnl—CDD—410797	Forkhead (FH) domain found in Forkhead box protein J2 (FOXJ2) and similar proteins
CDD	CDD	gnl—CDD—410798	Forkhead (FH) domain found in Forkhead box protein K (FOXK) subfamily
	CDD	gnl—CDD—410799	Forkhead (FH) domain found in Forkhead box protein L1 (FOXL1) and similar proteins
	CDD	gnl—CDD—410800	Forkhead (FH) domain found in Forkhead box protein L2 (FOXL2) and similar proteins
	CDD	gnl—CDD—410801	Forkhead (FH) domain found in Forkhead box protein L3 (FOXL3) and similar proteins

Table S2 continued from previous page

Gene family	PANTHER/CDD	ID	Description
CDD	gnl—CDD—410802	Forkhead (FH) domain found in Forkhead box protein L2 (FOXL2) and similar proteins	
CDD	gnl—CDD—410803	Forkhead (FH) domain found in the Forkhead box protein M (FOXMM) subfamily	
CDD	gnl—CDD—410804	Forkhead (FH) domain found in Forkhead box protein N1 (FOXNL) and similar proteins	
CDD	gnl—CDD—410805	Forkhead (FH) domain found in Forkhead box protein N2 (FOXNO) and similar proteins	
CDD	gnl—CDD—410806	Forkhead (FH) domain found in the Forkhead box protein O (FOXO) subfamily	
CDD	gnl—CDD—410807	Forkhead (FH) domain found in the Forkhead box protein P (FOXP) subfamily	
CDD	gnl—CDD—410808	Forkhead (FH) domain found in Forkhead box protein Q1 (FOXQ1) and similar proteins	
CDD	gnl—CDD—410809	Forkhead (FH) domain found in Forkhead box protein Q2 (FOXQ2) and similar proteins	
CDD	gnl—CDD—410810	Forkhead (FH) domain found in the Forkhead box protein R (FOXKR) subfamily	
CDD	gnl—CDD—410811	Forkhead (FH) domain found in Forkhead box protein S1 (FOXSI)	
CDD	gnl—CDD—410812	Forkhead (FH) domain found in Forkhead box protein A1 (FOXA1) and similar proteins	
CDD	gnl—CDD—410813	Forkhead (FH) domain found in Forkhead box protein A2 (FOXA2) and similar proteins	
CDD	gnl—CDD—410814	Forkhead (FH) domain found in Forkhead box protein A3 (FOXAS) and similar proteins	
CDD	gnl—CDD—410816	Forkhead (FH) domain found in Forkhead box protein B1 (FOXB1) and similar proteins	
CDD	gnl—CDD—410817	Forkhead (FH) domain found in Forkhead box protein B2 (FOXB2) and similar proteins	
CDD	gnl—CDD—410818	Forkhead (FH) domain found in Forkhead box protein C1 (FOXCL) and similar proteins	
CDD	gnl—CDD—410819	Forkhead (FH) domain found in Forkhead box protein C2 (FOXJ2) and similar proteins	
CDD	gnl—CDD—410820	Forkhead (FH) domain found in Forkhead box proteins FOXD1, FOXD2 and similar proteins	
CDD	gnl—CDD—410821	Forkhead (FH) domain found in Forkhead box protein D3 (FOXD3) and similar proteins	
CDD	gnl—CDD—410822	Forkhead (FH) domain found in Forkhead box protein D4 (FOXF4) and similar proteins	
CDD	gnl—CDD—410823	Forkhead (FH) domain found in Forkhead box protein F1 (FOXFL) and similar proteins	
CDD	gnl—CDD—410824	Forkhead (FH) domain found in Forkhead box protein F2 (FOXFL2) and similar proteins	
CDD	gnl—CDD—410825	Forkhead (FH) domain found in Forkhead box protein J2 (FOXJ2) and similar proteins	
CDD	gnl—CDD—410826	Forkhead (FH) domain found in Forkhead box protein J3 (FOXJ3) and similar proteins	
CDD	gnl—CDD—410827	Forkhead (FH) domain found in Forkhead box protein J1 (FOXJ1) and similar proteins	
CDD	gnl—CDD—410828	Forkhead (FH) domain found in Forkhead box protein K1 (FOXK1) and similar proteins	
CDD	gnl—CDD—410829	Forkhead (FH) domain found in Forkhead box protein K2 (FOXK2) and similar proteins	
CDD	gnl—CDD—410830	Forkhead (FH) domain found in Forkhead box protein N1 (FOXNN1)	
CDD	gnl—CDD—410831	Forkhead (FH) domain found in Forkhead box protein N4 (FOXNN4)	
CDD	gnl—CDD—410832	Forkhead (FH) domain found in Forkhead box protein N2 (FOXNN2)	
CDD	gnl—CDD—410833	Forkhead (FH) domain found in Forkhead box protein N3 (FOXNN3)	
CDD	gnl—CDD—410834	Forkhead (FH) domain found in Forkhead box protein O1 (FOXO1)	
CDD	gnl—CDD—410835	Forkhead (FH) domain found in Forkhead box protein O3 (FOXO3)	
CDD	gnl—CDD—410836	Forkhead (FH) domain found in Forkhead box protein O4 (FOXO4) and similar proteins	
CDD	gnl—CDD—410837	Forkhead (FH) domain found in Forkhead box protein O6 (FOXO6) and similar proteins	
CDD	gnl—CDD—410838	Forkhead (FH) domain found in Forkhead box protein P1 (FOXP1)	
CDD	gnl—CDD—410839	Forkhead (FH) domain found in Forkhead box protein P2 (FOXP2)	
CDD	gnl—CDD—410840	Forkhead (FH) domain found in Forkhead box protein P3 (FOXP3) and similar proteins	
PANTHER	gnl—CDD—410841	FORKHEAD BOX PROTEIN PTHR11829	
PANTHER	PTHR11829-SF142	FORKHEAD PROTEIN PTHR11829-SF142	
PANTHER	PTHR11829-SF156	FORKHEAD BOX PROTEIN E3 PTHR11829-SF206	
PANTHER	PTHR11829-SF209	FORKHEAD BOX PROTEIN Q1 PTHR11829-SF335	
PANTHER	PTHR11829-SF335	FORKHEAD BOX PROTEIN B1 PTHR11829-SF340	
PANTHER	PTHR11829-SF342	FORKHEAD BOX PROTEIN D2 PTHR11829-SF348	
PANTHER	PTHR11829-SF348	FORKHEAD BOX PROTEIN H1 PTHR11829-SF361	
PANTHER	PTHR11829-SF361	FORKHEAD BOX PROTEIN L2 PTHR11829-SF398	
PANTHER	PTHR11829-SF398	FORKHEAD BOX PROTEIN D3 PTHR11829-SF399	
PANTHER	PTHR11829-SF399	FORKHEAD BOX TRANSCRIPTION FACTOR FKH-9 PTHR11829-SF401	
PANTHER	PTHR13862	FORKHEAD BOX CL-A-RELATED PTHR13862-SF17	
PANTHER	PTHR13862-SF17	FORKHEAD BOX PROTEIN N4 PTHR13862-SF19	
PANTHER	PTHR13862-SF19	FORKHEAD BOX PROTEIN N2 PTHR13862-SF22	
PANTHER	PTHR13862-SF22	FORKHEAD BOX PROTEIN N3-LIKE PROTEIN PTHR13862-SF26	
PANTHER	PTHR45767	FORKHEAD BOX PROTEIN N2 PTHR45767-SF2	
PANTHER	PTHR45767-SF2	FORKHEAD BOX PROTEIN O PTHR45796	
PANTHER	PTHR45796	FORKHEAD BOX PROTEIN P, ISOFORM C PTHR45796-SF3	
PANTHER	PTHR45796-SF4	FORKHEAD BOX PROTEIN C P, ISOFORM C PTHR4581-SF3	
PANTHER	PTHR4581-SF4	FORKHEAD BOX PROTEIN K2 PTHR46078	
PANTHER	PTHR46078	FORKHEAD BOX PROTEIN K1 PTHR46262	
PANTHER	PTHR46262	FORKHEAD BOX PROTEIN J2 FAMILY MEMBER PTHR46262-SF2	
PANTHER	PTHR46617	FORKHEAD BOX PROTEIN BINIOU PTHR46617-G1	

Table S2 continued from previous page

Gene family	PANTHER/CDD	ID	Description
Fox	PANTHER	PTHR46617:SF3	FORKHEAD BOX PROTEIN G1
	PANTHER	PTHR46721	FORKHEAD BOX PROTEIN N1
	PANTHER	PTHR46721:SF2	FORKHEAD BOX N1
	PANTHER	PTHR46805	FORKHEAD BOX PROTEIN J1
	PANTHER	PTHR46878	FORKHEAD BOX PROTEIN M1
	PANTHER	PTHR46878:SF1	FORKHEAD BOX PROTEIN M1
	PANTHER	PTHR47316	FORKHEAD BOX PROTEIN H1
	PANTHER	PTHR47316:SF1	FORKHEAD BOX PROTEIN H1

**Supplementary Table S3. List of DSFGs from reference species used to assess the identity of DSFGs in molluses.** NCBI accession numbers are reported in parenthesis. Each row represents an orthology group.

<i>Homo sapiens</i>	<i>Drosophila melanogaster</i>	<i>Caenorhabditis elegans</i>	Group
<b>Dmrt gene family</b>			
<i>DMRT1</i> (NP_068770.2)	-	-	1
<i>DMRT2</i> (NP_006548.1)	<i>dmrt11E</i> (NP_511146.2)	-	2
<i>DMRT3</i> (NP_067063.1)	<i>dmrt93B</i> (NP_524428.1)	<i>dmd-4</i> (NP_510466.1)	3
<i>DMRT4/A1</i> (NP_071443.2)	<i>dmrt99b</i> (NP_524549.1)	<i>dmd-5</i> (NP_495138.2)	A1/2
<i>DMRT5/A2</i> (NP_115486.1)			
<i>DMRT6/B1</i> (NP_149056.1)	-	-	-
<i>DMRT7/C2</i> (NP_001035973.1)	-	-	-
<i>DMRT8/C1</i> (NP_149042.2)	-	-	-
-	<i>dsx</i> (NP_731197.1)	-	-
-	-	<i>mab3</i> (NP_001256882.1)	-
-	-	<i>dmd-3</i> (NP_001256883.1)	-
-	-	<i>dmd-6</i> (NP_001370045.1)	-
-	-	<i>dmd-7</i> (NP_741551.1)	-
-	-	<i>dmd-8</i> (NP_503176.2)	-
-	-	<i>dmd-9</i> (NP_500305.1)	-
-	-	<i>dmd-11</i> (NP_001379162.1)	-
-	-	<i>mab-23</i> (NP_001041089.1)	-
<b>Sox gene family</b>			
<i>SRY</i> (NP_003131.1)	-	-	A
<i>SOX3</i> (NP_005625.2)			
<i>SOX2</i> (NP_003097.1)	<i>dichaete</i> (NP_524066.1)	<i>sox3</i> (NP_510439.1)	B1
<i>SOX1</i> (NP_005977.2)	<i>soxN</i> (NP_524735.1)	<i>sox2</i> (NP_741836.1)	
<i>SOX14</i> (NP_004180.1)	<i>sox21a</i> (NP_648694.1)	-	B2
<i>SOX21</i> (NP_009015.1)	<i>sox21b</i> (NP_648695.1)		
<i>SOX11</i> (NP_003099.1)	<i>sox14</i> (NP_476894.1)	<i>sem-2</i> (NP_740846.1)	C
<i>SOX12</i> (NP_008874.2)			
<i>SOX4</i> (NP_003098.1)			
<i>SOX13</i> (NP_005677.2)			
<i>SOX5</i> (NP_008871.3)	<i>sox102f</i> (NP_726612.1)	<i>egl-13</i> (NP_001024918.1)	D
<i>SOX6</i> (NP_001139291.2)			
<i>SOX9</i> (NP_000337.1)			
<i>SOX8</i> (NP_055402.2)	<i>sox110b</i> (NP_651839.1)	-	E
<i>SOX10</i> (NP_008872.1)			
<i>SOX18</i> (NP_060889.1)			
<i>SOX7</i> (NP_113627.1)	<i>sox15</i> (NP_523739.2)	-	F
<i>SOX17</i> (NP_071899.1)			
<i>SOX15</i> (NP_008873.1)	-	-	G
<i>SOX30</i> (NP_848511.1)	-	-	H
<b>Fox gene family</b>			
<i>FOXA1/HNF-3<math>\alpha</math></i> (NP_004487.2)	<i>forkhead/fkh</i> (NP_524542.1)	<i>pha-4/Ce-fkh1</i> (NP_001041114.1)	A
<i>FOXA2/HNF-3<math>\beta</math></i> (NP_068556.2)			
<i>FOXA3/HNF-3<math>\gamma</math></i> (NP_004488.2)			
<i>FOXB1</i> (NP_036314.2)	<i>fd96Ca/fd4</i> (NP_524495.1)	<i>lin-31</i> (NP_494704.1)	B
<i>FOXB2</i> (NP_001013757.1)	<i>fd96Cb/fd5</i> (NP_524496.1)		
<i>FOXC1/MF1/FKHL7</i> (NP_001444.2)	<i>crocodile/fd1</i> (NP_524202.1)	-	C
<i>FOXC2/MFH1</i> (NP_005242.1)			
<i>FOXD1/FREAC4</i> (NP_004463.1)			
<i>FOXD2/FREAC9</i> (NP_004465.3)	<i>fd59A/fd3</i> (NP_523814.1)	<i>unc-130</i> (NP_496411.1)	D
<i>FOXD3</i> (NP_036315.1)			
<i>FOXD4</i> (NP_997188.2)			
<i>FOXE1/TITF2</i> (NP_004464.2)	-	-	E
<i>FOXE3</i> (NP_036318.1)			
<i>FOXF1</i> (NP_001442.2)	<i>binious/FoxF</i> (NP_523950.2)	<i>let-381/F26B1.7</i> (NP_491826.1)	F
<i>FOXF2</i> (NP_001443.1)			
<i>FOXG1/BF1/HBF2</i> (NP_005240.3)	<i>slp1</i> (NP_476730.1) <i>slp2</i> (NP_476834.1) <i>fd19B/cg9571</i> (NP_608369.1)	<i>fkh2/T14G12.4</i> (NP_508644.1)	G
<i>FOXH1/FAST1</i> (NP_003914.1)	-	-	H
<i>FOXI1/FREAC6/HFH3</i> (NP_036320.2)	-	-	I
<i>FOXJ1</i> (NP_001445.2)	-	-	J1
<i>FOXJ2</i> (XP_011519063.1)	-	-	J2
<i>FOXJ3</i> (XP_005270689.1)	-	-	J3
<i>FOXK1/ILF1</i> (NP_001032242.1)	<i>foxK/LD16137</i> (NP_001261701.1)	-	K
<i>FOXK2</i> (NP_004505.2)			
<i>FOXL1</i> (NP_005241.1)	<i>foxL1/fd2</i> (NP_523912.1)	-	L1
<i>FOXL2</i> (NP_075555.1)	-	-	L2
<i>FOXM1</i> (NP_001400854.1)	-	-	M

Table S3 continued from previous page

<i>Homo sapiens</i>	<i>Drosophila melanogaster</i>	<i>Caenorhabditis elegans</i>	Group
Fox gene family			
<i>FOXN1/WHN</i> ( <i>NP_001356298.1</i> ) <i>FOXN4</i> ( <i>NP_998761.2</i> )	<i>jumeau</i> ( <i>NP_524302.1</i> )	-	N1/4
<i>FOXN2/HTLF</i> ( <i>NP_001362376.1</i> ) <i>FOXN3/CHES1</i> ( <i>NP_001078940.1</i> )	<i>ches-1</i> ( <i>NP_511071.3</i> )	-	N2/3
<i>FOXO1</i> ( <i>NP_002006.2</i> ) <i>FOXO3</i> ( <i>NP_963853.1</i> ) <i>FOXO3B</i> ( <i>NP_001355064.1</i> )	-	<i>daf-16</i> ( <i>NP_001364785.1</i> )	O
<i>FOXP1</i> ( <i>NP_001231739.1</i> ) <i>FOXP2</i> ( <i>NP_683696.2</i> ) <i>FOXP3</i> ( <i>NP_054728.2</i> ) <i>FOXP4</i> ( <i>XP_011512591.1</i> )	<i>foxP/cg16899</i> ( <i>NP_001247011.1</i> )	<i>F26D12.1</i> ( <i>NP_001293813.1</i> )	P
<i>FOXQ/HFH11</i> ( <i>NP_150285.3</i> ) - <i>FOX51/FREAC10</i> ( <i>NP_004109.1</i> )	- <i>fd102C/cd11152</i> ( <i>NP_651951.1</i> ) - - - - -	- <i>fkh-10/C25A1.2</i> ( <i>NP_492676.2</i> ) - <i>PES-1</i> ( <i>NP_001023406.1</i> ) <i>B0286.5/FKH-6</i> ( <i>NP_494775.1</i> ) <i>F40H3.4/FKH-8</i> ( <i>NP_001254107.1</i> ) <i>C29F7.4/FKH-3</i> ( <i>NP_001294822.1</i> ) <i>K03C7.2/FKH-9</i> ( <i>NP_001024760.1</i> )	Q1 Q2 S - - - - -

**Supplementary Table S4. Genomic data of mammals used to retrieve DSFGs and compute AASD of SCOs.** For each species, the relative ID, taxonomic information, BUSCO statistics, NCBI accession number, and source publication are reported.

Species	ID	Class	Group	Order	Type	BUSCO statistics (mammalia.ncbi)	NCBI acc. no.	Reference
<i>Gallus gallus</i>	Ggal	Aves	Neognathae	Galliformes	Genome	C:99.0%[S:98.6%,D:0.4%],F:0.2%,M:0.8%	GCF_0166949485.2	Vertebrate Genome Project
<i>Chrysochloris asiatica</i>	Casi	Mammalia	Afrotheria	Afroscordida	Genome	C:98.9%[S:97.4%,D:0.6%],F:0.4%,M:0.9%	GCF_000296735.1	Vertebrate Genome Project
<i>Elephas maximus indicus</i>	Emax	Mammalia	Afrotheria	Proboscidea	Genome	C:98.9%[S:98.3%,D:0.6%],F:0.4%,M:0.7%	GCF_024166365.1	Vertebrate Genome Project
<i>Trichechus manatus latirostris</i>	Tman	Mammalia	Afrotheria	Tubulidentata	Genome	C:96.1%[S:95.7%,D:0.4%],F:1.8%,M:2.1%	GCF_000294395.1	Foote et al., 2015
<i>Orycteropus afer afer</i>	Oafe	Mammalia	Euarctioglires	Lagomorpha	Genome	C:98.3%[S:96.4%,D:1.9%],F:0.5%,M:1.6%	GCF_030433575.1	N/A
<i>Ocytrotus princeps</i>	Opi	Mammalia	Euarctioglires	Primates	Genome	C:97.3%[S:95.1%,D:2.7%],F:1.7%,M:1.2%	GCF_001604975.1	Vertebrate Genome Project
<i>Cebus imitator</i>	Cimi	Mammalia	Euarctioglires	Primates	Genome	C:99.6%[S:97.3%,D:2.3%],F:0.2%,M:0.2%	GCF_00001405.40	Orkin et al., 2021
<i>Homo sapiens</i>	Hsap	Mammalia	Euarctioglires	Primates	Genome	C:98.3%[S:97.2%,D:1.1%],F:0.4%,M:1.3%	GCF_002740605.2	Vertebrate Genome Project
<i>Lemur catta</i>	Lcat	Mammalia	Euarctioglires	Rodentia	Genome	C:96.4%[S:95.7%,D:0.7%],F:1.7%,M:1.9%	The Genome Sequencing Platform	
<i>Cavia porcellus</i>	Cpor	Mammalia	Euarctioglires	Rodentia	Genome	C:99.4%[S:98.7%,D:0.7%],F:0.2%,M:0.4%	GCF_000001035.27	Genome Reference Consortium
<i>Mus musculus</i>	Mms	Mammalia	Euarctioglires	Artiodactyla	Genome	C:99.1%[S:96.9%,D:2.2%],F:0.5%,M:0.6%	GCF_002686445.1	Mead et al., 2020
<i>Sexturus carolinensis</i>	Scar	Mammalia	Euarctioglires	Artiodactyla	Genome	C:98.7%[S:97.0%,D:1.7%],F:0.6%,M:0.7%	GCF_019923385.1	Deng et al., 2016
<i>Bubalus bubalis</i>	Bbul	Mammalia	Euarctioglires	Artiodactyla	Genome	C:98.4%[S:95.7%,D:2.7%],F:0.6%,M:1.0%	Genome 10K	
<i>Balaenoptera musculus</i>	Bmus	Mammalia	Euarctioglires	Artiodactyla	Genome	C:98.7%[S:98.3%,D:0.4%],F:0.7%,M:0.6%	GCF_000803125.2	Elbers et al., 2019
<i>Camelus dromedarius</i>	Cdro	Mammalia	Euarctioglires	Artiodactyla	Genome	C:96.4%[S:95.2%,D:3.5%],F:0.5%,M:0.8%	GCF_0300280280.51	Vertebrate Genome Project
<i>Phacochoerus africanus</i>	Paf	Mammalia	Euarctioglires	Artiodactyla	Genome	C:98.8%[S:98.3%,D:0.5%],F:0.6%,M:0.6%	GCF_016306355.1	N/A
<i>Tursiops truncatus</i>	Ttru	Mammalia	Euarctioglires	Artiodactyla	Genome	C:97.3%[S:95.2%,D:2.1%],F:1.1%,M:1.6%	GCF_011762395.1	Xiong et al., 2009
<i>Aitropoda melanoleuca</i>	Amel	Mammalia	Euarctioglires	Carnivora	Genome	C:97.3%[S:96.6%,D:0.7%],F:1.3%,M:1.4%	GCF_002007345.2	Fan et al., 2019
<i>Canis lupus familiaris</i>	Clup	Mammalia	Euarctioglires	Carnivora	Genome	C:98.5%[S:96.7%,D:1.8%],F:0.6%,M:0.9%	GCF_01110085.1	Wang et al., 2021
<i>Mirounga angustirostris</i>	Mang	Mammalia	Euarctioglires	Carnivora	Genome	C:96.7%[S:94.5%,D:2.2%],F:1.9%,M:1.4%	GCF_02228875.2	Moreno et al., 2024
<i>Panthera tigris</i>	Ptig	Mammalia	Euarctioglires	Carnivora	Genome	C:99.4%[S:98.9%,D:0.5%],F:0.3%,M:0.3%	GCF_018350195.1	Bredemeyer et al., 2023
<i>Desmodus rotundus</i>	Drot	Mammalia	Euarctioglires	Chiroptera	Genome	C:98.2%[S:97.2%,D:1.0%],F:0.5%,M:1.3%	GCF_022882495.1	Bat 1K
<i>Pteropus griseatus</i>	Pgr	Mammalia	Euarctioglires	Chiroptera	Genome	C:97.2%[S:96.9%,D:0.3%],F:1.1%,M:1.7%	GCF_002729225.1	Foare et al., 2020
<i>Rhinolophus ferrumequinum</i>	Rfer	Mammalia	Euarctioglires	Chiroptera	Genome	C:99.2%[S:97.9%,D:1.3%],F:0.3%,M:0.5%	Vertebrate Genome Project	
<i>Ceratotherium simum simum</i>	Csim	Mammalia	Euarctioglires	Perissodactyla	Genome	C:98.8%[S:98.6%,D:0.2%],F:0.9%,M:0.3%	GCF_000283315.1	N/A
<i>Equus quagga</i>	Mjav	Mammalia	Euarctioglires	Perissodactyla	Genome	C:98.5%[S:95.0%,D:3.5%],F:0.5%,M:1.0%	GCF_021613505.1	Vilstrup et al., 2013
<i>Sarcophilus harrisii</i>	Shar	Mammalia	Euarctioglires	Pholidota	Genome	C:95.7%[S:93.7%,D:2.0%],F:1.9%,M:2.4%	GCF_014570335.1	N/A
<i>Monodelphis domestica</i>	Mdom	Mammalia	Metatheria	Dasyuromorphia	Genome	C:95.5%[S:94.5%,D:1.0%],F:0.9%,M:3.6%	GCF_00263505.1	Stammnitz et al., 2023
<i>Ornithorhynchus anatinus</i>	Oana	Mammalia	Metatheria	Didelphimorphia	Genome	C:95.1%[S:92.3%,D:2.8%],F:0.9%,M:4.0%	GCF_027887165.1	Vertebrate Genome Project
<i>Dasypus novemcinctus</i>	Dnov	Mammalia	Prototheria	Monotremata	Genome	C:92.3%[S:91.2%,D:1.1%],F:1.4%,M:6.3%	GCF_004115125.2	zhou2021 platypus
<i>Choloepus didactylus</i>	Cdid	Mammalia	Xenarthra	Cingulata	Genome	C:96.9%[S:94.3%,D:2.6%],F:0.7%,M:2.4%	GCF_030445035.1	Vertebrate Genome Project
				Xenarthra	Genome	C:97.8%[S:91.9%,D:5.9%],F:0.7%,M:1.5%	GCF_015220235.1	Vertebrate Genome Project

**Supplementary Table S5. Genomic data of *Drosophila* used to retrieve DSFGs and compute AASD of SCOs.** For each species, the relative ID, taxonomic information, BUSCO statistics, NCBI accession number, and source publication are reported.

Species	ID	Family	Subgenus	Type	BUSCO statistics (diptera_odb10)	NCBI acc. no.	Reference
<i>Anopheles gambiae</i>					C:99.4%[S:99.1%,D:0.3%],F:0.1%,M:0.5%	GCF_943734735.2	Habtewold et al., 2023
<i>Drosophila sechellia</i>	Agam	Culicidae	Cellia	Genome	C:99.9%[S:99.3%,D:0.6%],F:0.0%,M:0.1%	GCF_004382195.2	Chakraborty et al., 2021
<i>Drosophila melanogaster</i>	Dsec	Drosophilidae	Sophophora	Genome	C:100.0%[S:99.7%,D:0.3%],F:0.0%,M:0.0%	GCF_000001215.4	Hoskins et al., 2015
<i>Drosophila erecta</i>	Dmel	Drosophilidae	Sophophora	Genome	C:99.9%[S:99.5%,D:0.4%],F:0.0%,M:0.1%	GCF_003286155.1	Dong et al., 2022
<i>Drosophila suzukii</i>	Dere	Drosophilidae	Sophophora	Genome	C:99.7%[S:96.5%,D:3.2%],F:0.1%,M:0.2%	GCF_013340165.1	Paris et al., 2020
<i>Drosophila elegans</i>	Dsuz	Drosophilidae	Sophophora	Genome	C:99.8%[S:99.5%,D:0.3%],F:0.1%,M:0.1%	GCF_018152505.1	Kim et al., 2021
<i>Drosophila serrata</i>	Dele	Drosophilidae	Sophophora	Genome	C:99.9%[S:97.5%,D:2.4%],F:0.0%,M:0.1%	GCF_002093755.2	Allen et al., 2017
<i>Drosophila kikkawai</i>	Dser	Drosophilidae	Sophophora	Genome	C:100.0%[S:99.1%,D:0.9%],F:0.0%,M:0.0%	GCF_018152535.1	Kim et al., 2021
<i>Drosophila bimaculata</i>	Dkik	Drosophilidae	Sophophora	Genome	C:99.9%[S:99.2%,D:0.7%],F:0.0%,M:0.1%	GCF_018153845.1	Kim et al., 2021
<i>Drosophila ananassae</i>	Dbip	Drosophilidae	Sophophora	Genome	C:99.6%[S:99.3%,D:0.3%],F:0.0%,M:0.4%	GCF_017639315.1	Tvedte et al., 2021
<i>Drosophila pseudoobscura</i>	Dana	Drosophilidae	Sophophora	Genome	C:99.7%[S:98.8%,D:0.9%],F:0.1%,M:0.2%	GCF_009870125.1	Liao et al., 2021
<i>Drosophila miranda</i>	Dpse	Drosophilidae	Sophophora	Genome	C:99.8%[S:85.6%,D:14.2%],F:0.1%,M:0.1%	GCF_003369915.1	Mahajan et al., 2018
<i>Drosophila willistoni</i>	Dmir	Drosophilidae	Sophophora	Genome	C:99.6%[S:98.4%,D:1.2%],F:0.0%,M:0.4%	GCF_018902025.1	Ranz et al., 2023
<i>Drosophila arizonae</i>	Dwil	Drosophilidae	Drosophila	Genome	C:95.7%[S:95.3%,D:0.4%],F:1.2%,M:3.1%	GCF_001654025.1	Sanchez-Flores et al., 2016
<i>Drosophila hydei</i>	Dari	Drosophilidae	Drosophila	Genome	C:99.7%[S:97.5%,D:2.2%],F:0.1%,M:0.2%	GCF_003285905.1	Dong et al., 2022
<i>Drosophila grimshawi</i>	Dhyd	Drosophilidae	Drosophila	Genome	C:99.9%[S:99.2%,D:0.7%],F:0.0%,M:0.1%	GCF_018153295.1	Kim et al., 2021
<i>Drosophila albomicans</i>	Dgrl	Drosophilidae	Drosophila	Genome	C:99.9%[S:99.1%,D:0.8%],F:0.0%,M:0.1%	GCF_009650485.2	Mai et al., 2020
<i>Drosophila busckii</i>	Dbus	Drosophilidae	Drosophila	Genome	C:98.1%[S:97.4%,D:0.7%],F:0.3%,M:1.6%	GCF_011750605.1	Renschler et al., 2019

**Supplementary Table S6. Complete set of DSFGs in bivalves.** For each gene, the species ID (Sp. ID) as in Supp. Tab. S1, the accession number (Gene ID), the Possvm-based annotation, and the CDD domains (including their Pssm-ID) are indicated.

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Airc	Contig6.279	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
	scaffold.235.403	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Apur	XP_0152698016..1	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Cang	EVM0027346..1	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Cari	XP_011441049..2	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Cgig	XP_0222333988..1	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Cvir	XP_041358115..1	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Gaeg	M00000013645	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Hbia	M00000045261	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Hruf	XP_046372338..2	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
	XP_046335704..1	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Meal	XP_05268518..1	Dmr1	Dmr1	Dmr1-OG4/NA	Doublesex DNA-binding motif (214606)	N/A
Mcor	CAAC5397186..1	Dmr1	Dmr1	Dmr1-OG4/NA	Doublesex DNA-binding motif (214606)	N/A
Medu	CAG2232556..1	Dmr1	Dmr1	Dmr1-OG4/NA	Doublesex DNA-binding motif (214606)	N/A
Mgal	VDI03798..1	Dmr1	Dmr1	Dmr1-OG4/NA	Doublesex DNA-binding motif (214606; partial)	-
Mmar	MMAM00000008302	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Mner	g120437.tl..1	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Obim	XP_014782618..1	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Oedu	XP_048736174..2	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Osin	XP_036368911..1	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Osin	XP_036366646..1	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Pcan	XP_029647701..2	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Pcan	XP_023090051..1	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Pmar	DN30637..c0.g1.i1.p1	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Pmax	XP_033733655..1	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Ppur	DN2292..c0..g1..i1.p1	Dmr1	Dmr1	Dmr1-OG4/NA	Doublesex DNA-binding motif (214606)	N/A
Pstr	KAK3599448..1	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Pvir	s01850..168	Dmr1	Dmr1	Dmr1-OG4/NA	Doublesex DNA-binding motif (214606)	N/A
Pyes	XP_021353714..1	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Sbro	EVM0020695..1	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Sglo	Sg011295	Dmr1	Dmr1	Dmr1-OG0/NA	Doublesex DNA-binding motif (214606)	N/A
Airc	Contig172..94	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Apec	DN21321..1..g1..i1..p1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Bglia	XP_053887190..1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Cpli	DN116454..c0..g1..i1..p1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Gaeg	XP_041359971..1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Hruf	XP_048255484..1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Mcor	CAAC5404148..1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Medu	CAG22325366..1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Mgal	VDA2971..1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Mmer	XP_045156965..2	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Mphi	scaf..30477..1..9	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Obim	XP_052832484..1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Osin	XP_029650766..2	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Pcor	DN165632..c0..g1..i1..p1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Pcor	DN115233..c0..g1..i1..p1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Pmar	DN28033..c0..g1..i1..p1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Pmax	XP_033738864..1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Pstr	KAK3603675..1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Pvir	s00097..235	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Pyes	XP_021368778..1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Rphi	XP_060589226..1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Sbro	EVM0001645..1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Scon	Chr8..1365	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Tgra	KA18306274..1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Tsqu	DN5151730..c0..g1..i1..p1	Dmr1	Dmr1	Dmr1-t2	Dmr1-t2	-
Acal	XP_005096932..1	Dmr1	Dmr1	Dmr1-t3	Dmr1-t3	-
Aire	Contig349..40	Dmr1	Dmr1	Dmr1-t3	Dmr1-t3	-
Amar	Amar12594..1	Dmr1	Dmr1	Dmr1-t3	Dmr1-t3	-
Apur	scaf..95..76	Dmr1	Dmr1	Dmr1-t3	Dmr1-t3	-
Bglia	XP_013077145..1	Dmr1	Dmr1	Dmr1-t3	Dmr1-t3	-
Cang	XP_052687934..1	Dmr1	Dmr1	Dmr1-t3	Dmr1-t3	-
Cari	EVM0028466..1	Dmr1	Dmr1	Dmr1-t3	Dmr1-t3	-

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Cgig	XP_011427033.2	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Cpli	DN37429_c0.g2.i1.p1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	N/A	-	
Csin	Hic.asm.11.1.174	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Cvir	XP_022317913.1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Dpol	XP_052246678.1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Gaeg	XP_041357684.1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Hbia	M00000034631	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270600)	-	
Hruf	XP_046367747.2	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Mare	XP_052777877.1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Mcal	XP_052765466.1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Mcor	CAAC5360634.1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Medu	AGP226664.1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Mgal	VID32052.1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Mmar	MMAM00000039146	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Mmer	XP_045157038.1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Mner	g243052.t1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Mphi	scat42165.0.7	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Oedu	XP_048761857.1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Pcan	XP_025110327.1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Pcor	DN33000_c0.g1.i1.p1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Pmax	XP_033737545.1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Pvir	s000978272	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Pyes	XP_0213777273.1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Rphi	XP_060576862.1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Sbro	EYMM0006488.1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Scon	Chr8.2435	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Sglo	Sg014397	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Tgra	KA18305800.1	Dmr <sub>t</sub> -3	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Acal	XP_005096931.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Airc	Contig349.42	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Amar	Amal2441	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Apec	DN10969_c0.g1.i1.p1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Bapl	Scaffold95.77	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Bglia	XP_013089080.2	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Cang	XP_052685010.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Cari	EWMM0008623.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Ccgig	NP_001295834.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Cpli	DN67594_c0.g1.i1.p1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Cpli	DN55944_c0.g1.i1.p1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Cpli	DN54583_c0.g1.i1.p1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Cpli	DN9274_c0.g1.i1.p1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Cpli	DN57594_c0.g2.i1.p1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Csin	Hic.asm.11.400	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Csin	Hic.asm.11.338	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Cvir	XP_022319926.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Dpol	XP_052252151.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Dpol	XP_052224782.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Dpol	XP_05274906.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Gaeg	XP_041357898.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Hbia	M00000038945	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Hbia	M00000038948	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Hbia	M00000038946	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Hbia	M00000019813	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Hbia	M00000038947	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Hruf	XP_046367734.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Lorb	DN147341_c0.g1.i1.p1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Lorb	DN40512_c0.g1.i1.p1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Mare	XP_052776891.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Mare	XP_052774906.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Mare	XP_052776885.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Mcal	XP_052092340.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Mchi	DNS35166_c0.g1.i1.p1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Medu	CAc5398878.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Medu	CAG2203978.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Mgal	VDI24477.1	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Mmar	MIMAM00000040448	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	N/A	-	
Mmar	MIMAM00000008515	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	
Mmar	MMAM00000016566	Dmr <sub>t</sub> -4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270600)	-	

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Mmar	MMAM00000047004	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Mmar	MMAM00000044361	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	N/A	-
Mmer	XP_045157593_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Mmer	XP_045157053_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	N/A	-
Mmer	XP_045157113_2	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600)	-
Mmod	DN116828_c0.g1.i1.p1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600)	-
Mmer	g34603.t1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	N/A	-
Mmer	g58531.t1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Mmer	g241174.t1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Mmer	g198270.t1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Mphi	scaf.68796.0.3	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Oedu	XP_048763391_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602, 270465)	-
Pcan	XP_025210328_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	N/A	-
Pcor	DN12587_c0.g3.i3.p1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600)	-
Pcor	DN1623_c0.g2.i1.p1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601)	-
Pcor	DN1623_c0.g1.i3.p1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Pcor	DN12587_c0.g1.i4.p1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Pgen	DN68344_c0.g1.i1.p1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600)	-
Pgen	DN35556_c0.g1.i2.p1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600)	-
Pmax	XP_0338737544_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270602, 270465)	-
Poku	DN35178_c0.g1.i2.p1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Pstr	KAK3612677_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600)	-
Pstr	KAK3583105_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	N/A	-
Pstr	KAK3583112_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Pstr	KAK3583110_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Pstr	KAK3583109_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Pstr	s00733g2	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Pvtr	XP_021377274_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270465)	-
Pyes	DN36973_c0.g1.i1.p1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	N/A	-
Rdec	XP_066600638_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Rphi	XP_066600746_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Rphi	XP_0605785674_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Sbro	EVM0004355_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Sbro	EVM0021940_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	N/A	-
Scon	Chr8.1999	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Scon	Chr8.2143	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Sgrn	Sgo106592	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Sgra	DN54078_c0.g1.i1.p1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Sgra	DN6659_c0.g1.i3.p1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Tgra	KA18305799_1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Tsqu	DN75749_c0.g1.i1.p1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Apec	DN3872_c0.g4.i1.p1	Dmrt	Dmrt-4/5	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Mchi	DN34711_c0.g1.i2.p1	Dmrt	N/A	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Scon	Chr8.738	Dmrt	N/A	Doublesex DNA-binding motif (214606)	CUE-like DMA domain (270553, 270601, 270600, 270602)	-
Acal	XP_00597243_2	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Airc	Contig6.157	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Amar	Amo08751	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Apcc	DN107972_c0.g1.i1.p1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Apur	scf0ffld.124.7	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Bglia	XP_013067134_2	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Cang	XP_052701295_1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Cari	EVM0004613_1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Chu	DN101169_c0.g1.i1.p1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Cgig	XP_011413445_1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Cpil	DN47094_c0.g1.i1.p1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Csin	Hic.asm.10.638	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Csin	Hic.asm.10.437	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Cvir	XP_022333322_1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Cvir	XP_022334050_1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Dpol	XP_052272379_1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Gaeg	XP_041352454_1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Hbia	MO000018167	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Hruf	XP_046371021_1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Mcal	XP_052169228_1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Mchi	DN23553_c0.g1.i1.p1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Mcor	CA5C574046_1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Medu	CA62201348_1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Mgal	VID17457.1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Mmar	MMAM00000008663	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552; partial)	-
Mmar	XP_045173733..1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872; partial); HNF3 C-terminal domain (430552)	-
Mmod	DN103780_c0.g1.i1.p1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872; partial)	-
Miner	g1922.i17.t1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872; partial)	-
Mphi	scat:4682.0.0	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872; partial)	-
Obim	XP_01478201..2	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872; partial)	-
Oedu	XP_04873259..1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872)	-
Pcan	XP_0283090786..1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552; partial)	-
Pcor	DN3042..e0..g1..i1..p1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872; partial); HNF3 C-terminal domain (430552)	-
Pgen	DN174637_c0.g1.i1.p1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872; partial)	-
Pmar	DN30866_c0.g1.i1.p1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872; partial)	-
Pmax	XP_033734080..1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872; partial)	-
Pstr	KAK357847..1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552; partial)	-
Pvir	s00068s447	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872); HNF3 C-terminal domain (430552)	-
Pyes	XP_021361791..1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872; partial)	-
Rphi	XP_060590755..1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872; partial)	-
Sbro	EVM0003194..1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872; partial)	-
Chr4.2670	Chr4..2670	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872; partial)	-
Scon	Sg008s464	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872; partial)	-
Sglo	DIN78052_c0..g1..i1..p1	Fox	Fox-A	Forkhead domain A1 (410812)	Forkhead N-terminal (369872; partial)	-
Sgra	XP_005089018..1	Fox	Fox-B	Forkhead domain B2 (410812)	Forkhead N-terminal (369872; partial); HNF3 C-terminal domain (430552)	-
Acal	XP_005089018..1	Fox	Fox-B	Forkhead domain B2 (410812)	Forkhead N-terminal (369872; partial); HNF3 C-terminal domain (430552)	-
Airc	Contig36..38	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Apur	XP_0130785204..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Bgla	XP_052700333..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Cang	EVMM0003536..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Cari	DN98613_c0..g1..i1..p1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Chru	XP_011445364..2	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Cgig	Csin	Hic..asm..16..1..347	Fox	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Cvif	XP_022334612..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Dpol	DPol	XP_052256324..1	Fox	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Dpol	XP_052281977..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Dpol	XP_041361159..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Gaeg	Hbia	MD000029836	Fox	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Hruf	XP_046358590..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Lorb	DNS5589..c3..g1..i1..p1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Mare	XP_052791461..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Mcal	XP_052100219..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Mcor	CAC5382565..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Medu	CAG2229716..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Mgal	V1D169670..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Minar	MMAM0000015629	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Mmer	XP_045215505..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Mmer	g250725..t1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Mphi	scat:10920..0..0	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Obim	XP_052832317..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Pcor	XP_048732871..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Oedu	XP_028653697..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Osin	XP_025075261..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Pcan	DN23979_c0..g1..i1..p1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Pcor	DN23979..c1..g1..i1..p1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Pstr	XP_033749587..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Pvir	s001..89s215	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Pyes	XP_060564412..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Rphi	Chr5..12	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Scon	Sg012012..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Sglo	KA18304921..1	Fox	Fox-B	Forkhead domain B2 (410817)	Forkhead domain B2 (410817)	N/A
Tgra	XP_005106277..1	Fox	Fox-C	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	N/A
Acal	Contig58..63	Fox	Fox-C	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	N/A
Airc	Arna17094	Fox	Fox-C	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	N/A
Amar	scat:577..50	Fox	Fox-C	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	N/A
Apur	Bgla	XP_055890240..1	Fox	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	N/A
Cang	XP_052715579..1	Fox	Fox-C	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	N/A
Cari	ENM0002771..1	Fox	Fox-C	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	N/A
Chu	DN96575..c0..g1..i1..p1	Fox	Fox-C	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	N/A
Cgig	XP_011417585..2	Fox	Fox-C	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	N/A

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)		Additional domains (Pssm-ID)	Notes
Cpli	DN157725_c0_g_1.i1.p1	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Csin	Hic.acm.17.1.357	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Csin	Hic.acm.17.1.443	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Cvir	XP_022346235.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Dpol	XP_052247387.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Gaeg	XP_041377087.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Hbia	M00000031958	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Hruf	XP_046372770.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Mare	XP_052819073.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Mcal	XP_052063314.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain C(410791)	N/A
Mehi	DN13809_c0_g_1.i1.p1	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Mcor	CAGC537404.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Medu	CAGZ2206844.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Mgal	VD122482.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Mmar	MMAM00000035616	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Mmer	XP_045194706.2	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Mner	g82158.t1	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Mphi	scf.69950.1.0	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Obim	XP_014786040.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Oedu	XP_048762038.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Osin	XP_029653806.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Pcan	XP_025115697.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Pcor	DN14158_c0_g_2.i1.p1	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Pmax	XP_033755061.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Pstr	KAK3590939.1	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Pvir	s020236.i2	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Pyes	XP_021346967.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Rphi	XP_060597004.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Sbro	EVM0022192.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Scon	Chr11.448	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Sglo	Sg009485	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain L1 (410801)	N/A
Tgra	KA183032551.i	Fox	Fox-C	Fox-C	Fox-C	Forkhead domain C2 (410801)	N/A
Airc	Contig1003.15	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Amar	Amar1686	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Apec	DN87882_c0_g_1.i1.p1	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Apur	scatfold.13962.i1	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Bglia	XP_013096936.i2	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Cang	XP_052688370.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Cari	EVM0005770.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Cgig	XP_0114446328.i2	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Cpli	DN23774_c0_g_1.i1.p1	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Csin	Hic.acm.11.1425	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Cvir	XP_022316146.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Dpol	XP_052256035.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Gaeg	XP_052256591.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Hbia	XP_041356731.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Mcal	M00000030583	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Hruf	XP_046329290.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Lorb	DN224803_c0_g_1.i1.p1	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Mare	XP_052777467.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Mcal	XP_052095202.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Mgal	VID107735.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Mcor	CAC5382691.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Medu	CAGZ2204666.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Medu	CAGZ2248150.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Medu	CAGZ2203862.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Mgal	VIDH30661	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D3 (410821)	N/A
Mgal	MAMAM0000002467	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Mgal	XP_045157253.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Mgal	g192986.t1	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Mgal	scf.69489.1.2	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Mphi	scf.42856.0.4	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Obim	XP_052829256.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A
Oedu	XP_048762457.i	Fox	Fox-D	Fox-D	Fox-D	Forkhead domain D4 (410822)	N/A

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Pcan	XP_0251101523_1	Fox	Fox-D	Forkhead domain D4 (410822)	N/A	-
Pcor	DN15187_c0.g1.i1.p1	Fox	Fox-D	Forkhead domain D4 (410822)	N/A	-
Pmar	DN31265_c0.g1.i1.p1	Fox	Fox-D	Forkhead domain D4 (410822)	N/A	-
Pmax	XP_033737945_1	Fox	Fox-D	Forkhead domain D4 (410822)	N/A	-
Pstr	KAK3592139_1	Fox	Fox-D	Forkhead domain D4 (410822)	N/A	-
Pvir	s01402691	Fox	Fox-D	Forkhead domain D4 (410822)	N/A	-
Pvir	s00097340	Fox	Fox-D	Forkhead domain D4 (410822)	N/A	-
Pyes	XP_021345225_1	Fox	Fox-D	Forkhead domain D4 (410822)	N/A	-
Rphi	XP_060585828_1	Fox	Fox-D	Forkhead domain D4 (410822)	N/A	-
Sbro	EVM0002351_1	Fox	Fox-D	Forkhead domain D4 (410822)	N/A	-
Scon	Chr8_2069	Fox	Fox-D	Forkhead domain D4 (410822)	N/A	-
Sglo	Sglo13024	Fox	Fox-D	Forkhead domain D4 (410822)	N/A	-
Tgra	KA18306624_1	Fox	Fox-D	Forkhead domain D4 (410822)	N/A	-
Airc	Contig089_19	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Amar	Amal2850	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Apar	scatfold_253_23	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Cang	XP_02688878_1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Cari	EVM0003839_1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Chu	DN108936_c0.g1.i1.p1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Cgig	XP_011444776_2	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Cvir	XP_022319236_1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Dpol	XP_012286560_1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Hbia	M00000038943	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Hruf	XP_046353578_2	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Mare	XP_052775423_1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Meal	XP_052075782_1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Mcor	CA.C5384360_1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Medu	CAG2217852_1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Medu	CAG2194171_1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Medu	CAG2199036_1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Mgal	V1DH0460_1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Mmar	MMAM00000033594	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Mmer	XP_045157592_2	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Mmod	DN117568_c0.g1.i1.p1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Mner	g241620_t1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Mphi	scat31587_0.4	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Oedu	XP_048762291_1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Pmar	DN27017_c0.g1.i1.p1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Pmax	XP_033737819_1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Pstr	KAK3583103_1	Fox	Fox-E	Forkhead domain E (410793)	N/A	-
Pvir	s00145g54	Fox	Fox-F	Forkhead domain F (410823)	N/A	-
Pyes	XP_021373858_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Rdec	DN24595_c4.g1.i1.p1	Fox	Fox-F	Forkhead domain F (410823)	N/A	-
Rphi	XP_060575687_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Sbro	EVMM0110028_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Sglo	Sglo009305	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Apur	XP_005105969_2	Fox	Fox-F	Forkhead domain F (410823)	N/A	-
Acal	Contig1133_18	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Airc	EVM0011190_1	Fox	Fox-F	Forkhead domain F (410823)	N/A	-
Amar	Amara39500	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Cgig	DN32615	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Apec	DN75342_c0.g1.i1.p1	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Apur	scatfold_860_37	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Cang	XP_053892380_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Cari	Contig1133_18	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Chia	Amara39500	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Cpili	DN7628_c0.g1.i1.p1	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Csin	Hicasm_17_158	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Cvir	XP_022335664_1	Fox	Fox-F	Forkhead domain F (410823)	N/A	-
Dpol	XP_052232755_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Gaeg	XP_041375666_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Hbia	M0000007664	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Hruf	XP_046372649_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Mcal	XP_052060477_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Mcor	CA.C5387332_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Medu	CA.G252876_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Mgal	VID121852_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	-
Mmar	MMAM0000030848	Fox	Fox-F	Forkhead domain F (410794)	N/A	-

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Mmer	XP_045194642_1	Fox	Fox-F	Fox-F	Forkhead domain F (410794)	N/A
Mmod	DN104261_c0_g1.i1.p1	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Muer	g106129_t1	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Mphi	scf40546_0.2	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Obim	XP_014777539_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Oedu	XP_048732202_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Pcor	XP_025116910_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Pcor	DN180603_c0_g1.i1.p1	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Pcor	DN129940_c0_g1.i1.p1	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Pmar	DN14344_c0_g1.i1.p1	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Pmax	XP_033755005_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Pstr	KAK3601654_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Pvir	s133835g10	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Pyes	XP_021356008_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Rphi	XP_0660601663_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Sbro	EVM0015186_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Scon	Chr11_927	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Scon	Chr11_810	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Sglo	Sgo05267	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Tgra	KA18302829_1	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Tsqu	DN137576_c0_g1.i1.p1	Fox	Fox-F	Forkhead domain F (410794)	N/A	
Acal	XP_005099252_2	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Acal	XP_005099253_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Airc	Contig025_38	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Amar	Anal0381	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Apec	DN1036_c0_g1.i1.p1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Apur	scatford_36470_28	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Bglia	XP_055879295_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Cang	XP_052699015_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Cari	EVM0001189_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Chu	DN104980_c0_g1.i2.p1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Cgig	XP_011427689_2	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Cpii	DN58419_c0_g1.i1.p1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Csin	Hic_asm_10_1034	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Cvri	XP_022334541_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Dpol	XP_052270224_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Dpol	XP_052270147_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Gaeg	XP_041354930_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Hbia	M0000035850	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Hruf	XP_046371537_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Hruf	MMAM0000030730	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Mcal	XP_048255351_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Mcor	CAAC54056396_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Medu	CAG2193433_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Mgal	V1D124297_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Mmar	MMAM0000030730	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Mmer	XP_045162348_2	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Mmod	DN60588_c0_g1.i1.p1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Mner	g133265_t1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Mphi	scf115017_0.3	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Obim	XP_032821454_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Oedu	XP_048737541_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Pcan	XP_025105677_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Pcan	XP_025106039_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Pcor	XP_025105724_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Pcor	DN81635_c0_g2.i1.p1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Pcor	DN81635_c0_g1.i1.p1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Pgen	DN112984_c0_g1.i1.p1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Pmar	DN28516_c0_g1.i1.p1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Pmax	XP_033734631_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Poku	DN41090_c0_g2.i1.p1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Pstr	KAK3604690_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Pvir	s0038335	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Pyes	XP_021363790_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Rphi	XP_060589805_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Sbro	EVM0011335_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Sbro	EVM0012606_1	Fox	Fox-G	Forkhead domain G (410795)	N/A	
Scon	Chr3_2805	Fox	Fox-G	Forkhead domain G (410795)	N/A	

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Sglo	Sglo14601	Fox	Fox-G	Forkhead domain G (410795)	N/A	-
Sgra	DN149488_c0.g1.i1.p1	Fox	Fox-G	Forkhead domain G (410795)	N/A	-
Airc	Contig18244.2	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Airc	Contig178.106	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Amar	Amar16564	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Amar	Amar05868	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Cang	XP_022684756.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Cari	EVM002137.7.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Cflu	DN138466_c0.g1.i1.p1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Cflu	XP_034313225.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Cgig	Hic.asm.14.811	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Csin	XP_022314590.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Cvir	MO0000018729	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Hbia	XP_048254913.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Hruf	XP_048255113.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Hruf	CAAC5408624.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Mcor	CAAC5397897.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Mcor	CAAC5397906.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Mcor	CAAC5403969.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Medu	CAAG2228803.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Medu	CAAG21188004.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Medu	CAAG2252853.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Medu	CAAG2202596.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Mgal	VDI162725.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Mgal	VDIH033947.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Mgal	VDI10844.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Minar	MMAM00000022684	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Mmer	XP_053378216.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Mmer	XP_045194303.2	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Mmer	XP_045198985.2	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Mner	g213542.tl	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Mphi	scat17525.0.4	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Mphi	scat28666.1.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Oedu	XP_048759429.2	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Pcan	XP_025075954.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Pcor	DN116957_c0.g1.i1.p1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Pmax	XP_032755807.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Pstr	KAK3603559.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Pvir	s75596633	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Pvir	s1002346131	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Rphi	XP_0605558970.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Rphi	XP_060567331.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Rphi	XP_060604067.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Sbro	EVM0016618.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Sbro	EVM0013817.1	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Scon	Chr11.1.359	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Scon	Chr2.2082	Fox	Fox-H	Forkhead domain H (410796)	N/A	-
Sglo	Sglo13003	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Acal	XP_005108651.1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Airc	Contig775.5	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Amar	Amar02822	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Apec	DN20109_c0.g1.i6.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Apur	XP_011445234.1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Bglia	DN65792_c0.g1.i1.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Cang	Hic.asm.0.1540	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Cang	XP_022319181.1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Cari	EVM0003558.1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Chu	DN127407_c0.g2.i1.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Cgig	XP_011445234.1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Cgig	XP_011445234.1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Cpii	MO000003225	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Csin	XP_046330175.1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Cvir	XP_022319268.1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Dpol	XP_052265485.1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Gaeg	XP_041362703.1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Hbia	DN146717.c0.g1.i1.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Hruf	DN146717.c0.g1.i1.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Lorb	XP_052764656.1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Mare	XP_052816854.1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Mare	XP_052764667_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Mare	XP_052775202_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Mare	XP_052775217_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Mare	XP_052775230_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Meal	XP_052065038_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Michi	DN41583_c0_g1.i15.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Mcor	CAC5405074_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Medu	CAG2242807_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Mgal	VDI12691_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Mmar	MMAM00000019873	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Mmer	XP_045212565_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Mmod	DN23659_c0_g1.i1.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Mner	g198765_t1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Mphi	scf.33310.2.9	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Obim	XP_052824622_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Oedu	XP_048763213_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Osin	XP_029638410_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Pcan	XP_025095423_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Pcor	DNN84891_c0_g1.i2.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Pcor	DNN480_c0_g2.i1.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Pgen	DNN8491_c0_g2.i2.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Pmar	DNN2399_c8_g1.i1.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Pmax	XP_03283751.c1.g1.i1.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Poku	DN19777_c2.g1.i1.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Ppur	DN2521_c0_g1.i4.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Pstr	KAK3579229_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Pvir	s01693e10	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Pyes	XP_021351058_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Rdec	DNN22834_c9.g7.i1.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Rphi	EVMM0018668_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Sbro	Chr1.3201	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Scon	Chr1.3198	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Sglo	Sglo00050	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Sgra	DN10939_c0_g1.i1.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Tgra	KA18318321_1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Tsqu	DN6625_c2.g1.i2.p1	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Airc	XP_005111247_3	Fox	Fox-J1	Forkhead domain J1 (410797)	N/A	-
Contig53_201	Fox	Fox-J1	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Amar	Ama34942	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Apur	Scafold_1.i1801.24	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Bglia	XP_013070049_1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Cang	XP_052716138_1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Cari	EVMM0008910_1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Chu	DN58808_c0_g1.i1.p1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Cgig	XP_0114222959_2	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Cpii	DNN731_c5.g1.i2.p1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Csin	Hic.asm.4..381	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Cvir	XP_022341777_1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Dpo1	XP_052285372_1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Gaeg	XP_041378546_1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Mcor	CAC5378041_1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Medu	CAG2221519_1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Hruf	XP_048247606_1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Mgal	XP_0527559824_1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Mare	XP_0527559802_1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Meal	XP_052082445_1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Michi	DN34970_c0_g1.i2.p1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Mmer	XP_053378821_1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Mmod	DN38610_c0_g1.i1.p1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Mphi	scf.40576.0..4	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Obim	XP_052833979_1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Oedu	XP_048739234_1	Fox	Fox-J1	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Osin	XP_029651657_1	Fox	Fox-J2/3	Forkhead domain J3 (410826)	N/A	-
Pcan	XP_0250502118_1	Fox	Fox-J2/3	Forkhead domain J3 (410826)	N/A	-
Pcor	DN2942_c0_g1.i5.p1	Fox	Fox-J2/3	Forkhead domain J3 (410826)	N/A	-
Pcor	DN2942_c0_g4.i3.p1	Fox	Fox-J2/3	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Pcor	DN2942_c0_g4.i3.p2	Fox	Fox-J2/3	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Pgen	DN5381_c0_g1.i7.p1	Fox	Fox-J2/3	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Pmar	DN1364_c0_g1.i3.p1	Fox	Fox-J2/3	Forkhead domain J3 (410826)	N/A	-
Pmax	XP_033763328_1	Fox	Fox-J2/3	Forkhead domain J3 (410826)	N/A	-
Poku	DN14771_c0_g2.i3.p1	Fox	Fox-J2/3	Forkhead domain J3 (410826)	N/A	-
Ppur	DN2181_c0_g1.i5.p1	Fox	Fox-J2/3	Forkhead domain J3 (410826)	N/A	-
Pstr	KAK3583417_1	Fox	Fox-J2/3	Forkhead domain J3 (410826)	N/A	-
Pvir	s00019g6	Fox	Fox-J2/3	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Pyes	XP_021374633_1	Fox	Fox-J2/3	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Rphi	XP_060577634_1	Fox	Fox-J2/3	Forkhead domain J3 (410826)	N/A	-
Sbro	EVM0013081_1	Fox	Fox-J2/3	Forkhead domain FOXJ2, FOXJ3 (410798)	N/A	-
Scon	Chr12_800	Fox	Fox-J2/3	Forkhead domain J3 (410826)	N/A	-
Sglo	Sg003279	Fox	Fox-J2/3	Forkhead domain J3 (410826)	N/A	-
DN1042_c0_g1.i2.p1	Fox	Fox-J2/3	Forkhead domain J3 (410826)	Forkhead domain J3 (410826)	N/A	-
Tgra	KA18320220_1	Fox	Fox-J2/3	Forkhead domain J3 (410826)	N/A	-
Tsqu	DN10376_c0_g1.i2.p1	Fox	Fox-J2/3	Forkhead domain J3 (410826)	N/A	-
Acal	XP_005092494_1	Fox-K	Fox-K	Forkhead associated (FHA) domain (410828)	Forkhead associated (FHA) domain (410828)	-
Apur	scaffold1855_43	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828)	-
Bglia	XP_013090285_1	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828)	-
Cang	XP_052688140_1	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828)	-
Cari	EYAM0009067_1	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828)	-
Cgig	DP64350_c0_g1.i1.p1	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828)	-
Cpli	DP22316096_1	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828)	-
Cvrl	XP_022316096_1	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828)	-
Gaeg	XP_041362451_1	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828)	-
Hbia	M00000008333	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828)	-
Hruf	XP_048248693_1	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828)	-
Mmar	MMAM0000012630	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828)	-
Mphi	scf.14580_0.11	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828; partial)	-
Obim	XP_014772374_1	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828)	-
Oedu	XP_048761057_2	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828)	-
Osin	XP_029646877_1	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828)	-
Rdec	DN21696_c3.g1.i1.p1	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828; partial)	-
Sglo	Sg005389	Fox	Fox-K	Forkhead domain K (410800)	Forkhead associated (FHA) domain (410828)	-
Acal	XP_012940028_1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Airc	Contig58_64	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Amar	Amal1791_4	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Apec	DNT4037_c0_g1.i1.p1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Apur	scffold122_2	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Bglia	XP_055890278_1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Cang	XP_052718686_1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Cari	EYAM0019009_1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Cgig	XP_011417586_2	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Cpli	DN157469_c0_g1.i1.p1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Csin	Hic.asm.17_1.125	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Cvrl	XP_022346240_1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Dpol	XP_052255043_1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Gaeg	XP_041375667_1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Hbia	M0000031057	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Hruf	XP_046344397_2	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Lorb	DN104458_c0_g1.i1.p1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Mare	XP_052817977_1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Mcal	XP_052063315_1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Mcor	CAC5374005_1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Medu	CAG2206845_1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Mgal	V1D122484_1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Mnar	VDHb7507_1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Mmod	MMAM0000028776	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Mner	XP_053402988_1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Muer	DN51324_c0_g1.i1.p1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Mphi	sg69550_0.0	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Obim	XP_014785001_1	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-
Oedu	XP_048762056_2	Fox	Fox-L1	Forkhead domain L1 (410801)	Forkhead domain L1 (410801)	-

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Pcan	XP_025077243_1	Fox	Fox-L1	Forkhead domain L1 (410801)	N/A	-
Pcor	DN28326_c0_g1.i1.p1	Fox	Fox-L1	Forkhead domain L1 (410801)	N/A	-
Pcor	DN187497_c0_g1.i1.p1	Fox	Fox-L1	Forkhead domain L1 (410801)	N/A	-
Pmar	DN30135_c0_g1.i1.p1	Fox	Fox-L1	Forkhead domain L1 (410801)	N/A	-
Pmax	XP_033755354_1	Fox	Fox-L1	Forkhead domain L1 (410801)	N/A	-
Ppur	DN73831_c0_g1.i1.p1	Fox	Fox-L1	Forkhead domain L1 (410801)	N/A	-
Pstr	KAKX59991_1	Fox	Fox-L1	Forkhead domain L1 (410801)	N/A	-
Pvir	s02023g.i1	Fox	Fox-L1	Forkhead domain L1 (410801)	N/A	-
Pyes	XP_021346965_1	Fox	Fox-L1	Forkhead domain L1 (410801)	N/A	-
Rphi	XP_060608039_1	Fox	Fox-L1	Forkhead domain L1 (410801)	N/A	-
Sbro	EVM0016190_1	Fox	Fox-L1	Forkhead domain L1 (410801)	N/A	-
Scon	Chr11.869	Fox	Fox-L1	Forkhead domain L1 (410801)	N/A	-
Sglo	Sglo098486	Fox	Fox-L1	Forkhead domain L1 (410801)	N/A	-
Acal	XP_005101910_2	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Airc	Contig551_34	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Amar	Amu34673	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Apur	sccaffold_84_159	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Bbla	XP_055865110_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Cang	XP_052718506_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Cari	EVM0021728_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Cflu	DN127322_c6_g2.i2.p1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Cgig	NP_001295827_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Cpli	DN75086_c5_g1.i2.p1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Csin	Hic.asm_4_274	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Cvir	XP_022345405_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Cvir	XP_022345173_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Dpol	XP_052212727_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Gaeg	XP_041378252_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Hbia	M00000035173	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Hruf	XP_048250285_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Lorb	DN129129_c0_g1.i1.p1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Mare	XP_052760992_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Mcal	XP_052082415_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Mcor	CAc5401149_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Medu	CAG2233672_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Mgal	VDI49865_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Mgal	VDI49864_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Mmar	MMAM0000016212	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Mmer	XP_045161614_2	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Mmod	DN2410_c0_g1.i1.p1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Miner	g832355_t1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Mphi	scat_50301_0.3	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Obim	XP_0147785648_2	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Oedu	XP_018729555_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Pcan	XP_025085354_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Pcor	DN35937_c0_g1.i1.p1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Pgen	DN134171_c0_g1.i1.p1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Pmar	DN32846_c0_g1.i1.p1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Pmax	XP_033724493_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Pstr	KAK3602726_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Pvir	s002466193	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Pyes	XP_0213535421_1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Rdec	DN21003_c0_g1.i2.p1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Acal	Contig281_47	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Rphi	EVM0017513_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Sbro	EVM0014371_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Scon	Chr12.1684	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Sglo	Sglo065363	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Tsqu	DN37_c29_g1.i1.p1	Fox	Fox-L2	Forkhead domain L2 (410802)	N/A	-
Acal	XP_005091040_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Rphi	XP_060586301_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Amar	Amu23426	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Apec	DN27027_c0_g1.i2.p1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Apur	sccaffold_17_163	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Bbla	XP_055899100_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Cang	XP_052711314_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Cari	EVM004311_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Cgig	XP_034303195_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Cpli	DN78931_c1.g1.i1.p1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Csin	Hic.asm.2.101	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Cvir	XP_022292787_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Dpol	XP_052270473_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Gaeg	XP_041365083_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Hbia	M00000027642	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Hruf	XP_048241610_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Mare	XP_052801997_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Mcal	XP_052066630_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Mchi	DN25972_c0.g1.i1.p1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Mcor	CAC5383890_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Medu	CAGC257106_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Mgal	VDH03464_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Mgal	VDH03462_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Mmar	MAMAM00000018109	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Mmer	XP_045177340_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Mmod	DN2607_c0.g2.i1.p1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Mphi	scf69935_0.10	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Obim	XP_052825413_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Oedu	XP_055996035_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Osin	XP_029633459_2	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Pcan	XP_025087495_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Pcor	DN55558_c0.g1.i1.p1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Pgen	DN145626_c0.g1.i1.p1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Pmar	DN30307_c0.g1.i1.p1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Pmax	XP_033751425_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Rdec	DN12531_c0.g1.i1.p1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Rphi	EVMM0009378.1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Sbro	Chr14_2061	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Scon	Sgt004456	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Sgra	DN11935_c0.g2.i3.p1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Pyes	XP_021371548_1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Tgra	DN20122_c0.g1.i1.p1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Tsqu	DN22139_c0.g1.i1.p1	Fox	Fox-N1/4	Forkhead domain N1 (410804)	N/A	-
Acal	XP_005409317_1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Airc	Contig117_153	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Amar	Ama09979	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Apec	DN11918_c0.g1.i10.p1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Apur	DN79231_c0.g1.i7.p1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Bgla	XP_013084125_2	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Cang	XP_052698143_1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Cari	EVMM0016469_1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Chu	DN125734_c1.g1.i18.p1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Cgig	XP_034324255_1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Cpli	DN79231_c0.g1.i7.p1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Csin	Hic.asm.10.136	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Cvir	XP_022331167_1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Dpol	XP_052270965_1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Gaeg	XP_041353111_1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Hbia	XP_046331344_1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Hruf	XP_052767715_1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Mare	XP_052107190_1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Mcal	DN39446_c1.g1.i1.p1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Mchi	CAC5378437_1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Mcor	CAG2235611_1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Medu	VID180286_1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Mgal	VID180289_1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Mmar	MMAM0000016688	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Mmer	XP_053376884_1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Mmod	DN2418_c0.g1.i31.p1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-
Mner	g156153_t2	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)		Additional domains (Pssm-ID)	Notes
Mphi	scat37509.0.5	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Obim	XP_0282822674..1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Oedu	XP_0487350022..1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Osin	XP_028633348..1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Pcan	XP_0285106088..1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Pcor	DN195240.c0.g..1..11..p1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Pcor	DN15191..c0.g..17..p1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Pcor	DN18451..c0..g13..p1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Pgen	DN14328..c0..g1..13..p1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Pmar	DN142157..c1..g..1..13..p1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Pmax	XP_033734749..1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Poku	DN17429..c4..g..1..12..p1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Ppur	DN5075..c0..g..1..1..p1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Pstr	KAK359505..1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Pvir	s00410g95	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Pyes	XP_021366964..1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Rdec	DN22296..c2..g..1..11..p1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Rphi	XP_060552999..1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Sglo	Sgo13452	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Sgra	DN13133..c0..g2..18..p1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Tgra	KA18308641..1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Tgra	KA18316727..1	Fox	Fox-N2/3	Forkhead domain N3 (410833)	N/A	-	-
Tsqu	DN75347..c0..g1..12..p1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Acal	XP_005112460..1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Airc	Contig116..24	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Amar	Ana07814	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Apec	DN23636..c0..g..1..11..p1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Apur	Scalfold..712..6	Fox	Fox-O	Forkhead domain O (410834)	N/A	-	-
Bglu	XP_013009525..2	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Cang	XP_052708272..1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Cari	EYAM00009968..1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Cflu	DN112955..c5..g..3..15..p1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Cgig	XP_011414359..1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Cpli	DN72415..c5..g..1..11..p1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Cvir	XP_0222287692..1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Cvir	XP_022287423..1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Dpol	XP_052342903..1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Gaeg	XP_041351714..1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Hria	MD0000024236	Fox	Fox-O	Forkhead domain O (410834)	N/A	-	-
Lorb	DN142512..c0..g..1..11..p1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Mcal	XP_052808484..1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Mchi	DN529098..c0..g..1..13..p1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Obiu	XP_052832486..1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Oedu	XP_048766900..1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Osin	XP_029657040..1	Fox	Fox-O	Forkhead domain O (410834)	N/A	-	-
Pcan	DN20008..c0..g..1..1..p1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Pcor	DN7549..c0..g..2..ii..p1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Pcor	DN187283..c1..i2..p1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Pcor	DN24871..c0..g..1..13..p1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Pgen	DN24871..c0..g..1..13..p2	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Pmar	DN44339..c0..g..5..ii..p1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Pmax	XP_033740844..1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Poku	DN10962..c1..g..2..ii..p1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Ppur	DN72510..c0..g..1..11..p1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Pstr	KAK357695..1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Pvir	s00079g102	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Pyes	XP_021377366..1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Rdec	DN30721..c0..g..1..1..p1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Rphi	XP_060589384..1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Scon	Chr2..491	Fox	Fox-O	Forkhead domain O (410806)	N/A	-	-
Sglo	Sglo18351	Fox	Fox-O	Forkhead domain O (410834)	N/A	-	-

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Sgra	DN5576_c0_g1.i1.p1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-
Tsqu	DN138852_c0_g1.i1.p1	Fox	Fox-O	Forkhead domain O (410806)	N/A	-
Airc	Contig3461.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Airc	Contig330.72	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Apir	scatfold.576.108	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Cang	XP_052675828.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Cari	EVM00015778.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Cgig	XP_011412452.2	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Cviro	XP_022300144.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Dpol	XP_052234997.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Dpol	XP_052327166.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Gaeg	XP_041362068.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Hbia	M0000010651	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Hruf	XP_048236781.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Hruf	XP_046328651.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Mcal	XP_052089402.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Mcor	CAAC5388114.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Medu	CAAG250347.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Mgal	VDI05564.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Mmar	M0MAM00000027087	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Mmar	XP_0151829632.2	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Mpho	DN9753_c0_g1.i1.p1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Mpho	scat66119.0.21	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Oedu	XP_048731527.2	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Pmax	DN32892_c0_g1.i1.p1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Pmax	XP_033727511.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Pstr	KAK3609024.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Pvir	s0129851	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Pyes	XP_021354438.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Rdec	DN23702_c2_g1.i2.p1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Bphi	XP_060566633.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Sbro	EVM0009544.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Scon	Chr1_409	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Sglo	Sg004484	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Tgra	KA18322379.1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Tsqu	DN207440_c0_g1.i1.p1	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Tsqu	XP_005106916.3	Fox	Fox-OG13/NA	Forkhead domain P (410807)	N/A	-
Acal	Contig85.21	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Airc	Ans19770	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Amar	scaffold.360.14	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Apur	XP_013071662.2	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Cang	XP_052676257.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Cari	EVM0001823.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Cgig	XP_011439389.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Cpili	DN35479_c0_g1.i1.p1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Csin	Hic_asm_12.159	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Cviro	XP_0222996913.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Dpol	XP_052252230.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Dpol	XP_052252240.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Gaeg	XP_041347345.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Hbia	M0000015843	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Hruf	XP_046382017.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Mcal	XP_052778846.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Mcor	XP_052062481.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Medu	CAAC54195.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Mgal	VDI050805.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Mmar	M0MAM0000012410	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Mpho	XP_045166371.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Obim	scat67833.0.2	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Oedu	XP_014772941.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Pmax	XP_025096321.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Pcor	DN7667_c0_g1.i1.p1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Pcor	DN34039_c0_g3.i1.p1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Pmax	XP_033744896.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	Annotated as Fox-Q2b	-
Pstr	KAK3589497.1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	N/A	-

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Pyes	XP_021371037_1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	N/A	Annotated as Fox-Q2b
Rphi	XP_060586724_1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	N/A	Annotated as Fox-Q2b
Sbro	EVM0000506_1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	N/A	Annotated as Fox-Q2b
Sbmo	EVM0006433_1	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	N/A	Annotated as Fox-Q2b
Scon	Chr1024	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	N/A	Annotated as Fox-Q2b
Sglo	Sg1024307	Fox	Fox-OG15/NA	Forkhead domain Q2 (410809)	N/A	Annotated as Fox-Q2b
Airc	Contig1425_7	Fox	Fox-OG16/NA	Forkhead domain Q2 (410809)	N/A	-
Apur	scatfold.604.173	Fox	Fox-OG16/NA	Forkhead domain Q2 (410809)	N/A	-
Cang	XP_052676026_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Cari	EVM0023364_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Chu	DN107758.c5.g1.i.12.p1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Cgig	XP_019927657_2	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Cpli	DNT0609.c0.g1.i.12.p1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Cvir	XP_022321288_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Cvir	XP_0222935893_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Dpol	XP_052222623_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Gaeg	XP_041365712_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Hbia	M0000030826	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Hruf	XP_046364140_2	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Lorb	DN243786.c0.g1.i.11.p1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Mare	XP_052802309_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Mcor	CAAC5370465_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Medu	CAG2202185_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Mgal	VDI24665_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Mmar	MAMAM0000032793	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Mmar	XP_045177123_2	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Mphi	scaf70200.0.4	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Obin	XP_014771053_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Oedu	XP_048728661_2	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Osin	XP_036362897_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Pmar	DN13205.c4.g1.i.12.p1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Pmax	XP_033752333_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Ppur	DN4666.c0.g1.i.2.p1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Pstr	KAK3581527_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Pvir	s1001194651	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Pyes	XP_021333413_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Rdec	DN22502.c0.g4.i.11.p1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Rphi	XP_060599562_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Sbro	EVM0002125_1	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Scon	Chr14.1.628	Fox	Fox-OG16/NA	Forkhead domain Q2 (410788)	N/A	-
Tsqu	DN2384.c1.g1.i.1.p1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Airc	Contig138_5	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Apec	DN23140.c1.g1.i.11.p1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Apur	scatfold.333.74	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Cang	XP_052712245_1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Cari	EVM0012682_1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Chu	DN113069.c3.g1.i.14.p1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Cpli	DNT6173.c0.g1.i.12.p1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Csin	Hic.acm.2.1802	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Csin	HC.asin.2.1455	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Cvir	XP_022286391_1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Dpol	XP_0522234537_1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Dpol	XP_041366058_1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Gaeg	Hbia	M0000014061	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-
Hruf	XP_046335487_2	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Mare	XP_0527089013_1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Meal	XP_052064572_1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Mcor	CAAC5370562_1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Medu	CAG2224977_1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Mmar	MAMAM0000037791	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Mmar	XP_053384136_1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Mmod	DN27089.c0.g1.i.1.p1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Muer	g144243.tl	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Oedu	XP_048731286_2	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Pcor	DN5679.c0.g2.i.1.p1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M
Pcor	DN5679.c0.g1.i.15.p1	Fox	Fox-OG2/NA	Forkhead domain M (410803)	-	Annotated as Fox-M

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Pcor	DN13056_c0_g1.i1.p1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Pcor	DN13056_c0_g2.i2.p1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Pgen	DN28413_c1.g1.i2.p1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Pmar	DN44947_c1.g2.i3.p1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Pmax	XP_033751305_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Poku	DN37223_c1.g1.i8.p1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Ppur	DN3451_c0_g1.i1.p1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Pstr	KAK3597624_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Pvir	s00219g11	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Pyes	XP_021377259_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Rdec	DN22152_c4.g4.i1.p1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Rphi	XP_060563990_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Sbro	EYVM0023670_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Sglo	SgI005561	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Sgra	DN54780_c0_g1.i1.p1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Tgra	KA18299135_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Tsqu	DN6434_c0_g1.i7.p1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-M	
Airc	Contig465_41	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Amar	Ama25933	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Apur	scatfold_867_41	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Cang	scatfold_381_16	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Dpol	XP_052700156_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Dpol	XP_052275604_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Cgig	XP_011434557_2	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Csin	Hic_acm_16_939	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Cvir	XP_022334263_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Dpol	XP_052278575_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Dpol	XP_052278576_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Hruf	XP_046341176_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mare	XP_052791890_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mare	XP_052791888_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mare	XP_0522791891_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mcal	XP_052098761_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mcal	XP_052279555_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mcor	CAAC5410385_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mcor	CAAC538023_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mcor	CAAC5379920_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mcor	CAAC5410389_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mcor	CAAC5410356_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mcor	CAAC5410381_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mcor	CAAC5410382_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mcor	CAAC5410383_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mcor	CAAC5410387_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mcor	CAAC5410380_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mcor	CAAC5410388_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Medu	CAcg2114460_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Medu	CAcg2194706_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Medu	CAcg2198066_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Medu	CAcg2198058_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Medu	CAcg2198055_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Medu	CAcg2198056_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Medu	CAcg2198063_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Medu	CAcg2198061_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Medu	CAcg2198057_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Medu	CAcg2198064_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Medu	CAcg2198065_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Medu	CAcg2198062_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mgal	VDI02550_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mgal	VDI39859_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mgal	VDI15906_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mgal	VDI02348_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mgal	VDI15903_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	
Mgal	VDI02347_1	Fox	Forkhead domain M (410803)	Fox-OG2/NA	Annotated as Fox-Q2c	

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Mgal	VDI15905_1	Fox	Forkhead domain Q2 (410809)	Forkhead domain Q2 (410809)	Annotated as Fox-Q2c	
Mgal	VDI125349_1	Fox	Fox-OOG28/NA	Fox-OOG28/NA	Annotated as Fox-Q2c	
Mgal	VDI15904_1	Fox	Fox-OOG28/NA	Fox-OOG28/NA	Annotated as Fox-Q2c	
Mmer	XP_053405097_1	Fox	Fox-OOG28/NA	Fox-OOG28/NA	Annotated as Fox-Q2c	
Mmod	DIN6982_c0_g1.i3.p1	Fox	Fox-OOG28/NA	Fox-OOG28/NA	Annotated as Fox-Q2c	
Mphi	scat.15444_0_1	Fox	Fox-OOG28/NA	Fox-OOG28/NA	Annotated as Fox-Q2c	
Pmar	XP_056092121_1	Fox	Fox-OOG28/NA	Fox-OOG28/NA	Annotated as Fox-Q2c	
Pmax	DN30963_c0_g1.i1.p1	Fox	Fox-OOG28/NA	Fox-OOG28/NA	Annotated as Fox-Q2c	
Pmax	XP_033749723_1	Fox	Fox-OOG28/NA	Fox-OOG28/NA	Annotated as Fox-Q2c	
Pvir	s03_13748	Fox	Fox-OOG28/NA	Fox-OOG28/NA	Annotated as Fox-Q2c	
Pyes	XP_021360588_1	Fox	Fox-OOG28/NA	Fox-OOG28/NA	Annotated as Fox-Q2c	
Rphi	XP_060558777_1	Fox	Fox-OOG28/NA	Fox-OOG28/NA	Annotated as Fox-Q2c	
Sbro	EVM0013029_1	Fox	Fox-OOG28/NA	Fox-OOG28/NA	Annotated as Fox-Q2c	
Scon	Chr5_396_1	Fox	Fox-OOG28/NA	Fox-OOG28/NA	Annotated as Fox-Q2c	
Sglo	sgl013625	Fox	Fox-OOG28/NA	Fox-OOG28/NA	Annotated as Fox-Q2c	
Acar	XP_005102249_1	Fox	Fox-OOG28/NA	Fox-OOG28/NA	Annotated as Fox-AB	
Airc	Contig536_41	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Amar	Ama26012	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Apur	scatfold.313_52	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Bglia	XP_055874345_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Cang	XP_052699279_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Hbia	MD000015535	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Cgig	XP_011441298_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Cvir	XP_022334408_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Dpol	XP_052275569_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Gaeq	XP_041375984_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Hbia	XP_000015535	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Hruf	XP_046356707_2	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Hruf	XP_046350710_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Hruf	XP_046350686_2	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Hruf	XP_046350709_2	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Hruf	XP_046350688_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Hruf	XP_046350660_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Hruf	XP_046350687_2	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Hruf	XP_046350714_2	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Hruf	XP_046350708_2	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Mare	XP_052795236_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Meal	XP_052102496_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Mcor	CAc5413394_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Medu	CAG2193762_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Mgal	VID135942_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Mmar	MMAM00000023830	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Mmer	XP_045215157_2	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Mphi	scat.7111_0_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Oedu	XP_048737442_2	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Pcan	XP_025078030_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Pmax	XP_038750900_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Pstr	KAK3601439_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Tgra	KAK3601419_1	Fox	Fox-OOG39/NA	Fox-OOG39/NA	Annotated as Fox-AB	
Airc	Contig120_132	Fox	Fox-P	Fox-P	Annotated as Fox-AB	
Pvir	s001896_393	Fox	Fox-P	Fox-P	Annotated as Fox-AB	
Pyes	XP_021357612_1	Fox	Fox-P	Fox-P	Annotated as Fox-AB	
Rphi	XP_060589081_1	Fox	Fox-P	Fox-P	Annotated as Fox-AB	
Sbro	EVM0003782_1	Fox	Fox-P	Fox-P	Annotated as Fox-AB	
Scon	Chr5_18	Fox	Fox-P	Fox-P	Annotated as Fox-AB	
Sglo	sgl004401	Fox	Fox-P	Fox-P	Annotated as Fox-AB	
Bglia	EVMM015348_1	Fox	Fox-P	Fox-P	Annotated as Fox-AB	
Cang	DPN101403_c0_g2_13_p1	Fox	Fox-P	Fox-P	Annotated as Fox-AB	
Cgig	DPN80534_6_g1_i4_p1	Fox	Fox-P	Fox-P	Annotated as Fox-AB	
Csin	Hic_asm_12_158	Fox	Fox-P	Fox-P	Annotated as Fox-AB	
Cvir	XP_022295865_1	Fox	Fox-P	Fox-P	Annotated as Fox-AB	

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)		Notes
					Forkhead domain (P2)		
Dpol	XP_052253868..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Gaeg	XP_041347582..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Hruf	XP_048239143..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Lorb	DN39652..c1..g1..i4..p1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Mare	XP_052780914..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Mcal	XP_052101037..1	Fox	Fox-P	Forkhead domain P2 (410839)	N/A	-	
Mgal	XP_052098374..1	Fox	Fox-P	Forkhead domain P2 (410839)	N/A	-	
Mgal	XP_052100969..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Mgal	XP_052061774..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Mchi	DN46200..c0..g1..i1..p1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Mcor	CAC5410517..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Medu	CAG2199156..1	Fox	Fox-P	Forkhead domain P2 (410839)	N/A	-	
Mgal	VDI14555..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Mgal	VDI10808..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Mgal	VDI10806..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Mmar	MMAM0000012411	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Mmer	XP_053376718..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Mmod	DN2151..c0..g1..i2..p1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Mner	s93347..t1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Mphi	XP_053998748..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Obim	XP_052822821..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Oedu	XP_036357858..1	Fox	Fox-P	Forkhead domain P (410807)	FOXP coiled-coil domain (465036)	-	
Osin	XP_025106713..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Pcan	DN1820..c0..g1..i2..p1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Pcen	DN35611..c0..g1..i6..p1	Fox	Fox-P	Forkhead domain P2 (410839)	N/A	-	
Pgen	DN41268..c0..g1..i1..p1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Pmar	XP_033745371..1	Fox	Fox-P	Forkhead domain P (410807)	FOXP coiled-coil domain (465036)	-	
Pmax	DN88526..c2..g1..i1..p1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Poku	DAK589495..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Ppur	Psr	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Pvir	s01329..g1..24	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Pvir	XP_02136304..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036; partial)	-	
Rdec	DN8026..c0..g1..i1..p1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Rphi	XP_060586741..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Sbro	EVMM0004295..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036; partial)	-	
Scon	Chr7..2129	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Scon	Chr7..2133	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Sglo	Sglo11345..1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Sgra	DN17101..c0..g1..i1..p1	Fox	Fox-P	Forkhead domain P2 (410839)	FOXP coiled-coil domain (465036)	-	
Tgra	KA18302344..1	Fox	Fox-P	Forkhead domain Q2 (410809)	N/A	-	
Acal	XP_00059459..2	Fox	Fox-Q2	Forkhead domain Q2 (410809)	N/A	-	
Airc	Contig1420..28	Fox	Fox-Q2	Forkhead domain Q2 (410809)	N/A	-	
Amar	Arna2905..1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Apar	scatfield..832..35	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Bbla	XP_055865367..1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Cang	XP_052699620..1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Cari	EVM0002665..1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Cgig	XP_011425762..2	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Cpli	DN105612..c0..g1..i1..p1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Csin	Hic..asm..16..4	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Cvrl	XP_022333968..1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Dpol	XP_0522486..c0..g2..i1..p1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Gaeg	XP_041363029..1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Gbaa	XP_041363041..1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Hbia	MD00003328..1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Hruf	XP_046373579..2	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Mcal	XP_052101305..1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Mchi	DN22486..c0..g2..i1..p1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Mcor	CAC5388792..1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Medu	CAG2191193..1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Mgal	VDI174621..1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Mmar	MMAM000000686..1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Mmer	XP_045215524..2	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	-	
Muer	s200553..t1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	N/A	-	
Mphi	scatf..22910..1..1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	N/A	-	
Obim	XP_014767584..1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	N/A	-	
Pcan	XP_025078472..1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	N/A	-	

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)		Additional domains (Pssm-ID)	Notes
Pmar	DN99235_c0_g1.i1.p1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	N/A	
	DN9466_c0_g1.i1.p1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	N/A	
Pmax	XP_033751003_1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	N/A	
Pstr	KAK3595133_1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	N/A	
Pvir	s0011523	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	N/A	
Pyes	XP_021343668_1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	N/A	
Rphi	XP_060571531_1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	N/A	
Sbro	EVM0023378_1	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	N/A	
Scon	Chr5_1074	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	N/A	
Scon	Chr5_2105	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	N/A	
Sglo	Sg1009183	Fox	Fox-Q2	Forkhead domain Q2 (410809)	Annotated as Fox-Q2a	N/A	
Acal	XP_005109004_3	Fox	N/A	Forkhead domain Q2 (41078)	-	-	
Airc	Contig879_9	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Amar	Amn25952	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Amar	Amn27375	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Cang	XP_052676682_1	Fox	N/A	Forkhead domain H (410796)	N/A	N/A	
Cang	XP_052680288_1	Fox	N/A	Forkhead domain H (410796)	N/A	N/A	
Cang	XP_052677368_1	Fox	N/A	Forkhead domain H (410796)	N/A	N/A	
Cari	EVM0001935_1	Fox	N/A	Forkhead domain H (410796)	N/A	N/A	
Cari	EVM0027332_1	Fox	N/A	Forkhead domain H (410796)	N/A	N/A	
Cgig	XP_034306826_1	Fox	N/A	Forkhead domain H (410796)	N/A	N/A	
Cgig	XP_01147567_2	Fox	N/A	Forkhead domain H (410796)	N/A	N/A	
Cpli	DN157619_c0_g1.i1.p1	Fox	N/A	Forkhead domain H (410796)	N/A	N/A	
Cvir	XP_022300767_1	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Cvir	XP_022300750_1	Fox	N/A	Forkhead domain Q2 (41078)	-	-	
Dpol	XP_052277921_1	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Dpol	XP_052227296_1	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Gaeg	XP_041366967_1	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Gaeg	XP_041375820_1	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Gaeg	XP_041347725_1	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Gaeg	XP_041375925_1	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Gaeg	XP_041375913_1	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Gaeg	XP_041375915_1	Fox	N/A	Forkhead domain N1 (410804)	N/A	N/A	
Hb2a	M00000018946	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Mare	XP_0527791886_1	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Mare	XP_052771066_1	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Meal	XP_032098820_1	Fox	N/A	Forkhead domain L1 (410801)	N/A	N/A	
Mcor	CAC5419379_1	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Medu	CAG2194707_1	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Medu	CAG2208945_1	Fox	N/A	Forkhead domain L1 (410801)	N/A	N/A	
Mgal	VDI15902_1	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Mgal	VDI15978_1	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Mnar	MMAM00000049704	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Mmer	XP_045216636_2	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Mmer	XP_045189131_2	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Mner	g159704_t1	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Mphi	scat_461_89_0_0	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Mphi	scat_15444_0_2	Fox	N/A	Forkhead domain Q2 (410809)	-	-	
Mphi	scat_27787_1_10	Fox	N/A	Forkhead domain L1 (410801)	N/A	N/A	
Obim	XP_014777604_1	Fox	N/A	Forkhead domain M (410803)	N/A	N/A	
Oedu	XP_048733629_2	Fox	N/A	Forkhead domain H (410796)	N/A	N/A	
Osin	XP_036359188_1	Fox	N/A	Forkhead domain M (410803)	N/A	N/A	
Osin	XP_029655092_1	Fox	N/A	Forkhead domain Q2 (41078)	N/A	N/A	
Pcor	DN89866_c4_g1.i1.p1	Fox	N/A	Forkhead domain L1 (41078)	N/A	N/A	
Pcor	DN155206_c2_g1.i1.p2	Fox	N/A	Forkhead domain FOX12, FOXJ3 (410798)	N/A	N/A	
Pcor	DN155905_c0_g1.i1.p1	Fox	N/A	Forkhead domain P (410798)	N/A	N/A	
Pmax	XP_033750561_1	Fox	N/A	Forkhead domain Q2 (410809)	N/A	N/A	
Pstr	KAK355306_1	Fox	N/A	Forkhead domain Q2 (410809)	N/A	N/A	
Pvir	s00585648	Fox	N/A	Forkhead domain L1 (410801)	N/A	N/A	
Pyes	XP_021348419_1	Fox	N/A	Forkhead domain Q2 (410809)	N/A	N/A	
Rdec	DN23525_c0_g1.i1.p1	Fox	N/A	Forkhead domain Q2 (410809)	N/A	N/A	
Rphi	XP_060585776_1	Fox	N/A	Forkhead domain Q2 (410809)	N/A	N/A	
Rphi	XP_060551131_1	Fox	N/A	Forkhead domain Q2 (410809)	N/A	N/A	
Scon	Chr5_397	Fox	N/A	Forkhead domain Q2 (410809)	N/A	N/A	
Sglo	Sg1021575	Fox	N/A	Forkhead domain Q2 (41078)	N/A	N/A	
Tsqu	DN23960_c0_g1.i1.p1	Sox	N/A	High mobility group box (438820)	Helix loop helix domain (197674)	N/A	
Acal	XP_0358244685_1	Sox	N/A	High mobility group box (438820)	High mobility group box (438820)	N/A	
	XP_012946205_1						

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Acal	XP_005105939..1	Sox	N/A	High mobility group box (438820)	N/A	
Apec	DN18806_c0.g1.i1.p1	Sox	N/A	High mobility group box (438820)	N/A	
Apec	DN108003_c0.g1.i1.p1	Sox	N/A	High mobility group box (438820)	N/A	
sccaffold..391..70		Sox	N/A	High mobility group box (438820)	N/A	
Apur	XP_013075241..2	Sox	N/A	High mobility group box (438820)	N/A	
Bglia	XP_013075156..1	Sox	N/A	High mobility group box (438820)	N/A	
Bglia	XP_052697278..1	Sox	N/A	High mobility group box (438820)	N/A	
Cang	XP_052713692..1	Sox	N/A	High mobility group box (438820)	N/A	
Cang	EVMM0018891..1	Sox	N/A	High mobility group box (438820)	N/A	
Cari	EVMM0005367..1	Sox	N/A	High mobility group box (438820)	N/A	
Cari	XP_011425869..2	Sox	N/A	High mobility group box (438820)	N/A	
Cgig	XP_034335819..1	Sox	N/A	High mobility group box (438820)	N/A	
Cgig	XP_022338075..8	Sox	N/A	High mobility group box (438820)	N/A	
Cvir	XP_022338079..1	Sox	N/A	High mobility group box (438820)	N/A	
Cvir	XP_052271004..1	Sox	N/A	High mobility group box (438820)	N/A	
Dpol	XP_041377139..1	Sox	N/A	High mobility group box (438820)	N/A	
Gaeg	M00000038049	Sox	N/A	High mobility group box (438820)	N/A	
Hbia	M00000004998	Sox	N/A	High mobility group box (438820)	N/A	
Hbia	XP_046328595..1	Sox	N/A	High mobility group box (438820)	N/A	
Hruf	XP_048238511..1	Sox	N/A	High mobility group box (438820)	N/A	
Hruf	CA5C384832..1	Sox	N/A	High mobility group box (438820)	N/A	
Mcor	CAAG293429..1	Sox	N/A	High mobility group box (438820)	N/A	
Medu	VDI78477..1	Sox	N/A	High mobility group box (438820)	N/A	
Mgal	DN113112_c0.g1.i1.p1	Sox	N/A	High mobility group box (438820)	N/A	
Mmod	Obim	XP_014776519..1	Sox	High mobility group box (438820)	N/A	
Oedu	XP_0487538250..2	Sox	N/A	High mobility group box (438820)	N/A	
Oedu	XP_048752144..2	Sox	N/A	High mobility group box (438820)	N/A	
Osin	XP_036358200..1	Sox	N/A	High mobility group box (438820)	N/A	
Osin	XP_029654541..1	Sox	N/A	High mobility group box (438820)	N/A	
Osin	XP_028656568..1	Sox	N/A	High mobility group box (438820)	N/A	
Osin	XP_028657644..1	Sox	N/A	High mobility group box (438820)	N/A	
Osin	XP_029656220..1	Sox	N/A	High mobility group box A, B and G (438837)	N/A	
Osin	XP_029655748..1	Sox	N/A	High mobility group box A, B and G (438837)	N/A	
Osin	XP_029655056..1	Sox	N/A	High mobility group box A, B and G (438837)	N/A	
Osin	XP_029655785..1	Sox	N/A	High mobility group box A, B and G (438837)	N/A	
Osin	XP_029656129..1	Sox	N/A	High mobility group box A, B and G (438837)	N/A	
Osin	XP_029654991..1	Sox	N/A	High mobility group box A, B and G (438837)	N/A	
Pcan	DN32781_c0.g1.i1.p1	Sox	N/A	High mobility group box (438820)	N/A	
Pcor	DN21964_c0.g1.i2.p1	Sox	N/A	High mobility group box (438820)	N/A	
Pcor	DN33290_c0.g1.i2.p1	Sox	N/A	High mobility group box (438820)	N/A	
Pmar	XP_035008_c0.g1.i4.p1	Sox	N/A	High mobility group box (438820)	N/A	
Pmax	DN639_c0.g1.i1..p1	Sox	N/A	High mobility group box (438820)	N/A	
Ppur	DN1784_c0.g1.i4.p1	Sox	N/A	High mobility group box (438820)	N/A	
Ppur	XP_025104729..1	Sox	N/A	High mobility group box (438820)	N/A	
Pyses	EVMM0018224..1	Sox	N/A	High mobility group box (438820)	N/A	
Sbro	Sg1009175	Sox	N/A	High mobility group box (438820)	N/A	
Sglo	Sg1012029	Sox	N/A	High mobility group box (438820)	N/A	
Tsqu	DN639_c0.g1.i1..p1	Sox	N/A	High mobility group box (438820)	N/A	
Acal	XP_005108230..1	Sox	N/A	High mobility group box B (438790)	N/A	
Acal	XP_035824438..1	Sox	N/A	High mobility group box B (438790)	N/A	
Apur	Contig449..126	Sox	N/A	High mobility group box B (438790)	N/A	
Airc	Contig144..115	Sox	N/A	High mobility group box B (438790)	N/A	
Amar	Am3..33032	Sox	N/A	High mobility group box B (438790)	N/A	
Amar	Am3..33828	Sox	N/A	High mobility group box B (438790)	N/A	
Apec	DN29410_c0.g1.i2..p1	Sox	N/A	High mobility group box B (438790)	N/A	
Apec	DN12297_c0.g1.i3..p1	Sox	N/A	High mobility group box B (438790)	N/A	
Apur	Sccaffold..15459..10	Sox	N/A	High mobility group box B (438790)	N/A	
Apur	Sccaffold..865..4	Sox	N/A	High mobility group box B (438790)	N/A	
Bglia	XP_013075432..1	Sox	N/A	High mobility group box B (438790)	N/A	
Bglia	XP_055868106..1	Sox	N/A	High mobility group box B (438790)	N/A	
Cang	XP_052706368..1	Sox	N/A	High mobility group box B (438790)	N/A	
Cang	XP_052705551..1	Sox	N/A	High mobility group box B (438790)	N/A	
Cari	EVMM026792..1	Sox	N/A	High mobility group box B (438790)	N/A	
Cari	EVMM0013865..1	Sox	N/A	High mobility group box B (438790)	N/A	
Cflu	DN118670_c2.g1.i2..p1	Sox	N/A	High mobility group box B (438790)	N/A	
Cflu	DN99542_c1.g1.i1..p1	Sox	N/A	High mobility group box B (438790)	N/A	
Cgig	XP_011433975..1	Sox	N/A	High mobility group box B (438790)	N/A	
Cgig	XP_011455662..1	Sox	N/A	High mobility group box B (438790)	N/A	

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Cpli	DN31343.c0.g1.i1.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Cpli	DN08511.c0.g1.i1.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Csin	Hic.asm.6.930	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Csin	Hic.B1/2	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Cvir	XP_022286516..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Cvir	XP_022343230..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Dpol	XP_052214544..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Dpol	XP_052217420..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Gaeg	XP_041353075..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Gaeg	XP_041357874..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
M00000092798	M00000033682	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Hbia	XP_046370193..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Hruf	XP_046326733..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Lorb	DN80278.c0.g1.i1.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Lorb	DN14..c4.g1.i1.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mare	XP_052784929..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mare	XP_052784720..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Meal	XP_052103617..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Meal	XP_052104911..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mchi	DN33632.c0.g1.i2.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mchi	DN45716.c0.g1.i1.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mcor	CA/C5401077..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Medu	CAG2229644..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Medu	CAG2206403..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mgal	VID13296..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mgal	VID169660..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mmar	MMAM00000041532	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mmar	MMAM00000023253	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mmer	XP_0452011594..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mmer	XP_045201080..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mmod	DN1981.c0.g1.i2.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mmod	DN78279.c0.g1.i1.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mner	g140596..t1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mner	g157489..t1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mphi	scaf_6349..0..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Mphi	scat_24206..0..4	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Obim	XP_014789771..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Obim	XP_014780771..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Oedu	XP_048746651..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Oedu	XP_048746663..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Osin	XP_029654000..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Osin	XP_029655838..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pcan	XP_025079293..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pcan	XP_025075598..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pcor	DN9087..c0.g1.i2.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pcor	DN14753..c0.g1.i2.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pcor	DN5688..c0.g3..i1.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pcor	DN5742..c0.g1.i2.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pgen	DN532..c0.g1..i1.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pgen	DN1199..c0.g1..i1.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pmar	DN30477..c0.g2..i1.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pmar	DN30459..c0.g1..i1.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pmax	XP_033760067..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pmax	XP_033753382..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Poku	DN51870..c0..i1..p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Poku	DN1067..c0..i1..p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Ppur	DN202737..c0..i1..p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pstr	KAK3586311..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pstr	KAK3589336..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pvir	s00319..159	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pvir	s00037281	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pyes	XP_021356125..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pyes	XP_02134413..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Pyes	XP_02137128..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Rdec	DN21447..c2..g7..i2..p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-
Rphi	XP_060556101..1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes	
Rphi	XP_060561544_1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-	
Sbro	EVM0016386_1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-	
Sbro	EVM0007529_1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-	
Scon	Chr9.1552	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-	
Scon	Chr9.1522	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-	
Scon	Chr9.1514	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-	
Sglo	Sgo10100	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-	
Sglo	Sgo1020107	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-	
Sgra	DN3782_c0_g1.i3.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-	
Sgra	DN357_c1_g1.i1.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-	
Tgra	KAI1831_0140_1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-	
Tgra	DN97880_c0_g1.i1.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-	
Tsqu	DNS0556_c0_g1.i2.p1	Sox	Sox-B1/2	High mobility group box B (438790)	N/A	-	
Airc	Contig80_70	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Amar	Am1.2726	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Apec	DN12286_c0_g3.i1.p1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Apur	scaf0016_61	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Cang	XP_052689209_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Cari	EVM0025846_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Chu	DN126276_c0_g1.i1.p1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Cgig	XP_011445203_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Cpil	DN19112_c0_g1.i1.p1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Hic.asm.11.1009	Sox	Sox-C	High mobility group box C (438838)	N/A	-		
Cvir	XP_022317619_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Dpol	XP_052257395_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Gaeg	XP_041355324_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Hbia	M00000037669	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Hruf	XP_046365064_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Lorb	DN14941_c0_g1.i1.p1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Mare	XP_05277703_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Mcal	XP_052087802_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Mchi	DN44798_c0_g4.i1.p1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Mcor	CAc5424030_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Medu	CAG21_89937_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Mgal	V1D14453_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Mgal	V1D14462_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Minar	MAMAM0000036315	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Minar	XP_045158937_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Mmod	DN104308_c0_g1.i1.p1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Muer	g26404_t1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Mphi	scaf.17954_1.5	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Oedu	XP_048762549_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Osin	DN3553_c2_g3.i1.p1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Pcan	DN788_c0_g1.i1.p1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Pcan	DN29124_c0_g2.i1.p1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Pmar	XP_033737425_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Poku	DN7105_c0_g2.i1.p1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Ppur	DN9859_c0_g1.i1.p1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Pstr	KAK3610995_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Scon	s00145_243	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Sgra	DNG6210_c0_g1.i1.p1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Pvir	XP_021356342_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Pyes	DNP52924_c0_g1.i1.p1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Rdec	Rphi	XP_060555827_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-
Rphi	EVM0006311_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Sbro	Chr8.1790	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Scon	Sgo00072	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Sgra	DNG6210_c0_g1.i1.p1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Tgra	KAI18306264_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Tgra	KA18306266_1	Sox	Sox-C	High mobility group box C (438838)	N/A	-	
Tsqu	DN11669_c1.g1.i2.p1	Sox	Sox-D	High mobility group box C (438838)	N/A	-	
Acal	XP_035824396_1	Sox	Sox-D	High mobility group box C (438838)	N/A	-	
Airc	Contig290_5_1	Sox	Sox-D	High mobility group box C (438838)	N/A	-	
Amar	Amz23921	Sox	Sox-D	High mobility group box C (438838)	N/A	-	
Apec	DN1990_c0_g1.i10.p1	Sox	Sox-D	High mobility group box C (438838)	N/A	-	

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Apur	scatfold_393.10	Sox	Sox-D	High mobility group box (438839)	N/A	-
BglA	XP_055899647..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Cari	EVM0012405..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Chu	DN124562..c0..g1..i1..p1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Cgig	XP_011425377..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Cpli	DN64448..c0..g1..i1..p1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Hic..asm..2..1656..	Hic..asm..2..1656..	Sox	Sox-D	High mobility group box (438839)	N/A	-
Csin	Hic..asm..2..1600..2	Sox	Sox-D	High mobility group box (438839)	N/A	-
Cvri	XP_022303926..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Dpol	XP_052213125..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Gaeg	XP_011367101..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Hbia	M0000014008	Sox	Sox-D	High mobility group box (438839)	N/A	-
Hruf	XP_046328046..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Lorb	DN5537..c0..g2..i3..p1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Mcal	XP_052800695..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Mcal	XP_0528065962..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Mchi	DN38691..c1..g1..i4..p1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Mcor	CAC536270..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Medu	CAG197887..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Mgal	VID147525..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Mgal	VID147529..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Mgal	VID147528..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Mgal	VID147527..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Mgal	VID147526..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Mgal	VID147530..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Mnar	MAMAM0000004319	Sox	Sox-D	High mobility group box (438839)	N/A	-
Miner	XP_053384959..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Mmod	DN588..c0..ef..i9..p1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Muer	g103147..42	Sox	Sox-D	High mobility group box (438839)	N/A	-
Mphi	scat..42..i1..3	Sox	Sox-D	High mobility group box (438839)	N/A	-
Obiu	XP_048728391..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Oedu	XP_029644081..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Osin	XP_029688657..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Pcan	DN353..c2..g2..i1..p1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Pcor	DN1386..c1..g1..i1..p1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Pcor	DN1357..c0..g1..i4..p1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Pcor	DN1386..c1..g2..i2..p1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Pgen	DN24654..c0..g1..i2..p1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Pmar	DN40112..c0..g1..i4..p1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Pmax	XP_033751614..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Poku	DN3371..c1..g4..i1..p1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Ppur	DN3331..c0..g1..i1..p1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Pstr	KAK3605548..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Pvir	s00219g102	Sox	Sox-D	High mobility group box (438839)	N/A	-
Pyes	XP_021365061..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Rdec	DN8093..c0..g1..i1..p1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Rphi	XP_060604110..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Sbro	EVM00000795..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Chir14..562..1	ScoD	Sox	Sox-D	High mobility group box (438839)	N/A	-
Sgra	DN6138..c0..g1..i6..p1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Tgra	KA18298781..1	Sox	Sox-D	High mobility group box (438839)	N/A	-
Tsqu	DN55031..c0..g1..i1..p1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Acal	XP_005102100..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Contig52..209	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-	-
Amar	ADM1107	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Apec	DN1330..c0..g1..i1..p1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Apur	APUR..48..7	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
BglA	XP_013091187..2	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Cang	XP_052689355..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Cari	EVM0005846..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Cflu	DN106407..c5..g2..i1..p1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Cgig	NP_001295801..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Cpli	DN71393..c0..g2..i1..p1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Cvri	DN71393..c0..g1..i2..p1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Csin	Hic..asm..0..35..3	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Dpol	XP_0522612895..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Gaeg	XP_041362638..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Hbia	M0000012324	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Hbia	M0000012325	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Hruf	XP_046359366..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Mare	XP_052786944..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Mare	XP_052786966..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Mcal	XP_052068536..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Mchi	DN12011..co.g1..i1..p1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Mcor	CAc5402442..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Medu	CAG2231021..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Mgal	VID182092..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Mgal	VID182090..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Mmar	MMAM00000042410	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Mmar	XP_045213795..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Mmod	DN78330..0..g1..i1..p1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Mphi	scat.25414..0..6	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Oedu	XP_056011913..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Pcan	XP_0250911262..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Pcor	DN96098..co.g1..i3..p1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Pcor	DN1274..co.g1..i1..p1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Pcor	DN30335..co.g1..i1..p1	Sox	Sox-E	High mobility group box E (438840)	N/A	-
Pmar	XP_033739301..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586)	-
Pmax	XP_033739301..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586)	-
Poku	DN87807..co.g1..i7..p1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Ppur	DN14600..co.g1..i1..p1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Pstr	KAK3600863..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Pstr	KAK3610785..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Pvir	s136484874	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Pyes	XP_021348843..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Rphi	XP_050604697..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Sbro	EVM0002110..1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Scon	Chr1..75	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Sglo	SG024297	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Sgra	DN22463..0..g1..i1..p1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Tsra	DN8973..c2..g1..i2..p1	Sox	Sox-E	High mobility group box E (438840)	Sox developmental protein N terminal (463586; partial)	-
Acal	XP_005107482..1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Airc	Contig80..101	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Amar	Anal16..16	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Apur	Scafold..546..32	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Bglia	XP_013074628..2	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Cang	XP_052685434..1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Cari	EVM00066323..1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Cflu	DN139006..co.g1..i1..p1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Cgig	XP_011448074..2	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Cpil	DN4414..co.g1..i1..p1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Csin	Hic..asm..1..5..4..9	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Cvpr	XP_022319962..1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Cvpr	XP_022314364..1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Dpol	XP_052274104..1	Sox	Sox-F	High mobility group box F (438820)	N/A	-
Gaeg	XP_041359436..1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Hbia	M0000015459	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Hruf	XP_046357912..1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Mare	XP_052774544..1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Mcal	XP_052274361..1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Mcor	MMAM00000025810	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Mmod	DNS80495..co.g1..i1..p1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Mphi	scat.61114..0..13	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Obim	XP_052826684..1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Oedu	XP_048764319..2	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Pcan	XP_025109598..1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Pcor	DN11375..co.g1..i1..p1	Sox	Sox-F	High mobility group box F (438841)	N/A	-

Table S6 continued from previous page

Sp ID	Gene ID	Group	Annotation	Main catalytic domain (Pssm-ID)	Additional domains (Pssm-ID)	Notes
Pcor	DN29649_c0_g1_i3_p1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Pcor	DN3688_c2_g1_i3_p1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Pgen	DN144332_c0_g1_i1_p1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Pmar	DN24748_c0_g1_i1_p1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Pmax	XP_03373287_1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Poku	DN41229_c1_g2_i1_p1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Pstr	KAK582343_1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Pvir	s00137g284	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Pyes	XP_021375109_1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Rdec	DN53443_c0_g1_i1_p1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Rphi	XP_06555438_1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Sbro	EVM0000861_1	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Scon	Chr8_897	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Sglo	Sg1005442	Sox	Sox-F	High mobility group box F (438841)	N/A	-
Sgra	DN5013_c3_g1_i7_p1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Airc	Contig1525_38	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Amar	Amz26724	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Apec	DN3182_c0_g1_i1_p1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Apur	scatford_768_3	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Cang	XP_05270370_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Cari	EVM001164_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Cgig	XP_011415853_3	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Ccli	DN80002_c0_g1_i1_p1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Csin	Hic_asm_15_1471	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Cvir	XP_022338738_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Dpol	XP_0522209448_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Gaeg	XP_041370217_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Gaeg	XP_041369137_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Hbia	M00000001184	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Hruf	XP_046355520_2	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Mcal	XP_052099860_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Mcor	CAc5408014_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Medu	CAG29257203_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Mgal	V1D10824_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Mgal	V1D10823_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Mmar	MMAM0000015662	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Mmer	XP_033407277_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Miner	g125234_t1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Mphi	scat12010_0_4	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Mphi	scf59202_0_9	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Obim	XP_0528352677_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Oedu	XP_056006679_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Osin	XP_036368794_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Pcor	DN186456_c0_g1_i1_p1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Pmar	DN40950_c1_g1_i2_p1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Pmax	XP_033755818_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Poku	DN16718_c0_g1_i6_p1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Ppur	DN7268_c0_g1_i1_p1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Pstr	KAK3582760_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Pvir	s004516108	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Pyes	XP_021340986_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Rdec	DN22482_c4_g1_i1_p1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Rphi	XP_060578490_1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Scon	Chr15_1899	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Sglo	Sg1010047	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-
Tsqu	DN874_c7_g1_i1_p1	Sox	Sox-H	High mobility group box Sox-30 (438842)	N/A	-

**Supplementary Table S7. Proportions of missing data in both DSFGs and bivalve species.**  
Bivalve species represented by transcriptomic data are highlighted with an asterisk (\*).

Species		Genes	
Species	% missing data (out of 33 DSFGs)	Group	% missing data (out of 43 bivalve species)
<i>A. irradians concentricus</i>	0.000000	<i>Dmrt-1L</i>	48.837209
<i>A. marissinica</i>	21.212121	<i>Dmrt-3</i>	30.232558
<i>A. pectinata*</i>	48.484848	<i>Dmrt-2</i>	55.813953
<i>A. purpuratus</i>	6.060606	<i>Dmrt-4/5</i>	6.976744
<i>C. angulata</i>	6.060606	<i>Fox-A</i>	13.953488
<i>C. ariakensis</i>	3.030303	<i>Fox-B</i>	30.232558
<i>C. fluminea*</i>	42.424242	<i>Fox-C</i>	23.255814
<i>C. gigas</i>	6.060606	<i>Fox-D</i>	20.930233
<i>C. plicata*</i>	21.212121	<i>Fox-E</i>	30.232558
<i>C. sinensis</i>	21.212121	<i>Fox-F</i>	18.604651
<i>C. virginica</i>	3.030303	<i>Fox-G</i>	16.279070
<i>D. polymorpha</i>	9.090909	<i>Fox-H</i>	41.860465
<i>H. bialata</i>	9.090909	<i>Fox-J1</i>	0.000000
<i>L. orbiculatus*</i>	63.636364	<i>Fox-J2/3</i>	9.302326
<i>M. arenaria</i>	21.212121	<i>Fox-L1</i>	18.604651
<i>M. californianus</i>	9.090909	<i>Fox-L2</i>	13.953488
<i>M. chinensis*</i>	57.575758	<i>Fox-N1/4</i>	6.976744
<i>M. coruscus</i>	0.000000	<i>Fox-N2/3</i>	6.976744
<i>M. edulis</i>	3.030303	<i>Fox-O</i>	13.953488
<i>M. galloprovincialis</i>	6.060606	<i>Fox-P</i>	9.302326
<i>M. margaritifera</i>	6.060606	<i>Fox-Q2</i>	30.232558
<i>M. mercenaria</i>	3.030303	<i>Fox-OG13/NA</i>	32.558140
<i>M. modiolus*</i>	36.363636	<i>Fox-OG15/NA</i>	34.883721
<i>M. nervosa</i>	27.272727	<i>Fox-OG16/NA</i>	30.232558
<i>M. philippinarum</i>	9.090909	<i>Fox-OG2/NA</i>	16.279070
<i>O. edulis</i>	6.060606	<i>Fox-OG28/NA</i>	39.534884
<i>P. coreanum*</i>	18.181818	<i>Fox-OG39/NA</i>	37.209302
<i>P. generosa*</i>	54.545455	<i>Sox-B1/2</i>	0.000000
<i>P. margaritifera*</i>	21.212121	<i>Sox-C</i>	0.000000
<i>P. maximus</i>	0.000000	<i>Sox-D</i>	4.651163
<i>P. okutanii*</i>	54.545455	<i>Sox-E</i>	9.302326
<i>P. purpuratus*</i>	54.545455	<i>Sox-F</i>	13.953488
<i>P. streckersoni</i>	6.060606	<i>Sox-H</i>	20.930233
<i>P. viridis</i>	3.030303		
<i>P. yessoensis</i>	3.030303		
<i>R. decussatus*</i>	51.515152		
<i>R. philippinarum</i>	3.030303		
<i>S. broughtonii</i>	12.121212		
<i>S. constricta</i>	12.121212		
<i>S. glomerata</i>	9.090909		
<i>S. grandis*</i>	54.545455		
<i>T. granosa</i>	42.424242		
<i>T. squamosa*</i>	48.484848		

**Supplementary Table S8. Complete set of DSFGs in mammals.** For each gene, the species ID (Sp. ID) as in Supp. Tab. S4, the accession number (Gene ID), and the Possvm-based annotation are indicated.

Species ID	Gene ID	Group	Annotation	Species ID	Gene ID	Group	Annotation
Hamp	XP_057601513.1	Dmrt	Dmrt-B1	Cimi	XP_017353003.1	Fox	Fox-C1
Hsap	NP_149056.1	Dmrt	Dmrt-B1	Oana	XP_028912285.1	Fox	Fox-C1
Oafe	XP_007944154.1	Dmrt	Dmrt-B1	Dnov	XP_058141148.1	Fox	Fox-C1
Amel	XP_034508822.1	Dmrt	Dmrt-B1	Bbub	XP_025120521.1	Fox	Fox-C1
Bbub	XP_025144486.2	Dmrt	Dmrt-B1	Drot	XP_053776986.1	Fox	Fox-C1
Casi	XP_006839931.1	Dmrt	Dmrt-B1	Rfer	XP_032971521.1	Fox	Fox-C1

Table S8 continued from previous page

Species ID	Gene ID	Group	Annotation	Species ID	Gene ID	Group	Annotation
Ppig	XP_039739094.1	Dmrt	Dmrt-B1	Mdom	XP_007488114.2	Fox	Fox-C1
Cimi	XP_017359902.1	Dmrt	Dmrt-B1	Casi	XP_006870626.1	Fox	Fox-C1
Mmus	XP_036020151.1	Dmrt	Dmrt-B1	Pafr	XP_047652737.1	Fox	Fox-C1
Bmus	XP_036712433.1	Dmrt	Dmrt-B1	Bmus	XP_036726614.1	Fox	Fox-C1
Oana	XP_039770576.1	Dmrt	Dmrt-B1	Hamp	XP_057555406.1	Fox	Fox-C1
Scar	XP_047395630.1	Dmrt	Dmrt-B1	Lcat	XP_045407522.1	Fox	Fox-C1
Mjav	XP_036877684.1	Dmrt	Dmrt-B1	Opri	XP_058524289.1	Fox	Fox-C1
Ptig	XP_042853169.1	Dmrt	Dmrt-B1	Cdid	XP_037699408.1	Fox	Fox-C1
Lcat	XP_045402503.1	Dmrt	Dmrt-B1	Hsap	NP_001444.2	Fox	Fox-C1
Emax	XP_049731502.1	Dmrt	Dmrt-B1	Mjav	XP_036867643.1	Fox	Fox-C1
Ttru	XP_004325875.3	Dmrt	Dmrt-B1	Scar	XP_047413053.1	Fox	Fox-C1
Mang	XP_045750547.1	Dmrt	Dmrt-B1	Cdro	XP_031291159.1	Fox	Fox-C1
Pafr	XP_047647463.1	Dmrt	Dmrt-B1	Mmus	NP_032618.2	Fox	Fox-C1
Opri	XP_004588846.2	Dmrt	Dmrt-B1	Emax	XP_049738278.1	Fox	Fox-C1
Cpor	XP_003464114.1	Dmrt	Dmrt-B1	Oafe	XP_007933566.1	Fox	Fox-C1
Dnov	XP_0581159531.1	Dmrt	Dmrt-B1	Mang	XP_045722432.1	Fox	Fox-C1
Cdid	XP_037665497.1	Dmrt	Dmrt-B1	Ttru	XP_019800382.2	Fox	Fox-C1
Equa	XP_046516609.1	Dmrt	Dmrt-B1	Oana	XP_028909462.1	Fox	Fox-C1
Mdom	XP_056671081.1	Dmrt	Dmrt-B1	Equa	XP_046496475.1	Fox	Fox-C1
Csim	XP_014649261.1	Dmrt	Dmrt-B1	Pgig	XP_039725399.1	Fox	Fox-C1
Cdro	XP_031321432.1	Dmrt	Dmrt-B1	Shar	XP_023352145.2	Fox	Fox-C1
Rfer	XP_032970385.1	Dmrt	Dmrt-B1	Ptig	XP_042841201.1	Fox	Fox-C1
Clup	XP_038393419.1	Dmrt	Dmrt-B1	Amel	XP_034516093.1	Fox	Fox-C1
Tman	XP_012409672.2	Dmrt	Dmrt-B1	Clup	XP_038439885.1	Fox	Fox-C1
Ggal	NP_001232910.1	Dmrt	Dmrt-B1	Scar	XP_047385678.1	Fox	Fox-C2
Oafe	XP_007954175.1	Dmrt	Dmrt-2	Mjav	XP_036849635.1	Fox	Fox-C2
Hsap	NP_870987.2	Dmrt	Dmrt-2	Cdro	XP_031292483.1	Fox	Fox-C2
Dnov	XP_004459445.2	Dmrt	Dmrt-2	Ttru	XP_033700545.1	Fox	Fox-C2
Shar	XP_003761708.1	Dmrt	Dmrt-2	Tman	XP_023588281.1	Fox	Fox-C2
Cpor	XP_005004778.1	Dmrt	Dmrt-2	Csim	XP_004437101.1	Fox	Fox-C2
Ttru	XP_019790310.1	Dmrt	Dmrt-2	Pafr	XP_047648754.1	Fox	Fox-C2
Ggal	NP_003643035.3	Dmrt	Dmrt-2	Mjav	XP_036849634.1	Fox	Fox-C2
Mdom	XP_001374295.1	Dmrt	Dmrt-2	Rfer	XP_032982997.1	Fox	Fox-C2
Oana	XP_028911116.1	Dmrt	Dmrt-2	Pgig	XP_039744659.1	Fox	Fox-C2
Tman	XP_004373589.1	Dmrt	Dmrt-2	Ggal	NP_001382975.1	Fox	Fox-C2
Equa	XP_046519727.1	Dmrt	Dmrt-2	Clup	XP_038394038.1	Fox	Fox-C2
Bmus	XP_036712720.1	Dmrt	Dmrt-2	Opri	XP_058530600.1	Fox	Fox-C2
Cdid	XP_037653375.1	Dmrt	Dmrt-2	Amel	XP_034495216.1	Fox	Fox-C2
Opri	XP_058513365.1	Dmrt	Dmrt-2	Bmus	XP_036688981.1	Fox	Fox-C2
Lcat	XP_045418965.1	Dmrt	Dmrt-2	Lcat	XP_045389072.1	Fox	Fox-C2
Cimi	XP_017373822.1	Dmrt	Dmrt-2	Cpor	XP_005008627.1	Fox	Fox-C2
Rfer	XP_032979206.1	Dmrt	Dmrt-2	Mmus	NP_038547.2	Fox	Fox-C2
Pgig	XP_039723855.1	Dmrt	Dmrt-2	Oafe	XP_007937406.1	Fox	Fox-C2
Ptig	XP_042820943.1	Dmrt	Dmrt-2	Cdid	XP_037671788.1	Fox	Fox-C2
Hamp	XP_057579943.1	Dmrt	Dmrt-2	Bbub	XP_006080917.3	Fox	Fox-C2
Pafr	XP_047623676.1	Dmrt	Dmrt-2	Emax	XP_049720645.1	Fox	Fox-C2
Cdro	XP_031305655.1	Dmrt	Dmrt-2	Rfer	XP_032982975.1	Fox	Fox-C2
Mmus	NP_665830.1	Dmrt	Dmrt-2	Equa	XP_046539000.1	Fox	Fox-C2
Csim	XP_014650903.1	Dmrt	Dmrt-2	Ptig	XP_042825379.1	Fox	Fox-C2
Casi	XP_006863797.1	Dmrt	Dmrt-2	Mang	XP_045746247.1	Fox	Fox-C2
Emax	XP_049752756.1	Dmrt	Dmrt-2	Cimi	XP_017399160.1	Fox	Fox-C2
Mjav	XP_017529816.2	Dmrt	Dmrt-2	Drot	XP_053772216.1	Fox	Fox-C2
Amel	XP_034503257.1	Dmrt	Dmrt-2	Hsap	NP_005242.1	Fox	Fox-C2
Clup	XP_038383091.1	Dmrt	Dmrt-2	Hamp	XP_057568935.1	Fox	Fox-C2
Mang	XP_045738142.1	Dmrt	Dmrt-2	Mdom	XP_001365891.1	Fox	Fox-C2
Bbub	XP_044795850.1	Dmrt	Dmrt-2	Dnov	XP_004450287.3	Fox	Fox-C2
Drot	XP_045047034.2	Dmrt	Dmrt-2	Casi	XP_006860181.1	Fox	Fox-C2
Scar	XP_047380961.1	Dmrt	Dmrt-2	Oana	XP_039769457.1	Fox	Fox-C2
Cdro	XP_031312888.1	Dmrt	Dmrt-C2	Shar	XP_031806033.1	Fox	Fox-C2
Clup	XP_038381591.1	Dmrt	Dmrt-C2	Oana	XP_028921129.1	Fox	Hnf-3g/Fox-A3
Csim	XP_004440164.1	Dmrt	Dmrt-C2	Hsap	NP_004488.2	Fox	Hnf-3g/Fox-A3
Casi	XP_006871451.1	Dmrt	Dmrt-C2	Pafr	XP_047645608.1	Fox	Hnf-3g/Fox-A3
Mang	XP_045746349.1	Dmrt	Dmrt-C2	Ptig	XP_042825317.1	Fox	Hnf-3g/Fox-A3
Bmus	XP_036690564.1	Dmrt	Dmrt-C2	Mjav	XP_017507570.1	Fox	Hnf-3g/Fox-A3
Lcat	XP_045387563.1	Dmrt	Dmrt-C2	Mmus	NP_032286.1	Fox	Hnf-3g/Fox-A3
Hamp	XP_057568752.1	Dmrt	Dmrt-C2	Casi	XP_006871384.1	Fox	Hnf-3g/Fox-A3
Cdid	XP_037676589.1	Dmrt	Dmrt-C2	Shar	XP_031817519.1	Fox	Hnf-3g/Fox-A3
Pafr	XP_047644735.1	Dmrt	Dmrt-C2	Oafe	XP_007941248.1	Fox	Hnf-3g/Fox-A3
Mmus	XP_017167796.1	Dmrt	Dmrt-C2	Ttru	XP_033701564.1	Fox	Hnf-3g/Fox-A3
Equa	XP_046507274.1	Dmrt	N/A	Mang	XP_045746361.1	Fox	Hnf-3g/Fox-A3
Oafe	XP_007941145.1	Dmrt	Dmrt-C2	Clup	XP_038384456.1	Fox	Hnf-3g/Fox-A3
Tman	XP_004388803.1	Dmrt	Dmrt-C2	Opri	XP_004597990.1	Fox	Hnf-3g/Fox-A3
Drot	XP_024433509.1	Dmrt	Dmrt-C2	Hamp	XP_057566925.1	Fox	Hnf-3g/Fox-A3
Oana	XP_007664048.1	Dmrt	Dmrt-C2	Rfer	XP_032982943.1	Fox	Hnf-3g/Fox-A3
Ttru	XP_033699919.1	Dmrt	Dmrt-C2	Dnov	XP_004481594.1	Fox	Hnf-3g/Fox-A3
Amel	XP_034494561.1	Dmrt	Dmrt-C2	Emax	XP_049757192.1	Fox	Hnf-3g/Fox-A3
Opri	XP_058531141.1	Dmrt	Dmrt-C2	Pgig	XP_039722184.1	Fox	Hnf-3g/Fox-A3
Pgig	XP_039731236.1	Dmrt	N/A	Tman	XP_004381648.1	Fox	Hnf-3g/Fox-A3
Lcat	XP_045393814.1	Dmrt	N/A	Mdom	XP_001364242.1	Fox	Hnf-3g/Fox-A3
Lcat	XP_045393815.1	Dmrt	N/A	Bmus	XP_036689576.1	Fox	Hnf-3g/Fox-A3
Lcat	XP_045393818.1	Dmrt	N/A	Equa	XP_046539522.1	Fox	Hnf-3g/Fox-A3
Scar	XP_047383853.1	Dmrt	Dmrt-C2	Bbub	XP_025125137.1	Fox	Hnf-3g/Fox-A3
Pgig	XP_039705918.1	Dmrt	Dmrt-C2	Csim	XP_004440228.1	Fox	Hnf-3g/Fox-A3
Hsap	XP_016882612.1	Dmrt	Dmrt-C2	Cimi	XP_017358010.1	Fox	Hnf-3g/Fox-A3
Equa	XP_046540061.1	Dmrt	N/A	Scar	XP_047384662.1	Fox	Hnf-3g/Fox-A3
Mjav	XP_036869866.1	Dmrt	Dmrt-C2	Cdid	XP_037675543.1	Fox	Hnf-3g/Fox-A3
Rfer	XP_032984641.1	Dmrt	Dmrt-C2	Lcat	XP_045387339.1	Fox	Hnf-3g/Fox-A3
Shar	XP_031819119.1	Dmrt	Dmrt-C2	Cdro	XP_010985920.2	Fox	Hnf-3g/Fox-A3
Emax	XP_049757407.1	Dmrt	Dmrt-C2	Cpor	XP_003464639.1	Fox	Hnf-3g/Fox-A3
Lcat	XP_045393817.1	Dmrt	N/A	Amel	XP_002928546.3	Fox	Hnf-3g/Fox-A3
Ptig	XP_042825393.1	Dmrt	Dmrt-C2	Drot	XP_024425331.1	Fox	Hnf-3g/Fox-A3
Bbub	XP_025124040.1	Dmrt	Dmrt-C2	Ggal	XP_025006986.1	Fox	Hnf-3a/Fox-A1
Cimi	XP_037585356.1	Dmrt	Dmrt-C2	Pafr	XP_047652638.1	Fox	Hnf-3a/Fox-A1
Cpor	XP_003464597.1	Dmrt	Dmrt-C2	Emax	XP_049755446.1	Fox	Hnf-3a/Fox-A1
Csim	XP_004443941.1	Dmrt	N/A	Cimi	XP_017386340.1	Fox	Hnf-3a/Fox-A1
Mdom	XP_007506399.1	Dmrt	Dmrt-C2	Csim	XP_004436444.1	Fox	Hnf-3a/Fox-A1
Dnov	XP_0581135474.1	Dmrt	Dmrt-C2	Cpor	XP_013010026.1	Fox	Hnf-3a/Fox-A1
Lcat	XP_045393811.1	Dmrt	N/A	Ttru	XP_004326816.1	Fox	Hnf-3a/Fox-A1
Hsap	NP_067063.1	Dmrt	Dmrt-3	Lcat	XP_045424245.1	Fox	Hnf-3a/Fox-A1
Rfer	XP_032979350.1	Dmrt	Dmrt-3	Scar	XP_047398474.1	Fox	Hnf-3a/Fox-A1
Cimi	XP_017373824.1	Dmrt	Dmrt-3	Mjav	XP_017504124.1	Fox	Hnf-3a/Fox-A1
Oafe	XP_007954185.1	Dmrt	Dmrt-3	Mang	XP_045757926.1	Fox	Hnf-3a/Fox-A1
Cdid	XP_037653502.1	Dmrt	Dmrt-3	Mmus	XP_017170451.1	Fox	Hnf-3a/Fox-A1

Table S8 continued from previous page

Species ID	Gene ID	Group	Annotation	Species ID	Gene ID	Group	Annotation
Mjav	XP_036847248.1	Dmrt	Dmrt-3	Casi	XP_006835393.1	Fox	Hnf-3a/Fox-A1
Mmus	XP_006527129.1	Dmrt	Dmrt-3	Dnov	XP_004453895.1	Fox	Hnf-3a/Fox-A1
Equa	XP_046521022.1	Dmrt	Dmrt-3	Cdro	XP_010982088.2	Fox	Hnf-3a/Fox-A1
Drot	XP_053780057.1	Dmrt	Dmrt-3	Clup	XP_038400478.1	Fox	Hnf-3a/Fox-A1
Hamp	XP_057581111.1	Dmrt	Dmrt-3	Hsap	NP_004487.2	Fox	Hnf-3a/Fox-A1
Lcat	XP_045418289.1	Dmrt	Dmrt-3	Ptig	XP_042845515.1	Fox	Hnf-3a/Fox-A1
Clup	XP_038381546.1	Dmrt	Dmrt-3	Pgig	XP_039712440.1	Fox	Hnf-3a/Fox-A1
Bmus	XP_036712559.1	Dmrt	Dmrt-3	Shar	XP_03756460.3	Fox	Hnf-3a/Fox-A1
Oana	XP_028910685.1	Dmrt	Dmrt-3	Amel	XP_019662027.2	Fox	Hnf-3a/Fox-A1
Casi	XP_006863796.1	Dmrt	Dmrt-3	Oafe	XP_007936759.1	Fox	Hnf-3a/Fox-A1
Scar	XP_047381281.1	Dmrt	Dmrt-3	Hamp	XP_057579078.1	Fox	Hnf-3a/Fox-A1
Ptig	XP_042820871.1	Dmrt	Dmrt-3	Bbub	XP_006074313.1	Fox	Hnf-3a/Fox-A1
Cpor	XP_003472245.1	Dmrt	Dmrt-3	Equa	XP_046511415.1	Fox	Hnf-3a/Fox-A1
Mdom	XP_001365681.1	Dmrt	Dmrt-3	Tman	XP_004376553.1	Fox	Hnf-3a/Fox-A1
Csim	XP_004440708.1	Dmrt	Dmrt-3	Rfer	XP_032965023.1	Fox	Hnf-3a/Fox-A1
Emax	XP_049752714.1	Dmrt	Dmrt-3	Drot	XP_053783670.1	Fox	Hnf-3a/Fox-A1
Opri	XP_004591889.2	Dmrt	Dmrt-3	Opri	XP_004584835.2	Fox	Hnf-3a/Fox-A1
Shar	XP_031799800.1	Dmrt	Dmrt-3	Mdom	XP_003339496.3	Fox	Hnf-3a/Fox-A1
Cdro	XP_031305659.1	Dmrt	Dmrt-3	Cdid	XP_037690784.1	Fox	Hnf-3a/Fox-A1
Bbub	XP_0251137363.1	Dmrt	Dmrt-3	Bmus	XP_036697173.1	Fox	Hnf-3a/Fox-A1
Pafr	XP_047625503.1	Dmrt	Dmrt-3	Oana	XP_028934365.1	Fox	Hnf-3a/Fox-A1
Mang	XP_045737886.1	Dmrt	Dmrt-3	Shar	XP_031825225.1	Fox	Fox-E1
Dnov	XP_004459447.2	Dmrt	Dmrt-3	Casi	XP_006863203.1	Fox	Fox-E1
Pgig	XP_039743957.1	Dmrt	Dmrt-3	Ggal	XP_040561338.1	Fox	Fox-E1
Ttru	XP_019789710.1	Dmrt	Dmrt-3	Oana	XP_007660203.2	Fox	Fox-E1
Ggal	XP_429193.2	Dmrt	Dmrt-3	Mdom	XP_001372714.1	Fox	Fox-E1
Amel	XP_034503143.1	Dmrt	Dmrt-3	Shar	XP_031824414.1	Fox	Fox-E1
Oafe	XP_007954174.1	Dmrt	Dmrt-1	Mang	XP_045737945.1	Fox	Fox-E1
Mdom	XP_001365618.1	Dmrt	Dmrt-1	Hamp	XP_057576142.1	Fox	Fox-E1
Tman	XP_004373524.1	Dmrt	Dmrt-1	Lcat	XP_045418451.1	Fox	Fox-E1
Cdro	XP_031305660.1	Dmrt	Dmrt-1	Equa	XP_046500267.1	Fox	Fox-E1
Bmus	XP_036711655.1	Dmrt	Dmrt-1	Hsap	NP_004464.2	Fox	Fox-E1
Shar	XP_031799789.1	Dmrt	Dmrt-1	Amel	XP_019650476.2	Fox	Fox-E1
Equa	XP_046519850.1	Dmrt	Dmrt-1	Csim	XP_004423293.1	Fox	Fox-E1
Bbub	XP_0251137364.1	Dmrt	Dmrt-1	Drot	XP_053778384.1	Fox	Fox-E1
Hsap	XP_006716795.1	Dmrt	Dmrt-1	Cimi	XP_017383334.1	Fox	Fox-E1
Pafr	XP_047625504.1	Dmrt	Dmrt-1	Cdro	XP_031305931.1	Fox	Fox-E1
Drot	XP_053774861.1	Dmrt	Dmrt-1	Rfer	XP_032978155.1	Fox	Fox-E1
Cpor	XP_023421841.1	Dmrt	Dmrt-1	Bbub	XP_006080990.4	Fox	Fox-E1
Mang	XP_045738144.1	Dmrt	Dmrt-1	Cdid	XP_037654444.1	Fox	Fox-E1
Emax	XP_049752111.1	Dmrt	Dmrt-1	Ttru	XP_019804445.2	Fox	Fox-E1
Clup	XP_038383087.1	Dmrt	Dmrt-1	Dnov	XP_004454020.1	Fox	Fox-E1
Ptig	XP_042820071.1	Dmrt	Dmrt-1	Clup	XP_038408835.1	Fox	Fox-E1
Mmus	NP_056641.2	Dmrt	Dmrt-1	Pafr	XP_047623959.1	Fox	Fox-E1
Opri	XP_058513291.1	Dmrt	Dmrt-1	Opri	XP_004581237.3	Fox	Fox-E1
Csim	XP_014650940.1	Dmrt	Dmrt-1	Scar	XP_047379621.1	Fox	Fox-E1
Pgig	XP_039744060.1	Dmrt	Dmrt-1	Pgig	XP_039711385.1	Fox	Fox-E1
Ggal	XP_040511578.1	Dmrt	Dmrt-1	Mmus	NP_899121.1	Fox	Fox-E1
Oana	XP_028910684.1	Dmrt	Dmrt-1	Emax	XP_049752918.1	Fox	Fox-E1
Amel	XP_011224113.2	Dmrt	Dmrt-1	Bmus	XP_036710285.1	Fox	Fox-E1
Ttru	XP_03714027.1	Dmrt	Dmrt-1	Ptig	XP_042821250.1	Fox	Fox-E1
Cdid	XP_037653737.1	Dmrt	Dmrt-1	Mjav	XP_036860072.1	Fox	Fox-E1
Hamp	XP_057579864.1	Dmrt	Dmrt-1	Ggal	NP_990282.3	Fox	Fox-D3
Rfer	XP_032979671.1	Dmrt	Dmrt-1	Drot	XP_053776629.1	Fox	Fox-D3
Casi	XP_006863795.1	Dmrt	Dmrt-1	Cdro	XP_031321142.1	Fox	Fox-D3
Scar	XP_047381742.1	Dmrt	Dmrt-1	Ttru	XP_019774777.1	Fox	Fox-D3
Mjav	XP_036847220.1	Dmrt	Dmrt-1	Oana	XP_007663453.2	Fox	Fox-D3
Lcat	XP_045418248.1	Dmrt	Dmrt-1	Amel	XP_034509500.1	Fox	Fox-D3
Cimi	XP_017373819.1	Dmrt	Dmrt-1	Clup	XP_038391469.1	Fox	Fox-D3
Dnov	XP_058158898.1	Dmrt	Dmrt-1	Cdid	XP_037682960.1	Fox	Fox-D3
Oana	XP_039766676.1	Dmrt	Dmrt-A1	Mmus	NP_034555.3	Fox	Fox-D3
Cimi	XP_017041520.1	Dmrt	Dmrt-A1	Cimi	XP_017397049.1	Fox	Fox-D3
Mdom	XP_001374014.2	Dmrt	Dmrt-A1	Casi	XP_006862529.1	Fox	Fox-D3
Mang	XP_045737916.1	Dmrt	Dmrt-A1	Emax	XP_049731927.1	Fox	Fox-D3
Bmus	XP_036712737.1	Dmrt	Dmrt-A1	Ptig	XP_042852934.1	Fox	Fox-D3
Tman	XP_004373619.1	Dmrt	Dmrt-A1	Scar	XP_047371747.1	Fox	Fox-D3
Hsap	NP_071443.2	Dmrt	Dmrt-A1	Bmus	XP_036709961.1	Fox	Fox-D3
Emax	XP_049751570.1	Dmrt	Dmrt-A1	Opri	XP_058516851.1	Fox	Fox-D3
Drot	XP_053780867.1	Dmrt	Dmrt-A1	Dnov	XP_004448976.1	Fox	Fox-D3
Scar	XP_047381594.1	Dmrt	Dmrt-A1	Mjav	XP_036880296.1	Fox	Fox-D3
Hamp	XP_057580644.1	Dmrt	Dmrt-A1	Tman	XP_004369185.1	Fox	Fox-D3
Ttru	XP_004321228.1	Dmrt	Dmrt-A1	Hsap	NP_036315.1	Fox	Fox-D3
Opri	XP_058528639.1	Dmrt	Dmrt-A1	Lcat	XP_045402384.1	Fox	Fox-D3
Rfer	XP_032978102.1	Dmrt	Dmrt-A1	Bbub	XP_025144350.1	Fox	Fox-D3
Amel	XP_034519591.1	Dmrt	Dmrt-A1	Mang	XP_045751362.1	Fox	Fox-D3
Clup	XP_038408464.1	Dmrt	Dmrt-A1	Hamp	XP_057560293.1	Fox	Fox-D3
Oafe	XP_007946600.1	Dmrt	Dmrt-A1	Pafr	XP_047648179.1	Fox	Fox-D3
Pafr	XP_047625485.1	Dmrt	Dmrt-A1	Shar	XP_031824242.1	Fox	Fox-D3
Pgig	XP_039736661.1	Dmrt	Dmrt-A1	Pgig	XP_039739159.1	Fox	Fox-D3
Cpor	XP_003478909.1	Dmrt	Dmrt-A1	Rfer	XP_032969240.1	Fox	Fox-D3
Mmus	NP_783578.1	Dmrt	Dmrt-A1	Mdom	XP_056670078.1	Fox	Fox-D3
Lcat	XP_045418524.1	Dmrt	Dmrt-A1	Scar	XP_047393167.1	Fox	Fox-N3
Bbub	XP_006067119.4	Dmrt	Dmrt-A1	Opri	XP_058511345.1	Fox	Fox-N3
Cdro	XP_010982933.2	Dmrt	Dmrt-A1	Cpor	XP_003472517.1	Fox	Fox-N3
Csim	XP_004428451.1	Dmrt	Dmrt-A1	Mmus	XP_006516295.1	Fox	Fox-N3
Mjav	XP_017516191.1	Dmrt	Dmrt-A1	Hsap	NP_001078940.1	Fox	Fox-N3
Ptig	XP_007098573.2	Dmrt	Dmrt-A1	Lcat	XP_045418142.1	Fox	Fox-N3
Casi	XP_006864558.1	Dmrt	Dmrt-A1	Cdro	XP_031310026.1	Fox	Fox-N3
Cdid	XP_037653778.1	Dmrt	Dmrt-A1	Rfer	XP_032965297.1	Fox	Fox-N3
Equa	XP_046520805.1	Dmrt	Dmrt-A1	Bbub	XP_045018818.1	Fox	Fox-N3
Dnov	XP_004456130.1	Dmrt	Dmrt-A1	Equa	XP_046503458.1	Fox	Fox-N3
Ggal	XP_015146712.1	Dmrt	Dmrt-A2	Pgig	XP_03973926.1	Fox	Fox-N3
Amel	XP_034508227.1	Dmrt	Dmrt-A2	Bmus	XP_036697221.1	Fox	Fox-N3
Opri	XP_004588718.2	Dmrt	Dmrt-A2	Ttru	XP_019775018.1	Fox	Fox-N3
Oana	XP_028902450.1	Dmrt	Dmrt-A2	Mjav	XP_036847649.1	Fox	Fox-N3
Bbub	XP_025145263.1	Dmrt	Dmrt-A2	Csim	XP_014645970.1	Fox	Fox-N3
Oafe	XP_007944067.1	Dmrt	Dmrt-A2	Drot	XP_024424133.1	Fox	Fox-N3
Dnov	XP_004476539.1	Dmrt	Dmrt-A2	Opri	XP_058516899.1	Fox	Fox-N3
Cdid	XP_037672757.1	Dmrt	Dmrt-A2	Cimi	XP_017380034.1	Fox	Fox-N3
Pgig	XP_039741660.1	Dmrt	Dmrt-A2	Bbub	XP_006061543.1	Fox	Fox-N3
Hsap	NP_115486.1	Dmrt	Dmrt-A2	Pafr	XP_047652286.1	Fox	Fox-N3
Mmus	NP_758500.2	Dmrt	Dmrt-A2	Hamp	XP_057589083.1	Fox	Fox-N3
Ttru	XP_019793397.2	Dmrt	Dmrt-A2	Cimi	XP_017372373.1	Fox	Fox-R2

Table S8 continued from previous page

Species ID	Gene ID	Group	Annotation	Species ID	Gene ID	Group	Annotation
Hamp	XP_057564703.1	Dmrt	Dmrt-A2	Hsap	NP_940853.1	Fox	Fox-R2
Csim	XP_014649324.1	Dmrt	Dmrt-A2	Mdom	XP_007478014.2	Fox	Fox-I2
Bmus	XP_036726057.1	Dmrt	Dmrt-A2	Mjav	XP_017495569.2	Fox	Fox-I2
Clup	XP_038413502.1	Dmrt	Dmrt-A2	Shar	XP_031812885.1	Fox	Fox-I2
Mjav	XP_036877643.1	Dmrt	Dmrt-A2	Cdro	XP_031317856.1	Fox	Fox-I2
Pafr	XP_047648562.1	Dmrt	Dmrt-A2	Cdid	XP_037661256.1	Fox	Fox-I2
Cdro	XP_031320961.1	Dmrt	Dmrt-A2	Hamp	XP_057593248.1	Fox	Fox-I2
Cimi	XP_017404203.1	Dmrt	Dmrt-A2	Ttru	XP_019795179.2	Fox	Fox-I2
Scar	XP_047394472.1	Dmrt	Dmrt-A2	Ptig	XP_042815793.1	Fox	Fox-I2
Mang	XP_045751377.1	Dmrt	Dmrt-A2	Bmus	XP_03685524.1	Fox	Fox-I2
Casi	XP_006839813.1	Dmrt	Dmrt-A2	Oafe	XP_007935078.1	Fox	Fox-I2
Tman	XP_023582760.1	Dmrt	Dmrt-A2	Bbub	XP_025129910.2	Fox	Fox-I2
Drot	XP_045060264.2	Dmrt	Dmrt-A2	Cimi	XP_017393376.1	Fox	Fox-I2
Lcat	XP_045402400.1	Dmrt	Dmrt-A2	Lcat	XP_045424382.1	Fox	Fox-I2
Rfer	XP_032969204.1	Dmrt	Dmrt-A2	Mmus	NP_899016.2	Fox	Fox-I2
Emax	XP_049731145.1	Dmrt	Dmrt-A2	Opri	XP_004579799.2	Fox	Fox-I2
Mdom	XP_001362692.2	Dmrt	Dmrt-A2	Scar	XP_047407992.1	Fox	Fox-I2
Cpor	XP_003463125.2	Dmrt	Dmrt-A2	Tman	XP_004375038.1	Fox	Fox-I2
Equa	XP_046516465.1	Dmrt	Dmrt-A2	Dnov	XP_004479387.2	Fox	Fox-I2
Ptig	XP_042853035.1	Dmrt	Dmrt-A2	Amel	XP_034517693.1	Fox	Fox-I2
Cimi	XP_017353253.1	Fox	N/A	Rfer	XP_032986720.1	Fox	Fox-I2
Cimi	XP_037600788.1	Fox	N/A	Equa	XP_036508191.1	Fox	Fox-I2
Tman	XP_023591424.1	Fox	Fox-M1	Casi	XP_006867537.1	Fox	Fox-I2
Bbub	XP_025138819.3	Fox	Fox-M1	Pafr	XP_047617388.1	Fox	Fox-I2
Casi	XP_006862752.1	Fox	Fox-M1	Csim	XP_014649530.1	Fox	Fox-I2
Emax	XP_049738130.1	Fox	Fox-M1	Drot	XP_045047252.2	Fox	Fox-I2
Oafe	XP_007935251.1	Fox	Fox-M1	Hsap	NP_997309.2	Fox	Fox-I2
Drot	XP_053777572.1	Fox	Fox-M1	Mang	XP_045732277.1	Fox	Fox-I2
Ttru	XP_019781496.1	Fox	Fox-M1	Cpor	XP_003479857.1	Fox	Fox-I2
Csim	XP_014649484.1	Fox	Fox-M1	Clup	XP_038433915.1	Fox	Fox-I2
Pafr	XP_047643777.1	Fox	Fox-M1	Ggal	NP_990523.3	Fox	Fox-D1
Oana	XP_028910212.1	Fox	Fox-M1	Lcat	XP_045421654.1	Fox	Fox-D1
Mang	XP_045727136.1	Fox	Fox-M1	Oana	XP_028933047.1	Fox	Fox-D1
Opri	XP_058511953.1	Fox	Fox-M1	Cimi	XP_017383122.1	Fox	Fox-D1
Pgig	XP_039725073.1	Fox	Fox-M1	Drot	XP_053781190.1	Fox	Fox-D1
Bmus	XP_036723013.1	Fox	Fox-M1	Pafr	XP_047630421.1	Fox	Fox-D1
Dnov	XP_058156702.1	Fox	Fox-M1	Clup	XP_038385555.1	Fox	Fox-D1
Mjav	XP_017520172.2	Fox	Fox-M1	Shar	XP_031822455.1	Fox	Fox-D1
Mang	XP_045727136.1	Fox	Fox-M1	Bmus	XP_036703665.1	Fox	Fox-D1
Opri	XP_058511953.1	Fox	Fox-M1	Mdom	XP_056680722.1	Fox	Fox-D1
Pgig	XP_039725073.1	Fox	Fox-M1	Ttru	XP_019784789.1	Fox	Fox-D1
Bmus	XP_036723013.1	Fox	Fox-M1	Cdid	XP_037657638.1	Fox	Fox-D1
Dnov	XP_058156702.1	Fox	Fox-M1	Mmus	NP_03268.2	Fox	Fox-D1
Mjav	XP_017520172.2	Fox	Fox-M1	Opri	XP_047412450.1	Fox	Fox-D1
Amel	XP_034501532.1	Fox	Fox-M1	Pftr	XP_047412450.1	Fox	Fox-D1
Equa	XP_046512461.1	Fox	Fox-M1	Hamp	XP_057554809.1	Fox	Fox-D1
Dnov	XP_012379785.1	Fox	Fox-M1	Mdom	XP_056680722.1	Fox	Fox-D1
Clup	XP_038432360.1	Fox	Fox-M1	Ttru	XP_019784789.1	Fox	Fox-D1
Mdom	XP_056655089.1	Fox	Fox-M1	Cdid	XP_037657638.1	Fox	Fox-D1
Rfer	XP_032974271.1	Fox	Fox-M1	Opri	XP_058512313.1	Fox	Fox-D1
Hamp	XP_057558870.1	Fox	Fox-M1	Pftr	XP_042823025.1	Fox	Fox-D1
Cdro	XP_031300170.1	Fox	Fox-M1	Hamp	XP_057554809.1	Fox	Fox-D1
Ptig	XP_042847499.1	Fox	Fox-M1	Mdom	XP_056680722.1	Fox	Fox-D1
Shar	XP_031795171.1	Fox	Fox-M1	Rfer	XP_032965774.1	Fox	Fox-D1
Ggal	XP_046799087.1	Fox	Fox-M1	Emax	XP_049721422.1	Fox	Fox-D1
Cdid	XP_037702327.1	Fox	Fox-M1	Bbub	XP_025126650.2	Fox	Fox-D1
Mmus	NP_032047.4	Fox	Fox-M1	Pgig	XP_039714034.1	Fox	Fox-D1
Cpor	XP_013000601.1	Fox	Fox-M1	Hamp	XP_057554809.1	Fox	Fox-D1
Scar	XP_047405644.1	Fox	Fox-M1	Mjav	XP_036882073.1	Fox	Fox-D1
Ggal	XP_015153061.3	Fox	Fox-O6	Ptig	XP_042844167.1	Fox	Fox-D1
Cimi	XP_017397276.1	Fox	Fox-O6	Mang	XP_045742895.1	Fox	Fox-D1
Mdom	XP_001381541.2	Fox	Fox-O6	Opri	XP_058512313.1	Fox	Fox-D1
Pafr	XP_047645485.1	Fox	Fox-O6	Pftr	XP_042823025.1	Fox	N/A
Mmus	NP_918949.1	Fox	Fox-O6	Cdro	XP_031325602.1	Fox	N/A
Emax	XP_049731901.1	Fox	Fox-O6	Dnov	XP_058140947.1	Fox	N/A
Scar	XP_047415675.1	Fox	Fox-O6	Cdid	XP_037666595.1	Fox	N/A
Ptig	XP_042852325.1	Fox	Fox-O6	Shar	XP_023357319.1	Fox	Fox-N3
Ttru	XP_033692835.1	Fox	Fox-O6	Mdom	XP_056666520.1	Fox	Fox-N3
Rfer	XP_032969239.1	Fox	Fox-O6	Oana	XP_028914735.1	Fox	Fox-N3
Tman	XP_023582811.1	Fox	Fox-O6	Ggal	XP_015143223.2	Fox	Fox-N3
Pgig	XP_039732650.1	Fox	Fox-O6	Ggal	XP_425185.2	Fox	Fox-II
Cdro	XP_010997575.2	Fox	Fox-O6	Cpor	XP_003473379.2	Fox	Fox-II
Bmus	XP_036684744.1	Fox	Fox-O6	Cdro	XP_031293629.1	Fox	Fox-II
Casi	XP_006867666.1	Fox	Fox-O6	Pgig	XP_039732217.1	Fox	Fox-II
Oana	XP_028936057.1	Fox	Fox-O6	Pafr	XP_047640916.1	Fox	Fox-II
Hsap	NP_001278210.2	Fox	Fox-O6	Lcat	XP_045408637.1	Fox	Fox-II
Clup	XP_038413629.1	Fox	Fox-O6	Cimi	XP_017403198.1	Fox	Fox-II
Oafe	XP_007950907.1	Fox	Fox-O6	Oafe	XP_007937941.1	Fox	Fox-II
Bbub	XP_044801109.1	Fox	Fox-O6	Csim	XP_004428597.1	Fox	Fox-II
Hamp	XP_057605223.1	Fox	Fox-O6	Cdid	XP_037655181.1	Fox	Fox-II
Cpor	XP_004999818.1	Fox	Fox-O6	Hsap	NP_036320.2	Fox	Fox-II
Dnov	XP_058159776.1	Fox	Fox-O6	Ptig	XP_007077705.2	Fox	Fox-II
Mang	XP_045753906.1	Fox	Fox-O6	Mdom	XP_001370284.1	Fox	Fox-II
Csim	XP_014639249.1	Fox	Fox-O6	Ttru	XP_033711041.1	Fox	Fox-II
Amel	XP_034509638.1	Fox	Fox-O6	Shar	XP_003756860.1	Fox	Fox-II
Shar	XP_031818190.1	Fox	Fox-O6	Rfer	XP_032952356.1	Fox	Fox-II
Opri	XP_04591579.1	Fox	Fox-O6	Mjav	XP_017510459.2	Fox	Fox-II
Mjav	XP_036858263.1	Fox	Fox-O6	Casi	XP_006864891.1	Fox	Fox-II
Lcat	XP_045404045.1	Fox	Fox-O6	Mmus	NP_076396.3	Fox	Fox-II
Drot	XP_053776876.1	Fox	Fox-O6	Scar	XP_047412385.1	Fox	Fox-II
Equa	XP_046516690.1	Fox	Fox-O6	Drot	XP_024432893.2	Fox	Fox-II
Cdid	XP_037673411.1	Fox	Fox-O6	Hamp	XP_057605245.1	Fox	Fox-II
Oana	XP_028918467.1	Fox	Fox-H1	Amel	XP_002931169.2	Fox	Fox-II
Mjav	XP_017499188.2	Fox	Fox-H1	Opri	XP_004587500.1	Fox	Fox-II
Scar	XP_053599327.1	Fox	Fox-H1	Mang	XP_045743383.1	Fox	Fox-II
Cpor	XP_003463743.1	Fox	Fox-H1	Equa	XP_046524041.1	Fox	Fox-II
Bmus	XP_036686706.1	Fox	Fox-H1	Bmus	XP_036703672.1	Fox	Fox-II
Mjav	XP_036859882.1	Fox	Fox-H1	Oana	XP_028905836.1	Fox	Fox-II
Shar	XP_031803442.1	Fox	Fox-H1	Tman	XP_004371270.1	Fox	Fox-II
Emax	XP_049710232.1	Fox	Fox-H1	Emax	XP_049714816.1	Fox	Fox-II
Dnov	XP_058132182.1	Fox	Fox-H1	Clup	XP_038389675.1	Fox	Fox-II
Rfer	XP_032982232.1	Fox	Fox-H1	Bbub	XP_006075587.1	Fox	Fox-II
Tman	XP_004387459.1	Fox	Fox-H1	Dnov	XP_004460267.2	Fox	Fox-II
Equa	XP_046516690.1	Fox	Fox-H1	Ggal	NP_990283.1	Fox	Fox-D2
Cdid	XP_037659123.1	Fox	Fox-H1	Emax	XP_049731922.1	Fox	Fox-D2
Mang	XP_045751338.1	Fox	Fox-H1	Cdro	XP_031320938.1	Fox	Fox-D2
Ttru	XP_033698705.1	Fox	Fox-H1	Ptig	XP_042852978.1	Fox	Fox-D2
Ptig	XP_042829602.1	Fox	Fox-H1	Cdid	XP_037683211.1	Fox	Fox-D2
Bbub	XP_006064665.1	Fox	Fox-H1	Clup	XP_038413506.1	Fox	Fox-D2

Table S8 continued from previous page

Species ID	Gene ID	Group	Annotation	Species ID	Gene ID	Group	Annotation
Mmus	NP_032015.1	Fox	Fox-H1	Mang	XP_045751380.1	Fox	Fox-D2
Clup	XP_038411781.1	Fox	Fox-H1	Oafe	XP_007944900.1	Fox	Fox-D2
Cimi	XP_017367179.1	Fox	Fox-H1	Drot	XP_053776598.1	Fox	Fox-D2
Drot	XP_024428211.1	Fox	Fox-H1	Equa	XP_046519295.1	Fox	Fox-D2
Amel	XP_034524715.1	Fox	Fox-H1	Hamp	XP_057551251.1	Fox	Fox-D2
Mang	XP_045751622.1	Fox	Fox-H1	Cpor	XP_003461518.1	Fox	Fox-D2
Oafe	XP_007954253.1	Fox	Fox-H1	Csim	XP_004438679.1	Fox	Fox-D2
Hsap	NP_003914.1	Fox	Fox-H1	Casi	XP_006839796.1	Fox	Fox-D2
Csim	XP_004443058.1	Fox	Fox-H1	Scar	XP_053599256.1	Fox	Fox-D2
Cdro	XP_010997045.2	Fox	Fox-H1	Shar	XP_031825021.1	Fox	Fox-D2
Casi	XP_006830949.1	Fox	Fox-H1	Cimi	XP_017365579.1	Fox	Fox-D2
Pafr	XP_047641595.1	Fox	Fox-H1	Mdom	XP_003340123.1	Fox	Fox-D2
Equa	XP_046497693.1	Fox	Fox-H1	Bmus	XP_036724630.1	Fox	Fox-D2
Pgig	XP_039702105.1	Fox	Fox-H1	Pafr	XP_047645114.1	Fox	Fox-D2
Hamp	XP_057592839.1	Fox	Fox-H1	Cdro	XP_031319575.1	Fox	Fox-D2
Mdom	XP_016288146.2	Fox	Fox-H1	Mjav	XP_036877843.1	Fox	Fox-D2
Lcat	XP_045417128.1	Fox	Fox-H1	Hsap	NP_004465.3	Fox	Fox-D2
Ggal	XP_001234496.5	Fox	Fox-O3/Fox-O3B	Dnov	XP_004447059.1	Fox	Fox-D2
Mdom	XP_001368493.2	Fox	Fox-O3/Fox-O3B	Amel	XP_034511495.1	Fox	Fox-D2
Bmus	XP_036727351.1	Fox	Fox-O3/Fox-O3B	Oafe	XP_007944917.1	Fox	Fox-D2
Equa	XP_046532634.1	Fox	Fox-O3/Fox-O3B	Bbub	XP_025144570.1	Fox	Fox-D2
Emax	XP_049752199.1	Fox	Fox-O3/Fox-O3B	Ttru	XP_019793278.1	Fox	Fox-D2
Hsap	NP_001355064.1	Fox	Fox-O3/Fox-O3B	Rfer	XP_032969197.1	Fox	Fox-D2
Shar	XP_023358910.2	Fox	Fox-O3/Fox-O3B	Tman	XP_004371825.1	Fox	Fox-D2
Pgig	XP_039725886.1	Fox	Fox-O3/Fox-O3B	Oana	XP_028902568.1	Fox	Fox-D2
Cdid	XP_037698233.1	Fox	Fox-O3/Fox-O3B	Opri	XP_004598554.3	Fox	Fox-D2
Cimi	XP_037599874.1	Fox	Fox-O3/Fox-O3B	Pgig	XP_039741502.1	Fox	Fox-D2
Bbub	XP_044780301.2	Fox	Fox-O3/Fox-O3B	Mmus	NP_032619.1	Fox	Fox-D2
Opri	XP_058515970.1	Fox	Fox-O3/Fox-O3B	Lcat	XP_045401421.1	Fox	Fox-D2
Pafr	XP_047618159.1	Fox	Fox-O3/Fox-O3B	Csim	XP_004433028.1	Fox	N/A
Mmus	NP_0627174.1	Fox	Fox-O3/Fox-O3B	Rfer	XP_032946066.1	Fox	N/A
Oafe	XP_007939014.1	Fox	Fox-O3/Fox-O3B	Pgig	XP_039715358.1	Fox	N/A
Clup	XP_038410808.1	Fox	Fox-O3/Fox-O3B	Emax	XP_049754885.1	Fox	Fox-N3
Dnov	XP_058163213.1	Fox	Fox-O3/Fox-O3B	Tman	XP_023587457.1	Fox	Fox-N3
Scar	XP_047415677.1	Fox	Fox-O3/Fox-O3B	Cdid	XP_037678324.1	Fox	N/A
Ptig	XP_042842859.1	Fox	Fox-O3/Fox-O3B	Dnov	XP_004462180.1	Fox	N/A
Mang	XP_045721230.1	Fox	Fox-O3/Fox-O3B	Bbub	XP_006068529.1	Fox	N/A
Casi	XP_006840010.1	Fox	Fox-O3/Fox-O3B	Mang	XP_045748670.1	Fox	N/A
Mjav	XP_036868157.1	Fox	Fox-O3/Fox-O3B	Pafr	XP_047620398.1	Fox	N/A
Cpor	XP_023419223.1	Fox	Fox-O3/Fox-O3B	Bmus	XP_036694953.1	Fox	N/A
Hsap	NP_963853.1	Fox	Fox-O3/Fox-O3B	Cdro	XP_010983165.1	Fox	N/A
Drot	XP_024407555.2	Fox	Fox-O3/Fox-O3B	Hamp	XP_057573282.1	Fox	N/A
Lcat	XP_045400090.1	Fox	Fox-O3/Fox-O3B	Rfer	XP_032949909.1	Fox	N/A
Ttru	XP_033723495.1	Fox	Fox-O3/Fox-O3B	Equa	XP_046528957.1	Fox	N/A
Oana	XP_001511165.3	Fox	Fox-O3/Fox-O3B	Pgig	XP_039707370.1	Fox	N/A
Csim	XP_014636634.1	Fox	Fox-O3/Fox-O3B	Amel	XP_002930558.1	Fox	N/A
Cdro	XP_031311979.1	Fox	Fox-O3/Fox-O3B	Mjav	XP_017524585.2	Fox	N/A
Tman	XP_023583398.1	Fox	Fox-O3/Fox-O3B	Clup	XP_038442440.1	Fox	N/A
Amel	XP_034524842.1	Fox	Fox-O3/Fox-O3B	Mdom	XP_007483490.1	Fox	N/A
Rfer	XP_032957487.1	Fox	Fox-O3/Fox-O3B	Shar	XP_023358615.2	Fox	N/A
Hamp	XP_057593622.1	Fox	Fox-O3/Fox-O3B	Ggal	NP_001382146.1	Fox	N/A
Shar	XP_023352140.2	Fox	Fox-Q1	Mjav	XP_036853444.1	Fox	N/A
Mang	XP_045721917.1	Fox	Fox-Q1	Pafr	XP_047615141.1	Fox	N/A
Oana	XP_039766532.1	Fox	Fox-Q1	Bmus	XP_036692188.1	Fox	N/A
Mdom	XP_007488113.2	Fox	Fox-Q1	Ttru	XP_033703709.1	Fox	N/A
Oana	XP_039766897.1	Fox	Fox-Q1	Bbub	XP_006053445.2	Fox	N/A
Lcat	XP_045407521.1	Fox	Fox-Q1	Oana	XP_039768943.1	Fox	Fox-N2
Csim	XP_014639916.1	Fox	Fox-Q1	Shar	XP_031806310.1	Fox	Fox-N2
Rfer	XP_032969223.1	Fox	Fox-Q1	Pgig	XP_039743145.1	Fox	Fox-N2
Emax	XP_049738363.1	Fox	Fox-Q1	Emax	XP_049726307.1	Fox	Fox-N2
Scar	XP_047414319.1	Fox	Fox-Q1	Casi	XP_006839498.1	Fox	Fox-N2
Drot	XP_024408099.2	Fox	Fox-Q1	Csim	XP_004436724.1	Fox	Fox-N2
Ggal	XP_015137671.3	Fox	Fox-Q1	Cdro	XP_031322832.1	Fox	Fox-N2
Clup	XP_038440419.1	Fox	Fox-Q1	Amel	XP_011215591.1	Fox	Fox-N2
Bmus	XP_036726386.1	Fox	Fox-Q1	Bmus	XP_036729273.1	Fox	Fox-N2
Hamp	XP_057556649.1	Fox	Fox-Q1	Tman	XP_023591138.1	Fox	Fox-N2
Ttru	XP_019800554.1	Fox	Fox-Q1	Dnov	XP_023444629.2	Fox	Fox-N2
Cdid	XP_037700042.1	Fox	Fox-Q1	Lcat	XP_045405515.1	Fox	Fox-N2
Cdro	XP_031291204.1	Fox	Fox-Q1	Hsap	XP_047300063.1	Fox	Fox-N2
Equa	XP_046496965.1	Fox	Fox-Q1	Opri	XP_058523737.1	Fox	Fox-N2
Mjav	XP_036867494.1	Fox	Fox-Q1	Cimi	XP_017404041.1	Fox	Fox-N2
Ptig	XP_042841203.1	Fox	Fox-Q1	Clup	XP_038407168.1	Fox	Fox-N2
Bbub	XP_025120519.3	Fox	Fox-Q1	Drot	XP_053780764.1	Fox	Fox-N2
Mmus	NP_032265.3	Fox	Fox-Q1	Mang	XP_045740997.1	Fox	Fox-N2
Cimi	XP_017352994.1	Fox	Fox-Q1	Scar	XP_047378662.1	Fox	Fox-N2
Opri	XP_058528571.1	Fox	Fox-Q1	Equa	XP_046519098.1	Fox	Fox-N2
Dnov	XP_058141147.1	Fox	Fox-Q1	Ptig	XP_042837276.1	Fox	Fox-N2
Oafe	XP_007933584.1	Fox	Fox-Q1	Pafr	XP_047638079.1	Fox	Fox-N2
Pgig	XP_039725409.1	Fox	Fox-Q1	Cdid	XP_037662960.1	Fox	Fox-N2
Casi	XP_006870628.1	Fox	Fox-Q1	Ggal	XP_046794679.1	Fox	Fox-N2
Pafr	XP_047651460.1	Fox	Fox-Q1	Oafe	XP_007952184.1	Fox	Fox-N2
Hsap	NP_150285.3	Fox	Fox-Q1	Mmus	XP_036016210.1	Fox	Fox-N2
Cimi	XP_037597639.1	Fox	Fox-G1	Mjav	XP_017530752.1	Fox	Fox-N2
Lcat	XP_045411043.1	Fox	Fox-G1	Hamp	XP_057599130.1	Fox	Fox-N2
Cdro	XP_010984479.2	Fox	Fox-G1	Cpor	XP_003473126.1	Fox	Fox-N2
Rfer	XP_032965585.1	Fox	Fox-G1	Bbub	XP_044781285.1	Fox	Fox-N2
Drot	XP_024407046.1	Fox	Fox-G1	Mdom	XP_001375500.1	Fox	Fox-N2
Pafr	XP_047651592.1	Fox	Fox-G1	Ttru	XP_033694467.1	Fox	Fox-N2
Oana	XP_039770060.1	Fox	Fox-G1	Rfer	XP_032981124.1	Fox	Fox-N2
Pgig	XP_039726380.1	Fox	Fox-G1	Emax	XP_049728576.1	Fox	Fox-R2
Dnov	XP_058149613.1	Fox	Fox-G1	Tman	XP_004391284.1	Fox	Fox-R2
Mdom	XP_001364896.1	Fox	Fox-G1	Shar	XP_031815965.1	Fox	Fox-R1
Mang	XP_045757915.1	Fox	Fox-G1	Drot	XP_024426718.1	Fox	Fox-R1
Mjav	XP_017501197.2	Fox	Fox-G1	Mdom	XP_001380644.1	Fox	Fox-R1
Emax	XP_049754324.1	Fox	Fox-G1	Oana	XP_028931756.1	Fox	Fox-R1
Ttru	XP_019779874.1	Fox	Fox-G1	Casi	XP_006834052.1	Fox	Fox-R1
Cdid	XP_037690644.1	Fox	Fox-G1	Equa	XP_046540933.1	Fox	Fox-R1
Hsap	NP_005240.3	Fox	Fox-G1	Mmus	NP_001028641.1	Fox	Fox-R1
Shar	XP_031808132.1	Fox	Fox-G1	Cpor	XP_003472699.1	Fox	Fox-R1
Ptig	XP_042845427.1	Fox	Fox-G1	Bbub	XP_044785576.1	Fox	Fox-R1
Scar	XP_047396708.1	Fox	Fox-G1	Pgig	XP_039722531.1	Fox	Fox-R1
Clup	XP_038400306.1	Fox	Fox-G1	Rfer	XP_032976774.1	Fox	Fox-R1
Hamp	XP_057580105.1	Fox	Fox-G1	Ptig	XP_042814206.1	Fox	Fox-R1

Table S8 continued from previous page

Species ID	Gene ID	Group	Annotation	Species ID	Gene ID	Group	Annotation
Casi	XP_006835434.1	Fox	Fox-G1	Bmus	XP_036721689.1	Fox	Fox-R1
Opri	XP_058521348.1	Fox	Fox-G1	Bmus	XP_036715648.1	Fox	Fox-R1
Bmus	XP_036700190.1	Fox	Fox-G1	Cdro	XP_031299399.1	Fox	Fox-R1
Bbub	XP_025126917.1	Fox	Fox-G1	Pafr	XP_047609433.1	Fox	Fox-R1
Mmus	NP_001153584.1	Fox	Fox-G1	Hsap	XP_016873064.1	Fox	Fox-R1
Ggal	NP_989659.2	Fox	Fox-O1	Cdid	XP_037695751.1	Fox	Fox-R1
Cimi	XP_017397972.1	Fox	Fox-O1	Csim	XP_004427297.1	Fox	Fox-R1
Rfer	XP_032960350.1	Fox	Fox-O1	Emax	XP_049713316.1	Fox	Fox-R1
Casi	XP_006873780.1	Fox	Fox-O1	Mjav	XP_017505694.2	Fox	Fox-R1
Pgig	XP_039695386.1	Fox	Fox-O1	Hsap	NP_859072.1	Fox	Fox-R1
Dnov	XP_058132895.1	Fox	Fox-O1	Lcat	XP_045412286.1	Fox	Fox-R1
Ttru	XP_019792245.1	Fox	Fox-O1	Dnov	XP_023443955.1	Fox	Fox-R1
Oana	XP_001512968.3	Fox	Fox-O1	Amel	XP_034521812.1	Fox	Fox-R1
Hsap	NP_002006.2	Fox	Fox-O1	Hamp	XP_057603312.1	Fox	Fox-R1
Mang	XP_045725787.1	Fox	Fox-O1	Tman	XP_004385753.1	Fox	Fox-R1
Bmus	XP_036687548.1	Fox	Fox-O1	Oafe	XP_007934779.1	Fox	Fox-R1
Amel	XP_034519677.1	Fox	Fox-O1	Clup	XP_038391436.1	Fox	Fox-R1
Scar	XP_047409558.1	Fox	Fox-O1	Scar	XP_047373976.1	Fox	Fox-R1
Mmus	NP_062713.2	Fox	Fox-O1	Opri	XP_058519782.1	Fox	Fox-R1
Opri	XP_012782130.2	Fox	Fox-O1	Ttru	XP_033717853.1	Fox	Fox-R1
Cdid	XP_037656524.1	Fox	Fox-O1	Oafe	XP_007947037.1	Fox	Fox-R1
Mjav	XP_036867814.1	Fox	Fox-O1	Mang	XP_045748841.1	Fox	Fox-R1
Cdro	XP_031321792.1	Fox	Fox-O1	Cimi	XP_017378961.1	Fox	Fox-R1
Lcat	XP_045423168.1	Fox	Fox-O1	Lcat	XP_045383006.1	Fox	Fox-K2
Clup	XP_038429080.1	Fox	Fox-O1	Casi	XP_006869677.1	Fox	Fox-K2
Cpor	XP_023416198.1	Fox	Fox-O1	Mmus	XP_011247520.1	Fox	Fox-K2
Drot	XP_024424320.2	Fox	Fox-O1	Drot	XP_024409582.3	Fox	Fox-K2
Mdom	XP_001368312.2	Fox	Fox-O1	Cpor	XP_013002484.1	Fox	Fox-K2
Shar	XP_003764601.3	Fox	Fox-O1	Tman	XP_023584943.1	Fox	Fox-K2
Csim	XP_004443237.1	Fox	Fox-O1	Scar	XP_047400835.1	Fox	Fox-K2
Tman	XP_023597991.1	Fox	Fox-O1	Opri	XP_058532091.1	Fox	Fox-K2
Emax	XP_049709381.1	Fox	Fox-O1	Oafe	XP_007957978.1	Fox	Fox-K2
Pafr	XP_047611907.1	Fox	Fox-O1	Hsap	NP_004505.2	Fox	Fox-K2
Bbub	XP_006065837.2	Fox	Fox-O1	Emax	XP_049715130.1	Fox	Fox-K2
Ptig	XP_007097795.2	Fox	Fox-O1	Cimi	XP_017380207.1	Fox	Fox-K2
Hamp	XP_057563333.1	Fox	Fox-O1	Hamp	XP_057574401.1	Fox	N/A
Oafe	XP_007943433.1	Fox	Fox-O1	Hamp	XP_057571581.1	Fox	N/A
Mang	XP_045725786.1	Fox	Fox-O1	Lcat	XP_045397194.1	Fox	Fox-K1
Ggal	XP_015134143.3	Fox	Fox-O4	Bmus	XP_036681953.1	Fox	Fox-K1
Casi	XP_006868595.1	Fox	Fox-O4	Pgig	XP_039734750.1	Fox	Fox-K1
Mdom	XP_056665439.1	Fox	Fox-O4	Equa	XP_046521295.1	Fox	Fox-K1
Csim	XP_004439923.1	Fox	Fox-O4	Hsap	NP_001032242.1	Fox	Fox-K1
Cdid	XP_037677588.1	Fox	Fox-O4	Emax	XP_049758808.1	Fox	Fox-K1
Hsap	NP_005929.2	Fox	Fox-O4	Oana	XP_028905334.1	Fox	Fox-K1
Emax	XP_049727631.1	Fox	Fox-O4	Opri	XP_012786071.2	Fox	Fox-K1
Ptig	XP_007099158.2	Fox	Fox-O4	Cimi	XP_017389267.1	Fox	Fox-K1
Drot	XP_053773393.1	Fox	Fox-O4	Drot	XP_053782069.1	Fox	Fox-K1
Bmus	XP_036696517.1	Fox	Fox-O4	Cpor	XP_0234240906.1	Fox	Fox-K1
Oafe	XP_007957073.1	Fox	Fox-O4	Dnov	XP_004454003.2	Fox	Fox-K1
Cimi	XP_017372253.1	Fox	Fox-O4	Scar	XP_047389516.1	Fox	Fox-K1
Mjav	XP_017525885.1	Fox	Fox-O4	Csim	XP_004440941.1	Fox	Fox-K1
Rfer	XP_032969557.1	Fox	Fox-O4	Mmus	NP_951031.2	Fox	Fox-K1
Pafr	XP_047621021.1	Fox	Fox-O4	Shar	XP_031796898.1	Fox	Fox-K1
Mmus	NP_061259.1	Fox	Fox-O4	Hamp	XP_057604533.1	Fox	Fox-K1
Mang	XP_045735285.1	Fox	Fox-O4	Mang	XP_045731846.1	Fox	Fox-K1
Pgig	XP_039697652.1	Fox	Fox-O4	Casi	XP_006859933.1	Fox	Fox-K1
Scar	XP_047393133.1	Fox	Fox-O4	Clup	XP_038395513.1	Fox	Fox-K1
Oana	XP_039768253.1	Fox	Fox-O4	Cdid	XP_037669880.1	Fox	Fox-K1
Dnov	XP_058146903.1	Fox	Fox-O4	Pafr	XP_047635407.1	Fox	Fox-K1
Cdid	XP_037678715.1	Fox	Fox-O4	Tman	XP_004380896.3	Fox	Fox-K1
Amel	XP_002929104.1	Fox	Fox-O4	Amel	XP_011219634.2	Fox	Fox-K1
Tman	XP_023590438.1	Fox	Fox-O4	Mjav	XP_036866572.1	Fox	Fox-K1
Clup	XP_038443581.1	Fox	Fox-O4	Ttru	XP_019802935.2	Fox	Fox-K1
Cpor	XP_013009181.1	Fox	Fox-O4	Cdro	XP_031327189.1	Fox	Fox-K1
Lcat	XP_045394627.1	Fox	Fox-O4	Mdom	XP_056662440.1	Fox	Fox-K1
Hamp	XP_057575075.1	Fox	Fox-O4	Ptig	XP_042828249.1	Fox	Fox-K1
Bbub	XP_006076443.1	Fox	Fox-O4	Rfer	XP_032951743.1	Fox	Fox-K1
Ttru	XP_033705152.1	Fox	Fox-O4	Oafe	XP_007941047.1	Fox	Fox-K1
Equa	XP_046529856.1	Fox	Fox-O4	Ggal	XP_015149844.1	Fox	Fox-K1
Opri	XP_004595290.1	Fox	Fox-O4	Bbub	XP_025131152.1	Fox	Fox-K1
Drot	XP_053773548.1	Fox	Fox-O4	Hsap	NP_001400854.1	Fox	Fox-M1
Cdro	XP_010977055.2	Fox	Fox-O4	Lcat	XP_045410542.1	Fox	N/A
Pgig	XP_039724773.1	Fox	Fox-J3	Ggal	XP_046760565.1	Fox	N/A
Mdom	XP_007492986.2	Fox	Fox-J3	Oana	XP_028912611.1	Fox	Fox-K2
Lcat	XP_045401678.1	Fox	Fox-J3	Clup	XP_038402133.1	Fox	N/A
Shar	XP_031818195.1	Fox	Fox-J3	Amel	XP_019662714.1	Fox	N/A
Equa	XP_046517148.1	Fox	Fox-J3	Mang	XP_045744166.1	Fox	N/A
Mjav	XP_036858181.1	Fox	Fox-J3	Equa	XP_046532308.1	Fox	N/A
Oafe	XP_007953745.1	Fox	Fox-J3	Ptig	XP_042846258.1	Fox	N/A
Mmus	NP_766287.1	Fox	Fox-J3	Oafe	XP_007942676.1	Fox	N/A
Drot	XP_024410422.2	Fox	Fox-J3	Casi	XP_006839645.1	Fox	N/A
Tman	XP_023582916.1	Fox	Fox-J3	Mmus	XP_011246143.1	Fox	Fox-R2
Dnov	XP_004477472.1	Fox	Fox-J3	Opri	XP_058514738.1	Fox	N/A
Hsap	XP_005270689.1	Fox	Fox-J3	Scar	XP_047391089.1	Fox	N/A
Csim	XP_004443185.1	Fox	Fox-J3	Ggal	XP_015137672.3	Fox	N/A
Scar	XP_047408133.1	Fox	Fox-J3	Ggal	XP_015141419.2	Fox	N/A
Cimi	XP_017397267.1	Fox	Fox-J3	Ggal	XP_015144532.3	Fox	N/A
Bmus	XP_036706277.1	Fox	Fox-J3	Ggal	XP_015136219.2	Fox	N/A
Cdid	XP_037683454.1	Fox	Fox-J3	Cpor	XP_004999985.1	Fox	N/A
Cpor	XP_004999975.1	Fox	Fox-J3	Cdid	XP_037667673.1	Sox	Sox-18
Oana	XP_028921226.1	Fox	Fox-J3	Opri	XP_004586020.2	Sox	Sox-18
Bbub	XP_044801100.1	Fox	Fox-J3	Clup	XP_038427314.1	Sox	Sox-18
Ttru	XP_019793563.1	Fox	Fox-J3	Scar	XP_047396502.1	Sox	Sox-18
Casi	XP_006839889.1	Fox	Fox-J3	Cpor	XP_003462830.1	Sox	Sox-18
Clup	XP_038413618.1	Fox	Fox-J3	Mjav	XP_036862488.1	Sox	Sox-18
Emax	XP_049734920.1	Fox	Fox-J3	Bmus	XP_036681681.1	Sox	Sox-18
Pafr	XP_047645481.1	Fox	Fox-J3	Cdro	XP_031289217.1	Sox	Sox-18
Mang	XP_054368911.1	Fox	Fox-J3	Lcat	XP_045384104.1	Sox	Sox-18
Opri	XP_058535693.1	Fox	Fox-J3	Cimi	XP_017371142.1	Sox	Sox-18
Amel	XP_011233273.1	Fox	Fox-J3	Equa	XP_046535409.1	Sox	Sox-18
Hamp	XP_057563297.1	Fox	Fox-J3	Bbub	XP_025119695.1	Sox	Sox-18
Ptig	XP_042852753.1	Fox	Fox-J3	Csim	XP_004430565.1	Sox	Sox-18
Cdro	XP_031320784.1	Fox	Fox-J3	Rfer	XP_032951103.1	Sox	Sox-18

Table S8 continued from previous page

Species ID	Gene ID	Group	Annotation	Species ID	Gene ID	Group	Annotation
Oana	XP_028910333.1	Fox	Fox-J2	Drot	XP_053783310.1	Sox	Sox-18
Hamp	XP_057559090.1	Fox	Fox-J2	Pgig	XP_039720817.1	Sox	Sox-18
Pafr	XP_047643563.1	Fox	Fox-J2	Ttru	XP_019806802.1	Sox	Sox-18
Ttru	XP_03372947.1	Fox	Fox-J2	Pafr	XP_047626512.1	Sox	Sox-18
Pgig	XP_039730188.1	Fox	Fox-J2	Mang	XP_045742130.1	Sox	Sox-18
Hsap	XP_011519063.1	Fox	Fox-J2	Mmus	NP_033262.2	Sox	Sox-18
Equa	XP_046526439.1	Fox	Fox-J2	Ptig	XP_042835603.1	Sox	Sox-18
Shar	XP_031794189.1	Fox	Fox-J2	Hsap	NP_060889.1	Sox	Sox-18
Scar	XP_047406645.1	Fox	Fox-J2	Hamp	XP_057559896.1	Sox	Sox-18
Cimi	XP_017367793.1	Fox	Fox-J2	Dnov	XP_058143328.1	Sox	Sox-18
Csim	XP_004438721.1	Fox	Fox-J2	Shar	XP_023350160.2	Sox	Sox-18
Emax	XP_049738321.1	Fox	Fox-J2	Oana	XP_028925867.1	Sox	Sox-18
Tman	XP_004387308.1	Fox	Fox-J2	Mdom	XP_007475557.3	Sox	Sox-18
Bmus	XP_036722276.1	Fox	Fox-J2	Ggal	NP_989640.1	Sox	Sox-18
Dnov	XP_004455323.1	Fox	Fox-J2	Ggal	XP_040551910.1	Sox	Sox-17
Mmus	NP_068699.1	Fox	Fox-J2	Opri	XP_004580649.2	Sox	Sox-17
Bbub	XP_0251138942.2	Fox	Fox-J2	Ggal	NP_001034415.2	Sox	Sox-17
Casi	XP_006875525.1	Fox	Fox-J2	Mdom	XP_001379706.1	Sox	Sox-17
Mjav	XP_017507797.1	Fox	Fox-J2	Cdid	XP_037658799.1	Sox	Sox-17
Drot	XP_053777818.1	Fox	Fox-J2	Lcat	XP_045416951.1	Sox	Sox-17
Mdom	XP_001364301.2	Fox	Fox-J2	Mmus	NP_001276393.1	Sox	Sox-17
Rfer	XP_032973158.1	Fox	Fox-J2	Tman	XP_004372647.1	Sox	Sox-17
Cdro	XP_031300018.1	Fox	Fox-J2	Rfer	XP_032981648.1	Sox	Sox-17
Opri	XP_004596477.2	Fox	Fox-J2	Cpor	XP_005002064.1	Sox	Sox-17
Cpor	XP_003470329.1	Fox	Fox-J2	Oafe	XP_007950410.1	Sox	Sox-17
Ptig	XP_042847649.1	Fox	Fox-J2	Oana	XP_028925260.1	Sox	Sox-17
Ggal	XP_046759046.1	Fox	Fox-J2	Mdom	XP_001368625.2	Sox	Sox-17
Amel	XP_034501007.1	Fox	Fox-J2	Amel	XP_002925690.2	Sox	Sox-17
Oafe	XP_007935347.1	Fox	Fox-J2	Scar	XP_047417202.1	Sox	Sox-17
Mang	XP_045729866.1	Fox	Fox-J2	Hsap	NP_071899.1	Sox	Sox-17
Lcat	XP_045410352.1	Fox	Fox-J2	Cimi	XP_017387282.1	Sox	Sox-17
Cdid	XP_037701021.1	Fox	Fox-J2	Mjav	XP_036855229.1	Sox	Sox-17
Clup	XP_038432619.1	Fox	Fox-J2	Shar	XP_003759826.2	Sox	Sox-17
Opri	XP_004592944.1	Fox	Fox-J1	Bbub	XP_025121191.1	Sox	Sox-17
Bmus	XP_036692694.1	Fox	Fox-J1	Casi	XP_006871910.1	Sox	Sox-17
Ttru	XP_033703524.1	Fox	Fox-J1	Shar	XP_031802254.1	Sox	Sox-17
Mdom	XP_016286122.1	Fox	Fox-J1	Clup	XP_038435168.1	Sox	Sox-17
Pafr	XP_047614731.1	Fox	Fox-J1	Csim	XP_004435766.1	Sox	Sox-17
Mjav	XP_017529912.2	Fox	Fox-J1	Emax	XP_049716401.1	Sox	Sox-17
Cpor	XP_003461484.1	Fox	Fox-J1	Cdro	XP_010987841.2	Sox	Sox-17
Mmus	XP_006532332.1	Fox	Fox-J1	Ptig	XP_042829097.1	Sox	Sox-17
Ptig	XP_042823405.1	Fox	Fox-J1	Bmus	XP_036685683.1	Sox	Sox-17
Equa	XP_046530684.1	Fox	Fox-J1	Pgig	XP_039724865.1	Sox	Sox-17
Clup	XP_038402010.1	Fox	Fox-J1	Mang	XP_045752504.1	Sox	Sox-17
Emax	XP_049715791.1	Fox	Fox-J1	Pafr	XP_047639872.1	Sox	Sox-17
Oana	XP_028912588.1	Fox	Fox-J1	Hamp	XP_057592218.1	Sox	Sox-17
Tman	XP_004374194.1	Fox	Fox-J1	Equa	XP_046498273.1	Sox	Sox-17
Drot	XP_045046069.2	Fox	Fox-J1	Dnov	XP_058131520.1	Sox	Sox-17
Hsap	NP_001445.2	Fox	Fox-J1	Drot	XP_024434049.3	Sox	Sox-17
Pgig	XP_039715456.1	Fox	Fox-J1	Ttru	XP_019780997.1	Sox	Sox-17
Cimi	XP_017391310.1	Fox	Fox-J1	Oana	XP_028909114.1	Sox	Sox-7
Rfer	XP_032946522.1	Fox	Fox-J1	Bbub	XP_006064379.2	Sox	Sox-7
Lcat	XP_045381684.1	Fox	Fox-J1	Csim	XP_014645382.1	Sox	Sox-7
Dnov	XP_004454510.2	Fox	Fox-J1	Mdom	XP_001373591.1	Sox	Sox-7
Ggal	NP_001308464.2	Fox	Fox-J1	Casi	XP_006864394.1	Sox	Sox-7
Hamp	XP_057572756.1	Fox	Fox-J1	Ptig	XP_042839510.1	Sox	Sox-7
Cdro	XP_031325009.1	Fox	Fox-J1	Tman	XP_004382257.1	Sox	Sox-7
Oafe	XP_007957820.1	Fox	Fox-J1	Clup	XP_038429590.1	Sox	Sox-7
Amel	XP_034497541.1	Fox	Fox-J1	Pgig	XP_039731999.1	Sox	Sox-7
Cdid	XP_037665215.1	Fox	Fox-J1	Ggal	XP_046795199.1	Sox	Sox-7
Shar	XP_031821724.1	Fox	Fox-J1	Cimi	XP_017395944.1	Sox	Sox-7
Scar	XP_047399365.1	Fox	Fox-J1	Mmus	NP_035576.1	Sox	Sox-7
Csim	XP_004432853.1	Fox	Fox-J1	Emax	XP_049722562.1	Sox	Sox-7
Casi	XP_006869617.1	Fox	Fox-J1	Amel	XP_011214939.2	Sox	Sox-7
Bbub	XP_006045356.1	Fox	Fox-J1	Rfer	XP_032989201.1	Sox	Sox-7
Mang	XP_045743961.1	Fox	Fox-J1	Equa	XP_046512534.1	Sox	Sox-7
Oana	XP_039768325.1	Fox	Fox-P3	Ttru	XP_033713462.1	Sox	Sox-7
Pgig	XP_039699607.1	Fox	Fox-P3	Cpor	XP_003479745.1	Sox	Sox-7
Csim	XP_014646458.1	Fox	Fox-P3	Shar	XP_003758123.1	Sox	Sox-7
Equa	XP_046528495.1	Fox	Fox-P3	Cdid	XP_037668898.1	Sox	Sox-7
Cpor	XP_023417629.1	Fox	Fox-P3	Hamp	XP_057576011.1	Sox	Sox-7
Rfer	XP_032973683.1	Fox	Fox-P3	Bmus	XP_036712190.1	Sox	Sox-7
Oafe	XP_007956564.1	Fox	Fox-P3	Scar	XP_047405527.1	Sox	Sox-7
Shar	XP_023363009.2	Fox	Fox-P3	Drot	XP_053767954.1	Sox	Sox-7
Bmus	XP_036695836.1	Fox	Fox-P3	Pafr	XP_047615897.1	Sox	Sox-7
Mdom	XP_056665055.1	Fox	Fox-P3	Hsap	NP_113627.1	Sox	Sox-7
Cimi	XP_037589291.1	Fox	Fox-P3	Mang	XP_045736052.1	Sox	Sox-7
Emax	XP_049728302.1	Fox	Fox-P3	Dnov	XP_058144023.1	Sox	Sox-7
Drot	XP_053773440.1	Fox	Fox-P3	Cdro	XP_031297983.1	Sox	Sox-7
Mjav	XP_036852243.1	Fox	Fox-P3	Lcat	XP_045391548.1	Sox	Sox-7
Mmus	NP_001186276.1	Fox	Fox-P3	Opri	XP_004579156.2	Sox	Sox-7
Ttru	XP_033704944.1	Fox	Fox-P3	Oafe	XP_007936906.1	Sox	Sox-7
Cdro	XP_037677418.1	Fox	Fox-P3	Mjav	XP_036854859.1	Sox	Sox-7
Bbub	XP_006073707.2	Fox	Fox-P3	Pafr	XP_047609577.1	Sox	Sox-30
Casi	XP_006876748.1	Fox	Fox-P3	Ptig	XP_007077656.3	Sox	Sox-30
Scar	XP_047391408.1	Fox	Fox-P3	Bmus	XP_036704125.1	Sox	Sox-30
Hamp	XP_057574707.1	Fox	Fox-P3	Oana	XP_039766108.1	Sox	Sox-30
Lcat	XP_045393738.1	Fox	Fox-P3	Cimi	XP_017372578.1	Sox	Sox-30
Pafr	XP_047620716.1	Fox	Fox-P3	Oafe	XP_007948121.1	Sox	Sox-30
Amel	XP_011221377.1	Fox	Fox-P3	Cdro	XP_031305277.1	Sox	Sox-30
Dnov	XP_023446473.1	Fox	Fox-P3	Emax	XP_049730174.1	Sox	Sox-30
Ptig	XP_007096061.2	Fox	Fox-P3	Ttru	XP_057558120.1	Sox	Sox-30
Opri	XP_058514303.1	Fox	Fox-P3	Lcat	XP_045408644.1	Sox	Sox-30
Clup	NP_001161933.1	Fox	Fox-P3	Shar	XP_031809461.1	Sox	Sox-30
Mang	XP_045729506.1	Fox	Fox-P3	Bbub	XP_006058618.4	Sox	Sox-30
Cdid	XP_037677000.1	Fox	Fox-P3	Drot	XP_024432555.2	Sox	Sox-30
Hsap	NP_054728.2	Fox	Fox-P3	Clup	XP_038390830.1	Sox	Sox-30
Cdro	XP_031301084.1	Fox	Fox-P3	Cpor	XP_003473352.1	Sox	Sox-30
Tman	XP_004376890.1	Fox	Fox-P3	Ttru	XP_019780387.1	Sox	Sox-30
Oana	XP_028909446.1	Fox	Fox-F2	Opri	XP_004587550.2	Sox	Sox-30
Oana	XP_028912286.1	Fox	Fox-F2	Mdom	XP_001379720.1	Sox	Sox-30
Opri	XP_012786079.2	Fox	Fox-L3	Equa	XP_046521338.1	Sox	Sox-30
Mdom	XP_007498348.2	Fox	Fox-L3	Csim	XP_004428646.1	Sox	Sox-30

Table S8 continued from previous page

Species ID	Gene ID	Group	Annotation	Species ID	Gene ID	Group	Annotation
Drot	XP_045040885.2	Fox	Fox-L3	Cdid	XP_037655441.1	Sox	Sox-30
Rfer	XP_032952385.1	Fox	Fox-L3	Ggal	XP_414564.1	Sox	Sox-30
Scar	XP_047390130.1	Fox	Fox-L3	Dnov	XP_058162315.1	Sox	Sox-30
Oana	XP_001511921.2	Fox	Fox-L3	Scar	XP_047412799.1	Sox	Sox-30
Ttru	XP_004328913.2	Fox	Fox-L3	Mmus	NP_775560.1	Sox	Sox-30
Dnov	XP_058141662.1	Fox	Fox-L3	Amel	XP_019650665.2	Sox	Sox-30
Mmus	NP_001182057.1	Fox	Fox-L3	Pgig	XP_039732492.1	Sox	Sox-30
Hsap	NP_001361767.1	Fox	Fox-L3	Mjav	XP_036881068.1	Sox	Sox-30
Emax	XP_049761077.1	Fox	Fox-L3	Mang	XP_045743539.1	Sox	Sox-30
Mang	XP_045728168.1	Fox	Fox-L3	Rfer	XP_032952708.1	Sox	Sox-30
Mjav	XP_036864912.1	Fox	Fox-L3	Tman	XP_004371314.1	Sox	Sox-30
Shar	XP_012398242.1	Fox	Fox-L3	Hsap	NP_848511.1	Sox	Sox-30
Cimi	XP_017363768.1	Fox	Fox-L3	Casi	XP_0068633994.1	Sox	Sox-30
Hamp	XP_057605555.1	Fox	Fox-L3	Amel	XP_034508949.1	Sox	Sox-30
Pafn	XP_047635063.1	Fox	Fox-L3	Scar	XP_047389638.1	Sox	Sox-8
Tman	XP_004386032.1	Fox	Fox-L3	Bmus	XP_036681892.1	Sox	Sox-8
Csim	XP_004440991.1	Fox	Fox-L3	Pafn	XP_047635134.1	Sox	Sox-8
Equa	XP_046523569.1	Fox	Fox-L3	Drot	XP_024409185.2	Sox	Sox-8
Lcat	XP_045398087.1	Fox	Fox-L3	Mjav	XP_017502547.2	Sox	Sox-8
Cpor	XP_003469985.1	Fox	Fox-L3	Clup	XP_038394815.1	Sox	Sox-8
Bmus	XP_036680082.1	Fox	Fox-L3	Rfer	XP_032957935.1	Sox	Sox-8
Bbub	XP_044792209.1	Fox	Fox-L3	Equa	XP_046523802.1	Sox	Sox-8
Clup	XP_038394791.1	Fox	Fox-L3	Amel	XP_034525488.1	Sox	Sox-8
Casi	XP_006859950.1	Fox	Fox-L3	Cimi	XP_017366444.1	Sox	Sox-8
Cdid	XP_037669698.1	Fox	Fox-L3	Csim	XP_004438275.1	Sox	Sox-8
Ptig	XP_042828254.1	Fox	Fox-L3	Bbub	XP_025130606.3	Sox	Sox-8
Ggal	XP_425229.3	Fox	Fox-L3	Lcat	XP_045397007.1	Sox	Sox-8
Cdro	XP_031327057.1	Fox	Fox-L3	Hamp	XP_057551434.1	Sox	Sox-8
Amel	XP_034525410.1	Fox	Fox-L3	Pgig	XP_039740781.1	Sox	Sox-8
Oafe	XP_007941358.1	Fox	Fox-L3	Hsap	NP_055402.2	Sox	Sox-8
Cdro	XP_031326970.1	Fox	Fox-P1	Mang	XP_045730580.1	Sox	Sox-8
Oafe	XP_007944266.1	Fox	Fox-P1	Cdro	XP_010989436.2	Sox	Sox-8
Ttru	XP_033721167.1	Fox	Fox-P1	Ptig	XP_042827523.1	Sox	Sox-8
Amel	XP_011217264.1	Fox	Fox-P1	Ttru	XP_033696379.1	Sox	Sox-8
Mdom	XP_007500148.1	Fox	Fox-P1	Bmus	XP_036722097.1	Sox	N/A
Opri	XP_058534843.1	Fox	Fox-P1	Ttru	XP_019791993.1	Sox	N/A
Shar	XP_031801298.1	Fox	Fox-P1	Oana	XP_028908350.1	Sox	Sox-4
Pafn	XP_047635064.1	Fox	Fox-P1	Opri	XP_058523457.1	Sox	Sox-4
Mjav	XP_036875087.1	Fox	Fox-P1	Ggal	NP_989815.2	Sox	Sox-4
Hsap	NP_001231739.1	Fox	Fox-P1	Casi	XP_006860848.1	Sox	Sox-4
Casi	XP_006874684.1	Fox	Fox-P1	Lcat	XP_045407834.1	Sox	Sox-4
Emax	XP_049718943.1	Fox	Fox-P1	Ttru	XP_033719780.1	Sox	Sox-4
Mjav	XP_036860335.1	Fox	Fox-P1	Cpor	XP_003468908.2	Sox	Sox-4
Pgig	XP_039707610.1	Fox	Fox-P1	Mdom	XP_007487941.1	Sox	Sox-4
Drot	XP_045048358.1	Fox	Fox-P1	Scar	XP_047414294.1	Sox	Sox-4
Mmus	XP_030110934.1	Fox	Fox-P1	Hsap	NP_003098.1	Sox	Sox-4
Tman	XP_023583661.1	Fox	Fox-P1	Bbub	XP_025121975.3	Sox	Sox-4
Cimi	XP_017398291.1	Fox	Fox-P1	Hamp	XP_057555091.1	Sox	Sox-4
Rfer	XP_032987798.1	Fox	Fox-P1	Mmus	NP_033264.2	Sox	Sox-4
Scar	XP_047390767.1	Fox	Fox-P1	Drot	XP_053776704.1	Sox	Sox-4
Ptig	XP_042835351.1	Fox	Fox-P1	Oafe	XP_007950280.1	Sox	Sox-4
Oana	XP_039766065.1	Fox	Fox-P1	Dnov	XP_012385001.3	Sox	Sox-4
Equa	XP_046523540.1	Fox	Fox-P1	Clup	XP_038440195.1	Sox	Sox-4
Clup	XP_038421798.1	Fox	Fox-P1	Equa	XP_046497050.1	Sox	Sox-4
Mang	XP_054360611.1	Fox	Fox-P1	Amel	XP_034516148.1	Sox	Sox-4
Bmus	XP_036724195.1	Fox	Fox-P1	Cdro	XP_031290990.1	Sox	Sox-4
Cpor	XP_005005313.1	Fox	Fox-P1	Emax	XP_049732340.1	Sox	Sox-4
Cdid	XP_037656169.1	Fox	Fox-P1	Csim	XP_004432138.1	Sox	Sox-4
Hamp	XP_057561364.1	Fox	Fox-P1	Mjav	XP_036850086.1	Sox	Sox-4
Lcat	XP_045385591.1	Fox	Fox-P1	Pgig	XP_039722376.1	Sox	Sox-4
Bbub	XP_044789782.1	Fox	Fox-P1	Ptig	XP_042841543.1	Sox	Sox-4
Csim	XP_004419976.1	Fox	Fox-P1	Bmus	XP_036725088.1	Sox	Sox-4
Dnov	XP_004482088.1	Fox	Fox-P1	Tman	XP_023594118.1	Sox	Sox-4
Ggal	XP_040520196.1	Fox	Fox-P1	Pafr	XP_047651758.1	Sox	Sox-4
Tman	XP_004385433.1	Fox	Fox-N1	Mang	XP_045722257.1	Sox	Sox-4
Pafn	XP_047614767.1	Fox	Fox-N1	Rfer	XP_032971451.1	Sox	Sox-4
Cdid	XP_037665986.1	Fox	Fox-N1	Cdid	XP_037700378.1	Sox	Sox-4
Casi	XP_006874270.1	Fox	Fox-N1	Cimi	XP_017399572.1	Sox	Sox-4
Emax	XP_049715494.1	Fox	Fox-N1	Shar	XP_03180246.1	Sox	Sox-4
Scar	XP_047401997.1	Fox	Fox-N1	Shar	XP_003775357.4	Sox	Sox-15
Oafe	XP_007935559.1	Fox	Fox-N1	Dnov	XP_004468798.1	Sox	Sox-15
Ggal	XP_415816.5	Fox	Fox-N1	Mdom	XP_056673573.1	Sox	Sox-15
Mmus	XP_006532328.1	Fox	Fox-N1	Oana	XP_039766818.1	Sox	Sox-15
Oana	XP_039770471.1	Fox	Fox-N1	Clup	XP_038392636.1	Sox	Sox-15
Mdom	XP_001375832.1	Fox	Fox-N1	Mjav	XP_017527261.1	Sox	Sox-15
Dnov	XP_004450005.2	Fox	Fox-N1	Cdid	XP_037664941.1	Sox	Sox-15
Shar	XP_031823539.1	Fox	Fox-N1	Pgig	XP_039703166.1	Sox	Sox-15
Cpor	XP_023420823.1	Fox	Fox-N1	Pafr	XP_047613485.1	Sox	Sox-15
Lcat	XP_045381339.1	Fox	Fox-N1	Bmus	XP_036692236.1	Sox	Sox-15
Bbub	XP_025136790.1	Fox	Fox-N1	Ttru	XP_033704061.1	Sox	Sox-15
Opri	XP_004593965.2	Fox	Fox-N1	Emax	XP_049714749.1	Sox	Sox-15
Bmus	XP_036693663.1	Fox	Fox-N1	Opri	XP_004594843.2	Sox	Sox-15
Drot	XP_04505001.2	Fox	Fox-N1	Cpor	XP_0034662299.1	Sox	Sox-15
Amel	XP_034496649.1	Fox	Fox-N1	Rfer	XP_032947056.1	Sox	Sox-15
Pgig	XP_039742030.1	Fox	Fox-N1	Cimi	XP_037595583.1	Sox	Sox-15
Rfer	XP_032947344.1	Fox	Fox-N1	Hamp	XP_057571542.1	Sox	Sox-15
Mang	XP_045744088.1	Fox	Fox-N1	Equa	XP_046531541.1	Sox	Sox-15
Equa	XP_046533715.1	Fox	Fox-N1	Ptig	XP_042823564.1	Sox	Sox-15
Cimi	XP_037599212.1	Fox	Fox-N1	Lcat	XP_045426075.1	Sox	Sox-15
Mjav	XP_017515466.1	Fox	Fox-N1	Cdro	XP_031323594.1	Sox	Sox-15
Hamp	XP_057571196.1	Fox	Fox-N1	Mang	XP_045746055.1	Sox	Sox-15
Hsap	XP_011523660.1	Fox	Fox-N1	Drot	XP_024420518.1	Sox	Sox-15
Ptig	XP_007076110.1	Fox	Fox-N1	Tman	XP_004376270.1	Sox	Sox-15
Ttru	XP_033703421.1	Fox	Fox-N1	Bbub	XP_006062923.4	Sox	Sox-15
Cdro	XP_031324225.1	Fox	Fox-N1	Hsap	NP_008873.1	Sox	Sox-15
Csim	XP_014645264.1	Fox	Fox-N1	Oafe	XP_007950453.1	Sox	Sox-15
Clup	XP_038404162.1	Fox	Fox-N1	Scar	XP_047400352.1	Sox	Sox-15
Shar	XP_003760278.3	Fox	Fox-F2	Mmus	NP_033261.1	Sox	Sox-15
Drot	XP_024408147.2	Fox	Fox-F2	Csim	XP_004433180.1	Sox	Sox-15
Mdom	XP_001378770.3	Fox	Fox-F2	Amel	XP_019660673.2	Sox	Sox-15
Casi	XP_006870627.1	Fox	Fox-F2	Casi	XP_006863559.1	Sox	Sox-15
Equa	XP_046496582.1	Fox	Fox-F2	Rfer	XP_032974648.1	Sox	N/A
Bmus	XP_036725085.1	Fox	Fox-F2	Pgig	XP_039738421.1	Sox	N/A

Table S8 continued from previous page

Species ID	Gene ID	Group	Annotation	Species ID	Gene ID	Group	Annotation
Lcat	XP_045408674.1	Fox	Fox-F2	Dnov	XP_058138887.1	Sox	N/A
Opri	XP_012785618.3	Fox	Fox-F2	Drot	XP_024431130.1	Sox	N/A
Tman	XP_012414385.1	Fox	Fox-F2	Ggal	XP_015151955.3	Sox	Sox-12
Mjav	XP_036867662.1	Fox	Fox-F2	Equa	XP_046533915.1	Sox	Sox-12
Cdro	XP_010975130.2	Fox	Fox-F2	Emax	XP_049725681.1	Sox	Sox-12
Hsap	NP_001443.1	Fox	Fox-F2	Rfer	XP_032951289.1	Sox	Sox-12
Emax	XP_049738291.1	Fox	Fox-F2	Cpor	XP_005006612.1	Sox	Sox-12
Pgig	XP_039725405.1	Fox	Fox-F2	Bmus	XP_036683115.1	Sox	Sox-12
Csim	XP_004419240.1	Fox	Fox-F2	Tman	XP_004370660.1	Sox	Sox-12
Cdid	XP_037700041.1	Fox	Fox-F2	Bbub	XP_025119497.1	Sox	Sox-12
Dnov	XP_023447171.2	Fox	Fox-F2	Mang	XP_045741538.1	Sox	Sox-12
Cimi	XP_017352993.2	Fox	Fox-F2	Opri	XP_058535679.1	Sox	Sox-12
Oafe	XP_007933567.1	Fox	Fox-F2	Cdid	XP_037667390.1	Sox	Sox-12
Ptig	XP_042841202.1	Fox	Fox-F2	Pafr	XP_047627358.1	Sox	Sox-12
Mmus	NP_034355.2	Fox	Fox-F2	Mmus	NP_035568.1	Sox	Sox-12
Scar	XP_047414417.1	Fox	Fox-F2	Csim	XP_014645525.1	Sox	Sox-12
Bbub	XP_025120520.3	Fox	Fox-F2	Ttru	XP_004331924.3	Sox	Sox-12
Mang	XP_045721916.1	Fox	Fox-F2	Scar	XP_047398506.1	Sox	Sox-12
Amel	XP_034516095.1	Fox	Fox-F2	Ptig	XP_042836628.1	Sox	Sox-12
Rfer	XP_032969138.1	Fox	Fox-F2	Shar	XP_031810003.1	Sox	Sox-12
Clup	XP_038439937.1	Fox	Fox-F2	Pgig	XP_039719928.1	Sox	Sox-12
Ttru	XP_004311990.3	Fox	Fox-F2	Casi	XP_006860781.1	Sox	Sox-12
Pafr	XP_047651459.1	Fox	Fox-F2	Hsap	NP_008874.2	Sox	Sox-12
Hamp	XP_057555617.1	Fox	Fox-F2	Oafe	XP_007932707.1	Sox	Sox-12
Ggal	XP_046759228.1	Fox	N/A	Mjav	XP_017521282.2	Sox	Sox-12
Mdom	XP_007491677.2	Fox	N/A	Clup	XP_038427830.1	Sox	Sox-12
Shar	XP_031817418.1	Fox	N/A	Dnov	XP_058142829.1	Sox	Sox-12
Oana	XP_028911073.1	Fox	Fox-L2	Hamp	XP_057559844.1	Sox	Sox-12
Rfer	XP_032988626.1	Fox	Fox-L2	Lcat	XP_045385476.1	Sox	Sox-12
Ggal	NP_001012630.1	Fox	Fox-L2	Drot	XP_024415008.3	Sox	Sox-12
Mdom	XP_007494005.2	Fox	Fox-L2	Cdro	XP_03128988.1	Sox	Sox-12
Emax	XP_049726166.1	Fox	Fox-L2	Cimi	XP_017370646.1	Sox	Sox-12
Hsap	NP_075555.1	Fox	Fox-L2	Mdom	XP_007474545.2	Sox	Sox-12
Amel	XP_034517711.1	Fox	Fox-L2	Oana	XP_028907426.1	Sox	Sox-11
Drot	XP_024421665.2	Fox	Fox-L2	Cdro	XP_031322369.1	Sox	Sox-11
Mang	XP_045719675.1	Fox	Fox-L2	Pafr	XP_047635202.1	Sox	Sox-11
Tman	XP_004381516.1	Fox	Fox-L2	Cdid	XP_037669298.1	Sox	Sox-11
Casi	XP_006846679.1	Fox	Fox-L2	Mdom	XP_007476236.1	Sox	Sox-11
Csim	XP_004419996.1	Fox	Fox-L2	Dnov	XP_058143908.1	Sox	Sox-11
Mmus	NP_036150.1	Fox	Fox-L2	Mang	XP_045739658.1	Sox	Sox-11
Pafr	XP_047639332.1	Fox	Fox-L2	Casi	XP_006864036.1	Sox	Sox-11
Bbub	XP_025146742.1	Fox	Fox-L2	Tman	XP_004373138.1	Sox	Sox-11
Opri	XP_012783325.2	Fox	Fox-L2	Hamp	XP_057599420.1	Sox	Sox-11
Shar	XP_031815678.1	Fox	Fox-L2	Pgig	XP_039710941.1	Sox	Sox-11
Hamp	XP_057595697.1	Fox	Fox-L2	Cimi	XP_017397486.1	Sox	Sox-11
Cimi	XP_017379188.1	Fox	Fox-L2	Lcat	XP_045404807.1	Sox	Sox-11
Cdid	XP_037683851.1	Fox	Fox-L2	Emax	XP_049759749.1	Sox	Sox-11
Equa	XP_046508064.1	Fox	Fox-L2	Ptig	XP_042835747.1	Sox	Sox-11
Mjav	XP_036847828.1	Fox	Fox-L2	Shar	XP_031807046.1	Sox	Sox-11
Cdro	XP_031299319.1	Fox	Fox-L2	Cpor	XP_003464995.1	Sox	Sox-11
Lcat	XP_045411920.1	Fox	Fox-L2	Mjav	XP_017522349.2	Sox	Sox-11
Clup	XP_038426115.1	Fox	Fox-L2	Scar	XP_047379435.1	Sox	Sox-11
Cpor	XP_023415973.1	Fox	Fox-L2	Opri	XP_004582643.2	Sox	Sox-11
Ptig	XP_042812852.1	Fox	Fox-L2	Clup	XP_038416510.1	Sox	Sox-11
Oafe	XP_007945699.1	Fox	Fox-L2	Drot	XP_024408195.2	Sox	Sox-11
Ttru	XP_019789813.1	Fox	Fox-L2	Ggal	NP_001382976.1	Sox	Sox-11
Pgig	XP_039692762.1	Fox	Fox-L2	Equa	XP_046516078.1	Sox	Sox-11
Bmus	XP_036707130.1	Fox	Fox-L2	Amel	XP_034515849.1	Sox	Sox-11
Scar	XP_047420926.1	Fox	Fox-L2	Hsap	NP_003099.1	Sox	Sox-11
Oana	XP_028902621.2	Fox	Fox-L2	Oafe	XP_007933317.1	Sox	Sox-11
Bmus	XP_036729925.1	Fox	Fox-L2	Bmus	XP_036729557.1	Sox	Sox-11
Lcat	XP_045404844.1	Fox	Fox-L2	Bbub	XP_025117398.3	Sox	Sox-11
Casi	XP_006875916.1	Fox	Fox-L2	Rfer	XP_032980233.1	Sox	Sox-11
Dnov	XP_004472353.2	Fox	Fox-L2	Ttru	XP_019798147.1	Sox	Sox-11
Emax	XP_049713621.1	Fox	Fox-L2	Csim	XP_014647619.1	Sox	Sox-11
Csim	XP_004437309.2	Fox	Fox-L2	Mmus	NP_033260.4	Sox	Sox-11
Cimi	XP_017392139.1	Fox	Fox-L2	Ggal	NP_989526.1	Sry	
Bbub	XP_025116942.2	Fox	Fox-L2	Shar	XP_023362772.2	Sox	Sox-3
Pgig	XP_039708536.1	Fox	Fox-L2	Ptig	XP_042831086.1	Sox	Sox-3
Cdid	XP_037663497.1	Fox	Fox-L2	Dnov	XP_058147425.1	Sox	Sox-3
Tman	XP_004390712.1	Fox	Fox-L2	Rfer	XP_032973530.1	Sox	Sox-3
Clup	NP_001129118.1	Fox	Fox-L2	Oafe	XP_007949915.1	Sox	Sox-3
Opri	XP_004590794.2	Fox	Fox-L2	Csim	XP_014651689.1	Sox	Sox-3
Pafr	XP_047635149.1	Fox	Fox-L2	Drot	XP_024434131.1	Sry	
Hamp	XP_057599300.1	Fox	Fox-L2	Hsap	NP_003131.1	Sry	
Mang	XP_045739697.1	Fox	Fox-L2	Equa	XP_046528503.1	Sox	Sox-3
Amel	XP_002922377.3	Fox	Fox-L2	Emax	XP_049728984.1	Sox	Sox-3
Ttru	XP_033694400.1	Fox	Fox-L2	Opri	XP_058514980.1	Sox	Sox-3
Scar	XP_047393964.1	Fox	Fox-L2	Ttru	XP_033705620.1	Sox	Sox-3
Hsap	NP_001129121.1	Fox	Fox-L2	Clup	XP_038442551.1	Sox	Sox-3
Mmus	NP_001094934.1	Fox	Fox-L2	Oana	XP_001520781.2	Sry	
Rfer	XP_032980088.1	Fox	Fox-L2	Hsap	NP_005625.2	Sox	Sox-3
Ptig	XP_042837881.1	Fox	Fox-L2	Mjav	XP_036859504.1	Sox	Sox-3
Equa	XP_046517777.1	Fox	Fox-L2	Mdom	XP_056665898.1	Sry	
Cdro	XP_010996180.2	Fox	Fox-L2	Emax	XP_049729443.1	Sry	
Mjav	XP_017536768.2	Fox	Fox-L2	Cdro	XP_031300930.1	Sox	Sox-3
Oafe	XP_007952295.1	Fox	Fox-L2	Pgig	XP_039735027.1	Sry	
Drot	XP_024408151.2	Fox	Fox-L2	Shar	XP_031801149.1	Sry	
Oana	XP_028931929.2	Fox	Fox-L1	Lcat	XP_045393696.1	Sox	Sox-3
Bbub	XP_0060407558.1	Fox	Fox-L1	Cimi	XP_017374345.1	Sry	
Mdom	XP_056674514.1	Fox	Fox-L1	Hamp	XP_057574656.1	Sox	Sox-3
Oafe	XP_007937407.1	Fox	Fox-L1	Mmus	NP_033263.2	Sox	Sox-3
Mjav	XP_036849629.1	Fox	Fox-L1	Amel	XP_034504946.1	Sox	Sox-3
Rfer	XP_032984676.1	Fox	Fox-L1	Cdid	XP_037678811.1	Sox	Sox-3
Hamp	XP_057569166.1	Fox	Fox-L1	Bmus	XP_036696323.1	Sox	Sox-3
Pgig	XP_039744660.1	Fox	Fox-L1	Ttru	XP_033706224.1	Sry	
Dnov	XP_058135467.1	Fox	Fox-L1	Bbub	XP_025132502.3	Sox	Sox-3
Csim	XP_004437100.1	Fox	Fox-L1	Cdid	XP_037680520.1	Sry	
Casi	XP_006860180.1	Fox	Fox-L1	Cpor	XP_003464754.2	Sox	Sox-3
Opri	XP_004584130.2	Fox	Fox-L1	Scar	XP_047393176.1	Sry	
Clup	XP_038394040.1	Fox	Fox-L1	Opri	XP_058515261.1	Sry	
Hsap	NP_005241.1	Fox	Fox-L1	Pafr	XP_047620863.1	Sox	Sox-3
Drot	XP_024411663.1	Fox	Fox-L1	Mmus	NP_035694.1	Sry	

Table S8 continued from previous page

Species ID	Gene ID	Group	Annotation	Species ID	Gene ID	Group	Annotation
Amel	XP_002913428.1	Fox	Fox-L1	Drot	XP_053773489.1	Sox	Sox-3
Cdro	XP_031292484.1	Fox	Fox-L1	Mdom	XP_007507433.1	Sox	Sox-3
Cimi	XP_017399161.1	Fox	Fox-L1	Scar	XP_047392867.1	Sox	Sox-3
Cdid	XP_037671791.1	Fox	Fox-L1	Oafe	XP_007957280.1	Sox	Sry
Ttru	XP_004312143.1	Fox	Fox-L1	Cimi	XP_017355427.1	Sox	Sox-3
Rfer	XP_032984241.1	Fox	Fox-L1	Ttru	XP_033706225.1	Sox	Sry
Ptig	XP_042825075.1	Fox	Fox-L1	Mang	XP_045726266.1	Sox	Sox-3
Bmus	XP_036688240.1	Fox	Fox-L1	Casi	XP_006875033.1	Sox	Sox-3
Tman	XP_004377800.1	Fox	Fox-L1	Opri	XP_058515255.1	Sox	Sry
Mang	XP_045746246.1	Fox	Fox-L1	Mjav	XP_036857050.1	Sox	Sry
Shar	XP_031806032.1	Fox	Fox-L1	Dnov	XP_058142693.1	Sox	Sox-8
Mmus	NP_032050.2	Fox	Fox-L1	Casi	XP_006873911.1	Sox	Sox-8
Paf	XP_047648182.1	Fox	Fox-L1	Cdid	XP_0337670405.1	Sox	Sox-8
Equa	XP_046536521.1	Fox	Fox-L1	Emax	XP_049759218.1	Sox	Sox-8
Scar	XP_047385703.1	Fox	Fox-L1	Tman	XP_023584110.1	Sox	Sox-8
Ggal	XP_001231599.5	Fox	Fox-L1	Oafe	XP_007955685.1	Sox	Sox-8
Emax	XP_049720646.1	Fox	Fox-L1	Mjav	XP_017510601.1	Sox	N/A
Lcat	XP_045389091.1	Fox	Fox-L1	Clup	XP_038432983.1	Sox	N/A
Ggal	XP_004948470.2	Fox	Fox-P4	Ptig	XP_042847878.1	Sox	N/A
Casi	XP_006860505.1	Fox	Fox-P4	Mang	XP_054363126.1	Sox	N/A
Equa	XP_046496143.1	Fox	Fox-P4	Amel	XP_019650885.1	Sox	N/A
Cdid	XP_037698927.1	Fox	Fox-P4	Ggal	NP_989664.1	Sox	Sox-1
Mdom	XP_056674181.1	Fox	Fox-P4	Mjav	XP_036880873.1	Sox	Sox-1
Oafe	XP_007934458.1	Fox	Fox-P4	Amel	XP_034519706.1	Sox	Sox-1
Opri	XP_058519292.1	Fox	Fox-P4	Lcat	XP_045422689.1	Sox	Sox-1
Mmus	XP_017173175.1	Fox	Fox-P4	Oana	XP_028904097.1	Sox	Sox-1
Rfer	XP_032956861.1	Fox	Fox-P4	Opri	XP_058526612.1	Sox	Sox-1
Hsap	XP_011512591.1	Fox	Fox-P4	Cdid	XP_037656244.1	Sox	Sox-1
Ptig	XP_042842255.1	Fox	Fox-P4	Scar	XP_047410448.1	Sox	Sox-1
Ttru	XP_033720169.1	Fox	Fox-P4	Mdom	XP_007501341.1	Sox	Sox-1
Mang	XP_045722826.1	Fox	Fox-P4	Equa	XP_046519750.1	Sox	Sox-1
Mjav	XP_036861021.1	Fox	Fox-P4	Ttru	XP_019776483.2	Sox	Sox-1
Cimi	XP_017402725.1	Fox	Fox-P4	Clup	XP_038425993.1	Sox	Sox-1
Clup	XP_038409938.1	Fox	Fox-P4	Cdro	XP_031322193.1	Sox	Sox-1
Shar	XP_031820748.1	Fox	Fox-P4	Shar	XP_031814590.1	Sox	Sox-1
Lcat	XP_045397843.1	Fox	Fox-P4	Oafe	XP_007940578.1	Sox	Sox-1
Pgig	XP_039710787.1	Fox	Fox-P4	Rfer	XP_032959657.1	Sox	Sox-1
Drot	XP_053769706.1	Fox	Fox-P4	Mmus	NP_033259.2	Sox	Sox-1
Bmus	XP_036725368.1	Fox	Fox-P4	Dnov	XP_058132394.1	Sox	Sox-1
Amel	XP_002914563.3	Fox	Fox-P4	Bmus	XP_036687021.1	Sox	Sox-1
Csim	XP_004424223.1	Fox	Fox-P4	Paf	XP_047612043.1	Sox	Sox-1
Oana	XP_028925009.1	Fox	Fox-P4	Drot	XP_045042640.2	Sox	Sox-1
Hamp	XP_057555834.1	Fox	Fox-P4	Ptig	XP_042814197.1	Sox	Sox-1
Paf	XP_047653356.1	Fox	Fox-P4	Cimi	XP_017357713.1	Sox	Sox-1
Cpor	XP_003473920.1	Fox	Fox-P4	Mang	XP_045723818.1	Sox	Sox-1
Scar	XP_047414507.1	Fox	Fox-P4	Casi	XP_006851579.1	Sox	Sox-1
Emax	XP_049747581.1	Fox	Fox-P4	Cpor	XP_005007254.3	Sox	Sox-1
Tman	XP_004379494.1	Fox	Fox-P4	Hsap	NP_005977.2	Sox	Sox-1
Dnov	XP_058162686.1	Fox	Fox-P4	Hamp	XP_057563730.1	Sox	Sox-1
Bbub	XP_025125991.2	Fox	Fox-P4	Pgig	XP_039721974.1	Sox	Sox-1
Cdro	XP_031290589.1	Fox	Fox-P4	Bbub	XP_025118363.1	Sox	Sox-1
Bmus	XP_036718879.1	Fox	Fox-P2	Emax	XP_049723386.1	Sox	Sox-1
Ttru	XP_019780690.1	Fox	Fox-P2	Mdom	XP_001373727.1	Sox	N/A
Lcat	XP_045420440.1	Fox	Fox-P2	Shar	XP_003762017.1	Sox	N/A
Pgig	XP_039716577.1	Fox	Fox-P2	Hamp	XP_057557078.1	Sox	Sox-5
Csim	XP_004418850.1	Fox	Fox-P2	Equa	XP_046497249.1	Sox	Sox-5
Shar	XP_031794706.1	Fox	Fox-P2	Bbub	XP_045021534.1	Sox	Sox-5
Equa	XP_046526335.1	Fox	Fox-P2	Csim	XP_014646859.1	Sox	Sox-5
Cdro	XP_031311093.1	Fox	Fox-P2	Cdro	XP_031299822.1	Sox	Sox-5
Amel	XP_034526782.1	Fox	Fox-P2	Tman	XP_023593782.1	Sox	Sox-5
Ptig	XP_042826228.1	Fox	Fox-P2	Emax	XP_049741103.1	Sox	Sox-5
Opri	XP_058510875.1	Fox	Fox-P2	Oafe	XP_007944956.1	Sox	Sox-5
Cpor	XP_013013912.1	Fox	Fox-P2	Paf	XP_047643337.1	Sox	Sox-5
Casi	XP_006859359.1	Fox	Fox-P2	Casi	XP_006867010.1	Sox	Sox-5
Oana	XP_028928988.1	Fox	Fox-P2	Mdom	XP_001368820.2	Sox	Sox-2
Paf	XP_047618861.1	Fox	Fox-P2	Csim	XP_004424739.1	Sox	Sox-2
Oafe	XP_007942246.1	Fox	Fox-P2	Emax	XP_049730807.1	Sox	Sox-2
Hsap	NP_683696.2	Fox	Fox-P2	Casi	XP_006869879.1	Sox	Sox-2
Mdom	XP_007504226.1	Fox	Fox-P2	Ptig	XP_042855191.1	Sox	Sox-2
Ggal	XP_0250007337.1	Fox	Fox-P2	Ggal	NP_990519.3	Sox	Sox-2
Dnov	XP_058153204.1	Fox	Fox-P2	Clup	XP_038439395.1	Sox	Sox-2
Bbub	XP_044802585.1	Fox	Fox-P2	Equa	XP_046515240.1	Sox	Sox-2
Cimi	XP_017390531.1	Fox	Fox-P2	Hamp	XP_057595873.1	Sox	Sox-2
Rfer	XP_032954948.1	Fox	Fox-P2	Hsap	NP_003097.1	Sox	Sox-2
Hamp	XP_057586867.1	Fox	Fox-P2	Shar	XP_031813587.1	Sox	Sox-2
Mmus	XP_036021645.1	Fox	Fox-P2	Mang	XP_045719633.1	Sox	Sox-2
Mjav	XP_036883368.1	Fox	Fox-P2	Scar	XP_047420710.1	Sox	Sox-2
Emax	XP_049750594.1	Fox	Fox-P2	Tman	XP_023592536.1	Sox	Sox-2
Cdid	XP_037692269.1	Fox	Fox-P2	Bbub	XP_006056297.2	Sox	Sox-2
Drot	XP_053782699.1	Fox	Fox-P2	Rfer	XP_032989610.1	Sox	Sox-2
Scar	XP_047417295.1	Fox	Fox-P2	Lcat	XP_045395708.1	Sox	Sox-2
Mang	XP_045720337.1	Fox	Fox-P2	Opri	XP_058518203.1	Sox	Sox-2
Tman	XP_004382624.1	Fox	Fox-P2	Oana	XP_028928807.1	Sox	Sox-2
Clup	XP_038413332.1	Fox	Fox-P2	Bmus	XP_036705414.1	Sox	Sox-2
Cdid	XP_037688024.1	Fox	N/A	Mmus	NP_035573.3	Sox	Sox-2
Dnov	XP_058148549.1	Fox	N/A	Cdid	XP_037695755.1	Sox	Sox-2
Ggal	XP_414186.5	Fox	Fox-F1	Pgig	XP_039706608.1	Sox	Sox-2
Emax	XP_049720665.1	Fox	Fox-F1	Oafe	XP_007934094.1	Sox	Sox-2
Mdom	XP_001365832.4	Fox	Fox-F1	Paf	XP_047643165.1	Sox	Sox-2
Hsap	NP_001442.2	Fox	Fox-F1	Dnov	XP_004477690.1	Sox	Sox-2
Clup	XP_038394042.1	Fox	Fox-F1	Cimi	XP_017393930.1	Sox	Sox-2
Mjav	XP_017511561.2	Fox	Fox-F1	Drot	XP_053774095.1	Sox	Sox-2
Cimi	XP_017391155.2	Fox	Fox-F1	Amel	XP_034518206.1	Sox	Sox-2
Bbub	XP_006047564.2	Fox	Fox-F1	Cdro	XP_010974669.2	Sox	Sox-2
Dnov	XP_058135594.1	Fox	Fox-F1	Mjav	XP_017496634.2	Sox	Sox-2
Drot	XP_053772163.1	Fox	Fox-F1	Ttru	XP_004311832.3	Sox	Sox-2
Tman	XP_004377797.2	Fox	Fox-F1	Oana	XP_028931165.1	Sox	Sox-14
Bmus	XP_036688072.1	Fox	Fox-F1	Mdom	XP_007493983.1	Sox	Sox-14
Oana	XP_028931561.1	Fox	Fox-F1	Scar	XP_047419052.1	Sox	Sox-14
Paf	XP_047645925.1	Fox	Fox-F1	Cdid	XP_037683894.1	Sox	Sox-14
Scar	XP_047385701.1	Fox	Fox-F1	Bmus	XP_036707689.1	Sox	Sox-14
Lcat	XP_045389195.1	Fox	Fox-F1	Mjav	XP_017531607.1	Sox	Sox-14
Shar	XP_003758519.2	Fox	Fox-F1	Casi	XP_006846671.1	Sox	Sox-14

Table S8 continued from previous page

Species ID	Gene ID	Group	Annotation	Species ID	Gene ID	Group	Annotation
Amel	XP_002913450.2	Fox	Fox-F1	Pafn	XP_047610665.1	Sox	Sox-14
Opri	XP_058530279.1	Fox	Fox-F1	Hamp	XP_057595112.1	Sox	Sox-14
Pgig	XP_039744668.1	Fox	Fox-F1	Amel	XP_002923375.1	Sox	Sox-14
Casi	XP_006860182.1	Fox	Fox-F1	Csim	XP_004419391.1	Sox	Sox-14
Mmus	NP_034556.2	Fox	Fox-F1	Cpor	XP_003476788.1	Sox	Sox-14
Cdid	XP_037671853.1	Fox	Fox-F1	Hsap	NP_004180.1	Sox	Sox-14
Cpor	XP_003460772.2	Fox	Fox-F1	Clup	XP_038426113.1	Sox	Sox-14
Equa	XP_046538978.1	Fox	Fox-F1	Bbub	XP_025147052.1	Sox	Sox-14
Ptig	XP_007082882.2	Fox	Fox-F1	Rfer	XP_032988974.1	Sox	Sox-14
Csim	XP_014648135.1	Fox	Fox-F1	Equa	XP_046511648.1	Sox	Sox-14
Cdro	XP_031292478.1	Fox	Fox-F1	Drot	XP_024422013.1	Sox	Sox-14
Rfer	XP_032985403.1	Fox	Fox-F1	Lcat	XP_045395088.1	Sox	Sox-14
Mang	XP_045746573.2	Fox	Fox-F1	Opri	XP_004588396.1	Sox	Sox-14
Hamp	XP_057568154.1	Fox	Fox-F1	Pgig	XP_039692761.1	Sox	Sox-14
Oafe	XP_007937340.1	Fox	Fox-F1	Ptig	XP_042813213.1	Sox	Sox-14
Ttru	XP_033709961.1	Fox	Fox-F1	Ttru	XP_019789673.1	Sox	Sox-14
Mdom	XP_056669097.1	Fox	N/A	Cimi	XP_017383705.1	Sox	Sox-14
Shar	XP_031805673.1	Fox	N/A	Emax	XP_049726472.1	Sox	Sox-14
Ggal	XP_425714.5	Fox	Fox-S1	Mang	XP_045719678.1	Sox	Sox-14
Clup	XP_038427866.1	Fox	Fox-S1	Oafe	XP_007935393.1	Sox	Sox-14
Hsap	NP_004109.1	Fox	Fox-S1	Ggal	NP_990092.1	Sox	Sox-14
Pgig	XP_039719969.1	Fox	Fox-S1	Tman	XP_004381489.1	Sox	Sox-14
Bmus	XP_036682343.1	Fox	Fox-S1	Dnov	XP_004453833.1	Sox	Sox-14
Mdom	XP_001364156.2	Fox	Fox-S1	Shar	XP_012400053.1	Sox	Sox-14
Amel	XP_002918325.1	Fox	Fox-S1	Mmus	NP_035570.1	Sox	Sox-14
Shar	XP_031809982.1	Fox	Fox-S1	Cdro	XP_010975059.1	Sox	Sox-14
Bbub	XP_025119474.1	Fox	Fox-S1	Ggal	NP_001383603.1	Sox	Sox-21
Casi	XP_006860769.1	Fox	Fox-S1	Mang	XP_045726097.1	Sox	Sox-21
Cdro	XP_010992198.2	Fox	Fox-S1	Rfer	XP_032959615.1	Sox	Sox-21
Mang	XP_045741514.1	Fox	Fox-S1	Bmus	XP_036687740.1	Sox	Sox-21
Pafn	XP_047627330.1	Fox	Fox-S1	Dnov	XP_058132486.1	Sox	Sox-21
Dnov	XP_004464085.1	Fox	Fox-S1	Equa	XP_046520974.1	Sox	Sox-21
Opri	XP_004585763.2	Fox	Fox-S1	Drot	XP_053772754.1	Sox	Sox-21
Cpor	XP_003476682.1	Fox	Fox-S1	Opri	XP_058526666.1	Sox	Sox-21
Mjav	XP_017509202.2	Fox	Fox-S1	Cdid	XP_037656473.1	Sox	Sox-21
Csim	XP_004442549.2	Fox	Fox-S1	Csim	XP_014637259.1	Sox	Sox-21
Rfer	XP_032950864.1	Fox	Fox-S1	Shar	XP_031814415.1	Sox	Sox-21
Equa	XP_046536068.1	Fox	Fox-S1	Ttru	XP_033699347.1	Sox	Sox-21
Scar	XP_047396619.1	Fox	Fox-S1	Tman	XP_004371490.1	Sox	Sox-21
Tman	XP_023581200.1	Fox	Fox-S1	Mmus	NP_808421.1	Sox	Sox-21
Lcat	XP_045385439.1	Fox	Fox-S1	Oana	XP_001512271.3	Sox	Sox-21
Emax	XP_049725743.1	Fox	Fox-S1	Clup	XP_038425312.1	Sox	Sox-21
Mmus	NP_034356.1	Fox	Fox-S1	Ptig	XP_042813926.1	Sox	Sox-21
Ttru	XP_033696063.1	Fox	Fox-S1	Mjav	XP_036880860.1	Sox	Sox-21
Ptig	XP_007073610.2	Fox	Fox-S1	Amel	XP_034519700.1	Sox	Sox-21
Hamp	XP_057559118.1	Fox	Fox-S1	Casi	XP_006832065.1	Sox	Sox-21
Drot	XP_024415665.3	Fox	Fox-S1	Hamp	XP_057563720.1	Sox	Sox-21
Cimi	XP_017370299.1	Fox	Fox-S1	Cdro	XP_031322066.1	Sox	Sox-21
Oafe	XP_007932722.1	Fox	Fox-S1	Mang	XP_045726110.1	Sox	Sox-21
Cdid	XP_037667492.1	Fox	Fox-S1	Pgig	XP_039693387.1	Sox	Sox-21
Ggal	XP_015135808.1	Fox	Fox-B2	Scar	XP_047410577.1	Sox	Sox-21
Pafn	XP_047625507.1	Fox	Fox-B2	Pafn	XP_047612344.1	Sox	Sox-21
Shar	XP_012398564.2	Fox	Fox-B2	Oafe	XP_007942198.1	Sox	Sox-21
Csim	XP_004443968.1	Fox	Fox-B2	Cimi	XP_017399045.1	Sox	Sox-21
Scar	XP_047379799.1	Fox	Fox-B2	Mdom	XP_007501455.1	Sox	Sox-21
Cdro	XP_031305772.1	Fox	Fox-B2	Bbub	XP_025118466.2	Sox	Sox-21
Equa	XP_046519944.1	Fox	Fox-B2	Lcat	XP_045423694.1	Sox	Sox-21
Bbub	XP_006067725.3	Fox	Fox-B2	Hsap	NP_009015.1	Sox	Sox-21
Cimi	XP_017398753.1	Fox	Fox-B2	Emax	XP_049709053.1	Sox	Sox-21
Mmus	NP_032049.1	Fox	Fox-B2	Scar	XP_047405480.1	Sox	Sox-5
Pgig	XP_039724438.1	Fox	Fox-B2	Opri	XP_058511892.1	Sox	Sox-5
Lcat	XP_045418554.1	Fox	Fox-B2	Scar	XP_047405480.1	Sox	Sox-5
Casi	XP_006834992.1	Fox	Fox-B2	Mmus	XP_006506994.1	Sox	Sox-5
Mdom	XP_003341714.1	Fox	Fox-B2	Bmus	XP_036693736.1	Sox	Sox-9
Mjav	XP_036847034.1	Fox	Fox-B2	Hamp	XP_057573119.1	Sox	Sox-9
Oana	XP_001508095.2	Fox	Fox-B2	Ggal	NP_989612.1	Sox	Sox-9
Mang	XP_045737891.1	Fox	Fox-B2	Drot	XP_0244229046.1	Sox	Sox-9
Drot	XP_053776114.1	Fox	Fox-B2	Tman	XP_004374235.1	Sox	Sox-9
Dnov	XP_023440163.2	Fox	Fox-B2	Cimi	XP_017391090.1	Sox	Sox-9
Opri	XP_058513381.1	Fox	Fox-B2	Clup	NP_001002978.1	Sox	Sox-9
Ttru	XP_019789709.1	Fox	Fox-B2	Dnov	XP_058140625.1	Sox	Sox-9
Cdid	XP_037654560.1	Fox	Fox-B2	Opri	XP_004593006.1	Sox	Sox-9
Tman	XP_023593539.1	Fox	Fox-B2	Casi	XP_006872359.1	Sox	Sox-9
Amel	XP_034502902.1	Fox	Fox-B2	Cdro	XP_031304465.1	Sox	Sox-9
Bmus	XP_036712759.1	Fox	Fox-B2	Mjav	XP_017519930.2	Sox	Sox-9
Clup	XP_038381543.1	Fox	Fox-B2	Csim	XP_00432762.1	Sox	Sox-9
Emax	XP_049751166.1	Fox	Fox-B2	Ptig	XP_042822833.1	Sox	Sox-9
Hsap	NP_001013757.1	Fox	Fox-B2	Equa	XP_046530969.1	Sox	Sox-9
Rfer	XP_032978089.1	Fox	Fox-B2	Emax	XP_049717106.1	Sox	Sox-9
Hamp	XP_057580851.1	Fox	Fox-B2	Oana	XP_001506094.2	Sox	Sox-9
Oafe	XP_007936768.1	Fox	Fox-B2	Rfer	XP_032945891.1	Sox	Sox-9
Ptig	XP_042819403.1	Fox	Fox-B2	Ttru	XP_033703168.1	Sox	Sox-9
Oana	XP_001518148.2	Fox	Fox-B1	Amel	XP_019660087.2	Sox	Sox-9
Clup	XP_038436513.1	Fox	Fox-B1	Mang	XP_045743969.1	Sox	Sox-9
Shar	XP_003755656.2	Fox	Fox-B1	Pafn	XP_047614296.1	Sox	Sox-9
Ttru	XP_033708504.1	Fox	Fox-B1	Scar	XP_047398689.1	Sox	Sox-9
Mjav	XP_017523883.1	Fox	Fox-B1	Bbub	XP_025135172.2	Sox	Sox-9
Casi	XP_006831648.1	Fox	Fox-B1	Shar	XP_003768593.1	Sox	Sox-9
Rfer	XP_032964940.1	Fox	Fox-B1	Hsap	NP_000337.1	Sox	Sox-9
Tman	XP_004374676.1	Fox	Fox-B1	Mmus	NP_035567.3	Sox	Sox-9
Hsap	NP_036314.2	Fox	Fox-B1	Oafe	XP_007957879.1	Sox	Sox-9
Amel	XP_034516791.1	Fox	Fox-B1	Cdid	XP_037664998.1	Sox	Sox-9
Pgig	XP_039720931.1	Fox	Fox-B1	Lcat	XP_045425568.1	Sox	Sox-9
Cdro	XP_031308829.1	Fox	Fox-B1	Pgig	XP_039715158.1	Sox	Sox-9
Opri	XP_004578087.1	Fox	Fox-B1	Ggal	XP_015139949.1	Sox	Sox-10
Mang	XP_045755904.1	Fox	Fox-B1	Mmus	NP_035567.1	Sox	Sox-10
Cimi	XP_017387405.1	Fox	Fox-B1	Csim	XP_004437902.1	Sox	Sox-10
Dnov	XP_012376939.1	Fox	Fox-B1	Hsap	NP_008872.1	Sox	Sox-10
Equa	XP_046508174.1	Fox	Fox-B1	Mjav	XP_036862891.1	Sox	Sox-10
Mdom	XP_001365592.1	Fox	Fox-B1	Oafe	XP_00793944.1	Sox	Sox-10
Scar	XP_047396474.1	Fox	Fox-B1	Opri	XP_00458957.1	Sox	Sox-10
Emax	XP_049708649.1	Fox	Fox-B1	Scar	XP_047407584.1	Sox	Sox-10
Bmus	XP_036700853.1	Fox	Fox-B1	Mang	XP_045728505.1	Sox	Sox-10

Table S8 continued from previous page

Species ID	Gene ID	Group	Annotation	Species ID	Gene ID	Group	Annotation
Oafe	XP_007956505.1	Fox	Fox-B1	Cdro	XP_031319106.1	Sox	Sox-10
Bbub	XP_006043154.1	Fox	Fox-B1	Shar	XP_031797111.1	Sox	Sox-10
Ggal	XP_004943811.1	Fox	Fox-B1	Cimi	XP_017364446.1	Sox	Sox-10
Cpor	XP_013000887.1	Fox	Fox-B1	Ttru	XP_033721970.1	Sox	Sox-10
Ptig	XP_042843621.1	Fox	Fox-B1	Dnov	XP_004466267.1	Sox	Sox-10
Mmus	NP_071773.2	Fox	Fox-B1	Ptig	XP_042849176.1	Sox	Sox-10
Csim	XP_004421699.1	Fox	Fox-B1	Tman	XP_004373905.1	Sox	Sox-10
Hamp	XP_057575814.1	Fox	Fox-B1	Cpor	XP_003470533.2	Sox	Sox-10
Lcat	XP_045411390.1	Fox	Fox-B1	Hamp	XP_057599512.1	Sox	Sox-10
Cdid	XP_037691321.1	Fox	Fox-B1	Oana	XP_028934366.1	Sox	Sox-10
Drot	XP_024424310.1	Fox	Fox-B1	Amel	XP_034500270.1	Sox	Sox-10
Paf	XP_047625395.1	Fox	Fox-B1	Casi	XP_006865257.1	Sox	Sox-10
Oana	XP_028925820.1	Fox	Hnf-3b/Fox-A2	Bbub	XP_006071428.1	Sox	Sox-10
Mdom	XP_001382097.1	Fox	Hnf-3b/Fox-A2	Cdid	XP_037702433.1	Sox	Sox-10
Clup	XP_038427348.1	Fox	Hnf-3b/Fox-A2	Lcat	XP_045408856.1	Sox	Sox-10
Amel	XP_002924170.1	Fox	Hnf-3b/Fox-A2	Equa	XP_046502255.1	Sox	Sox-10
Cimi	XP_017396970.1	Fox	Hnf-3b/Fox-A2	Pgig	XP_039733644.1	Sox	Sox-10
Mjav	XP_036878344.1	Fox	Hnf-3b/Fox-A2	Clup	XP_038406550.1	Sox	Sox-10
Ttru	XP_019783169.2	Fox	Hnf-3b/Fox-A2	Rfer	XP_032973012.1	Sox	Sox-10
Mang	XP_045739756.1	Fox	Hnf-3b/Fox-A2	Emax	XP_049740365.1	Sox	Sox-10
Shar	XP_003758186.1	Fox	Hnf-3b/Fox-A2	Drot	XP_024434928.2	Sox	Sox-10
Emax	XP_049725634.1	Fox	Hnf-3b/Fox-A2	Paf	XP_047642247.1	Sox	Sox-10
Scar	XP_047397766.1	Fox	Hnf-3b/Fox-A2	Mdom	XP_001381534.3	Sox	Sox-10
Pgig	XP_039730271.1	Fox	Hnf-3b/Fox-A2	Bmus	XP_036720803.1	Sox	Sox-10
Equa	XP_046535478.1	Fox	Hnf-3b/Fox-A2	Hsap	XP_011519134.2	Sox	Sox-5
Hsap	NP_068556.2	Fox	Hnf-3b/Fox-A2	Lcat	XP_045410158.1	Sox	Sox-5
Mmus	NP_001277994.1	Fox	Hnf-3b/Fox-A2	Oana	XP_028925314.1	Sox	Sox-13
Lcat	XP_045385044.1	Fox	Hnf-3b/Fox-A2	Paf	XP_047607942.1	Sox	Sox-13
Drot	XP_045050691.1	Fox	Hnf-3b/Fox-A2	Shar	XP_003767655.1	Sox	Sox-13
Cdid	XP_037668003.1	Fox	Hnf-3b/Fox-A2	Ggal	XP_015154614.2	Sox	Sox-13
Tman	XP_004376477.1	Fox	Hnf-3b/Fox-A2	Mdom	XP_056671743.1	Sox	Sox-13
Casi	XP_006860717.1	Fox	Hnf-3b/Fox-A2	Ttru	XP_033708141.1	Sox	Sox-13
Oafe	XP_007948556.1	Fox	Hnf-3b/Fox-A2	Cpor	XP_013013443.1	Sox	Sox-13
Paf	XP_047627543.1	Fox	Hnf-3b/Fox-A2	Bmus	XP_036684089.1	Sox	Sox-13
Hamp	XP_057558835.1	Fox	Hnf-3b/Fox-A2	Mang	XP_045363913.1	Sox	Sox-13
Dnov	XP_023439993.2	Fox	Hnf-3b/Fox-A2	Tman	XP_023586088.1	Sox	Sox-13
Bmus	XP_036681859.1	Fox	Hnf-3b/Fox-A2	Hamp	XP_057585117.1	Sox	Sox-13
Rfer	XP_032950275.1	Fox	Hnf-3b/Fox-A2	Cdro	XP_031294667.1	Sox	Sox-13
Cpor	XP_003476486.1	Fox	Hnf-3b/Fox-A2	Drot	XP_024430737.2	Sox	Sox-13
Opri	XP_004585707.2	Fox	Hnf-3b/Fox-A2	Dnov	XP_058130986.1	Sox	Sox-13
Bbub	XP_044783908.1	Fox	Hnf-3b/Fox-A2	Mjav	XP_017506787.2	Sox	Sox-13
Cdro	XP_031290193.1	Fox	Hnf-3b/Fox-A2	Amel	XP_034522993.1	Sox	Sox-13
Csim	XP_004442004.1	Fox	Hnf-3b/Fox-A2	Mmus	XP_011246247.1	Sox	Sox-13
Ptig	XP_042835681.1	Fox	Hnf-3b/Fox-A2	Scar	XP_047376454.1	Sox	Sox-13
Ggal	XP_046794381.1	Fox	Hnf-3b/Fox-A2	Ptig	XP_007086583.2	Sox	Sox-13
Oana	XP_028902569.1	Fox	Fox-E3	Clup	XP_038441905.1	Sox	Sox-13
Emax	XP_049731498.1	Fox	Fox-E3	Oafe	XP_007954219.1	Sox	Sox-13
Hsap	NP_036318.1	Fox	Fox-E3	Opri	XP_058525035.1	Sox	Sox-13
Mang	XP_045751381.1	Fox	Fox-E3	Emax	XP_049714298.1	Sox	Sox-13
Hamp	XP_057562286.1	Fox	Fox-E3	Equa	XP_046537707.1	Sox	Sox-13
Drot	XP_053777352.1	Fox	Fox-E3	Casi	XP_006834281.1	Sox	Sox-13
Ttru	XP_033709174.1	Fox	Fox-E3	Hsap	NP_005677.2	Sox	Sox-13
Pafr	XP_047647470.1	Fox	Fox-E3	Cimi	XP_037593260.1	Sox	Sox-13
Ptig	XP_042852491.1	Fox	Fox-E3	Bbub	XP_025133526.2	Sox	Sox-13
Clup	XP_038413565.1	Fox	Fox-E3	Cdid	XP_037680910.1	Sox	Sox-13
Rfer	XP_032971209.1	Fox	Fox-E3	Lcat	XP_045391966.1	Sox	Sox-13
Bmus	XP_036690899.1	Fox	Fox-E3	Csim	XP_004425199.1	Sox	Sox-13
Pgig	XP_039712722.1	Fox	Fox-E3	Rfer	XP_032948937.1	Sox	Sox-13
Bbub	XP_0251144571.1	Fox	Fox-E3	Pgig	XP_039737775.1	Sox	Sox-13
Cdro	XP_031321430.1	Fox	Fox-E3	Ggal	XP_025006442.1	Sox	Sox-6
Tman	XP_023582802.1	Fox	Fox-E3	Emax	XP_049746477.1	Sox	Sox-6
Pgig	XP_039738800.1	Fox	Fox-E3	Ttru	XP_019804010.1	Sox	Sox-6
Mmus	NP_056573.1	Fox	Fox-E3	Hamp	XP_057604501.1	Sox	Sox-6
Mjav	XP_036877848.1	Fox	Fox-E3	Opri	XP_058519285.1	Sox	Sox-6
Casi	XP_006839917.1	Fox	Fox-E3	Rfer	XP_032975662.1	Sox	Sox-6
Dnov	XP_058160703.1	Fox	Fox-E3	Shar	XP_031797909.1	Sox	Sox-6
Lcat	XP_045401422.1	Fox	Fox-E3	Bbub	XP_045018884.1	Sox	Sox-6
Cimi	XP_017365578.1	Fox	Fox-E3	Dnov	XP_058161675.1	Sox	Sox-6
Cdid	XP_037672792.1	Fox	Fox-E3	Oana	XP_039767308.1	Sox	Sox-6
Opri	XP_004598546.1	Fox	Fox-E3	Hsap	NP_001139291.2	Sox	Sox-6
Scar	XP_047375686.1	Fox	Fox-E3	Bmus	XP_036717651.1	Sox	Sox-6
Oana	XP_001516678.2	Fox	Fox-D4	Mjav	XP_036882886.1	Sox	Sox-6
Pgig	XP_039744744.1	Fox	Fox-D4	Amel	XP_034501998.1	Sox	Sox-6
Hamp	XP_057576439.1	Fox	Fox-D4	Equa	XP_046493708.1	Sox	Sox-6
Paf	XP_047625506.1	Fox	Fox-D4	Mdom	XP_007497087.2	Sox	Sox-6
Hsap	NP_001119806.1	Fox	Fox-D4	Cpor	XP_023419042.1	Sox	Sox-6
Emax	XP_049752498.1	Fox	Fox-D4	Pgig	XP_039707456.1	Sox	Sox-6
Cdid	XP_037706680.1	Fox	Fox-D4	Oafe	XP_007955001.1	Sox	Sox-6
Hsap	NP_954714.2	Fox	Fox-D4	Cdid	XP_037695265.1	Sox	Sox-6
Rfer	XP_032979406.1	Fox	Fox-D4	Casi	XP_06865846.1	Sox	Sox-6
Ttru	XP_033714036.1	Fox	Fox-D4	Cimi	XP_17381104.1	Sox	Sox-6
Ptig	XP_042820135.1	Fox	Fox-D4	Mmus	XP_036008726.1	Sox	Sox-6
Cdro	XP_010998361.2	Fox	Fox-D4	Ptig	XP_042815591.1	Sox	Sox-6
Mang	XP_045739597.1	Fox	Fox-D4	Paf	XP_047632165.1	Sox	Sox-6
Scar	XP_047380997.1	Fox	Fox-D4	Tman	XP_023593283.1	Sox	Sox-6
Drot	XP_053785851.1	Fox	Fox-D4	Drot	XP_053780580.1	Sox	Sox-6
Hsap	NP_001078945.1	Fox	Fox-D4	Mang	XP_054368007.1	Sox	Sox-6
Amel	XP_034503132.1	Fox	Fox-D4	Csim	XP_014642191.1	Sox	Sox-6
Oafe	XP_007954771.1	Fox	Fox-D4	Scar	XP_047372506.1	Sox	Sox-6
Equa	XP_046519849.1	Fox	Fox-D4	Lcat	XP_045413549.1	Sox	Sox-6
Bbub	XP_025137372.3	Fox	Fox-D4	Cdro	XP_010975526.2	Sox	Sox-6
Dnov	XP_023444807.2	Fox	Fox-D4	Clup	XP_038425054.1	Sox	Sox-6
Bmus	XP_036712053.1	Fox	Fox-D4	Emax	XP_049725376.1	Sox	Sox-18
Hsap	NP_036316.1	Fox	Fox-D4	Casi	XP_006873390.1	Sox	N/A
Opri	XP_004591937.2	Fox	Fox-D4	Oafe	XP_007952883.1	Sox	N/A
Cimi	XP_017398902.1	Fox	Fox-D4	Mmus	NP_035577.1	Sox	Sox-8
Mdom	XP_001373972.1	Fox	Fox-D4	Opri	XP_004596857.2	Sox	N/A
Mmus	NP_032048.1	Fox	Fox-D4	Cpor	XP_003478479.1	Sox	N/A
Hsap	NP_997188.2	Fox	Fox-D4	Oana	XP_028913325.1	Sox	Sox-8
Hsap	NP_954586.4	Fox	Fox-D4	Ggal	NP_990062.2	Sox	N/A
Csim	XP_014650927.1	Fox	Fox-D4	Cpor	XP_023421073.1	Sox	N/A
Cpor	XP_013010584.1	Fox	Fox-D4	Cimi	XP_017378281.2	Sox	N/A
Mjav	XP_036847234.1	Fox	Fox-D4	Cdid	XP_037702077.1	Sox	N/A

Table S8 continued from previous page

Species ID	Gene ID	Group	Annotation	Species ID	Gene ID	Group	Annotation
Casi	XP_006863895.1	Fox	Fox-D4	Shar	XP_031794321.1	Sox	N/A
Mang	XP_045757341.1	Fox	N/A	Mdom	XP_016280870.1	Sox	N/A
Clup	XP_038401447.1	Fox	N/A	Ggal	XP_040553151.1	Sox	N/A
Amel	XP_019655962.1	Fox	N/A	Oana	XP_028913966.1	Sox	Sox-5
Clup	XP_038439326.1	Fox	N/A	Lcat	XP_045422920.1	Sox	N/A
Ggal	NP_001382914.1	Fox	Fox-C1				

**Supplementary Table S9. Complete set of DSFGs in *Drosophila*.** For each gene, the species ID (Sp. ID) as in Supp. Tab. S5, the accession number (Gene ID), and the Possvm-based annotation are indicated.

Species ID	Gene ID	Group	Annotation	Species ID	Gene ID	Group	Annotation
Agam	XP_061505148.1	Dmrt	Dsx	Dbus	XP_017845192.1	Fox	Slp-1
Dere	XP_026836830.1	Dmrt	Dsx	Dgri	XP_001988485.1	Fox	Slp-1
Dkik	XP_041630485.1	Dmrt	Dsx	Dhyd	XP_023179855.2	Fox	Slp-1
Dmir	XP_017142875.1	Dmrt	Dsx	Dari	XP_017860496.1	Fox	Slp-1
Dgri	XP_043071113.1	Dmrt	Dsx	Dana	XP_001961572.2	Fox	Slp-1
Dsec	XP_002038750.1	Dmrt	Dsx	Dmir	XP_017155124.1	Fox	Slp-1
Dser	XP_020809854.1	Dmrt	Dsx	Dbip	XP_017105354.1	Fox	Slp-1
Dele	XP_017119779.1	Dmrt	Dsx	Dmel	NP_476730.1	Fox	Slp-1
Dalb	XP_034117252.2	Dmrt	Dsx	Dere	XP_001968625.1	Fox	Slp-1
Dwil	XP_023035845.1	Dmrt	Dsx	Dsec	XP_002037770.1	Fox	Slp-1
Dhyd	XP_023178918.2	Dmrt	Dsx	Dkik	XP_017021730.1	Fox	Slp-1
Dmel	NP_001262353.1	Dmrt	Dsx	Dpse	XP_001356670.4	Fox	Slp-1
Dpse	XP_033235910.1	Dmrt	Dsx	Dwil	XP_002065500.1	Fox	Slp-1
Dari	XP_017874634.1	Dmrt	Dsx	Dser	XP_020800881.1	Fox	Slp-1
Dbus	XP_017847641.1	Dmrt	Dsx	Dsuz	XP_016927184.1	Fox	Slp-1
Dbip	XP_017088683.2	Dmrt	Dsx	Agam	XP_061514780.1	Fox	Slp-2
Dsuz	XP_036675224.1	Dmrt	Dsx	Dalb	XP_034100740.1	Fox	Slp-2
Dsuz	XP_036672758.1	Dmrt	Dsx	Dele	XP_017110569.1	Fox	Slp-2
Dana	XP_014766033.1	Dmrt	Dsx	Dere	XP_001968626.1	Fox	Slp-2
Dbus	XP_017844894.1	Dmrt	Dmrt-99B	Dbus	XP_017845289.1	Fox	Slp-2
Dmel	NP_524549.1	Dmrt	Dmrt-99B	Dsec	XP_002037771.1	Fox	Slp-2
Dere	XP_001981330.1	Dmrt	Dmrt-99B	Dpse	XP_001356669.3	Fox	Slp-2
Dwil	XP_023034529.1	Dmrt	Dmrt-99B	Dmir	XP_017151517.1	Fox	Slp-2
Dkik	XP_017021883.1	Dmrt	Dmrt-99B	Dkik	XP_017021724.1	Fox	Slp-2
Dser	XP_020811324.1	Dmrt	Dmrt-99B	Dana	XP_001961573.1	Fox	Slp-2
Dalb	XP_051862400.1	Dmrt	Dmrt-99B	Dari	XP_017860500.1	Fox	Slp-2
Dsuz	XP_036674364.1	Dmrt	Dmrt-99B	Dser	XP_020800889.1	Fox	Slp-2
Dsec	XP_002037224.1	Dmrt	Dmrt-99B	Dbip	XP_017105682.2	Fox	Slp-2
Dele	XP_017131043.1	Dmrt	Dmrt-99B	Dhyd	XP_023179875.2	Fox	Slp-2
Dbip	XP_017099960.2	Dmrt	Dmrt-99B	Dsuz	XP_016926388.1	Fox	Slp-2
Dmir	XP_017145289.1	Dmrt	Dmrt-99B	Dwil	XP_023031666.1	Fox	Slp-2
Dpse	XP_001357766.3	Dmrt	Dmrt-99B	Dgri	XP_001988487.3	Fox	Slp-2
Dhyd	XP_023160825.2	Dmrt	Dmrt-99B	Dmel	NP_476834.1	Fox	Slp-2
Dari	XP_017869153.1	Dmrt	Dmrt-99B	Agam	XP_317309.5	Fox	Fd-3/Fd-59A
Dgri	XP_001996117.1	Dmrt	Dmrt-99B	Dbip	XP_017090886.2	Fox	Fd-3/Fd-59A
Dhyd	XP_023160826.2	Dmrt	Dmrt-99B	Dpse	XP_001361889.1	Fox	Fd-3/Fd-59A
Dana	XP_001964762.1	Dmrt	Dmrt-99B	Dmel	NP_523814.1	Fox	Fd-3/Fd-59A
Agam	XP_061501728.1	Dmrt	Dmrt-93B	Dsuz	XP_016927872.1	Fox	Fd-3/Fd-59A
Dalb	XP_034117959.2	Dmrt	Dmrt-93B	Dele	XP_017132051.1	Fox	Fd-3/Fd-59A
Dpse	XP_001360059.2	Dmrt	Dmrt-93B	Dser	XP_020803123.1	Fox	Fd-3/Fd-59A
Dkik	XP_017036725.1	Dmrt	Dmrt-93B	Dwil	XP_002061322.3	Fox	Fd-3/Fd-59A
Dbip	XP_017102685.2	Dmrt	Dmrt-93B	Dgri	XP_001987215.1	Fox	Fd-3/Fd-59A
Dbus	XP_017844858.1	Dmrt	Dmrt-93B	Dana	XP_001960803.1	Fox	Fd-3/Fd-59A
Dsuz	XP_036672900.1	Dmrt	Dmrt-93B	Dmir	XP_033248262.1	Fox	Fd-3/Fd-59A
Dser	XP_020817755.1	Dmrt	Dmrt-93B	Dalb	XP_034105309.1	Fox	Fd-3/Fd-59A
Dari	XP_017874225.1	Dmrt	Dmrt-93B	Dhyd	XP_030081270.1	Fox	Fd-3/Fd-59A
Dgri	XP_001990371.1	Dmrt	Dmrt-93B	Dmir	XP_017150782.2	Fox	Fd-3/Fd-59A
Dwil	XP_002073560.1	Dmrt	Dmrt-93B	Dkik	XP_017024427.1	Fox	Fd-3/Fd-59A
Dmel	NP_524428.1	Dmrt	Dmrt-93B	Dari	XP_017865679.1	Fox	Fd-3/Fd-59A
Dmir	XP_017140144.1	Dmrt	Dmrt-93B	Dsec	XP_002039987.1	Fox	Fd-3/Fd-59A
Dsec	XP_002044276.1	Dmrt	Dmrt-93B	Dere	XP_001976310.1	Fox	Fd-3/Fd-59A
Dele	XP_017118537.1	Dmrt	Dmrt-93B	Dbus	XP_017836437.1	Fox	Fd-3/Fd-59A
Dana	XP_001954937.1	Dmrt	Dmrt-93B	Agam	XP_315933.4	Fox	Crocodile
Dere	XP_001979344.2	Dmrt	Dmrt-93B	Dalb	XP_034109751.1	Fox	Crocodile
Dmel	NP_511146.2	Dmrt	Dmrt-11E	Dele	XP_017121437.1	Fox	Crocodile
Dsec	XP_002042926.2	Dmrt	Dmrt-11E	Dmel	NP_524202.1	Fox	Crocodile
Dwil	XP_002075263.1	Dmrt	Dmrt-11E	Dsec	XP_002040809.1	Fox	Crocodile
Dpse	XP_001355530.3	Dmrt	Dmrt-11E	Dsuz	XP_016933536.1	Fox	Crocodile
Dbip	XP_017089150.2	Dmrt	Dmrt-11E	Dari	XP_017864431.1	Fox	Crocodile
Dser	XP_020808350.1	Dmrt	Dmrt-11E	Dmir	XP_017139058.1	Fox	Crocodile
Dmir	XP_017136076.1	Dmrt	Dmrt-11E	Dhyd	XP_023173236.2	Fox	Crocodile
Dana	XP_032309371.1	Dmrt	Dmrt-11E	Dbip	XP_017095284.1	Fox	Crocodile
Dere	XP_015010652.2	Dmrt	Dmrt-11E	Dwil	XP_002061802.1	Fox	Crocodile
Dbus	XP_017850127.1	Dmrt	Dmrt-11E	Dser	XP_020817581.1	Fox	Crocodile
Dsuz	XP_016924080.2	Dmrt	Dmrt-11E	Dbus	XP_017841359.1	Fox	Crocodile
Dalb	XP_034099986.1	Dmrt	Dmrt-11E	Dpse	XP_001354188.2	Fox	Crocodile
Dgri	XP_043071903.1	Dmrt	Dmrt-11E	Dkik	XP_017029613.1	Fox	Crocodile
Dele	XP_017112743.2	Dmrt	Dmrt-11E	Dana	XP_001958355.1	Fox	Crocodile
Dhyd	XP_023180022.1	Dmrt	Dmrt-11E	Dgri	XP_001983802.1	Fox	Crocodile
Dkik	XP_017034619.2	Dmrt	Dmrt-11E	Dere	XP_001973629.1	Fox	Crocodile
Dari	XP_017869421.1	Dmrt	Dmrt-11E	Dmir	XP_017144084.1	Fox	Fkh
Agam	XP_310668.5	Dmrt	N/A	Dari	XP_017873677.1	Fox	Fkh
Dkik	XP_017022631.1	Fox	Fox-3F	Dpse	XP_033238940.1	Fox	Fkh
Dsec	XP_002037072.2	Fox	Fox-3F	Dwil	XP_002070503.2	Fox	Fkh
Dser	XP_020800708.1	Fox	Fox-3F	Dsuz	XP_036673542.1	Fox	Fkh
Dbip	XP_017087408.2	Fox	Fox-3F	Dalb	XP_051862733.1	Fox	Fkh
Dere	XP_026837609.1	Fox	Fox-3F	Dbip	XP_017092849.2	Fox	Fkh
Dana	XP_014759802.1	Fox	Fox-3F	Dbus	XP_017849701.2	Fox	Fkh
Dpse	XP_033238487.1	Fox	Fox-3F	Dsec	XP_002043027.2	Fox	Fkh
Dmir	XP_017145653.1	Fox	Fox-3F	Dgri	XP_001989593.3	Fox	Fkh
Dmel	NP_001356931.1	Fox	Fox-3F	Dana	XP_001955055.2	Fox	Fkh
Dele	XP_017112308.1	Fox	Fox-3F	Dkik	XP_017037933.1	Fox	Fkh
Dsuz	XP_016942113.1	Fox	Fox-3F	Dser	XP_020811220.1	Fox	Fkh

Table S9 continued from previous page

Species ID	Gene ID	Group	Annotation	Species ID	Gene ID	Group	Annotation
Dwil	XP_046868406.1	Fox	N/A	Dere	XP_001981453.3	Fox	Fkh
Dbus	XP_033150245.1	Fox	N/A	Agam	XP_061497286.1	Fox	Fkh
Dhyd	XP_030079965.1	Fox	N/A	Dmel	NP_001263038.1	Fox	Fkh
Dari	XP_017869433.1	Fox	N/A	Dele	XP_017131720.1	Fox	Fkh
Dgri	XP_043071916.1	Fox	N/A	Dhyd	XP_023165610.2	Fox	Fkh
Dalb	XP_051860438.1	Fox	N/A	Dalb	XP_034114947.1	Fox	Fd-5/Fd-96Cb
Agam	XP_061503465.1	Fox	Ches-1	Dele	XP_017123797.1	Fox	Fd-5/Fd-96Cb
Dbip	XP_017091473.2	Fox	Ches-1	Dgri	XP_001990839.1	Fox	Fd-5/Fd-96Cb
Dkik	XP_017022654.1	Fox	Ches-1	Dana	XP_001953691.1	Fox	Fd-5/Fd-96Cb
Dalb	XP_034097387.1	Fox	Ches-1	Dere	XP_001981873.2	Fox	Fd-5/Fd-96Cb
Dhyd	XP_023174414.2	Fox	Ches-1	Dkik	XP_017020777.1	Fox	Fd-5/Fd-96Cb
Dana	XP_032308510.1	Fox	Ches-1	Dwil	XP_023036287.1	Fox	Fd-5/Fd-96Cb
Dwil	XP_023030939.1	Fox	Ches-1	Dbip	XP_017103917.2	Fox	Fd-5/Fd-96Cb
Dser	XP_020805864.1	Fox	Ches-1	Dser	XP_020817575.1	Fox	Fd-5/Fd-96Cb
Dgri	XP_032593899.1	Fox	Ches-1	Dsec	XP_002043869.2	Fox	Fd-5/Fd-96Cb
Dbus	XP_017852827.2	Fox	Ches-1	Dsuz	XP_016923664.2	Fox	Fd-5/Fd-96Cb
Dari	XP_017869569.1	Fox	Ches-1	Dmel	NP_524496.1	Fox	Fd-5/Fd-96Cb
Dpse	XP_033241472.1	Fox	Ches-1	Agam	XP_061497721.1	Fox	Fd-4/Fd-96Ca
Dmir	XP_033252440.1	Fox	Ches-1	Dana	XP_001953690.1	Fox	Fd-4/Fd-96Ca
Agam	XP_061505879.1	Fox	Jumeau	Dmel	NP_001287516.1	Fox	Fd-4/Fd-96Ca
Dalb	XP_034114927.1	Fox	Jumeau	Dsuz	XP_016923665.2	Fox	Fd-4/Fd-96Ca
Dgri	XP_001990048.1	Fox	Jumeau	Dpse	XP_033232491.1	Fox	Fd-4/Fd-96Ca
Dmir	XP_033244968.1	Fox	Jumeau	Dser	XP_020817686.1	Fox	Fd-4/Fd-96Ca
Dsec	XP_02031751.2	Fox	Jumeau	Dele	XP_017123798.1	Fox	Fd-4/Fd-96Ca
Dser	XP_020814051.1	Fox	Jumeau	Dkik	XP_017020776.1	Fox	Fd-4/Fd-96Ca
Dari	XP_017867184.1	Fox	Jumeau	Dwil	XP_002072971.1	Fox	Fd-4/Fd-96Ca
Dkik	XP_017026689.1	Fox	Jumeau	Dalb	XP_034114767.1	Fox	Fd-4/Fd-96Ca
Dana	XP_001952947.1	Fox	Jumeau	Dgri	XP_001990840.1	Fox	Fd-4/Fd-96Ca
Dmir	XP_033244970.1	Fox	Jumeau	Dbus	XP_017846227.1	Fox	Fd-4/Fd-96Ca
Dere	XP_001980679.1	Fox	Jumeau	Dbip	XP_017103916.2	Fox	Fd-4/Fd-96Ca
Dele	XP_017119152.1	Fox	Jumeau	Dhyd	XP_023172915.1	Fox	Fd-4/Fd-96Ca
Dsuz	XP_016944421.1	Fox	Jumeau	Dari	XP_017856594.1	Fox	Fd-4/Fd-96Ca
Dhyd	XP_023174343.2	Fox	Jumeau	Dsec	XP_002043867.1	Fox	Fd-4/Fd-96Ca
Dwil	XP_023034158.2	Fox	Jumeau	Dere	XP_001981874.1	Fox	Fd-4/Fd-96Ca
Dpse	XP_0323233400.1	Fox	Jumeau	Dmir	XP_017141748.1	Fox	Fd-4/Fd-96Ca
Dpse	XP_015038643.2	Fox	Jumeau	Agam	XP_061507474.1	Fox	N/A
Dmir	XP_033244969.1	Fox	Jumeau	Agam	XP_061497423.1	Fox	N/A
Dmel	NP_524302.1	Fox	Jumeau	Agam	XP_312480.5	Fox	N/A
Dbip	XP_017100743.2	Fox	Jumeau	Dana	XP_044571024.1	Sox	Sox-21B
Dbus	XP_017845708.1	Fox	Jumeau	Dari	XP_017862175.1	Sox	Sox-21B
Dbip	XP_043069566.1	Fox	Hcm-1	Dhyd	XP_030080409.1	Sox	Sox-21B
Dsec	XP_002044492.1	Fox	Hcm-1	Dbus	XP_033149417.1	Sox	Sox-21B
Dere	XP_026838997.1	Fox	Hcm-1	Dere	XP_001972712.1	Sox	Sox-21B
Dhyd	XP_023165574.2	Fox	Hcm-1	Dmir	XP_033243279.1	Sox	Sox-21B
Dele	XP_041564577.1	Fox	Hcm-1	Dmel	NP_001261829.1	Sox	Sox-21B
Dsuz	XP_036678464.1	Fox	Hcm-1	Dele	XP_017128640.1	Sox	Sox-21B
Dalb	XP_051862243.1	Fox	Hcm-1	Dpse	XP_002134931.3	Sox	Sox-21B
Dana	XP_032311767.1	Fox	Hcm-1	Dwil	XP_023034010.1	Sox	Sox-21B
Dpse	XP_033237266.1	Fox	Hcm-1	Dsec	XP_002030465.1	Sox	Sox-21B
Dmel	NP_726538.1	Fox	Hcm-1	Dalb	XP_034107607.1	Sox	Sox-21B
Dmir	XP_017155587.1	Fox	Hcm-1	Dbip	XP_017090122.2	Sox	Sox-21B
Dser	XP_020809944.1	Fox	Hcm-1	Dhyd	XP_023171730.2	Sox	Sox-14
Dbus	XP_017853536.1	Fox	Hcm-1	Dsec	XP_061509694.1	Sox	Sox-21B
Dwil	XP_023053694.1	Fox	Hcm-1	Dkik	XP_017034105.1	Sox	Sox-21B
Dkik	XP_017037372.1	Fox	Hcm-1	Dgri	XP_032598957.1	Sox	Sox-21B
Dwil	XP_002072563.1	Fox	Fd-102C	Dser	XP_020816310.1	Sox	Sox-21B
Dhyd	XP_023163080.1	Fox	Fd-102C	Dgri	XP_032590737.1	Sox	Sox-14
Dbip	XP_017098296.2	Fox	Fd-102C	Dhyd	XP_023171730.2	Sox	Sox-14
Dmel	NP_651951.1	Fox	Fd-102C	Dbus	XP_017836915.1	Sox	Sox-14
Dsuz	XP_036671526.1	Fox	Fd-102C	Dpse	XP_001360869.4	Sox	Sox-14
Dele	XP_017126055.1	Fox	Fd-102C	Dkik	XP_017018688.1	Sox	Sox-14
Dbus	XP_017853523.1	Fox	Fd-102C	Dbip	XP_017090231.2	Sox	Sox-14
Dsec	XP_002043702.1	Fox	Fd-102C	Dser	XP_020817997.1	Sox	Sox-14
Agam	XP_061505343.1	Fox	Fd-102C	Dmir	XP_033246935.1	Sox	Sox-14
Dere	XP_001982706.1	Fox	Fd-102C	Dsec	XP_002040195.2	Sox	Sox-14
Dmir	XP_017155765.1	Fox	Fd-102C	Dmel	NP_001286801.1	Sox	Sox-14
Dana	XP_032308054.1	Fox	Fd-102C	Dmir	XP_017149652.2	Sox	Sox-14
Dalb	XP_034112224.1	Fox	Fd-102C	Dsuz	XP_016926392.1	Sox	Sox-14
Dpse	XP_015044407.1	Fox	Fd-102C	Dari	XP_017868666.1	Sox	Sox-14
Dkik	XP_041632589.1	Fox	Fd-102C	Dalb	XP_034106594.1	Sox	Sox-14
Dgri	XP_043071810.1	Fox	Fd-102C	Dere	XP_001976516.1	Sox	Sox-14
Dser	XP_020809941.1	Fox	Fd-102C	Dwil	XP_02061086.2	Sox	Sox-14
Dari	XP_017869379.1	Fox	Fd-102C	Agam	XP_061506760.1	Sox	Sox-14
Agam	XP_061502996.1	Fox	Fox-P	Dele	XP_017125359.1	Sox	Sox-14
Dhyd	XP_023172914.1	Fox	Fox-P	Dana	XP_001959457.2	Sox	Sox-14
Dbus	XP_033148941.1	Fox	Fox-P	Dpse	XP_001353547.1	Sox	Dichaete
Dbip	XP_043067838.1	Fox	Fox-P	Dele	XP_017128756.1	Sox	Dichaete
Dgri	XP_032593352.2	Fox	Fox-P	Dmir	XP_017138810.1	Sox	Dichaete
Dana	XP_044571198.1	Fox	Fox-P	Dmel	NP_524066.1	Sox	Dichaete
Dalb	XP_051862673.1	Fox	Fox-P	Dere	XP_001972713.1	Sox	Dichaete
Dmir	XP_033244694.1	Fox	Fox-P	Dalb	XP_034104310.1	Sox	Dichaete
Dmel	NP_001247011.1	Fox	Fox-P	Dsec	XP_002030467.1	Sox	Dichaete
Dere	XP_015009985.2	Fox	Fox-P	Dene	XP_017128756.1	Sox	Dichaete
Dele	XP_017117570.1	Fox	Fox-P	Dpse	XP_017089891.1	Sox	Dichaete
Dsuz	XP_016930987.1	Fox	Fox-P	Dagam	XP_061514536.1	Sox	Dichaete
Dpse	XP_033241607.1	Fox	Fox-P	Dbus	XP_017841081.1	Sox	Dichaete
Dwil	XP_046865977.1	Fox	Fox-P	Dkik	XP_017034102.1	Sox	Dichaete
Dsec	XP_032577368.1	Fox	Fox-P	Dwil	XP_002061712.1	Sox	Dichaete
Dser	XP_020813734.1	Fox	Fox-P	Dari	XP_017864416.1	Sox	Dichaete
Dwil	XP_002072142.2	Fox	fox-O	Dhyd	XP_023173062.1	Sox	Dichaete
Agam	XP_061497073.1	Fox	fox-O	Dser	XP_020815889.1	Sox	Dichaete
Dari	XP_017859446.1	Fox	fox-O	Dsuz	XP_016933433.1	Sox	Dichaete
Dsuz	XP_036674215.1	Fox	fox-O	Dere	XP_001975666.1	Sox	Sox-15
Dsuz	XP_036672729.1	Fox	fox-O	Dele	XP_017133202.1	Sox	Sox-15
Dgri	XP_032595778.1	Fox	fox-O	Dbus	XP_017835743.2	Sox	Sox-15
Dbip	XP_017091959.2	Fox	fox-O	Dwil	XP_023030517.1	Sox	Sox-15
Dmir	XP_017143568.1	Fox	fox-O	Dalb	XP_051860622.1	Sox	Sox-15
Dalb	XP_034113157.1	Fox	fox-O	Dana	XP_001954979.1	Sox	Sox-15
Dsec	XP_032576567.1	Fox	fox-O	Dpse	XP_017090318.2	Sox	Sox-15
Dhyd	XP_023171342.1	Fox	fox-O	Dari	XP_017867550.1	Sox	Sox-15
Dbus	XP_033149836.1	Fox	fox-O	Dmir	XP_017149731.1	Sox	Sox-15
Dere	XP_026839669.1	Fox	fox-O	Dsec	XP_002033808.1	Sox	Sox-15
Dkik	XP_017026758.1	Fox	fox-O	Dagam	XP_061509362.1	Sox	Sox-15

Table S9 continued from previous page

Species ID	Gene ID	Group	Annotation	Species ID	Gene ID	Group	Annotation
Dser	XP_020817293.1	Fox	fox-O	Dmel	NP_523739.2	Sox	Sox-15
Dpse	XP_033236382.1	Fox	fox-O	Dser	XP_020802427.1	Sox	Sox-15
Dele	XP_041565035.1	Fox	fox-O	Dhyd	XP_030081376.1	Sox	Sox-15
Dana	XP_032311428.1	Fox	fox-O	Dsuz	XP_016927906.1	Sox	Sox-15
Dmel	NP_650330.3	Fox	fox-O	Dpse	XP_001361762.2	Sox	Sox-15
Agam	XP_001688749.2	Fox	Biniov	Dkik	XP_017019703.1	Sox	Sox-15
Dmel	NP_523950.2	Fox	Biniov	Dgri	XP_001987303.1	Sox	Sox-15
Dalb	XP_034105454.1	Fox	Biniov	Dsec	XP_002030464.1	Sox	Sox-21A
Dser	XP_020817612.1	Fox	Biniov	Dgri	XP_001984256.1	Sox	Sox-21A
Dele	XP_017121425.1	Fox	Biniov	Dsuz	XP_016934639.1	Sox	Sox-21A
Dkik	XP_017017381.1	Fox	Biniov	Dmir	XP_017138458.1	Sox	Sox-21A
Dana	XP_001958356.2	Fox	Biniov	Dkik	XP_017034345.1	Sox	Sox-21A
Dari	XP_017864430.1	Fox	Biniov	Dbus	XP_017840298.2	Sox	Sox-21A
Dbus	XP_017841230.1	Fox	Biniov	Dari	XP_017862178.1	Sox	Sox-21A
Dwil	XP_002061803.1	Fox	Biniov	Dalb	XP_034108985.1	Sox	Sox-21A
Dgri	XP_001983803.1	Fox	Biniov	Dser	XP_020816311.1	Sox	Sox-21A
Dere	XP_001971514.2	Fox	Biniov	Agam	XP_061513404.1	Sox	Sox-21A
Dsuz	XP_036672243.1	Fox	Biniov	Dele	XP_017128154.1	Sox	Sox-21A
Dpse	XP_002134730.3	Fox	Biniov	Dana	XP_032310341.1	Sox	Sox-21A
Dhyd	XP_030081482.1	Fox	Biniov	Dhyd	XP_023173058.2	Sox	Sox-21A
D bip	XP_017095285.2	Fox	Biniov	Dere	XP_001972711.1	Sox	Sox-21A
D mir	XP_017139097.1	Fox	Biniov	D bip	XP_04307009.1	Sox	Sox-21A
D sec	XP_002035641.1	Fox	Biniov	D wil	XP_023033998.1	Sox	Sox-21A
Agam	XP_061503012.1	Fox	Fd-2/Fox-L1	D mel	NP_001261827.1	Sox	Sox-21A
D mel	NP_001246609.1	Fox	Fd-2/Fox-L1	D pse	XP_0323239763.1	Sox	Sox-21A
D hyd	XP_023178350.2	Fox	Fd-2/Fox-L1	D wil	XP_002066518.2	Sox	Sox-N
D bus	XP_017840363.1	Fox	Fd-2/Fox-L1	D ser	XP_020807900.1	Sox	Sox-N
D kik	XP_017019394.1	Fox	Fd-2/Fox-L1	D suz	XP_036678170.1	Sox	Sox-N
D ser	XP_020807718.1	Fox	Fd-2/Fox-L1	D bus	XP_033150665.1	Sox	Sox-N
Dari	XP_017863042.1	Fox	Fd-2/Fox-L1	D sec	XP_002036270.2	Sox	Sox-N
D sec	XP_002035260.1	Fox	Fd-2/Fox-L1	D gri	XP_001993541.2	Sox	Sox-N
D ere	XP_001971815.2	Fox	Fd-2/Fox-L1	D ele	XP_017121213.1	Sox	Sox-N
D alb	XP_051860220.1	Fox	Fd-2/Fox-L1	D dari	XP_017860249.1	Sox	Sox-N
D ana	XP_001956070.1	Fox	Fd-2/Fox-L1	D alb	XP_051858933.1	Sox	Sox-N
D suz	XP_016932589.2	Fox	Fd-2/Fox-L1	D kik	XP_017024634.1	Sox	Sox-N
D ele	XP_017126783.1	Fox	Fd-2/Fox-L1	Agam	XP_061516479.1	Sox	Sox-N
D gri	XP_001983216.1	Fox	Fd-2/Fox-L1	D mir	XP_017154597.1	Sox	Sox-N
D pse	XP_001352427.1	Fox	Fd-2/Fox-L1	D bip	XP_017099296.2	Sox	Sox-N
D wil	XP_002062380.1	Fox	Fd-2/Fox-L1	D ere	XP_026834908.1	Sox	Sox-N
D bip	XP_017098174.2	Fox	Fd-2/Fox-L1	D mel	NP_001260269.1	Sox	Sox-N
D mir	XP_017137846.1	Fox	Fd-2/Fox-L1	D hyd	XP_023172210.2	Sox	Sox-N
D ere	XP_001977848.1	Fox	Fd-19B	D pse	XP_001355808.4	Sox	Sox-N
D hyd	XP_030079337.1	Fox	Fd-19B	D ana	XP_001962613.2	Sox	Sox-N
D mel	NP_608369.1	Fox	Fd-19B	D bus	XP_017845374.1	Sox	Sox-100B
D sec	XP_002039497.1	Fox	Fd-19B	D dari	XP_017874440.1	Sox	Sox-100B
D suz	XP_016923819.2	Fox	Fd-19B	Agam	XP_061509035.1	Sox	Sox-100B
D ele	XP_041565246.1	Fox	Fd-19B	D hyd	XP_023168522.1	Sox	Sox-100B
D pse	XP_033239022.1	Fox	Fd-19B	D kik	XP_017037366.1	Sox	Sox-100B
D mir	XP_017134931.2	Fox	Fd-19B	D ana	XP_001954720.1	Sox	Sox-100B
Dari	XP_017869476.1	Fox	Fd-19B	D ele	XP_017122478.1	Sox	Sox-100B
D gri	XP_043072169.1	Fox	Fd-19B	D ser	XP_0208111779.1	Sox	Sox-100B
D alb	XP_034101064.2	Fox	Fd-19B	D bip	XP_017088796.2	Sox	Sox-100B
D wil	XP_046867097.1	Fox	Fd-19B	D mir	XP_017145320.1	Sox	Sox-100B
D ser	XP_020808484.1	Fox	Fd-19B	D alb	XP_051863654.1	Sox	Sox-100B
D kik	XP_017017955.1	Fox	Fd-19B	D mel	NP_651839.1	Sox	Sox-100B
D bus	XP_033150229.1	Fox	Fd-19B	D sec	XP_002037434.1	Sox	Sox-100B
Agam	XP_061512060.1	Fox	N/A	D ere	XP_001981122.1	Sox	Sox-100B
D alb	XP_034098689.1	Fox	N/A	D suz	XP_016935593.2	Sox	Sox-100B
D gri	XP_032595015.1	Fox	N/A	D wil	XP_023035523.1	Sox	Sox-100B
D alb	XP_034098925.1	Fox	Slp-1	D gri	XP_032597554.1	Sox	Sox-100B
Dele	XP_017110574.1	Fox	Slp-1	D pse	XP_001357577.3	Sox	Sox-100B

**Supplementary Table S10. All the enriched GO terms for Group 1 and Group 2 genes of bivalves, mammals, and *Drosophila*.**

Dataset	Group of genes	GO-ID	Term	Annotated	Significant	Expected	classicFisher
Bivalvia	Group 1 + Group 2	GO:0060255	regulation of macromolecule metabolic process	737	59	31.91	0.04525
Bivalvia	Group 1 + Group 2	GO:0080090	regulation of primary metabolic process	673	53	29.14	0.01818
Bivalvia	Group 1 + Group 2	GO:0019219	regulation of nucleobase-containing compound metabolic process	541	41	23.42	0.02388
Bivalvia	Group 1 + Group 2	GO:0063531	DNA-templated transcription	571	39	24.72	0.03767
Bivalvia	Group 1 + Group 2	GO:0032774	RNA biosynthetic process	579	39	25.07	0.04490
Bivalvia	Group 1 + Group 2	GO:0051252	regulation of RNA metabolic process	517	37	22.38	0.02719
Bivalvia	Group 1 + Group 2	GO:0006355	regulation of DNA-templated transcription	490	35	21.22	0.03751
Bivalvia	Group 1 + Group 2	GO:2001141	regulation of RNA biosynthetic process	491	35	21.26	0.03844
Bivalvia	Group 1 + Group 2	GO:0069507	response to stress	370	33	16.02	0.01949
Bivalvia	Group 1 + Group 2	GO:0032502	developmental process	261	27	11.3	0.04445
Bivalvia	Group 1 + Group 2	GO:0006468	protein phosphorylation	345	23	14.94	0.02483
Bivalvia	Group 1 + Group 2	GO:0031325	positive regulation of cellular metabolic process	125	17	5.41	0.00801
Bivalvia	Group 1 + Group 2	GO:0010604	positive regulation of macronucleole metabolic process	151	17	6.54	0.04047
Bivalvia	Group 1 + Group 2	GO:0051172	negative regulation of nitrogen compound metabolic process	117	16	5.07	0.00814
Bivalvia	Group 1 + Group 2	GO:0051173	positive regulation of nitrogen compound metabolic process	137	15	5.93	0.02454
Bivalvia	Group 1 + Group 2	GO:0006310	DNA recombination	66	14	2.86	0.00087
Bivalvia	Group 1 + Group 2	GO:0048513	animal organ development	83	12	3.59	0.04088
Bivalvia	Group 1 + Group 2	GO:0010629	negative regulation of gene expression	78	11	3.38	0.00048
Bivalvia	Group 1 + Group 2	GO:0023051	regulation of signaling pathway	133	11	5.76	0.02872
Bivalvia	Group 1 + Group 2	GO:0045934	negative regulation of nucleobase-containing compound metabolic process	64	11	2.77	0.03637
Bivalvia	Group 1 + Group 2	GO:0009605	response to external stimulus	90	11	3.39	0.04544
Bivalvia	Group 1 + Group 2	GO:0044419	biological process involved in interspecies interaction between organisms	63	11	2.73	0.04761
Bivalvia	Group 1 + Group 2	GO:0006915	apoptotic process	95	10	4.11	0.00768
Bivalvia	Group 1 + Group 2	GO:0009966	regulation of signal transduction	120	10	5.12	0.03451
Bivalvia	Group 1 + Group 2	GO:0006417	regulation of translation	52	9	2.25	0.00033
Bivalvia	Group 1 + Group 2	GO:0045892	negative regulation of DNA-templated transcription	59	9	2.55	0.02968
Bivalvia	Group 1 + Group 2	GO:1802879	negative regulation of RNA biosynthetic process	59	9	2.55	0.02968
Bivalvia	Group 1 + Group 2	GO:0009607	response to biotic stimulus	55	9	2.38	0.03211
Bivalvia	Group 1 + Group 2	GO:0051253	positive regulation of RNA metabolic process	61	9	2.64	0.03719
Bivalvia	Group 1 + Group 2	GO:0006952	defense response	58	9	2.51	0.04163
Bivalvia	Group 1 + Group 2	GO:0006302	double-strand break repair	52	9	2.25	0.04860
Bivalvia	Group 1 + Group 2	GO:0080134	regulation of response to stress	52	9	2.25	0.04860
Bivalvia	Group 1 + Group 2	GO:0010564	regulation of cell cycle process	43	8	1.86	0.00669
Bivalvia	Group 1 + Group 2	GO:0042981	regulation of apoptotic process	70	8	3.03	0.01024
Bivalvia	Group 1 + Group 2	GO:0043067	regulation of programmed cell death	72	8	3.12	0.01205
Bivalvia	Group 1 + Group 2	GO:0048584	positive regulation of response to stimulus	61	8	2.64	0.03998
Bivalvia	Group 1 + Group 2	GO:0071310	cellular response to organic substance	52	7	2.25	0.00664
Bivalvia	Group 1 + Group 2	GO:0010628	positive regulation of gene expression	34	7	1.47	0.02662
Bivalvia	Group 1 + Group 2	GO:0045944	positive regulation of transcription by RNA polymerase II	38	6	1.65	0.00535
Bivalvia	Group 1 + Group 2	GO:1901987	regulation of cell cycle phase transition	29	6	1.26	0.02368
Bivalvia	Group 1 + Group 2	GO:2000779	regulation of double-strand break repair	11	6	0.48	0.02430
Bivalvia	Group 1 + Group 2	GO:0051247	positive regulation of protein metabolic process	54	6	2.34	0.02818
Bivalvia	Group 1 + Group 2	GO:0051248	negative regulation of protein metabolic process	55	6	2.38	0.03053
Bivalvia	Group 1 + Group 2	GO:1901690	import into cell	52	6	2.42	0.03303
Bivalvia	Group 1 + Group 2	GO:1902531	regulation of intracellular signal transduction	59	6	2.55	0.04118
Bivalvia	Group 1 + Group 2	GO:0044770	cell cycle phase transition	35	6	1.52	0.04666
Bivalvia	Group 1 + Group 2	GO:0000122	negative regulation of transcription by RNA polymerase II	31	5	1.34	0.00988
Bivalvia	Group 1 + Group 2	GO:0006402	mRNA catabolic process	35	5	1.52	0.01640
Bivalvia	Group 1 + Group 2	GO:1901980	regulation of mitotic cell cycle phase transition	18	5	0.78	0.02436
Bivalvia	Group 1 + Group 2	GO:1901698	response to nitrogen compound	41	5	1.78	0.03072
Bivalvia	Group 1 + Group 2	GO:0006401	RNA catabolic process	56	6	2.42	0.04013
Bivalvia	Group 1 + Group 2	GO:0030155	regulation of cell adhesion	11	4	0.48	0.00089
Bivalvia	Group 1 + Group 2	GO:0048568	embryonic organ development	12	4	0.52	0.00129
Bivalvia	Group 1 + Group 2	GO:0007517	muscle organ development	13	4	0.56	0.00180
Bivalvia	Group 1 + Group 2	GO:0051607	defense response to virus	13	4	0.56	0.00180
Bivalvia	Group 1 + Group 2	GO:0050759	regulation of double-strand break repair via homologous recombination	5	4	0.22	0.00537
Bivalvia	Group 1 + Group 2	GO:0006401	ribosomal small subunit biogenesis	21	4	0.91	0.01148
Bivalvia	Group 1 + Group 2	GO:0042274	determination of left-right symmetry	44	4	1.21	0.03102
Bivalvia	Group 1 + Group 2	GO:0043066	negative regulation of apoptotic process	28	4	1.26	0.03480
Bivalvia	Group 1 + Group 2	GO:0043069	chromosome organization involved in meiotic cell cycle	29	4	1.26	0.00247
Bivalvia	Group 1 + Group 2	GO:0045132	meiotic chromosome segregation	8	3	0.35	0.00382
Bivalvia	Group 1 + Group 2	GO:0042326	negative regulation of phosphorylation	10	3	0.0768	

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Bivalvia	Group 1 + Group 2	GO:0008285	negative regulation of cell population proliferation	10	3	0.43	0.00768
Bivalvia	Group 1 + Group 2	GO:0022604	regulation of cell morphogenesis	10	3	0.43	0.00768
Bivalvia	Group 1 + Group 2	GO:0001894	tissue homeostasis	10	3	0.43	0.00768
Bivalvia	Group 1 + Group 2	GO:0003007	heart morphogenesis	10	3	0.43	0.00768
Bivalvia	Group 1 + Group 2	GO:0051093	negative regulation of developmental process	11	3	0.48	0.01022
Bivalvia	Group 1 + Group 2	GO:0001501	skeletal system development	11	3	0.48	0.01022
Bivalvia	Group 1 + Group 2	GO:0042327	positive regulation of phosphorylation	12	3	0.52	0.01320
Bivalvia	Group 1 + Group 2	GO:0010562	positive regulation of phosphorus metabolic process	13	3	0.56	0.01662
Bivalvia	Group 1 + Group 2	GO:0045010	actin nucleation	13	3	0.56	0.01662
Bivalvia	Group 1 + Group 2	GO:0045937	positive regulation of phosphate metabolic process	13	3	0.56	0.01662
Bivalvia	Group 1 + Group 2	GO:0007127	meiosis I	14	3	0.61	0.02049
Bivalvia	Group 1 + Group 2	GO:0031400	negative regulation of protein modification process	14	3	0.61	0.02049
Bivalvia	Group 1 + Group 2	GO:0061982	meiosis I cell cycle process	14	3	0.61	0.02049
Bivalvia	Group 1 + Group 2	GO:0097190	apoptotic signaling pathway	16	3	0.69	0.02958
Bivalvia	Group 1 + Group 2	GO:0040008	regulation of growth	16	3	0.69	0.02958
Bivalvia	Group 1 + Group 2	GO:0051345	positive regulation of hydrolase activity	17	3	0.74	0.03479
Bivalvia	Group 1 + Group 2	GO:0010257	NADH dehydrogenase complex assembly	17	3	0.74	0.03479
Bivalvia	Group 1 + Group 2	GO:0032981	mitochondrial respiratory chain complex I assembly	17	3	0.74	0.03479
Bivalvia	Group 1 + Group 2	GO:0032880	regulation of protein localization	18	3	0.78	0.04045
Bivalvia	Group 1 + Group 2	GO:0005976	polysaccharide metabolic process	18	3	0.78	0.04045
Bivalvia	Group 1 + Group 2	GO:0048729	tissue morphogenesis	19	3	0.82	0.04655
Bivalvia	Group 1 + Group 2	GO:0018022	peptidyl-lysine methylation	19	3	0.82	0.04655
Bivalvia	Group 1 + Group 2	GO:0000041	transition metal ion transport	19	3	0.82	0.04655
Bivalvia	Group 1 + Group 2	GO:0032488	Cdc42 protein signal transduction	2	0.09	0.00187	0.00187
Bivalvia	Group 1 + Group 2	GO:0022600	digestive system process	2	0.09	0.00187	0.00187
Bivalvia	Group 1 + Group 2	GO:0007697	nuclear migration	2	0.09	0.00187	0.00187
Bivalvia	Group 1 + Group 2	GO:0032282	negative regulation of actin filament bundle assembly	2	0.09	0.00187	0.00187
Bivalvia	Group 1 + Group 2	GO:1905168	negative regulation of double-strand break repair via homologous recombination	2	0.09	0.00187	0.00187
Bivalvia	Group 1 + Group 2	GO:0002064	epithelial cell development	3	3	0.13	0.00544
Bivalvia	Group 1 + Group 2	GO:0061583	tracheola morphogenesis	3	2	0.13	0.00544
Bivalvia	Group 1 + Group 2	GO:0010833	regulation of myotote differentiation	3	2	0.13	0.00544
Bivalvia	Group 1 + Group 2	GO:0010833	telomere maintenance via telomere lengthening	3	2	0.13	0.00544
Bivalvia	Group 1 + Group 2	GO:0000959	mitochondrial RNA metabolic process	3	2	0.13	0.00544
Bivalvia	Group 1 + Group 2	GO:0033617	mitochondrial cytochrome c oxidase assembly	3	2	0.13	0.00544
Bivalvia	Group 1 + Group 2	GO:2000179	positive regulation of neural precursor cell proliferation	3	2	0.13	0.00544
Bivalvia	Group 1 + Group 2	GO:0050777	negative regulation of immune response	3	2	0.13	0.00544
Bivalvia	Group 1 + Group 2	GO:0007095	mitotic G2 DNA damage checkpoint signaling	3	2	0.13	0.00544
Bivalvia	Group 1 + Group 2	GO:2000736	regulation of stem cell differentiation	3	2	0.13	0.00544
Bivalvia	Group 1 + Group 2	GO:0016233	telomere capping	3	2	0.13	0.00544
Bivalvia	Group 1 + Group 2	GO:0045910	negative regulation of DNA recombination	3	2	0.13	0.00544
Bivalvia	Group 1 + Group 2	GO:0051701	biological process involved in interaction with host	3	2	0.13	0.00544
Bivalvia	Group 1 + Group 2	GO:0046677	response to antibiotic	4	4	0.17	0.01058
Bivalvia	Group 1 + Group 2	GO:0046620	regulation of organ growth	4	4	0.17	0.01058
Bivalvia	Group 1 + Group 2	GO:0030514	negative regulation of BMP signaling pathway	4	4	0.17	0.01058
Bivalvia	Group 1 + Group 2	GO:1901678	iron coordination entity transport	4	4	0.17	0.01058
Bivalvia	Group 1 + Group 2	GO:0007519	skeletal muscle tissue development	4	4	0.17	0.01058
Bivalvia	Group 1 + Group 2	GO:0032507	maintenance of protein location in cell	4	4	0.17	0.01058
Bivalvia	Group 1 + Group 2	GO:0007416	synapse assembly	4	4	0.17	0.01058
Bivalvia	Group 1 + Group 2	GO:0098781	ncRNA transcription	4	4	0.17	0.01058
Bivalvia	Group 1 + Group 2	GO:0035023	regulation of Rho protein signal transduction	4	4	0.17	0.01058
Bivalvia	Group 1 + Group 2	GO:0003190	atrioventricular valve formation	4	4	0.17	0.01058
Bivalvia	Group 1 + Group 2	GO:0003181	atrioventricular valve morphogenesis	5	2	0.22	0.01712
Bivalvia	Group 1 + Group 2	GO:0090101	negative regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	5	2	0.22	0.01712
Bivalvia	Group 1 + Group 2	GO:0010001	glial cell differentiation	5	2	0.22	0.01712
Bivalvia	Group 1 + Group 2	GO:00090288	homologous chromosome pairing at meiosis	5	2	0.22	0.01712
Bivalvia	Group 1 + Group 2	GO:0030510	negative regulation of cellular response to growth factor stimulus	5	2	0.22	0.01712
Bivalvia	Group 1 + Group 2	GO:0061371	determination of heart left/right asymmetry	5	2	0.22	0.01712
Bivalvia	Group 1 + Group 2	GO:0042063	gliogenesis	5	2	0.22	0.01712
Bivalvia	Group 1 + Group 2	GO:0007162	endothelial reticulum unfolded protein response	5	2	0.22	0.01712
Bivalvia	Group 1 + Group 2	GO:0042026	protein refolding	5	2	0.22	0.01712
Bivalvia	Group 1 + Group 2	GO:0007129	heart valve morphogenesis	5	2	0.22	0.01712
Bivalvia	Group 1 + Group 2	GO:0003188	atrioventricular valve development	5	2	0.22	0.01712
Bivalvia	Group 1 + Group 2	GO:0003179	embryonic heart tube morphogenesis	5	2	0.22	0.01712
Bivalvia	Group 1 + Group 2	GO:0003143	embryonic heart tube development	5	2	0.22	0.01712

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Bivalvia	Group 1 + Group 2	GO:00390287	regulation of cellular response to growth factor stimulus	5	2	0.22	0.01712
Bivalvia	Group 1 + Group 2	GO:00605358	skeletal muscle organ development	5	2	0.22	0.01712
Bivalvia	Group 1 + Group 2	GO:00346260	cellular response to unfolded protein	5	2	0.22	0.01712
Bivalvia	Group 1 + Group 2	GO:00305059	BMP signaling pathway	6	2	0.26	0.02496
Bivalvia	Group 1 + Group 2	GO:00063660	transcription by RNA polymerase I	6	2	0.26	0.02496
Bivalvia	Group 1 + Group 2	GO:0045185	maintenance of protein location	6	2	0.26	0.02496
Bivalvia	Group 1 + Group 2	GO:0045143	homologous chromosome segregation	6	2	0.26	0.02496
Bivalvia	Group 1 + Group 2	GO:0001889	liver development	6	2	0.26	0.02496
Bivalvia	Group 1 + Group 2	GO:0071772	response to BMP	6	2	0.26	0.02496
Bivalvia	Group 1 + Group 2	GO:0071773	cellular response to BMP stimulus	6	2	0.26	0.02496
Bivalvia	Group 1 + Group 2	GO:0090992	regulation of transmembrane receptor protein serine/threonine kinase signalling pathway	6	2	0.26	0.02496
Bivalvia	Group 1 + Group 2	GO:0003170	heart valve development	6	2	0.26	0.02496
Bivalvia	Group 1 + Group 2	GO:0003012	muscle system process	6	2	0.26	0.02496
Bivalvia	Group 1 + Group 2	GO:0030490	maturation of SSU-rRNA	6	2	0.26	0.02496
Bivalvia	Group 1 + Group 2	GO:0018105	peptidyl-serine phosphorylation	7	2	0.3	0.03395
Bivalvia	Group 1 + Group 2	GO:0032465	regulation of cytokinesis	7	2	0.3	0.03395
Bivalvia	Group 1 + Group 2	GO:0061448	connective tissue development	7	2	0.3	0.03395
Bivalvia	Group 1 + Group 2	GO:0018209	peptidyl-serine modification	7	2	0.3	0.03395
Bivalvia	Group 1 + Group 2	GO:0032102	negative regulation of response to external stimulus	7	2	0.3	0.03395
Bivalvia	Group 1 + Group 2	GO:0061008	hepatociliary system development	7	2	0.3	0.03395
Bivalvia	Group 1 + Group 2	GO:0048638	regulation of developmental growth	7	2	0.3	0.03395
Bivalvia	Group 1 + Group 2	GO:0010212	response to ionizing radiation	7	2	0.3	0.03395
Bivalvia	Group 1 + Group 2	GO:0034329	cell junction assembly	7	2	0.3	0.03395
Bivalvia	Group 1 + Group 2	GO:1801652	response to peptide	8	2	0.35	0.04400
Bivalvia	Group 1 + Group 2	GO:0048732	gland development	8	2	0.35	0.04400
Bivalvia	Group 1 + Group 2	GO:0033157	regulation of intracellular protein transport	8	2	0.35	0.04400
Bivalvia	Group 1 + Group 2	GO:0051302	regulation of cell division	8	2	0.35	0.04400
Bivalvia	Group 1 + Group 2	GO:0001822	kidney development	8	2	0.35	0.04400
Bivalvia	Group 1 + Group 2	GO:0050808	synapse organization	8	2	0.35	0.04400
Bivalvia	Group 1 + Group 2	GO:0035967	cellular response to topologically incorrect protein	8	2	0.35	0.04400
Bivalvia	Group 1 + Group 2	GO:0031032	actomyosin structure organization	8	2	0.35	0.04400
Bivalvia	Group 1 + Group 2	GO:0001503	ossification	8	2	0.35	0.04400
Bivalvia	Group 1 + Group 2	GO:0000271	polysaccharide biosynthetic process	8	2	0.35	0.04400
Bivalvia	Group 1 + Group 2	GO:0008593	regulation of Notch signaling pathway	8	2	0.35	0.04400
Bivalvia	Group 1 + Group 2	GO:0045596	negative regulation of cell differentiation	8	2	0.35	0.04400
Bivalvia	Group 1 + Group 2	GO:0050691	regulation of defense response to virus by host	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:1800220	semaphorin-pepxin signaling pathway involved in bone trabecula morphogenesis	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0046621	negative regulation of organ growth	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0048799	animal organ maturation	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0046402	O antigen metabolic process	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0055272	response to host immune response	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0015743	malate transport	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0043628	regulatory ncRNA 3'-end processing	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0046622	inhibitory synapse assembly	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0009633	mitochondrial RNA processing	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0002230	positive regulation of defense response to virus by host	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0048742	regulation of skeletal muscle fiber development	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0048743	positive regulation of skeletal muscle fiber development	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:1802041	regulation of extrinsic apoptotic signaling pathway via death domain receptors	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0075136	response to host	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:1802285	semaphorin-pepxin signaling pathway involved in neuron projection guidance	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0033687	osteoblast proliferation	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:1802287	semaphorin-pepxin signaling pathway involved in axon guidance	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0033688	regulation of osteoblast proliferation	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0033689	negative regulation of osteoblast proliferation	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0070212	protein poly-ADP-ribosylation	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0072013	malate transmembrane transport	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0071423	lipopolysaccharide metabolic process	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:008653	beta-lactam antibiotic metabolic process	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0030653	beta-lactam antibiotic catabolic process	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0030655	negative regulation of exit from mitosis	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0008625	extrinsic apoptotic signaling pathway via death domain receptors	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0043144	(s)RNA processing	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0031848	protection from non-homologous end joining at telomere	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0010526	retrotransposon silencing	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0061668	mitochondrial ribosome assembly	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0035418	protein localization to synapse	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0061430	bone trabecula morphogenesis	1	1	0.04	0.04330

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Bivalvia	Group 1 + Group 2	GO:0006356	regulation of transcription by RNA polymerase I	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:1901269	lipooligosaccharide metabolic process	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0035622	intrahepatic bile duct development	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0016444	somatic cell DNA recombination	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:1901271	lipooligosaccharide biosynthetic process	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0052173	response to defenses of other organism	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0003382	epithelial cell morphogenesis	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0070977	bone maturation	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0031571	mitotic G1 DNA damage checkpoint signaling	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0043247	telomere maintenance in response to DNA damage	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0010669	epithelial structure maintenance	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0009301	snRNA transcription	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0042149	cellular response to glucose starvation	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:1902914	regulation of protein polyubiquitination	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:1902915	negative regulation of protein polyubiquitination	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0032196	transposition	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0032197	retrotransposition	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0034964	box H/ACA sno(s)RNA processing	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0007168	receptor guanylyl cyclase signaling pathway	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0017001	antibiotic catabolic process	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0030277	maintenance of gastrointestinal epithelium	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:000495	box H/ACA sno(s)RNA 3'-end processing	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:1903513	endoplasmic reticulum to cytosol transport	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:1990481	mRNA pseudouridine synthesis	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0044481	mitotic G1/S transition checkpoint signaling	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0051155	positive regulation of striated muscle cell differentiation	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0007140	male meiotic nuclear division	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0050772	positive regulation of axogenesis	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0090646	mitochondrial tRNA processing	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0099172	presynapse organization	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0032065	maintenance of protein location in cell cortex	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:1904152	regulation of retrograde protein transport, ER to cytosol	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0097222	mitochondrial mRNA polyadenylation	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0051149	positive regulation of muscle cell differentiation	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0020033	antigenic variation	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0052200	response to host defenses	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0016074	ossification involved in bone maturation	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0043931	miRNA-mediated gene silencing by mRNA destabilization	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0035279	common bile duct development	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0061009	presynapse assembly	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0093054	regulation of skeletal muscle tissue development	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0048641	positive regulation of skeletal muscle tissue development	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0048643	regulation of muscle tissue development	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:1901861	positive regulation of muscle tissue development	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:1901863	regulation of muscle organ development	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0048634	snRNA 3'-end processing	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0031120	box H/ACA sno(s)RNA metabolic process	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0031126	functional cell wall polysaccharide biosynthetic process	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0033979	nuclear matrix anchoring at nuclear membrane	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0071947	protein deubiquitination involved in ubiquitin-dependent protein catabolic process	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0042783	fungal-type cell wall polysaccharide metabolic process	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0030970	exansion of host immune response	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0051274	retrograde protein transport, ER to cytosol	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0051278	beta-glucan biosynthetic process	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0009103	fungal-type cell wall polysaccharide biosynthetic process	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0009292	lipopolysaccharide biosynthetic process	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0033979	lactam catalytic process	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0042783	(1->3)-beta-D-glucan metabolic process	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0030970	antibiotic metabolic process	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0016999	lactam metabolic process	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0072338	adhesion of symbiont to host cell	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:1902414	protein localization to cell junction	1	0.04	0.04330	
Bivalvia	Group 1 + Group 2	GO:0009272	fungal-type cell wall biogenesis	1	0.04	0.04330	

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Bivalvia	Group 1 + Group 2	GO:0009243	O antigen biosynthetic process	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0043578	nuclear matrix organization	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0009245	lipid A biosynthetic process	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0044406	adhesion or symbiosis to host	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0032978	protein insertion into membrane from inner side	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0032979	protein insertion into mitochondrial inner membrane from matrix	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0046493	lipid A metabolic process	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0072695	regulation of DNA recombination at telomere	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:180044	regulation of protein K63-linked ubiquitination	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0035871	protein K11-linked deubiquitination	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:190045	negative regulation of protein K63-linked ubiquitination	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0048239	negative regulation of DNA recombination at telomere	1	1	0.04	0.04330
Bivalvia	Group 1 + Group 2	GO:0045132	meiotic chromosome segregation	64	11	2.63	0.00145
Drosophila	Group 1 + Group 2	GO:0000819	sister chromatid segregation	140	11	5.75	0.02927
Drosophila	Group 1 + Group 2	GO:0070192	chromosome organization involved in meiotic cell cycle	54	9	2.22	0.00849
Drosophila	Group 1 + Group 2	GO:0007131	reciprocal meiotic recombination	37	7	1.52	0.00066
Drosophila	Group 1 + Group 2	GO:0007143	female meiotic nuclear division	54	6	2.22	0.02270
Drosophila	Group 1 + Group 2	GO:0035967	cellular response to topologically incorrect protein	44	5	1.81	0.03334
Drosophila	Group 1 + Group 2	GO:0035966	response to topologically incorrect protein	47	5	1.93	0.04266
Drosophila	Group 1 + Group 2	GO:0007141	male meiosis I	13	4	0.53	0.00150
Drosophila	Group 1 + Group 2	GO:0140543	positive regulation of piRNA transcription	12	3	6.9e-05	0.02169
Drosophila	Group 1 + Group 2	GO:0010526	retrotransposon silencing	15	3	0.62	0.02264
Drosophila	Group 1 + Group 2	GO:0007130	synaptonemal complex assembly	19	3	0.33	0.00331
Drosophila	Group 1 + Group 2	GO:0030719	P granule organization	10	3	0.41	0.00666
Drosophila	Group 1 + Group 2	GO:0035966	cellular response to misfolded protein	11	3	0.45	0.00888
Drosophila	Group 1 + Group 2	GO:0071218	response to misfolded protein	12	3	0.49	0.01149
Drosophila	Group 1 + Group 2	GO:0051788	meiosis II	12	3	0.49	0.01149
Drosophila	Group 1 + Group 2	GO:0007135	centromere complex assembly	15	3	0.62	0.02169
Drosophila	Group 1 + Group 2	GO:0034505	male germ-line cyst formation	19	3	0.78	0.04094
Drosophila	Group 1 + Group 2	GO:0007130	negative regulation of entry into reproductive diapause	2	2	0.08	0.00168
Drosophila	Group 1 + Group 2	GO:0030719	histone H4-K16 acetylation	5	2	0.21	0.01551
Drosophila	Group 1 + Group 2	GO:0051382	entry into diapause	2	2	0.29	0.03084
Drosophila	Group 1 + Group 2	GO:0055116	response to odorant	2	2	0.21	0.03084
Drosophila	Group 1 + Group 2	GO:0071712	meiotic sister chromatid cohesion	6	2	0.25	0.02264
Drosophila	Group 1 + Group 2	GO:0061963	ER-associated misfolded protein catabolic process	6	2	0.25	0.02264
Drosophila	Group 1 + Group 2	GO:0043984	regulation of entry into reproductive diapause	6	2	0.25	0.02264
Drosophila	Group 1 + Group 2	GO:0055115	male germ-line cyst formation	6	2	0.25	0.02264
Drosophila	Group 1 + Group 2	GO:1900834	negative regulation of entry into reproductive diapause	7	2	0.29	0.03084
Drosophila	Group 1 + Group 2	GO:0051177	histone H4-K16 acetylation	7	2	0.33	0.04002
Drosophila	Group 1 + Group 2	GO:0042795	entry into reproductive diapause	8	2	0.33	0.04002
Drosophila	Group 1 + Group 2	GO:0022611	dormancy process	8	2	0.33	0.04002
Drosophila	Group 1 + Group 2	GO:0009301	snRNA transcription	8	2	0.33	0.04002
Drosophila	Group 1 + Group 2	GO:0045144	meiotic sister chromatid segregation	8	2	0.33	0.04002
Drosophila	Group 1 + Group 2	GO:0001015	snRNA transcription by RNA polymerase II	1	1	0.04	0.04109
Drosophila	Group 1 + Group 2	GO:0010778	meiotic DNA repair synthesis involved in reciprocal meiotic recombination	1	1	0.04	0.04109
Drosophila	Group 1 + Group 2	GO:0072765	centromere localization	1	1	0.04	0.04109
Drosophila	Group 1 + Group 2	GO:0009302	sn(s)RNA transcription	8	2	0.33	0.04109
Drosophila	Group 1 + Group 2	GO:0051308	male meiosis chromosome separation	8	2	0.33	0.04109
Drosophila	Group 1 + Group 2	GO:0098653	centromere clustering	1	1	0.04	0.04109
Drosophila	Group 1 + Group 2	GO:0051415	microtubule nucleation by interphase microtubule organizing center	1	1	0.04	0.04109
Mammalia	Group 1 + Group 2	GO:0006955	immune response	1297	145	4.02	0.02661
Mammalia	Group 1 + Group 2	GO:0098542	defense response to other organism	853	112	31.58	0.02801
Mammalia	Group 1 + Group 2	GO:0045087	innate immune response	647	82	23.95	8.5e-10
Mammalia	Group 1 + Group 2	GO:0009181	regulation of cytokine production	630	51	23.33	0.04426
Mammalia	Group 1 + Group 2	GO:0042742	defense response to bacterium	233	45	8.63	1.7e-07
Mammalia	Group 1 + Group 2	GO:0006954	inflammatory response	642	45	23.77	0.01735
Mammalia	Group 1 + Group 2	GO:0019221	cytokine-mediated signaling pathway	382	44	14.14	3.9e-07
Mammalia	Group 1 + Group 2	GO:0002250	adaptive immune response	342	44	12.66	1.3e-07
Mammalia	Group 1 + Group 2	GO:0098542	positive regulation of cytokine production	478	31	17.7	0.02801
Mammalia	Group 1 + Group 2	GO:0070661	lymphocyte mediated immunity	402	41	14.88	0.02723
Mammalia	Group 1 + Group 2	GO:0002449	regulation of leukocyte proliferation	308	37	11.4	0.04426
Mammalia	Group 1 + Group 2	GO:0042110	T cell activation	432	35	15.99	0.02864
Mammalia	Group 1 + Group 2	GO:0048232	male gamete generation	491	34	9.52	1.9e-07
Mammalia	Group 1 + Group 2	GO:0007283	spermatogenesis	478	31	17.7	0.02801
Mammalia	Group 1 + Group 2	GO:0001819	leukocyte proliferation	273	29	10.11	0.01285
Mammalia	Group 1 + Group 2	GO:0002697	regulation of immune effector process	221	29	8.18	0.04833
Mammalia	Group 1 + Group 2	GO:0070663	regulation of inflammatory response	212	25	7.85	0.01870
Mammalia	Group 1 + Group 2	GO:0050727	positive regulation of defense response	300	24	11.11	0.02335
Mammalia	Group 1 + Group 2	GO:0031349	immune response-regulating cell surface receptor signaling pathway	240	24	8.89	0.01239
Mammalia	Group 1 + Group 2	GO:0002768		177	22	6.55	0.00336

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Mammalia	Group 1 + Group 2	GO:0050829	defense response to Gram-negative bacterium	66	17	2.44	1.7e-10
Mammalia	Group 1 + Group 2	GO:0071222	cellular response to lipopolysaccharide	164	17	6.07	0.00012
Mammalia	Group 1 + Group 2	GO:0010466	negative regulation of peptidase activity	163	16	6.04	0.00036
Mammalia	Group 1 + Group 2	GO:0002429	immune response-activating cell surface receptor signaling pathway	164	16	6.07	0.00243
Mammalia	Group 1 + Group 2	GO:1803555	regulation of tumor necrosis factor superfamily cytokine production	137	16	5.07	0.01244
Mammalia	Group 1 + Group 2	GO:0071706	tumor necrosis factor superfamily cytokine production	137	16	5.07	0.01244
Mammalia	Group 1 + Group 2	GO:0070665	positive regulation of leukocyte proliferation	132	16	4.89	0.02765
Mammalia	Group 1 + Group 2	GO:0045089	positive regulation of innate immune response	113	16	4.18	0.03224
Mammalia	Group 1 + Group 2	GO:0071356	cellular response to tumor necrosis factor	175	15	6.48	0.00219
Mammalia	Group 1 + Group 2	GO:0002695	negative regulation of leukocyte activation	148	15	5.48	0.01151
Mammalia	Group 1 + Group 2	GO:0002456	T cell mediated immunity	82	15	3.04	0.01605
Mammalia	Group 1 + Group 2	GO:0002705	positive regulation of leukocyte mediated immunity	113	15	4.18	0.01837
Mammalia	Group 1 + Group 2	GO:0032680	regulation of tumor necrosis factor production	133	15	4.92	0.03262
Mammalia	Group 1 + Group 2	GO:0032640	tumor necrosis factor production	133	15	4.92	0.03262
Mammalia	Group 1 + Group 2	GO:0050866	negative regulation of cell activation	165	15	6.11	0.04048
Mammalia	Group 1 + Group 2	GO:0031341	regulation of cell killing	71	14	2.63	0.00628
Mammalia	Group 1 + Group 2	GO:0001818	negative regulation of cytokine production	225	14	8.33	0.04065
Mammalia	Group 1 + Group 2	GO:0050820	defense response to Gram-positive bacterium	87	13	3.22	1.8e-05
Mammalia	Group 1 + Group 2	GO:0002286	T cell activation involved in immune response	94	13	3.48	0.02934
Mammalia	Group 1 + Group 2	GO:0050909	sensory perception of taste	52	13	1.93	0.04957
Mammalia	Group 1 + Group 2	GO:0007259	cell surface receptor signaling pathway via JAK-STAT	134	12	4.96	0.00411
Mammalia	Group 1 + Group 2	GO:0001973	antigen receptor-mediated signaling pathway	130	12	4.81	0.01823
Mammalia	Group 1 + Group 2	GO:00019731	antibacterial humoral response	40	11	1.48	0.00707
Mammalia	Group 1 + Group 2	GO:0030101	natural killer cell activation	66	11	2.44	2.7e-05
Mammalia	Group 1 + Group 2	GO:0442102	positive regulation of T cell proliferation	85	11	3.15	0.00029
Mammalia	Group 1 + Group 2	GO:0050852	defense response to fungus	48	11	1.78	0.00080
Mammalia	Group 1 + Group 2	GO:0071346	cellular response to type II interferon	98	11	3.63	0.00099
Mammalia	Group 1 + Group 2	GO:0008584	male gonad development	120	11	4.44	0.00748
Mammalia	Group 1 + Group 2	GO:0002820	negative regulation of adaptive immune response	42	11	1.56	0.00768
Mammalia	Group 1 + Group 2	GO:0043123	positive regulation of canonical NF-kappaB signal transduction	162	11	2.6	0.03860
Mammalia	Group 1 + Group 2	GO:0042100	B cell proliferation	76	11	2.81	0.04485
Mammalia	Group 1 + Group 2	GO:0001580	detection of chemical stimulus involved in sensory perception of bitter taste	29	10	1.07	4.7e-08
Mammalia	Group 1 + Group 2	GO:0030593	neutrophil chemotaxis	80	10	2.96	0.00071
Mammalia	Group 1 + Group 2	GO:0006936	complement activation	47	10	1.74	0.00078
Mammalia	Group 1 + Group 2	GO:0032760	positive regulation of tumor necrosis factor production	84	10	3.11	0.00105
Mammalia	Group 1 + Group 2	GO:0071347	cellular response to interleukin-1	84	10	3.11	0.00105
Mammalia	Group 1 + Group 2	GO:0050729	positive regulation of inflammatory response	122	10	4.52	0.01515
Mammalia	Group 1 + Group 2	GO:0002823	negative regulation of adaptive immune response [...]	39	10	1.44	0.03055
Mammalia	Group 1 + Group 2	GO:0050688	regulation of defense response to virus	62	10	2.3	0.03820
Mammalia	Group 1 + Group 2	GO:0002718	cytokine production involved in immune response	98	10	3.63	0.04352
Mammalia	Group 1 + Group 2	GO:0002367	binding of sperm to zona pellucida	35	9	1.3	3.7e-06
Mammalia	Group 1 + Group 2	GO:0007739	antimicrobial humoral immune response mediated by antimicrobial peptide	64	9	2.37	0.00055
Mammalia	Group 1 + Group 2	GO:0016184	killing of cells of another organism	64	9	2.37	0.00055
Mammalia	Group 1 + Group 2	GO:0031640	negative regulation of T cell activation	96	9	3.55	0.00913
Mammalia	Group 1 + Group 2	GO:0050868	regulation of leukocyte mediated cytotoxicity	59	9	2.18	0.01359
Mammalia	Group 1 + Group 2	GO:0001910	cilium movement involved in cell motility	114	9	4.22	0.02562
Mammalia	Group 1 + Group 2	GO:0001959	fusion of sperm to egg plasma membrane involved in single fertilization	119	9	4.41	0.03259
Mammalia	Group 1 + Group 2	GO:0030317	negative regulation of viral genome replication	39	8	1.44	7.5e-05
Mammalia	Group 1 + Group 2	GO:0002218	activation of innate immune responses	119	9	4.41	0.03259
Mammalia	Group 1 + Group 2	GO:0032757	positive regulation of interleukin-8 production	120	8	1.67	0.00022
Mammalia	Group 1 + Group 2	GO:0072678	regulation of B cell proliferation	54	8	2	0.00078
Mammalia	Group 1 + Group 2	GO:0006759	regulation of response to cytokine stimulus	124	9	4.59	0.04078
Mammalia	Group 1 + Group 2	GO:0062924	cilium movement involved in cell motility	128	8	2.41	0.00158
Mammalia	Group 1 + Group 2	GO:0007342	fusion of sperm to egg plasma membrane involved in single fertilization	25	8	0.93	2.1e-06
Mammalia	Group 1 + Group 2	GO:0045071	negative regulation of viral genome replication	69	8	2.55	0.00385
Mammalia	Group 1 + Group 2	GO:0002920	regulation of humoral immune response	36	8	1.33	0.00512
Mammalia	Group 1 + Group 2	GO:0043303	mast cell degranulation	43	8	1.59	0.00550
Mammalia	Group 1 + Group 2	GO:0002251	organ or tissue specific immune response	22	8	0.81	0.00652
Mammalia	Group 1 + Group 2	GO:0002886	regulation of myeloid leukocyte mediated immunity	45	8	1.67	0.00784
Mammalia	Group 1 + Group 2	GO:0043030	regulation of macrophage activation	46	8	2.41	0.00265
Mammalia	Group 1 + Group 2	GO:0002526	acute inflammatory response	83	8	1.7	0.00922
Mammalia	Group 1 + Group 2	GO:0032649	regulation of type II interferon production	88	8	3.26	0.01162
Mammalia	Group 1 + Group 2	GO:0032609	type II interferon production	88	8	3.26	0.01615

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Mammalia	Group 1 + Group 2	GO:0050691	regulation of defense response to virus by host	37	8	1.37	0.04975
Mammalia	Group 1 + Group 2	GO:0042267	natural killer cell mediated cytotoxicity	44	7	1.63	0.00107
Mammalia	Group 1 + Group 2	GO:0048247	lymphocyte chemotaxis	49	7	1.81	0.00204
Mammalia	Group 1 + Group 2	GO:0042119	neutrophil activation	32	7	1.18	0.00786
Mammalia	Group 1 + Group 2	GO:0032945	negative regulation of mononuclear cell proliferation	63	7	2.33	0.00845
Mammalia	Group 1 + Group 2	GO:0002720	positive regulation of cytokine production involved in immune response	66	7	2.44	0.01083
Mammalia	Group 1 + Group 2	GO:0032456	response to interferon-beta	26	7	0.96	0.01252
Mammalia	Group 1 + Group 2	GO:0050853	B cell receptor signaling pathway	45	7	1.67	0.01683
Mammalia	Group 1 + Group 2	GO:0098586	cellular response to virus	78	7	2.89	0.02523
Mammalia	Group 1 + Group 2	GO:0046425	regulation of receptor signaling pathway via JAK-STAT	78	7	2.89	0.02523
Mammalia	Group 1 + Group 2	GO:0002385	mucosal immune response	19	7	0.7	0.03253
Mammalia	Group 1 + Group 2	GO:0032755	positive regulation of interleukin-6 production	83	7	3.07	0.03395
Mammalia	Group 1 + Group 2	GO:1801222	regulation of non-canonical NF-kappaB signal transduction	89	7	3.3	0.04679
Mammalia	Group 1 + Group 2	GO:0048245	eosinophil chemotaxis	17	6	0.63	2.2e-05
Mammalia	Group 1 + Group 2	GO:0048240	sperm capacitation	25	6	0.93	0.00024
Mammalia	Group 1 + Group 2	GO:0070269	pyroptotic inflammatory response	26	6	0.96	0.00031
Mammalia	Group 1 + Group 2	GO:0002230	positive regulation of defense response to virus by host	27	6	1	0.00038
Mammalia	Group 1 + Group 2	GO:0031295	T cell costimulation	27	6	1	0.00038
Mammalia	Group 1 + Group 2	GO:0002825	regulation of T-helper 1 type immune response	29	6	1.07	0.00058
Mammalia	Group 1 + Group 2	GO:0002691	regulation of cellular extravasation	38	6	1.41	0.00252
Mammalia	Group 1 + Group 2	GO:0140374	antiviral innate immune response	41	6	1.52	0.00374
Mammalia	Group 1 + Group 2	GO:0002639	positive regulation of immunoglobulin production	44	6	1.63	0.00355
Mammalia	Group 1 + Group 2	GO:0032731	positive regulation of interleukin-1 beta production	46	6	1.7	0.00667
Mammalia	Group 1 + Group 2	GO:1804894	positive regulation of receptor signaling pathway via STAT	46	6	1.7	0.00667
Mammalia	Group 1 + Group 2	GO:0034113	heterotypic cell-cell adhesion	47	6	1.74	0.00742
Mammalia	Group 1 + Group 2	GO:0002548	monocyte chemotaxis	50	6	1.85	0.01000
Mammalia	Group 1 + Group 2	GO:0001961	positive regulation of cytokine-mediated signaling pathway	50	6	1.85	0.01000
Mammalia	Group 1 + Group 2	GO:0030449	regulation of complement activation	18	6	0.67	0.01079
Mammalia	Group 1 + Group 2	GO:0042531	regulation of tyrosine phosphorylation of STAT protein	51	6	1.89	0.01099
Mammalia	Group 1 + Group 2	GO:0060337	type I interferon-mediated signaling pathway	52	6	1.93	0.01205
Mammalia	Group 1 + Group 2	GO:0071357	cellular response to type I interferon	52	6	1.93	0.01205
Mammalia	Group 1 + Group 2	GO:0001914	regulation of T cell mediated cytotoxicity	25	6	0.93	0.01256
Mammalia	Group 1 + Group 2	GO:0051293	establishment of spindle localization	54	6	1.85	0.01437
Mammalia	Group 1 + Group 2	GO:0051873	regulation of complement activation	23	6	0.85	0.01626
Mammalia	Group 1 + Group 2	GO:0060760	positive regulation of tyrosine phosphorylation of STAT protein	57	6	2.11	0.01841
Mammalia	Group 1 + Group 2	GO:0150077	type I interferon-mediated signaling pathway	28	6	1.04	0.01949
Mammalia	Group 1 + Group 2	GO:0034340	regulation of type I interferon	58	6	2.15	0.01991
Mammalia	Group 1 + Group 2	GO:0051653	spindle localization	59	6	2.18	0.02150
Mammalia	Group 1 + Group 2	GO:0001895	retina homeostasis	60	6	2.22	0.02316
Mammalia	Group 1 + Group 2	GO:0071260	cellular response to mechanical stimulus	60	6	2.22	0.02316
Mammalia	Group 1 + Group 2	GO:0050672	negative regulation of lymphocyte proliferation	62	6	2.3	0.02676
Mammalia	Group 1 + Group 2	GO:0032729	regulation of type II interferon production	62	6	2.3	0.02676
Mammalia	Group 1 + Group 2	GO:0042509	regulation of tyrosine phosphorylation of STAT protein	62	6	2.3	0.02676
Mammalia	Group 1 + Group 2	GO:0007260	tyrosine phosphorylation of STAT protein	66	6	2.44	0.03500
Mammalia	Group 1 + Group 2	GO:0002227	innate immune response in mucosa	11	5	0.41	2.6e-05
Mammalia	Group 1 + Group 2	GO:0002830	positive regulation of type 2 immune response	14	5	0.52	0.00010
Mammalia	Group 1 + Group 2	GO:0035455	response to interferon-alpha	14	5	0.52	0.00010
Mammalia	Group 1 + Group 2	GO:00333005	regulation of tyrosine phosphorylation of STAT protein	17	5	0.63	0.00029
Mammalia	Group 1 + Group 2	GO:0061760	antifungal innate immune response	17	5	0.63	0.00029
Mammalia	Group 1 + Group 2	GO:0035458	cellular response to interferon-beta	21	5	0.78	0.00085
Mammalia	Group 1 + Group 2	GO:0046596	regulation of viral entry into host cell	31	5	1.15	0.00523
Mammalia	Group 1 + Group 2	GO:0007340	acrosome reaction	32	5	1.18	0.00601
Mammalia	Group 1 + Group 2	GO:0006968	cellular defense response	34	5	1.26	0.00782
Mammalia	Group 1 + Group 2	GO:0043666	microglial cell activation	45	5	1.67	0.02467
Mammalia	Group 1 + Group 2	GO:0002701	negative regulation of tumor necrosis factor production	46	5	1.7	0.02686
Mammalia	Group 1 + Group 2	GO:0140632	positive regulation of dephosphorylation	47	5	1.74	0.02918
Mammalia	Group 1 + Group 2	GO:1903556	negative regulation of tumor necrosis factor superfamily cytokine production	49	5	1.74	0.02918
Mammalia	Group 1 + Group 2	GO:0032653	regulation of lymphocyte apoptotic process	50	5	1.81	0.03417
Mammalia	Group 1 + Group 2	GO:0032613	T cell extravasation	13	4	0.48	0.03686
Mammalia	Group 1 + Group 2	GO:0043666	regulation of phosphoprotein phosphatase activity	15	4	0.56	0.00183
Mammalia	Group 1 + Group 2	GO:0002710	negative regulation of T cell mediated immunity	17	4	0.63	0.00301
Mammalia	Group 1 + Group 2	GO:0045624	T-helper 1 cell lineage commitment	17	4	0.63	0.00301
Mammalia	Group 1 + Group 2	GO:1903659	positive regulation of Thelper cell differentiation	17	4	0.63	0.00301
Mammalia	Group 1 + Group 2	GO:1903659	regulation of complement-dependent cytotoxicity	5	4	0.19	0.00398

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Mammalia	Group 1 + Group 2	GO:0008228	opsonization	19	4	0.7	0.00462
Mammalia	Group 1 + Group 2	GO:0043302	positive regulation of leukocyte degranulation	20	4	0.74	0.00561
Mammalia	Group 1 + Group 2	GO:0001916	positive regulation of T cell mediated cytotoxicity	20	4	0.74	0.00561
Mammalia	Group 1 + Group 2	GO:0001562	response to protozoan infection	21	4	0.78	0.00673
Mammalia	Group 1 + Group 2	GO:0002717	positive regulation of natural killer cell mediated immunity	21	4	0.78	0.00673
Mammalia	Group 1 + Group 2	GO:1903901	negative regulation of viral life cycle	21	4	0.78	0.00673
Mammalia	Group 1 + Group 2	GO:0035821	modulation of process of another organism	21	4	0.78	0.00673
Mammalia	Group 1 + Group 2	GO:0043304	regulation of mast cell degranulation	23	4	0.85	0.00940
Mammalia	Group 1 + Group 2	GO:0032740	positive regulation of interleukin-17 production	23	4	0.85	0.00940
Mammalia	Group 1 + Group 2	GO:0043032	positive regulation of macrophage activation	23	4	0.85	0.00940
Mammalia	Group 1 + Group 2	GO:007498	interleukin-1-mediated signaling pathway	24	4	0.89	0.01096
Mammalia	Group 1 + Group 2	GO:0010922	positive regulation of phosphatase activity	26	4	0.96	0.01454
Mammalia	Group 1 + Group 2	GO:0019884	antigen processing and presentation of exogenous antigen	28	4	1.04	0.01880
Mammalia	Group 1 + Group 2	GO:0002446	neutrophil mediated immunity	29	4	1.07	0.02118
Mammalia	Group 1 + Group 2	GO:0032743	positive regulation of interleukin-2 production	30	4	1.11	0.02375
Mammalia	Group 1 + Group 2	GO:2000352	negative regulation of endothelial cell apoptotic process	30	4	1.11	0.02375
Mammalia	Group 1 + Group 2	GO:1900225	regulation of NLRP3 inflammasome complex assembly	31	4	1.11	0.02375
Mammalia	Group 1 + Group 2	GO:0043320	response to exogenous dsRNA	31	4	1.15	0.02649
Mammalia	Group 1 + Group 2	GO:0032733	positive regulation of interleukin-10 production	32	4	1.18	0.02942
Mammalia	Group 1 + Group 2	GO:0046006	regulation of activated T cell proliferation	32	4	1.18	0.02942
Mammalia	Group 1 + Group 2	GO:0032814	regulation of natural killer cell activation	32	4	1.22	0.03252
Mammalia	Group 1 + Group 2	GO:0044546	NLRP3 inflammasome complex assembly	33	4	1.26	0.03582
Mammalia	Group 1 + Group 2	GO:0050798	activated T cell proliferation	34	4	1.26	0.03930
Mammalia	Group 1 + Group 2	GO:0046636	negative regulation of alpha-beta T cell activation	35	4	1.3	0.04297
Mammalia	Group 1 + Group 2	GO:0043331	response to dsRNA	36	4	1.33	0.04297
Mammalia	Group 1 + Group 2	GO:0035307	positive regulation of protein dephosphorylation	36	4	1.33	0.04297
Mammalia	Group 1 + Group 2	GO:0030890	positive regulation of B cell proliferation	37	4	1.37	0.04682
Mammalia	Group 1 + Group 2	GO:0038095	Fc-epsilon receptor signaling pathway	5	3	0.19	0.00048
Mammalia	Group 1 + Group 2	GO:0045098	negative regulation of complement activation, classical pathway	5	3	0.19	0.00048
Mammalia	Group 1 + Group 2	GO:0038156	interleukin-3-mediated signaling pathway	7	3	0.26	0.00158
Mammalia	Group 1 + Group 2	GO:0051838	cytosis by host of symbiont cells	8	3	0.3	0.00246
Mammalia	Group 1 + Group 2	GO:0006977	DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	8	3	0.3	0.00246
Mammalia	Group 1 + Group 2	GO:0097527	necroptotic signaling pathway	8	3	0.3	0.00246
Mammalia	Group 1 + Group 2	GO:0006924	activation-induced cell death of T cells	10	3	0.37	0.00499
Mammalia	Group 1 + Group 2	GO:1802563	regulation of neutrophil activation	10	3	0.37	0.00499
Mammalia	Group 1 + Group 2	GO:0010526	retrotransposon silencing	10	3	0.37	0.00499
Mammalia	Group 1 + Group 2	GO:0032754	positive regulation of interleukin-5 production	10	3	0.37	0.00499
Mammalia	Group 1 + Group 2	GO:2000551	regulation of T-helper 2 cell cytokine production	10	3	0.37	0.00499
Mammalia	Group 1 + Group 2	GO:1901731	positive regulation of platelet aggregation	10	3	0.37	0.00499
Mammalia	Group 1 + Group 2	GO:0032736	positive regulation of interleukin-13 production	11	3	0.41	0.00667
Mammalia	Group 1 + Group 2	GO:0007343	egg activation	11	3	0.41	0.00667
Mammalia	Group 1 + Group 2	GO:0035723	interleukin-15-mediated signaling pathway	11	3	0.41	0.00667
Mammalia	Group 1 + Group 2	GO:0150079	negative regulation of neuroinflammation response	12	3	0.44	0.00865
Mammalia	Group 1 + Group 2	GO:0060046	regulation of acrosome reaction	13	3	0.48	0.01094
Mammalia	Group 1 + Group 2	GO:0002323	natural killer cell activation involved in immune response	13	3	0.48	0.01094
Mammalia	Group 1 + Group 2	GO:0030898	negative regulation of B cell proliferation	13	3	0.48	0.01094
Mammalia	Group 1 + Group 2	GO:0006957	complement activation, alternative pathway	14	3	0.52	0.01354
Mammalia	Group 1 + Group 2	GO:0043306	positive regulation of mast cell degranulation	14	3	0.52	0.01354
Mammalia	Group 1 + Group 2	GO:1903027	regulation of opsonization	14	3	0.59	0.01973
Mammalia	Group 1 + Group 2	GO:0023035	CD40 signaling pathway	14	3	0.52	0.01354
Mammalia	Group 1 + Group 2	GO:0033008	positive regulation of mast cell activation involved in immune response	15	3	0.59	0.01973
Mammalia	Group 1 + Group 2	GO:0045064	T-helper 2 cell differentiation	15	3	0.56	0.01647
Mammalia	Group 1 + Group 2	GO:0043031	negative regulation of macrophage activation	16	3	0.56	0.01647
Mammalia	Group 1 + Group 2	GO:0002730	regulation of dendritic cell cytokine production	16	3	0.59	0.01973
Mammalia	Group 1 + Group 2	GO:0032516	positive regulation of phosphoprotein phosphatase activity	16	3	0.59	0.01973
Mammalia	Group 1 + Group 2	GO:0046597	chronic inflammatory response	17	3	0.63	0.02331
Mammalia	Group 1 + Group 2	GO:0043045	epigenetic programming of gene expression	17	3	0.63	0.02331
Mammalia	Group 1 + Group 2	GO:0002371	dendritic cell cytokine production	18	3	0.67	0.02722
Mammalia	Group 1 + Group 2	GO:1901538	changes to DNA methylation involved in embryo development	18	3	0.67	0.02722
Mammalia	Group 1 + Group 2	GO:0002888	positive regulation of myeloid leukocyte mediated immunity	19	3	0.7	0.03145
Mammalia	Group 1 + Group 2	GO:0045342	type I interferon-mediated signaling pathway	19	3	0.7	0.03145
Mammalia	Group 1 + Group 2	GO:0040119	positive regulation of type I interferon-mediated signaling pathway	19	3	0.7	0.03145
Mammalia	Group 1 + Group 2	GO:0060340	type II interferon-mediated signaling pathway	19	3	0.7	0.03145
Mammalia	Group 1 + Group 2	GO:0060333					

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Mammalia	Group 1 + Group 2	GO:0032753	positive regulation of interleukin-4 production	20	3	0.74	0.03602
Mammalia	Group 1 + Group 2	GO:0045063	T-helper 1 cell differentiation	21	3	0.78	0.04090
Mammalia	Group 1 + Group 2	GO:0042104	positive regulation of activated T cell proliferation	22	3	0.81	0.04609
Mammalia	Group 1 + Group 2	GO:0002726	positive regulation of T cell cytokine production	22	3	0.81	0.04609
Mammalia	Group 1 + Group 2	GO:1803660	negative regulation of complement-dependent cytotoxicity	2	0.07	0.00137	
Mammalia	Group 1 + Group 2	GO:0097528	execution phase of necrosis	2	0.07	0.00137	
Mammalia	Group 1 + Group 2	GO:0032987	positive regulation of myeloid dendritic cell activation	2	0.07	0.00137	
Mammalia	Group 1 + Group 2	GO:0009609	response to symbiotic bacterium	4	0.15	0.00781	
Mammalia	Group 1 + Group 2	GO:0048006	antigen processing and presentation, endogenous lipid antigen via MHC class Ib	4	0.15	0.00781	
Mammalia	Group 1 + Group 2	GO:0050859	negative regulation of B-cell receptor signaling pathway	4	0.15	0.00781	
Mammalia	Group 1 + Group 2	GO:0002765	immune response-inhibiting signal transduction	5	0.19	0.01270	
Mammalia	Group 1 + Group 2	GO:0036015	response to interleukin-3	5	0.19	0.01270	
Mammalia	Group 1 + Group 2	GO:0036016	cellular response to interleukin-3	5	0.19	0.01270	
Mammalia	Group 1 + Group 2	GO:2000416	regulation of eosinophil migration	5	0.19	0.01270	
Mammalia	Group 1 + Group 2	GO:0033634	regulation of cell-cell adhesion mediated by integrin	5	0.19	0.01270	
Mammalia	Group 1 + Group 2	GO:0045630	positive regulation of T-helper 2 cell differentiation	5	0.19	0.01270	
Mammalia	Group 1 + Group 2	GO:0043313	regulation of neutrophil degranulation	6	0.22	0.01859	
Mammalia	Group 1 + Group 2	GO:0033141	positive regulation of peptidyl-serine phosphorylation of STAT protein	6	0.22	0.01859	
Mammalia	Group 1 + Group 2	GO:0002826	negative regulation of T-helper 1 type immune response	6	0.22	0.01859	
Mammalia	Group 1 + Group 2	GO:0106300	protein-DNA covalent cross-linking repair	6	0.22	0.01859	
Mammalia	Group 1 + Group 2	GO:0045341	MHC class I biosynthetic process	6	0.22	0.01859	
Mammalia	Group 1 + Group 2	GO:0045343	regulation of MHC class I biosynthetic process	6	0.22	0.01859	
Mammalia	Group 1 + Group 2	GO:0048007	antigen processing and presentation, exogenous lipid antigen via MHC class Ib	6	0.22	0.01859	
Mammalia	Group 1 + Group 2	GO:0060322	positive regulation of response to type II interferon	6	0.22	0.01859	
Mammalia	Group 1 + Group 2	GO:0060325	positive regulation of type II interferon-mediated signaling pathway	6	0.22	0.01859	
Mammalia	Group 1 + Group 2	GO:2000659	regulation or interleukin-1-mediated signaling pathway	6	0.22	0.01859	
Mammalia	Group 1 + Group 2	GO:0061300	MHC class I biosynthetic process	6	0.22	0.01859	
Mammalia	Group 1 + Group 2	GO:0060545	positive regulation of necroptotic process	6	0.22	0.01859	
Mammalia	Group 1 + Group 2	GO:0043307	regulation of interleukin-1 mediated signaling pathway	7	0.26	0.02539	
Mammalia	Group 1 + Group 2	GO:0002733	antigen processing and presentation, exogenous lipid antigen via MHC class Ib	7	0.26	0.02539	
Mammalia	Group 1 + Group 2	GO:0002735	positive regulation of response to type II interferon	7	0.26	0.02539	
Mammalia	Group 1 + Group 2	GO:2000391	positive regulation of myeloid dendritic cell cytokine production	7	0.26	0.02539	
Mammalia	Group 1 + Group 2	GO:0010940	positive regulation of neutrophil extravasation	7	0.26	0.02539	
Mammalia	Group 1 + Group 2	GO:0060467	negative regulation of necrotic cell death	7	0.26	0.02539	
Mammalia	Group 1 + Group 2	GO:0060468	negative regulation of fertilization	7	0.26	0.02539	
Mammalia	Group 1 + Group 2	GO:0071358	prevention of polyspermy	7	0.26	0.02539	
Mammalia	Group 1 + Group 2	GO:0034342	cellular response to type III interferon	7	0.26	0.02539	
Mammalia	Group 1 + Group 2	GO:1903896	type III interferon-mediated signaling pathway	7	0.26	0.02539	
Mammalia	Group 1 + Group 2	GO:0002372	myeloid dendritic cell cytokine production	7	0.26	0.02539	
Mammalia	Group 1 + Group 2	GO:0062100	positive regulation of programmed necrotic cell death	7	0.26	0.02539	
Mammalia	Group 1 + Group 2	GO:0001781	neutrophil apoptosis	7	0.26	0.02539	
Mammalia	Group 1 + Group 2	GO:0043312	neutrophil degranulation	8	0.3	0.03304	
Mammalia	Group 1 + Group 2	GO:0042796	snRNA transcription by RNA polymerase III	8	0.3	0.03304	
Mammalia	Group 1 + Group 2	GO:0034342	response to type III interferon	8	0.3	0.03304	
Mammalia	Group 1 + Group 2	GO:1903589	positive regulation of blood vessel endothelial cell proliferation involved in sprouting angiogenesis	8	0.3	0.03304	
Mammalia	Group 1 + Group 2	GO:2000389	regulation of neutrophil extravasation	8	0.3	0.03304	
Mammalia	Group 1 + Group 2	GO:0002674	negative regulation of acute inflammatory response	8	0.3	0.03304	
Mammalia	Group 1 + Group 2	GO:0062638	negative regulation of immunoglobulin production	8	0.3	0.03304	
Mammalia	Group 1 + Group 2	GO:0032650	regulation of interleukin-1 alpha production	8	0.3	0.03304	
Mammalia	Group 1 + Group 2	GO:0032610	interleukin-1-alpha production	8	0.3	0.03304	
Mammalia	Group 1 + Group 2	GO:0035771	interleukin-4-mediated signaling pathway	8	0.3	0.03304	
Mammalia	Group 1 + Group 2	GO:0045869	negative regulation of single stranded viral RNA replication via double stranded DNA intermediate	8	0.3	0.03304	
Mammalia	Group 1 + Group 2	GO:0071352	cellular response to interleukin-2	8	0.3	0.03304	
Mammalia	Group 1 + Group 2	GO:2000553	positive regulation of T-helper 2 cell cytokine production	8	0.3	0.03304	
Mammalia	Group 1 + Group 2	GO:0038110	interleukin-2-mediated signaling pathway	8	0.3	0.03304	
Mammalia	Group 1 + Group 2	GO:0033004	negative regulation of mast cell activation	8	0.3	0.03304	
Mammalia	Group 1 + Group 2	GO:0070669	response to interleukin-2	9	0.3	0.04145	
Mammalia	Group 1 + Group 2	GO:0061518	microglial cell proliferation	9	0.3	0.04145	
Mammalia	Group 1 + Group 2	GO:0033139	regulation of peptidyl-serine phosphorylation of STAT protein	9	0.3	0.04145	
Mammalia	Group 1 + Group 2	GO:0044406	adhesion of symbiont to host	9	0.3	0.04145	
Mammalia	Group 1 + Group 2	GO:0010918	positive regulation of mitochondrial membrane potential	9	0.3	0.04145	
Mammalia	Group 1 + Group 2	GO:0033632	regulation of cell-cell adhesion mediated by integrin	9	0.3	0.04145	
Mammalia	Group 1 + Group 2	GO:0045625	regulation of T-helper 1 cell differentiation	9	0.3	0.04145	
Mammalia	Group 1 + Group 2	GO:0045628	regulation of T-helper 2 cell differentiation	9	0.3	0.04145	
Mammalia	Group 1 + Group 2	GO:1901625	cellular response to ergosterol	9	0.3	0.04145	
Mammalia	Group 1 + Group 2	GO:0043497	regulation of protein heterodimerization activity	1	0.04	0.03702	
Mammalia	Group 1 + Group 2	GO:1902569	negative regulation of activation of Janus kinase activity	1	0.04	0.03702	
Mammalia	Group 1 + Group 2	GO:0071660	positive regulation of IP-10 production	1	0.04	0.03702	
Mammalia	Group 1 + Group 2	GO:0044355	clearance of foreign intracellular DNA	1	0.04	0.03702	
Mammalia	Group 1 + Group 2	GO:0071652	positive regulation of chemokine (C-C motif) ligand 1 production	1	0.04	0.03702	
Mammalia	Group 1 + Group 2	GO:0071654	positive regulation of chemokine (C-C motif) ligand 1 production	1	0.04	0.03702	

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Mammalia	Group 1 + Group 2	GO:0002736	regulation of plasmacytoid dendritic cell cytokine production	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0002737	negative regulation of plasmacytoid dendritic cell cytokine production	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0071610	chemokine (C-C motif) ligand 1 production	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0051673	disruption of plasma membrane integrity in another organism	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0150073	regulation of protein-glutamine gamma-glutamyltransferase activity	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0150074	positive regulation of protein-glutamine gamma-glutamyltransferase activity	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0032759	positive regulation of TNF- $\alpha$ production	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:1905151	negative regulation of voltage-gated sodium channel activity	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0160006	FC receptor-mediated immune complex endocytosis	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0032723	positive regulation of connective tissue growth factor production	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:1905154	negative regulation of membrane invagination	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0032745	negative regulation of interleukin-21 production	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0016068	type I hypersensitivity	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:1904784	NLRP1 inflammasome complex assembly	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0035397	helper T cell enhancement of adaptive immune response	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:1905223	epicardium morphogenesis	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0050902	leukocyte adhesive activation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0002518	lymphocyte chemotaxis across high endothelial venule	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0034156	negative regulation of toll-like receptor 7 signaling pathway	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:1904848	negative regulation of cell chemotaxis to fibroblast growth factor	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0120042	negative regulation of macrophage proliferation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0034125	negative regulation of MyD88-dependent toll-like receptor signaling pathway	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0060097	cytoskeletal rearrangement involved in phagocytosis, engulfment	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:2000422	regulation of eosinophil chemotaxis	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:2000424	positive regulation of eosinophil chemotaxis	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0009073	positive regulation of protein homodimerization activity	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0051977	lysophospholipid transport	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0032665	regulation of interleukin-21 production	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0032679	regulation of TRAIL production	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0032625	interleukin-21 production	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:2000422	TRAIL production	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0001971	negative regulation of activation of membrane attack complex	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0036496	regulation or translational initiation by eIF2 alpha, dephosphorylation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0001497	eIF2alpha dephosphorylation in response to endoplasmic reticulum stress	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:2000229	regulation of pancreatic stellate cell proliferation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0051041	positive regulation of calcium-independent cell-cell adhesion	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:2000231	positive regulation of pancreatic stellate cell proliferation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:1900450	negative regulation of glutamate receptor signaling pathway	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0018003	peptidyl-lysine N6-acetyltransferase-induced eIF2 alpha dephosphorylation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:1903916	regulation of endoplasmic reticulum stress-induced eIF2 alpha dephosphorylation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:1903917	positive regulation of endoplasmic reticulum stress-induced eIF2 alpha dephosphorylation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0099046	clearance of foreign intracellular nucleic acids	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0150140	regulation of CD86 production	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0150142	positive regulation of CD86 production	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0150143	regulation or CD80 production	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0150145	positive regulation of CD80 production	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0140121	Lewy body formation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0140122	regulation of Lewy body formation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0140123	negative regulation of Lewy body formation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:1902310	positive regulation of peptidyl-serine dephosphorylation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:2000545	acylglycerol acyl-chain remodeling	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:1905675	negative regulation of adaptative immune memory response	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0090320	regulation of chylomicron remnant clearance	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0090321	positive regulation of chylomicron remnant clearance	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0140121	regulation of chylomicron remnant clearance	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0070246	natural killer cell apoptosis	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0070247	regulation of natural killer cell apoptosis	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:1901251	positive regulation of endothelial cell chemotaxis to fibroblast growth factor	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:1901247	negative regulation of phagocytosis, engulfment	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:1901249	regulation of lung goblet cell differentiation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0030186	melatonin in metabolic process	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0030187	pancreatic stellate cell proliferation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0072343	regulation of cellular defense response	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0010185	regulation of cellular defense response	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0001186	hepatocyte immune response	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:1903284	response to 11-deoxycorticosterone	1	1	0.04	0.03702

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Mammalia	Group 1 + Group 2	GO:0002373	plasmacytoid dendritic cell cytokine production	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0098784	biofilm matrix organization	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0098786	B cell antigen processing and presentation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0002450	negative regulation of mitotic centrosome separation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0046603	plasmacytoid dendritic cell antigen processing and presentation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0002470	plasmacytoid dendritic cell differentiation	1	1	0.04	0.03702
Mammalia	Group 1 + Group 2	GO:0073139	glomerular parietal epithelial cell differentiation	1	1	0.04	0.03702
Bivalvia	Group 1	GO:0080090	regulation of primary metabolic process	673	47	23.83	0.02305
Bivalvia	Group 1	GO:0019219	regulation of nucleobase-containing compound metabolic process	541	36	19.16	0.02702
Bivalvia	Group 1	GO:0051252	regulation of RNA metabolic process	517	32	18.31	0.03415
Bivalvia	Group 1	GO:006950	response to stress	370	29	13.1	0.02704
Bivalvia	Group 1	GO:0033554	negative regulation of nitrogen compound metabolic process	275	21	9.74	0.04194
Bivalvia	Group 1	GO:0051172	positive regulation of cellular metabolic process	117	15	4.14	0.00749
Bivalvia	Group 1	GO:0031325	DNA recombination	125	15	4.43	0.01000
Bivalvia	Group 1	GO:0006310	positive regulation of nitrogen compound metabolic process	66	13	2.34	0.00059
Bivalvia	Group 1	GO:0051173	negative regulation of gene expression	137	13	4.85	0.03359
Bivalvia	Group 1	GO:0010629	animal organ development	78	11	2.76	0.01719
Bivalvia	Group 1	GO:0048513	negative regulation of nucleobase-containing compound metabolic process	83	11	2.94	0.01719
Bivalvia	Group 1	GO:0045934	apoptotic process	64	11	2.27	0.01863
Bivalvia	Group 1	GO:0069195	positive regulation of DNA-templated transcription	95	9	3.36	0.00621
Bivalvia	Group 1	GO:0045892	negative regulation of RNA biosynthetic process	59	9	2.09	0.01512
Bivalvia	Group 1	GO:1902679	negative regulation of RNA metabolic process	59	9	2.09	0.01512
Bivalvia	Group 1	GO:0051253	negative regulation of RNA metabolic process	61	9	2.16	0.01918
Bivalvia	Group 1	GO:0006417	regulation of translation	52	8	1.84	0.00043
Bivalvia	Group 1	GO:0051726	regulation of cell cycle	75	8	2.66	0.03991
Bivalvia	Group 1	GO:0065009	regulation of molecular function	114	8	4.04	0.04846
Bivalvia	Group 1	GO:0045893	positive regulation of DNA-templated transcription	67	7	2.37	0.00908
Bivalvia	Group 1	GO:0042981	regulation of apoptotic process	70	7	2.48	0.01145
Bivalvia	Group 1	GO:0043067	regulation of programmed cell death	72	7	2.55	0.01326
Bivalvia	Group 1	GO:0000122	negative regulation of transcription by RNA polymerase II	31	5	1.1	0.00426
Bivalvia	Group 1	GO:0006402	mRNA catabolic process	35	5	1.24	0.00725
Bivalvia	Group 1	GO:0045944	positive regulation of transcription by RNA polymerase II	38	5	1.35	0.01030
Bivalvia	Group 1	GO:0071310	cellular response to organic substance	52	5	1.84	0.03586
Bivalvia	Group 1	GO:0009628	response to abiotic stimulus	53	5	1.88	0.03850
Bivalvia	Group 1	GO:0051248	negative regulation of protein metabolic process	55	5	1.95	0.04413
Bivalvia	Group 1	GO:0030155	regulation of cell adhesion	11	4	0.39	0.00041
Bivalvia	Group 1	GO:0048568	embryonic organ development	12	4	0.42	0.00061
Bivalvia	Group 1	GO:0051607	defense response to virus	13	4	0.46	0.00085
Bivalvia	Group 1	GO:0015669	regulation of double-strand break repair via homologous recombination	5	4	0.00359	0.00359
Bivalvia	Group 1	GO:0000902	cell morphogenesis	31	4	1.1	0.02274
Bivalvia	Group 1	GO:0000280	nuclear division	38	4	1.35	0.04409
Bivalvia	Group 1	GO:0050769	positive regulation of neurogenesis	5	3	0.18	0.00442
Bivalvia	Group 1	GO:0001819	positive regulation of cytokine production	7	3	0.25	0.00138
Bivalvia	Group 1	GO:0070192	chromosome organization involved in meiotic cell cycle	7	3	0.25	0.00138
Bivalvia	Group 1	GO:0007368	regulation of left/right symmetry	7	3	0.25	0.00138
Bivalvia	Group 1	GO:0001894	tissue homeostasis	10	3	0.35	0.00437
Bivalvia	Group 1	GO:0003007	heart morphogenesis	10	3	0.35	0.00437
Bivalvia	Group 1	GO:0008285	negative regulation of cell population proliferation	10	3	0.35	0.00437
Bivalvia	Group 1	GO:0051093	negative regulation of developmental process	11	3	0.39	0.00585
Bivalvia	Group 1	GO:0001501	meiotic system development	11	3	0.39	0.00585
Bivalvia	Group 1	GO:0007517	muscle organ development	13	3	0.46	0.00963
Bivalvia	Group 1	GO:0061982	meiosis I cell cycle process	14	3	0.5	0.01194
Bivalvia	Group 1	GO:0007127	meiosis I	14	3	0.5	0.02412
Bivalvia	Group 1	GO:0010257	NADH dehydrogenase complex assembly	19	3	0.67	0.02791
Bivalvia	Group 1	GO:0032488	mitochondrial respiratory chain complex I assembly	20	3	0.71	0.03200
Bivalvia	Group 1	GO:0022600	ribosomal small subunit biogenesis	21	3	0.74	0.02063
Bivalvia	Group 1	GO:0022603	regulation of anatomical structure morphogenesis	23	3	0.81	0.04603
Bivalvia	Group 1	GO:0140053	mitochondrial gene expression	23	3	0.81	0.04603
Bivalvia	Group 1	GO:0005976	positive regulation of double-strand break repair via homologous recombination	18	3	0.64	0.00125
Bivalvia	Group 1	GO:0048729	Cdc42 protein signal transduction	19	3	0.67	0.00125
Bivalvia	Group 1	GO:0035295	digestive system process	20	3	0.71	0.00125
Bivalvia	Group 1	GO:0042274	negative regulation of actin filament bundle assembly	22	2	0.07	0.00125
Bivalvia	Group 1	GO:0033617	mitochondrial cytochrome c oxidase assembly	22	2	0.11	0.00366
Bivalvia	Group 1	GO:0045910	negative regulation of DNA recombination	3	2	0.11	0.00366

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Bivalvia	Group 1	GO:2000179	positive regulation of neural precursor cell proliferation	3	2	0.11	0.00366
Bivalvia	Group 1	GO:0061383	trabecula morphogenesis	3	2	0.11	0.00366
Bivalvia	Group 1	GO:0051701	biological process involved in interaction with host	3	2	0.11	0.00366
Bivalvia	Group 1	GO:0010833	telomere maintenance via telomere lengthening	3	2	0.11	0.00366
Bivalvia	Group 1	GO:0007095	mitotic G2 DNA damage checkpoint signaling	3	2	0.11	0.00366
Bivalvia	Group 1	GO:0000959	mitochondrial RNA metabolic process	3	2	0.11	0.00366
Bivalvia	Group 1	GO:0002064	epithelial cell development	3	2	0.11	0.00366
Bivalvia	Group 1	GO:0010001	glial cell differentiation	4	2	0.14	0.00714
Bivalvia	Group 1	GO:0046620	regulation of organ growth	4	2	0.14	0.00714
Bivalvia	Group 1	GO:0003190	atrioventricular valve formation	4	2	0.14	0.00714
Bivalvia	Group 1	GO:0098781	ncRNA transcription	4	2	0.14	0.00714
Bivalvia	Group 1	GO:0030514	negative regulation of BMP signaling pathway	4	2	0.14	0.00714
Bivalvia	Group 1	GO:0035023	regulation of Rho protein signal transduction	4	2	0.14	0.00714
Bivalvia	Group 1	GO:0007416	synapse assembly	4	2	0.14	0.00714
Bivalvia	Group 1	GO:0001947	heart looping	5	2	0.18	0.01163
Bivalvia	Group 1	GO:0042026	protein refolding	5	2	0.18	0.01163
Bivalvia	Group 1	GO:0003143	embryonic heart tube morphogenesis	5	2	0.18	0.01163
Bivalvia	Group 1	GO:0061371	determination of heart left/right asymmetry	5	2	0.18	0.01163
Bivalvia	Group 1	GO:0035050	embryonic heart tube development	5	2	0.18	0.01163
Bivalvia	Group 1	GO:0007162	negative regulation of cell adhesion	5	2	0.18	0.01163
Bivalvia	Group 1	GO:0007129	homologous chromosome pairing at meiosis	5	2	0.18	0.01163
Bivalvia	Group 1	GO:0006360	transcription by RNA polymerase I	6	2	0.21	0.01704
Bivalvia	Group 1	GO:0001889	liver development	6	2	0.21	0.01704
Bivalvia	Group 1	GO:0045143	homologous chromosome segregation	6	2	0.21	0.01704
Bivalvia	Group 1	GO:0061008	hepatobiliary system development	6	2	0.25	0.02330
Bivalvia	Group 1	GO:0010212	response to ionizing radiation	7	2	0.25	0.02330
Bivalvia	Group 1	GO:0030490	maturaton of SSU-rRNA	7	2	0.25	0.02330
Bivalvia	Group 1	GO:0061448	connective tissue development	7	2	0.25	0.02330
Bivalvia	Group 1	GO:0048732	gland development	8	2	0.28	0.03036
Bivalvia	Group 1	GO:0001822	kidney development	8	2	0.28	0.03036
Bivalvia	Group 1	GO:0045596	negative regulation of cell differentiation	8	2	0.28	0.03036
Bivalvia	Group 1	GO:0000271	polysaccharide biosynthetic process	8	2	0.28	0.03036
Bivalvia	Group 1	GO:0001503	ossification	8	2	0.28	0.03036
Bivalvia	Group 1	GO:0045132	meiotic chromosome segregation	8	2	0.28	0.03036
Bivalvia	Group 1	GO:0016073	srNA metabolic process	9	2	0.32	0.03813
Bivalvia	Group 1	GO:0072001	renal system development	9	2	0.32	0.03813
Bivalvia	Group 1	GO:0060562	epithelial tube morphogenesis	9	2	0.32	0.03813
Bivalvia	Group 1	GO:0048562	embryonic organ morphogenesis	9	2	0.32	0.03813
Bivalvia	Group 1	GO:0042326	negative regulation of phosphorylation	10	2	0.35	0.04657
Bivalvia	Group 1	GO:0001934	positive regulation of protein phosphorylation	10	2	0.35	0.04657
Bivalvia	Group 1	GO:0022604	regulation or cell morphogenesis	10	2	0.35	0.04657
Bivalvia	Group 1	GO:0034314	Arp2/3 complex-mediated actin nucleation	10	2	0.35	0.04657
Bivalvia	Group 1	GO:0010669	epithelial structure maintenance	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0046621	negative regulation of organ growth	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0097222	mitochondrial mRNA polyadenylation	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0061668	common bile duct development	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0048799	regulation of transcription by RNA polymerase I	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0044406	regulation of protein K63-linked ubiquitination	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0043247	telomere maintenance in response to DNA damage	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0061009	common bile duct development	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0006356	regulation of transcription by RNA polymerase I	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0031120	regulation of protein K63-linked ubiquitination	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0044650	semaphorin-plexin signaling pathway involved in neuron projection guidance	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0096646	negative regulation of protein K63-linked ubiquitination	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0016074	semaphorin-plexin signaling pathway involved in axon guidance	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0032978	protein insertion into membrane from inner side	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0032979	snRNA pseudouridine synthesis	1	1	0.04	0.03541
Bivalvia	Group 1	GO:1900044	inhibitory synapse assembly	1	1	0.04	0.03541
Bivalvia	Group 1	GO:1902285	adhesion of symbiont to host cell	1	1	0.04	0.03541
Bivalvia	Group 1	GO:1900045	mitochondrial tRNA processing	1	1	0.04	0.03541
Bivalvia	Group 1	GO:1902287	sno(s)RNA metabolic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0050772	protein insertion into membrane from inner side	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0045943	lipopolysaccharide metabolic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0045943	positive regulation of transcription by RNA polymerase I	1	1	0.04	0.03541

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Bivalvia	Group 1	GO:0052572	response to host immune response	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0051274	beta-glucan biosynthetic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0051278	fungal-type cell wall polysaccharide biosynthetic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:1900220	semaphorin-peptidin signaling pathway involved in bone trabecula morphogenesis	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0008625	extrinsic apoptotic signaling pathway via death domain receptors	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0009301	snRNA transcription	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0015743	maleate transport	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0032065	maintenance of protein location in cell cortex	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0032196	retrotransposition	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0032197	cellular response to glucose starvation	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0042149	response to defenses of other organism	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0052173	(1->3)-beta-D-glucan metabolic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0006074	(1->3)-beta-D-glucan biosynthetic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0006075	response to host	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0075136	presynapse assembly	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0099054	protein localization to synapse	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0035418	ossification involved in bone maturation	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0043931	functional-type cell wall polysaccharide metabolic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0071966	presynapse organization	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0099172	box H/ACA sno(s)RNA metabolic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0033979	regulation of protein polyubiquitination	1	1	0.04	0.03541
Bivalvia	Group 1	GO:1902914	negative regulation of protein polyubiquitination	1	1	0.04	0.03541
Bivalvia	Group 1	GO:1902915	protein ubiquitination involved in ubiquitin-dependent protein catabolic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0071947	protein ubiquitination involved in ubiquitin-dependent protein catabolic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0050691	regulation of defense response to virus by host	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0046493	lipid A metabolic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0017001	antibiotic catabolic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0071423	malate transmembrane transport	1	1	0.04	0.03541
Bivalvia	Group 1	GO:1901269	lipooligosaccharide metabolic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0070212	protein poly-ADP-ribosylation	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0070213	protein auto-ADP-ribosylation	1	1	0.04	0.03541
Bivalvia	Group 1	GO:1902414	protein localization to cell junction	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0001100	negative regulation of exit from mitosis	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0048144	fibroblast proliferation	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0048145	regulation of fibroblast proliferation	1	1	0.04	0.03541
Bivalvia	Group 1	GO:1990481	negative regulation of fibroblast proliferation	1	1	0.04	0.03541
Bivalvia	Group 1	GO:1901271	lipooligosaccharide biosynthetic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0035622	intrahepatic bile duct development	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0030653	beta-lactam antibiotic metabolic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0030655	beta-lactam antibiotic catabolic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0052200	response to host defenses	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0048147	mRNA pseudouridine synthesis	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0009103	lipopolysaccharide biosynthetic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0048634	regulation of muscle organ development	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0010526	retrovirion silencing	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0046402	O antigen metabolic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0042753	evasion of host immune response	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0043144	antibiotic metabolic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0031848	protection from non-homologous end joining at telomere	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0061430	bone trabecula morphogenesis	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0020033	antigenic variation	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0009633	mitochondrial RNA processing	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0007168	receptor guanylyl cyclase signaling pathway	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0043628	regulatory ncRNA 3'-end processing	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0016999	lactam catabolic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0072340	regulation of extrinsic apoptotic signaling pathway via death domain receptors	1	1	0.04	0.03541
Bivalvia	Group 1	GO:1902041	fungal-type cell wall biogenesis	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0009272	lactam metabolic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0072338	miRNA-mediated gene silencing by mRNA destabilization	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0035279	box H/ACA sno(s)RNA processing	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0034964	box H/ACA sno(s)RNA 3'-end processing	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0000495	male meiotic nuclear division	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0003382	epithelial cell morphogenesis	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0009243	O antigen biosynthetic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0009245	lipid A biosynthetic process	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0036387	osteoblast proliferation	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0033688	negative regulation of osteoblast proliferation	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0033689	negative regulation of osteoblast proliferation	1	1	0.04	0.03541

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Bivalvia	Group 1	GO:0030277	maintenance of gastrintestinal epithelium	1	1	0.04	0.03541
Bivalvia	Group 1	GO:0048239	negative regulation of DNA recombination at telomere	1	0.04	0.03541	
Drosophila	Group 1	GO:0051276	chromosome organization	331	18	11.07	0.02925
Drosophila	Group 1	GO:0045132	meiotic chromosome segregation	64	7	2.14	0.00333
Drosophila	Group 1	GO:0007131	reciprocal meiotic recombination	37	6	1.24	0.00130
Drosophila	Group 1	GO:0031146	SCF-dependent proteasomal ubiquitin-dependent protein catabolic process	57	5	1.91	0.04125
Drosophila	Group 1	GO:0007141	male meiosis I	13	4	0.43	0.00069
Drosophila	Group 1	GO:1902275	regulation of chromatin organization	35	4	1.17	0.02847
Drosophila	Group 1	GO:0042078	genom-line stem cell division	36	4	1.2	0.03121
Drosophila	Group 1	GO:0010526	retrotransposon silencing	8	3	0.27	0.00183
Drosophila	Group 1	GO:0071218	cellular response to misfolded protein	12	3	0.4	0.00652
Drosophila	Group 1	GO:0034508	centromere complex assembly	19	3	0.64	0.02415
Drosophila	Group 1	GO:0007080	mitotic metaphase chromosome alignment	23	3	0.77	0.04004
Drosophila	Group 1	GO:0007020	microtubule nucleation	23	3	0.77	0.04004
Drosophila	Group 1	GO:0031445	regulation of heterochromatin formation	25	3	0.84	0.04953
Drosophila	Group 1	GO:0120261	regulation of heterochromatin organization	25	3	0.84	0.04953
Drosophila	Group 1	GO:0048136	male germline cyst formation	2	2	0.07	0.00112
Drosophila	Group 1	GO:0061964	negative regulation of entry into reproductive diapause	5	2	0.17	0.01043
Drosophila	Group 1	GO:0051382	kinetochore assembly	5	2	0.17	0.01043
Drosophila	Group 1	GO:0055116	entry into reproductive diapause	6	2	0.2	0.01531
Drosophila	Group 1	GO:0071712	ER-associated misfolded protein catabolic process	6	2	0.2	0.01531
Drosophila	Group 1	GO:0061963	regulation of entry into reproductive diapause	6	2	0.2	0.01531
Drosophila	Group 1	GO:0043984	histone H4-K16 acetylation	7	2	0.23	0.02096
Drosophila	Group 1	GO:0055115	entry into diapause	7	2	0.23	0.02096
Drosophila	Group 1	GO:1990084	response to odorant	7	2	0.23	0.02096
Drosophila	Group 1	GO:0042795	snoRNA transcription by RNA polymerase II	8	2	0.27	0.02733
Drosophila	Group 1	GO:0022611	dormancy process	8	2	0.27	0.02733
Drosophila	Group 1	GO:0009301	snoRNA transcription	8	2	0.27	0.02733
Drosophila	Group 1	GO:0071786	endoplasmic reticular tubular network organization	9	2	0.3	0.03438
Drosophila	Group 1	GO:0030007	intracellular potassium ion homeostasis	9	2	0.3	0.03438
Drosophila	Group 1	GO:0051383	kinetochore assembly	9	2	0.3	0.03438
Drosophila	Group 1	GO:0040020	regulation of meiotic nuclear division	9	2	0.3	0.03438
Drosophila	Group 1	GO:0043967	histone H4 acetylation	9	2	0.3	0.03438
Drosophila	Group 1	GO:0036376	sodium ion export across plasma membrane	9	2	0.3	0.03438
Drosophila	Group 1	GO:0006883	intracellular sodium ion homeostasis	9	2	0.3	0.03438
Drosophila	Group 1	GO:0001015	snoRNA transcription by RNA polymerase II	1	1	0.03	0.03245
Drosophila	Group 1	GO:0010778	meiotic DNA repair synthesis involved in reciprocal meiotic recombination	1	1	0.03	0.03245
Drosophila	Group 1	GO:0009302	sno(s)RNA transcription	1	1	0.03	0.03245
Drosophila	Group 1	GO:0051308	male meiosis chromosome separation	1	1	0.03	0.03245
Drosophila	Group 1	GO:0051415	microtubule nucleation by interphase microtubule organizing center	1	1	0.03	0.03245
Mammalia	Group 1	GO:0006935	immune response	1297	120	38.75	0.00270
Mammalia	Group 1	GO:0045087	positive regulation of immune response	647	69	19.33	1.6e-06
Mammalia	Group 1	GO:0050778	innate immune response	419	47	12.52	0.00677
Mammalia	Group 1	GO:0051308	adaptive immune response	342	39	10.22	1.4e-05
Mammalia	Group 1	GO:0051415	regulation of leukocyte activation	456	39	13.62	0.00521
Mammalia	Group 1	GO:0001819	positive regulation of cytokine production	402	37	12.01	0.00251
Mammalia	Group 1	GO:0019221	cytokine-mediated signaling pathway	382	35	8.16	2.6e-05
Mammalia	Group 1	GO:0042742	defense response to bacterium	233	32	9.96	0.00036
Mammalia	Group 1	GO:0042110	T cell activation	432	31	12.91	0.01328
Mammalia	Group 1	GO:0002697	regulation of immune effector process	308	31	9.2	0.01876
Mammalia	Group 1	GO:0051607	defense response to virus	257	28	7.68	4.9e-06
Mammalia	Group 1	GO:0002274	myeloid leukocyte activation	195	26	5.83	0.02936
Mammalia	Group 1	GO:0002703	immune response-regulating cell surface receptor signaling pathway	186	25	5.56	0.01358
Mammalia	Group 1	GO:0070661	leukocyte proliferation	273	24	8.16	0.02709
Mammalia	Group 1	GO:0031349	positive regulation of defense response	240	23	7.17	0.00568
Mammalia	Group 1	GO:1903131	mononuclear cell differentiation	392	23	11.71	0.04822
Mammalia	Group 1	GO:0050727	regulation of inflammatory response	300	21	8.96	0.00247
Mammalia	Group 1	GO:0006959	humoral immune response	177	21	5.29	0.04903
Mammalia	Group 1	GO:0002758	immune response-regulating cell surface receptor signaling pathway	177	20	5.29	0.00180
Mammalia	Group 1	GO:0001906	cell killing	153	20	4.57	0.04491
Mammalia	Group 1	GO:0002699	positive regulation of immune effector process	208	20	6.21	0.04774
Mammalia	Group 1	GO:0030217	T cell differentiation	255	18	7.62	0.04410
Mammalia	Group 1	GO:0045089	positive regulation of innate immune response	113	16	3.38	0.01390
Mammalia	Group 1	GO:0071222	cellular response to lipopolysaccharide	164	15	4.49	0.00112
Mammalia	Group 1	GO:1903555	regulation of tumor necrosis factor superfamily cytokine production	137	15	4.09	0.00448
Mammalia	Group 1	GO:0010466	negative regulation of peptidase activity	163	14	4.87	0.00039
Mammalia	Group 1	GO:0002429	immune response-activating cell surface receptor signaling pathway	164	14	4.9	0.00518
Mammalia	Group 1	GO:0032680	regulation of tumor necrosis factor production	133	14	3.97	0.01423
Mammalia	Group 1	GO:0032640	tumor necrosis factor production	133	14	3.97	0.01423

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Mammalia	Group 1	GO:0002705	positive regulation of leukocyte mediated immunity	113	14	3.38	0.03945
Mammalia	Group 1	GO:0050829	defense response to Gram-negative bacterium	66	13	1.97	6.1e-08
Mammalia	Group 1	GO:0002444	myeloid leukocyte mediated immunity	90	13	2.69	8.0e-05
Mammalia	Group 1	GO:0051091	positive regulation of DNA-binding transcription factor activity	227	13	6.78	0.01938
Mammalia	Group 1	GO:0002456	T cell mediated immunity	82	13	2.45	0.03948
Mammalia	Group 1	GO:0031341	regulation of cell killing	71	12	2.12	0.00420
Mammalia	Group 1	GO:0002655	negative regulation of leukocyte activation	148	12	4.42	0.01775
Mammalia	Group 1	GO:0002286	T cell activation involved in immune response	94	12	2.81	0.01940
Mammalia	Group 1	GO:0050866	leukocyte degranulation	62	10	1.85	0.00050
Mammalia	Group 1	GO:0001818	natural killer cell activation	165	12	4.93	0.03509
Mammalia	Group 1	GO:0050830	negative regulation of cytokine production	225	12	6.72	0.03811
Mammalia	Group 1	GO:0002275	defense response to Gram-positive bacterium	87	11	2.46	5.4e-05
Mammalia	Group 1	GO:0050909	myeloid cell activation involved in immune response	83	11	2.48	0.00097
Mammalia	Group 1	GO:0050909	sensory perception of taste	52	11	1.55	0.02888
Mammalia	Group 1	GO:0042102	positive regulation of T cell proliferation	85	10	2.54	0.00022
Mammalia	Group 1	GO:0043299	leukocyte degranulation	62	10	1.85	0.01940
Mammalia	Group 1	GO:0030101	natural killer cell activation	66	10	1.97	0.00092
Mammalia	Group 1	GO:0097636	cell surface receptor signaling pathway via STAT	139	10	4.15	0.03889
Mammalia	Group 1	GO:0016064	immunoglobulin mediated immune response	107	10	3.32	0.03212
Mammalia	Group 1	GO:0019724	B cell mediated immunity	108	10	3.23	0.03362
Mammalia	Group 1	GO:0032760	positive regulation of tumor necrosis factor production	84	9	2.51	0.00089
Mammalia	Group 1	GO:0006936	complement activation	47	9	1.4	0.00998
Mammalia	Group 1	GO:0050729	positive regulation of inflammatory response	122	9	3.64	0.01092
Mammalia	Group 1	GO:0051250	negative regulation of lymphocyte activation	124	9	3.7	0.01206
Mammalia	Group 1	GO:0050832	defense response to fungus	48	9	1.43	0.01262
Mammalia	Group 1	GO:0002823	negative regulation of adaptive immune response [...]	39	9	1.43	0.01741
Mammalia	Group 1	GO:0007239	cell surface receptor signaling pathway via JAK-SSTAT	134	9	4.4	0.01917
Mammalia	Group 1	GO:0002820	positive regulation of NF-kappaB transcription factor activity	136	9	4.06	0.02990
Mammalia	Group 1	GO:0042100	negative regulation of adaptive immune response	42	9	1.25	0.02592
Mammalia	Group 1	GO:0009620	B cell proliferation	76	9	2.27	0.02608
Mammalia	Group 1	GO:0035036	response to fungus	58	9	1.73	0.03229
Mammalia	Group 1	GO:0071887	sperm-egg recognition	46	9	1.37	0.03988
Mammalia	Group 1	GO:0007342	leukocyte apoptotic process	98	9	2.93	0.04681
Mammalia	Group 1	GO:0001580	fusion of sperm to egg plasma membrane involved in single fertilization	25	8	0.75	4.1e-07
Mammalia	Group 1	GO:0002218	detection of chemical stimulus involved in sensory perception of bitter taste	29	8	0.87	1.5e-06
Mammalia	Group 1	GO:0002649	activation of innate immune response	45	8	1.34	4.9e-05
Mammalia	Group 1	GO:0045576	regulation of type II interferon production	88	8	2.63	0.00475
Mammalia	Group 1	GO:0071346	mast cell activation	54	8	1.61	0.00632
Mammalia	Group 1	GO:0001959	cellular response to type II interferon	98	8	2.93	0.00900
Mammalia	Group 1	GO:0030317	regulation of cytokine-mediated signaling pathway	114	8	3.41	0.02094
Mammalia	Group 1	GO:0097722	flagellated sperm motility	119	8	3.55	0.02628
Mammalia	Group 1	GO:0008584	sperm motility	119	8	3.55	0.02828
Mammalia	Group 1	GO:0046546	male gonad development	120	8	3.58	0.02746
Mammalia	Group 1	GO:0060759	regulation of primary male sexual characteristics	121	8	3.61	0.02867
Mammalia	Group 1	GO:0062994	cilium movement involved in cell motility	124	8	3.37	0.03251
Mammalia	Group 1	GO:0007339	binding of sperm to zona pellucida	128	8	3.82	0.03817
Mammalia	Group 1	GO:0019731	antibacterial humoral response	35	7	1.05	6.6e-05
Mammalia	Group 1	GO:0032757	positive regulation of interleukin-8 production	40	7	1.19	0.00016
Mammalia	Group 1	GO:0032722	positive regulation of chemoattractant production	54	7	1.61	0.00107
Mammalia	Group 1	GO:0002251	organ or tissue specific immune response	60	7	1.79	0.00199
Mammalia	Group 1	GO:0098586	cellular response to virus	22	7	0.66	0.00357
Mammalia	Group 1	GO:0030533	neutrophil chemotaxis	78	7	2.33	0.00864
Mammalia	Group 1	GO:0043303	mast cell degranulation	80	7	2.39	0.00988
Mammalia	Group 1	GO:0002279	mast cell activation involved in immune response	43	7	1.28	0.01012
Mammalia	Group 1	GO:1904892	regulation of receptor signaling pathway via STAT	44	7	1.31	0.00018
Mammalia	Group 1	GO:0032886	neutrophil activation	83	7	0.96	0.00032
Mammalia	Group 1	GO:0045071	negative regulation of viral genome replication	39	6	1.17	0.00097
Mammalia	Group 1	GO:0002711	positive regulation of T cell mediated immunity	44	6	1.31	0.00185
Mammalia	Group 1	GO:0050691	regulation of immunoglobulin production	37	7	0.3360	0.03462
Mammalia	Group 1	GO:0035303	natural killer cell mediated cytotoxicity	103	6	0.87	0.00018
Mammalia	Group 1	GO:0042267	positive regulation of interleukin-1 beta production	44	6	1.31	0.00234
Mammalia	Group 1	GO:0032731	type I interferon-mediated signaling pathway	46	6	1.37	0.00438
Mammalia	Group 1	GO:0060337	regulation of B cell proliferation	52	6	1.55	0.00528
Mammalia	Group 1	GO:0030888	establishment of spindle localization	54	6	1.61	0.00528
Mammalia	Group 1	GO:0051293		54	6		

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Mammalia	Group 1	GO:0032720	positive regulation of type II interferon production	62	6	1.85	0.01035
Mammalia	Group 1	GO:0010921	regulation of phosphatase activity	65	6	1.94	0.01285
Mammalia	Group 1	GO:0002720	positive regulation of cytokine production involved in immune response	66	6	1.97	0.01379
Mammalia	Group 1	GO:0002385	mucosal immune response	19	6	0.57	0.02179
Mammalia	Group 1	GO:2000106	regulation of leukocyte apoptotic process	75	6	2.24	0.02449
Mammalia	Group 1	GO:0046425	regulation of receptor signaling pathway via JAK-STAT	78	6	2.33	0.02902
Mammalia	Group 1	GO:0032755	regulation of interleukin-6 production	83	6	2.48	0.03773
Mammalia	Group 1	GO:0061760	antifungal innate immune response	17	5	0.51	0.00011
Mammalia	Group 1	GO:0035458	cellular response to interferon-beta	21	5	0.63	0.00032
Mammalia	Group 1	GO:0048240	sperm capacitation	25	5	0.75	0.00075
Mammalia	Group 1	GO:0001914	regulation of T cell mediated cytotoxicity	25	5	0.75	0.00075
Mammalia	Group 1	GO:0002230	positive regulation of defense response to virus by host	27	5	0.81	0.00109
Mammalia	Group 1	GO:0002830	positive regulation of type 2 immune response	14	5	0.42	0.00192
Mammalia	Group 1	GO:0046596	regulation of viral entry into host cell	31	5	0.93	0.00208
Mammalia	Group 1	GO:0001774	microglial cell activation	35	5	1.05	0.00360
Mammalia	Group 1	GO:0140632	canonical inflammasome complex assembly	35	5	1.05	0.00360
Mammalia	Group 1	GO:0006953	acute-phase response	38	5	1.14	0.00518
Mammalia	Group 1	GO:0140374	antiviral innate immune response	41	5	1.22	0.00719
Mammalia	Group 1	GO:0001912	positive regulation of leukocyte mediated cytotoxicity	42	5	1.25	0.00719
Mammalia	Group 1	GO:0032720	negative regulation of tumor necrosis factor production	47	5	1.4	0.01273
Mammalia	Group 1	GO:0034113	heterotypic cell-cell adhesion	47	5	1.4	0.01273
Mammalia	Group 1	GO:1903556	negative regulation of tumor necrosis factor superfamily cytokine production	49	5	1.46	0.01508
Mammalia	Group 1	GO:0070228	regulation of lymphocyte apoptotic process	50	5	1.49	0.01635
Mammalia	Group 1	GO:0001961	positive regulation of cytokine-mediated signaling pathway	50	5	1.49	0.01635
Mammalia	Group 1	GO:0060760	positive regulation of response to cytokine stimulus	57	5	1.7	0.02732
Mammalia	Group 1	GO:0038634	signal transduction in absence of ligand	59	5	1.76	0.03114
Mammalia	Group 1	GO:0097192	extrinsic apoptotic signaling pathway in absence of ligand	59	5	1.76	0.03114
Mammalia	Group 1	GO:0071260	cellular response to mechanical stimulus	60	5	1.79	0.03317
Mammalia	Group 1	GO:0031640	killing of cells of another organism	64	5	1.91	0.04210
Mammalia	Group 1	GO:0061844	antimicrobial humoral immune response mediated by antimicrobial peptide	64	5	1.91	0.04210
Mammalia	Group 1	GO:0002227	innate immune response in mucosa	11	4	0.33	0.00022
Mammalia	Group 1	GO:0035455	response to interferon-alpha	14	4	0.42	0.00062
Mammalia	Group 1	GO:0048245	eosinophil chemotaxis	17	4	0.51	0.00137
Mammalia	Group 1	GO:0072540	T-helper 17 cell lineage commitment	17	4	0.51	0.00137
Mammalia	Group 1	GO:0002717	positive regulation of natural killer cell mediated immunity	21	4	0.63	0.00314
Mammalia	Group 1	GO:0032740	positive regulation of interleukin-17 production	23	4	0.69	0.00443
Mammalia	Group 1	GO:0043032	positive regulation of macrophage activation	23	4	0.69	0.00443
Mammalia	Group 1	GO:0070438	interleukin-1-mediated signaling pathway	24	4	0.72	0.00519
Mammalia	Group 1	GO:0070269	pyroptotic inflammatory response	26	4	0.78	0.00697
Mammalia	Group 1	GO:0031295	T cell costimulation	27	4	0.81	0.00800
Mammalia	Group 1	GO:0019884	antigen processing and presentation of exogenous antigen	28	4	0.84	0.00911
Mammalia	Group 1	GO:0002446	neutrophil mediated immunity	29	4	0.87	0.01033
Mammalia	Group 1	GO:2000352	negative regulation of endothelial cell apoptotic process	30	4	0.9	0.01164
Mammalia	Group 1	GO:0032743	positive regulation of interleukin-2 production	30	4	0.9	0.01164
Mammalia	Group 1	GO:1900225	regulation of NLRP3 inflammasome complex assembly	30	4	0.9	0.01164
Mammalia	Group 1	GO:0043330	response to exogenous RNA	31	4	0.93	0.01306
Mammalia	Group 1	GO:0046006	regulation of activated T cell proliferation	32	4	0.96	0.01458
Mammalia	Group 1	GO:0032814	regulation of natural killer cell activation	32	4	0.96	0.01458
Mammalia	Group 1	GO:0044546	NLRP3 inflammasome complex assembly	33	4	0.99	0.01621
Mammalia	Group 1	GO:0050798	activated T cell proliferation	34	4	1.02	0.01795
Mammalia	Group 1	GO:0006968	cellular defense response	34	4	1.02	0.01795
Mammalia	Group 1	GO:0002701	negative regulation of production of molecular mediator of immune response	35	4	1.05	0.01980
Mammalia	Group 1	GO:0043331	response to dsRNA	36	4	1.08	0.02176
Mammalia	Group 1	GO:0002691	regulation of cellular extravasation	38	4	1.14	0.02604
Mammalia	Group 1	GO:2000351	regulation of endothelial cell apoptotic process	43	4	1.28	0.03879
Mammalia	Group 1	GO:0032653	regulation of interleukin-10 production	45	4	1.34	0.04473
Mammalia	Group 1	GO:0032613	interleukin-10 production	45	4	1.34	0.04473
Mammalia	Group 1	GO:0045624	positive regulation of T helper cell differentiation	17	4	0.51	0.04786
Mammalia	Group 1	GO:0043666	regulation of phosphoprotein phosphatase activity	46	4	1.37	0.04788
Mammalia	Group 1	GO:1904894	positive regulation of receptor signaling pathway via STAT	46	4	1.37	0.04788
Mammalia	Group 1	GO:0038095	Fc-epsilon receptor signaling pathway	5	3	0.15	0.00025
Mammalia	Group 1	GO:0045959	negative regulation of complement activation, classical pathway	5	3	0.15	0.00025
Mammalia	Group 1	GO:0051838	cytolysis by host of symbiont cells	8	3	0.24	0.00133
Mammalia	Group 1	GO:0006924	necrotic protein signaling pathway	8	3	0.24	0.00133
Mammalia	Group 1	GO:1901731	activation-induced cell death of T cells	10	3	0.3	0.00272
Mammalia	Group 1	GO:2000551	positive regulation of platelet aggregation	10	3	0.3	0.00272
Mammalia	Group 1	GO:0007343	egg activation	11	3	0.33	0.00365
Mammalia	Group 1	GO:0035723	interleukin-15-mediated signaling pathway	11	3	0.33	0.00365

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Mammalia	Group 1	GO:015:0079	negative regulation of neuroinflammatory response	12	3	0.36	0.00476
Mammalia	Group 1	GO:0002323	natural killer cell activation involved in immune response	13	3	0.39	0.00606
Mammalia	Group 1	GO:0032197	retrotransposition	13	3	0.39	0.00606
Mammalia	Group 1	GO:0060046	regulation of acrosome reaction	13	3	0.39	0.00606
Mammalia	Group 1	GO:0072683	T cell extravasation	13	3	0.39	0.00606
Mammalia	Group 1	GO:0043306	positive regulation of mast cell degranulation	14	3	0.42	0.00754
Mammalia	Group 1	GO:0006957	complement activation, alternative pathway	14	3	0.42	0.00754
Mammalia	Group 1	GO:0050835	regulation of B cell receptor signaling pathway	14	3	0.42	0.00754
Mammalia	Group 1	GO:0023035	CD40 signaling pathway	14	3	0.42	0.00754
Mammalia	Group 1	GO:1903027	regulation of opsonization	14	3	0.42	0.00754
Mammalia	Group 1	GO:0002710	negative regulation of T cell mediated immunity	15	3	0.45	0.00922
Mammalia	Group 1	GO:0043031	negative regulation of macrophage activation	15	3	0.45	0.00922
Mammalia	Group 1	GO:0046597	negative regulation of viral entry into host cell	16	3	0.48	0.01110
Mammalia	Group 1	GO:0002888	positive regulation of myeloid leukocyte mediated immunity	17	3	0.51	0.01318
Mammalia	Group 1	GO:0002827	positive regulation of T-helper 1 type immune response	18	3	0.54	0.01348
Mammalia	Group 1	GO:0045346	regulation of MHC class II biosynthetic process	18	3	0.54	0.01548
Mammalia	Group 1	GO:0042832	defense response to protozoan	19	3	0.57	0.01798
Mammalia	Group 1	GO:0040019	positive regulation of embryonic development	19	3	0.57	0.01798
Mammalia	Group 1	GO:0060340	positive regulation of type I interferon-mediated signaling pathway	19	3	0.57	0.01798
Mammalia	Group 1	GO:0060333	type II interferon-mediated signaling pathway	19	3	0.57	0.01798
Mammalia	Group 1	GO:0045342	MHC class II biosynthetic process	19	3	0.6	0.02069
Mammalia	Group 1	GO:0032753	positive regulation of interleukin-4 production	20	3	0.6	0.02069
Mammalia	Group 1	GO:0001916	positive regulation of T cell mediated cytotoxicity	21	3	0.63	0.02362
Mammalia	Group 1	GO:1903901	negative regulation of viral life cycle	21	3	0.63	0.02362
Mammalia	Group 1	GO:0035821	modulation of process of another organism	21	3	0.63	0.02362
Mammalia	Group 1	GO:0002726	positive regulation of T cell cytokine production	22	3	0.66	0.02675
Mammalia	Group 1	GO:0042104	positive regulation of activated T cell proliferation	22	3	0.66	0.02675
Mammalia	Group 1	GO:0010743	regulation of macrophage derived foam cell differentiation	25	3	0.75	0.03742
Mammalia	Group 1	GO:0001562	regulation of T cell mediated cytotoxicity	25	3	0.75	0.03742
Mammalia	Group 1	GO:0032634	interleukin-4 production	25	3	0.75	0.03742
Mammalia	Group 1	GO:0071354	cellular response to interleukin-6	26	3	0.78	0.04140
Mammalia	Group 1	GO:0010922	positive regulation of phosphatase activity	26	3	0.78	0.04140
Mammalia	Group 1	GO:1901919	regulation of cilium-dependent cell motility	28	3	0.84	0.04995
Mammalia	Group 1	GO:0070741	response to interleukin-6	28	3	0.84	0.04995
Mammalia	Group 1	GO:0060295	regulation of cilium movement involved in cell motility	28	3	0.84	0.04995
Mammalia	Group 1	GO:0042269	regulation of natural killer cell mediated cytotoxicity	28	3	0.84	0.04995
Mammalia	Group 1	GO:0032634	execution phase of necrosis	2	2	0.06	0.00089
Mammalia	Group 1	GO:1903660	negative regulation of complement-dependent cytotoxicity	2	2	0.06	0.00089
Mammalia	Group 1	GO:0030887	positive regulation of myeloid dendritic cell activation	2	2	0.06	0.00089
Mammalia	Group 1	GO:0009609	response to symbiotic bacterium	4	4	0.12	0.00513
Mammalia	Group 1	GO:0048006	antigen processing and presentation, endogenous lipid antigen via MHC class Ib	4	4	0.12	0.00513
Mammalia	Group 1	GO:0002765	immune response-inhibiting signal transduction	5	5	0.15	0.00839
Mammalia	Group 1	GO:0045630	positive regulation of T helper 2 cell differentiation	5	5	0.15	0.00839
Mammalia	Group 1	GO:1903313	regulation of neutrophil degranulation	6	6	0.18	0.01233
Mammalia	Group 1	GO:0106300	protein-DNA covalent cross-linking repair	6	6	0.18	0.01233
Mammalia	Group 1	GO:0060545	positive regulation of necroticotic process	6	6	0.18	0.01233
Mammalia	Group 1	GO:0002826	negative regulation of T-helper 1 type immune response	6	6	0.18	0.01233
Mammalia	Group 1	GO:0031341	positive regulation of peptidyl-serine phosphorylation of STAT protein	6	6	0.18	0.01233
Mammalia	Group 1	GO:0060332	positive regulation of response to type II interferon	6	6	0.18	0.01233
Mammalia	Group 1	GO:0060335	positive regulation of type II interferon-mediated signaling pathway	6	6	0.18	0.01233
Mammalia	Group 1	GO:0043313	antigen processing and presentation, exogenous lipid antigen via MHC class Ib	6	6	0.18	0.01233
Mammalia	Group 1	GO:0002733	regulation of myeloid dendritic cell cytokine production	7	2	0.21	0.01693
Mammalia	Group 1	GO:0010940	positive regulation of necroticotic cell death	7	2	0.21	0.01693
Mammalia	Group 1	GO:0042796	snRNA transcription by RNA polymerase III	7	2	0.21	0.01693
Mammalia	Group 1	GO:0043312	neutrophil degranulation	7	2	0.21	0.01693
Mammalia	Group 1	GO:1903589	negative regulation of blood vessel endothelial cell proliferation involved in sprouting angiogenesis	8	2	0.24	0.02213
Mammalia	Group 1	GO:0006977	DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	8	2	0.24	0.02213

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Mammalia	Group 1	GO:0038110	interleukin-2-mediated signaling pathway	8	2	0.24	0.02213
Mammalia	Group 1	GO:0002638	negative regulation of immunoglobulin production	8	2	0.24	0.02213
Mammalia	Group 1	GO:0032650	regulation of interleukin-1 alpha production	8	2	0.24	0.02213
Mammalia	Group 1	GO:0035771	interleukin-4-mediated signaling pathway	8	2	0.24	0.02213
Mammalia	Group 1	GO:0032610	interleukin-1 alpha production	8	2	0.24	0.02213
Mammalia	Group 1	GO:0071352	cellular response to interleukin-2	8	2	0.24	0.02213
Mammalia	Group 1	GO:2800553	positive regulation of T-helper 2 cell cytokine production	8	2	0.24	0.02213
Mammalia	Group 1	GO:0034342	response to type III interferon	8	2	0.24	0.02213
Mammalia	Group 1	GO:0044406	adhesion of symbiont to host	9	2	0.27	0.02789
Mammalia	Group 1	GO:0045625	regulation of T-helper 1 cell differentiation	9	2	0.27	0.02789
Mammalia	Group 1	GO:0032754	regulation of peptidyl-serine phosphorylation of STAT protein	9	2	0.27	0.02789
Mammalia	Group 1	GO:0061518	microglial cell proliferation	9	2	0.27	0.02789
Mammalia	Group 1	GO:0070669	response to interleukin-2	9	2	0.27	0.02789
Mammalia	Group 1	GO:0010918	positive regulation of mitochondrial membrane potential	9	2	0.27	0.02789
Mammalia	Group 1	GO:1902563	regulation of neutrophil activation	10	2	0.3	0.03419
Mammalia	Group 1	GO:0002430	complement receptor mediated signaling pathway	10	2	0.3	0.03419
Mammalia	Group 1	GO:0032754	positive regulation of interleukin-5 production	10	2	0.3	0.03419
Mammalia	Group 1	GO:0010526	retrotransposon silencing	10	2	0.3	0.03419
Mammalia	Group 1	GO:0010528	regulation of transposition	10	2	0.3	0.03419
Mammalia	Group 1	GO:0010529	negative regulation of neutrophil activation	10	2	0.3	0.03419
Mammalia	Group 1	GO:1902563	regulation of neuroinflammation response	11	2	0.33	0.04097
Mammalia	Group 1	GO:0150078	positive regulation of neuroinflammation response	11	2	0.33	0.04097
Mammalia	Group 1	GO:0097011	cellular response to granulocyte macrophage colony-stimulating factor stimulus	11	2	0.33	0.04097
Mammalia	Group 1	GO:0097012	cellular response to granulocyte macrophage colony-stimulating factor	11	2	0.33	0.04097
Mammalia	Group 1	GO:0002732	positive regulation of dendritic cell cytokine production	11	2	0.33	0.04097
Mammalia	Group 1	GO:0032736	positive regulation of interleukin-13 production	11	2	0.33	0.04097
Mammalia	Group 1	GO:001857	positive regulation of cellular respiration	11	2	0.33	0.04097
Mammalia	Group 1	GO:0061517	macrophage proliferation	11	2	0.33	0.04097
Mammalia	Group 1	GO:1900226	negative regulation of NLRP3 inflammasome complex assembly	11	2	0.33	0.04097
Mammalia	Group 1	GO:0051770	positive regulation of nitric-oxide synthase biosynthetic process	11	2	0.33	0.04097
Mammalia	Group 1	GO:0032306	regulation of prostaglandin secretion	12	2	0.36	0.04822
Mammalia	Group 1	GO:0032308	positive regulation of prostaglandin secretion	12	2	0.36	0.04822
Mammalia	Group 1	GO:0042501	serine phosphorylation of STAT protein	12	2	0.36	0.04822
Mammalia	Group 1	GO:0060330	regulation of response to type II interferon	12	2	0.36	0.04822
Mammalia	Group 1	GO:0060334	regulation of type II interferon-mediated signaling pathway	12	2	0.36	0.04822
Mammalia	Group 1	GO:0018003	peptidyl-Lysine N6-acetylation	1	1	0.03	0.02987
Mammalia	Group 1	GO:0002384	hepatic immune response	1	1	0.03	0.02987
Mammalia	Group 1	GO:1905223	epicardium morphogenesis	1	1	0.03	0.02987
Mammalia	Group 1	GO:002450	B cell antigen processing and presentation	1	1	0.03	0.02987
Mammalia	Group 1	GO:0002470	plasmacytoid dendritic cell antigen processing and presentation	1	1	0.03	0.02987
Mammalia	Group 1	GO:0072343	pancreatic stellate cell proliferation	1	1	0.03	0.02987
Mammalia	Group 1	GO:0010185	regulation of cellular defense response	1	1	0.03	0.02987
Mammalia	Group 1	GO:0010186	positive regulation of cellular defense response	1	1	0.03	0.02987
Mammalia	Group 1	GO:0098784	biofilm matrix organization	1	1	0.03	0.02987
Mammalia	Group 1	GO:0098786	biofilm matrix disassembly	1	1	0.03	0.02987
Mammalia	Group 1	GO:0016068	type I hypersensitivity	1	1	0.03	0.02987
Mammalia	Group 1	GO:0032759	helper T cell enhancement of adaptive immune response	1	1	0.03	0.02987
Mammalia	Group 1	GO:0072343	glomerular parietal epithelial cell differentiation	1	1	0.03	0.02987
Mammalia	Group 1	GO:0010185	Lewy body formation	1	1	0.03	0.02987
Mammalia	Group 1	GO:0140121	regulation of Lewy body formation	1	1	0.03	0.02987
Mammalia	Group 1	GO:0140122	negative regulation of Lewy body formation	1	1	0.03	0.02987
Mammalia	Group 1	GO:0140123	melatonin metabolic process	1	1	0.03	0.02987
Mammalia	Group 1	GO:0030186	melatonin biosynthetic process	1	1	0.03	0.02987
Mammalia	Group 1	GO:0030187	positive regulation of IP-10 production	1	1	0.03	0.02987
Mammalia	Group 1	GO:0071660	cytoskeletal rearrangement involved in phagocytosis, engulfment	1	1	0.03	0.02987
Mammalia	Group 1	GO:0060097	positive regulation of calcium-independent cell-cell adhesion	1	1	0.03	0.02987
Mammalia	Group 1	GO:0051041	regulation of eosinophil chemotaxis	1	1	0.03	0.02987
Mammalia	Group 1	GO:2000422	lysophospholipid transport	1	1	0.03	0.02987
Mammalia	Group 1	GO:1901247	negative regulation of lung elated cell differentiation	1	1	0.03	0.02987
Mammalia	Group 1	GO:1901249	regulation of lung goblet cell differentiation	1	1	0.03	0.02987
Mammalia	Group 1	GO:0070246	natural killer cell apoptotic process	1	1	0.03	0.02987
Mammalia	Group 1	GO:0070247	regulation of natural killer cell apoptotic process	1	1	0.03	0.02987
Mammalia	Group 1	GO:1902310	positive regulation of peptidyl-serine dephosphorylation	1	1	0.03	0.02987
Mammalia	Group 1	GO:2000229	regulation of pancreatic stellate cell proliferation	1	1	0.03	0.02987

Table S10 continued from previous page

Dataset	Group of genes	GO.ID	Term	Annotated	Significant	Expected	classicFisher
Mammalia	Group 1	GO:1904784	NLRP1 inflammasome complex assembly	1	1	0.03	0.02987
Mammalia	Group 1	GO:2000231	positive regulation of pancreatic stellate cell proliferation	1	1	0.03	0.02987
Mammalia	Group 1	GO:0150140	regulation of CD86 production	1	1	0.03	0.02987
Mammalia	Group 1	GO:0150142	positive regulation of CD86 production	1	1	0.03	0.02987
Mammalia	Group 1	GO:0150143	regulation of CD80 production	1	1	0.03	0.02987
Mammalia	Group 1	GO:0150145	positive regulation of CD80 production	1	1	0.03	0.02987
Mammalia	Group 1	GO:0032665	regulation of interleukin-21 production	1	1	0.03	0.02987
Mammalia	Group 1	GO:0032679	regulation of TRAIL production	1	1	0.03	0.02987
Mammalia	Group 1	GO:0032625	interleukin-21 production	1	1	0.03	0.02987
Mammalia	Group 1	GO:0032639	TRAIL production	1	1	0.03	0.02987
Mammalia	Group 1	GO:0093220	regulation of chylomicron remnant clearance	1	1	0.03	0.02987
Mammalia	Group 1	GO:0090321	positive regulation of chylomicron remnant clearance pathway	1	1	0.03	0.02987
Mammalia	Group 1	GO:1900450	negative regulation of glutamate receptor signaling pathway	1	1	0.03	0.02987
Mammalia	Group 1	GO:0120042	negative regulation of macrophage proliferation	1	1	0.03	0.02987
Mammalia	Group 1	GO:1903916	regulation of endoplasmic reticulum stress-induced eIF2 alpha dephosphorylation	1	1	0.03	0.02987
Mammalia	Group 1	GO:1903917	positive regulation of endoplasmic reticulum stress-induced eIF2 alpha dephosphorylation	1	1	0.03	0.02987
Mammalia	Group 1	GO:0036496	regulation of translational initiation by eIF2 alpha dephosphorylation	1	1	0.03	0.02987
Mammalia	Group 1	GO:0036497	eIF2alpha dephosphorylation in response to endoplasmic reticulum stress	1	1	0.03	0.02987
Mammalia	Group 1	GO:1901625	cellular response to ergosterol	1	1	0.03	0.02987
Mammalia	Group 1	GO:0160006	Fc receptor-mediated immune complex endocytosis	1	1	0.03	0.02987
Mammalia	Group 1	GO:0036153	triglyceride acyl-chain remodeling	1	1	0.03	0.02987
Mammalia	Group 1	GO:0036155	acylglycerol acyl-chain remodeling	1	1	0.03	0.02987
Mammalia	Group 1	GO:0046603	negative regulation of mitotic centrosome separation	1	1	0.03	0.02987
Mammalia	Group 1	GO:0050902	leukocyte adhesive activation	1	1	0.03	0.02987
Mammalia	Group 1	GO:0034156	negative regulation of toll-like receptor 7 signaling pathway	1	1	0.03	0.02987
Mammalia	Group 1	GO:1905151	negative regulation of voltage-gated sodium channel activity	1	1	0.03	0.02987

# Data Availability

All the supplementary figures and supplementary tables, as well as high-resolution figures, will be accessible online at my GitHub personal page.