Supporting Information:

Evolution of male pregnancy associated with remodelling of canonical vertebrate immunity in seahorses and pipefishes

Olivia Roth[1], Monica Hongrø Solbakken[2], Ole Kristian Tørresen[2], Till Bayer[1], Michael Matschiner[2,3], Helle Tessand Baalsrud[2], Siv Nam Khang Hoff[2], Marine Servane Ono Brieuc[2], David Haase[1], Reinhold Hanel[4], Thorsten B.H. Reusch[1,*] & Sissel Jentoft[2,*]

1 GEOMAR Helmholtz Centre for Ocean Research Kiel, Marine Evolutionary Ecology, Düsternbrooker Weg 20, D-24105 Kiel, Germany

2 Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, Blindernveien 3, NO-0371 Oslo, Norway

3 Department of Palaeontology and Museum, University of Zürich, Karl-Schmid-Strasse 4, CH-8006 Zürich, Switzerland

4 Thünen Institute of Fisheries Ecology, Herwigstrasse 31, D-27572 Bremerhaven, Germany

* equal contribution

corresponding author: Olivia Roth: oroth@geomar.de

Table of contents

Table of tables

Table of figures

Table of SI Appendix Data Sets

1. Sample collection for Syngnathiformes genomes and DNA extraction

Syngnathiformes specimens were collected nearshore with handnets or at local fish markets in Germany, France, Cape Verde and Indonesia, or in the Eastern Atlantic off Angola aboard RV Dr. Fridtjof Nansen and off Madeira aboard RV Walther Herwig III between the years 2002 and 2014. A detailed list of sampling location for each specimen can be found in SI Appendix Table S1. Muscle tissues were stored at 4°C in 96 % ethanol before DNA extraction. Genomic DNA was extracted using the "High salt DNA extraction" method according to Phill Watts (https://www.liverpool.ac.uk/~kempsj/IsolationofDNA.pdf).

SI Appendix Table S1: Origin of samples used for genome sequencing, reproduction strategies and sex roles of each species.

species	sampling location	sampling date	sex roles	development
Aeoliscus strigatus	Kepulauan Seribu, Java Sea, Indonesia: 5°37'S/106°35'E	14.08.12	conventional	pelagic eggs and larvae
Dactylopterus volitans	Calvi, Mediterranean Sea, France: 42°33'35''N/8°43'39''E	03.09.02	conventional	pelagic eggs and larvae
Doryrhamphus dactyliophorus	Kepulauan Seribu, Java Sea, Indonesia: 5°37'S/106°35'E	14.08.12	sex-role reversed	male pregnancy, trunk brooder
Entelurus aequoreus	*RV Walther Herwig III*, Central Eastern Atlantic, Portugal: 42°22'84"N/012°35'18"W	03.03.10	sex-role reversed	male pregnancy, trunk brooder
Fistularia tabacaria	*RV Dr Fridtjof Nansen*, Central Eastern Atlantic, Angola: 09°27'S/13°04'E	16.03.03	conventional	pelagic eggs and larvae
Hippocampous whitei	Kepulauan Seribu, Java Sea, Indonesia: 5°37'S/106°35'E	14.08.12	conventional, monogamous	male pregnancy, tail brooder
Hippocampus comes	ref(2)		conventional, monogamous	male pregnancy, tail brooder
Hippocampus kuda	Kepulauan Seribu, Java Sea, Indonesia: 5°37'S/106°35'E	14.08.12	conventional, monogamous	male pregnancy, tail brooder
Macoramphosus scolopax	*RV L'Europe*, Mediterranean Sea, France: 42°34'12''N/3°32'24''E	23.06.13	conventional	pelagic eggs and larvae
Mullus surmuletus	Calvi, Mediterranean Sea, France: 42°34'48''N/8°43'28''E	03.09.12	conventional	pelagic eggs and larvae
Nerophis ophidion	Strande, Baltic Sea, Germany : 54°25'54'' N/ 10°10'15'' E	16.06.13	sex-role reversed	male pregnancy, trunk brooder
Syngnathus rostellatus	Strande, Baltic Sea, Germany : 54°25'54'' N/ 10°10'15'' E	16.06.13	sex-role reversed	male pregnancy, tail brooder
Syngnathus scovelli	ref(3)		sex-role reversed	male pregnancy, tail brooder
Syngnathus typhle	Strande, Baltic Sea, Germany : 54°25'54'' N/ 10°10'15'' E	16.06.13	sex-role reversed	male pregnancy, tail brooder

2. Syngnathiformes genomes

2.1. Fragmentation and Library preparation

Genomic DNA concentration was standardized to 50ng μl^{-1} with Qiagen Elution Buffer if necessary and subjected to sonication (Covaris S220, Life Technologies) to obtain DNA fragments of a mean length of 300 bp using the following settings: 200 cycles for 90s with w-peak at 105. All sequencing libraries were constructed following the Illumina TruSeq Sample Prep v2 low-Throughput Protocol according to manufaturer's instructions (Illumina).

2.2. Genome Sequencing

For *Syngnathus typhle* we made a high coverage and contiguous reference genome. To do so, two paired end libraries (180 bp and 300 bp insert size, 100 bp read length) were sequenced on the Illumina platform using HiSeq 2000 instruments at the Norwegian Sequencing Centre at the University of Oslo (http://www.sequencing.uio.no). In addition, an 8 kbp mate pair library using a 454 paired end approach was created and sequenced on the HiSeq 2000 instrument. The 3 kbp and 10 kbp Illumina mate pair libraries were sequenced at Genome Quebec at McGill University (http://www.genomequebec.com/en/home.html). The other 11 Syngnathiformes genomes were sequenced using a single library of 300-600 bp insert size for the different species, following the procedure outlined in refs(4, 5) with 150 bp read lengths.

2.3. Genome assemblies

For our main model species, *Syngnathus typhle*, the assembly strategy followed the recommended sequencing dataset for ALLPATHS-LG(6) using version R42780. First, any adapter sequences in the reads were removed with cutadapt(7) version 1.2.1, before assembly was performed. The 180 bp insert size library was used as the fragment library (used in the initial construction of unipaths), and all the other libraries were used as "jumping" libraries (used to expand unipaths and scaffold). *S. typhle* was assembled based on 35 Gbp of the 180 bp insert size library, 37 Gbp of the 300 bp insert size library, 62 Gbp of the 3 kbp insert size library, 31 Gbp of the 8 kbp library and 15 Gbp of the 10 kbp library.

The assemblies for all other Syngnathiformes species were performed as described in ref(5), but using Celera Assembler(8) version 8.3rc2 (ref(8),). In brief, reads were error-corrected and trimmed for adapters with merTrim from the Celera Assembler software using a *k*-mer size of 22 and Illumina adapters as options. Celera Assembler was run with the following options: merThreshold=0, merDistinct=0.9995, merTotal=0.995, unitigger=bogart, doOBT=0, doToggle=0; default settings were used for all other parameters. In addition, PreQC was run on the sequencing reads to estimate genome size(9). Each assembly based on one sequencing library for each species of approximately 350 bp insert size and 150 bp long reads (SI Appendix Table S2).

In order to assess the completeness of the genome assemblies, we run CEGMA v2.4.010312 (ref(10, 11)) and BUSCO v2.0 (ref(12)) with an actinopterygii standardized eukaryotic gene set. Basic statistics such as N50 contig and scaffold lengths were provided by the assemblathon_stats.pl script from Assemblathon 2 (ref(13)).

As expected, because of the long insert size libraries of *S. typhle,* the N50 scaffold lengths are much longer than those of the other species (3 Mbp). *Doryramphus dactylophorus* is the only species with longer contigs than *S. typhle*. We ran BUSCO(12) to validate the correctness of the genome assemblies. *Nerophis ophidion* and *Entelurus aequoreus* have much larger genomes than the other species, and their genome assemblies are fragmented compared to the other species, with only 33.6 and 21.6 percentage of complete BUSCO genes found respectively. The other species range from 73.8 (*Mullus surmuletus*) to 93.9 (*S. typhle*) percentage complete BUSCO genes found. This implies that the results in terms of the gene repertoire in *N. ophidion* and *E. aequeorus* need to be treated with caution.

SI Appendix Table S2: Summary statistics for the different genome assemblies. Genome sizes were estimated based on statistics from PreQC, sequencing coverage are sequencing reads in bp divided by estimated genome size. Assembly size is the actual size of the assembly using coverage as denominator. The BUSCO scores are divided into categories of complete, fragmented and missing genes.

Species	Estimated genome size inMbp (EGS)	Sequencing coverage based on EGS	Assembly size in Mbp (AS)	Sequencing coverage based on AS	N50 scaffold (bp)	N50 contig (bp)	BUSCO complete (percentage)	BUSCO fragmented (percentage)	BUSCO missing (percentage)
Aeoliscus strigatus	403	58.6	381	62.0	115785	15935	93.5	3.8	2.7
Macroramphorus scolopax	507	36.4	417	44.2	41786	13365	89.8	6.0	4.3
Mullus surmuletus	569	33.5	469	40.7	17184	7238	73.8	14.9	11.2
Dactylopterus volitans	498	33.7	577	29.1	17099	8271	74.1	13.6	12.3
Fistularia tabacaria	762	29.5	593	38.0	107234	17669	90.8	5.5	3.6
Doryramphus dactylophorus	651	31.3	619	32.9	75189	27594	87.1	7.8	5.1
Syngnathus rostellatus	347	55.0	283	67.5	87573	14947	89.0	5.9	5.1
Syngnathus typhle	NA	NA	315	578.4	3046963	25809	93.9	3.1	3.0
Hippocampus kuda	478	39.4	445	42.4	31191	10411	83.9	8.8	7.3
Hippocampus whitei	461	40.1	433	42.8	40759	10333	86.0	7.6	6.4
Nerophis ophidion	1581	13.1	976	21.2	6841	5245	33.6	25.6	40.8
Entelurus aequoreus	1834	34.4	557	113.2	3916	3393	21.6	27.1	51.3

2.4. Assembly of transcriptome for *Syngnathus typhle*

We used reference transcriptomes to aid the gene annotation of the *S. typhle* genome. Details on how transcriptomes were sequenced and analysed can be found under 7.1. To remove TruSeq Adapters RNA-seq reads were trimmed with cutadapt 1.5 at a quality threshold of Phred ≥20. The trimmed reads were then assembled with Trinity r20140717 (ref(14)).

2.5. Annotation of the *Syngnathus typhle* genome

Two different *ab initio* gene predictors were trained. GeneMark-ES v2.3e (ref(15)) on the genome assembly and SNAP v20131129 (ref(16)) on the genes found by CEGMA. MAKER v2.31 (refs(17, 18)) used the trained gene predictors, the Trinity transcriptome assembly, all repeats in RepBase(19) as called by MAKER and proteins from UniProtKB/SwissProt r2014_9 (ref(20)) for a first pass annotation of the genome assembly(21). The result of the first pass was used to retrain SNAP and train AUGUSTUS v3.0.2 (ref(22)) and a second iteration was performed using the same set-up. The protein sequences from final output of MAKER were BLASTed against the UniProtKB/SwissProt proteins and InterProScan v5.4-47 (ref(23)) was used to classify protein domains in the protein sequences. This information was transferred to all output of MAKER. This annotated 19,668 gene models. InterProScan(23) was run on the predicted proteins of these and gene names were allocated based on match with proteins in UniProt/SwissProt.

3. Syngnathiformes Phylogeny

3.1. Taxon selection and fossil calibration

To generate a time-calibrated phylogeny of Syngnathiformes, we compiled a genomic dataset that also included multiple representatives of diverse outgroup lineages, which allowed us to calibrate divergence times with a large number of fossil constraints. This dataset includes genome assemblies of 14 representatives of Syngnathiformes from which 12 are newly sequenced (this study), and the genome assemblies of *Syngnathus scovelli*, and *Hippocampus comes* were taken from ref(3) and ref(2) respectively. Moreover, 56 genome assemblies of non-Syngnathiformes species from 36 different teleost orders (SI Appendix Table S3) and *Lepisosteus oculatus* as an outgroup were included(24). For details, see

3.2. Selection of phylogenetic markers

To identify a set of orthologous markers for phylogenetic inference, we followed the analysis pipeline described in ref(4) with minor updates and modifications. The initial unfiltered set of markers included all 32,266 annotated zebrafish genes of release 87 of the ENSEMBL database(25). As in ref(4), we excluded genes that had less than five exons with a length of at least 150 bp, genes that could not be assigned to an ENSEMBL gene tree, or genes for which teleost fishes did not form a monophyletic group in the ENSEMBL gene tree. We also excluded all genes for which duplicates were present or homologs were missing in one or more of the ENSEMBL data sets of the eight species zebrafish (*Danio rerio*), cave fish (*Astyanax mexicanus*), cod (*Gadus morhua*), stickleback (*Gasterosteus aculeatus*), tilapia (*Oreochromis niloticus*), medaka (*Oryzias latipes*), fugu (*Takifugu rubripes*), and Amazon molly (*Poecilia formosa*). Using the same approach as in ref(4), we also removed exons sequences for which the tblastn bitscore(26) among known orthologs in the eight species was not at least 1.5 times as high as the highest bitscore with a potential paralog. After these filtering steps, 3,061 exons of 382 genes remained as queries for the identification of orthologs from all 71 assemblies.

3.3. Ortholog identification

Each of the 3,061 exons was used as a query in tblastn searches against each of the 71 assemblies to identify putative orthologs. Hits were filtered by exon-specific bitscore tresholds as described in ref(4). Furthermore, exon sequences were excluded if pairwise dN/dS ratios with the orthologous zebrafish sequence were greater than 0.25. We also removed unreliably aligned codons, all third codon positions, as well as exon alignments with (i) GC-content standard deviation above 0.05 (ref(27)) length shorter than 120 bp, or (iii) more than 10 missing sequences. We further filtered genes by exon tree congruence and clock-like evolution as described in ref(4). Based on this strict filtering procedure, we identified a set of 138 genes with a total alignment length of 108,372 bp and an overall proportion of missing data of 6.8%.

3.4. Bayesian phylogenetic inference

Phylogenetic clade support and divergence times were estimated in two separate

sets of Bayesian analyses with the software BEAST2 v.2.4.5 (ref(28)). As in ref(4), the first round of analyses was conducted to assess node support throughout the phylogeny; thus no topology or time constraints were applied except a single constraint on the root age, which was arbitrarily placed at 320 Ma. Fossil constraints, coupled with monophyly constraints on the respective clades, were then used in the second set of analyses to estimate the timeline of diversification in Syngnathiformes. All fossil constraints were applied with the CladeAge(29) package for BEAST2, with model parameters set as described in ref(30). Following ref(29), we identified the oldest fossil records of morphologically recognizable clades within our taxon set, resulting in 34 fossil constraints as listed below (1.3.4.). For each of the two sets of BEAST2 analyses, three replicate runs were performed, each with a chain length of 800 million MCMC generations of which the first 10% were discarded as burn-in. Maximum Clade Credibility (MCC) summary trees were produced with TreeAnnotator v.2.2.0.

3.5. Results Phylogeny

BEAST2 analyses based on 138 concatenated genes resulted in a strongly supported species tree consistent with recent phylogenomic studies(31) (SI Appendix Figure S1). All nodes received full support (BPP 1.0) except for the monophyly of Eurypterygia (combining Aulopiformes with Ctenosquamata within Neoteleostei[57]), which was supported by BPP 0.85. The monophyly of Eurypterygia, however, is not controversial as it has been supported by previous phylogenies(24, 31-33) as well as morphological synapomorphies(31). In agreement with other recent studies(24, 34), our phylogeny supports a sister group relationship between Scombriformes (tunas, mackerels, butterfishes, etc.) and Syngnathiformes. Within Syngnathiformes, we recover the monophyly of the families Syngnathidae (represented in our taxon set by nine species) and Centriscidae (represented by *Macroramphosus scolopax* and *Aeoliscus strigatus*) as well as the subfamilies Syngnathinae and Nerophinae (both represented by three species) as proposed in ref(35). We further find strong support for the suborder Syngnathoidei as defined in ref(31), represented in our phylogeny by members of three out of its five families (Syngnathidae, Fistulariidae, and Centriscidae; Aulostomidae and Solenostomidae were not analyzed). The suborders Dactylopteroidei and Mulloidei appear as sister taxa in our phylogeny;

however, both could be more closely related to the fourth syngnathiform suborder Callionymoidei(31), which is not represented in our phylogeny.

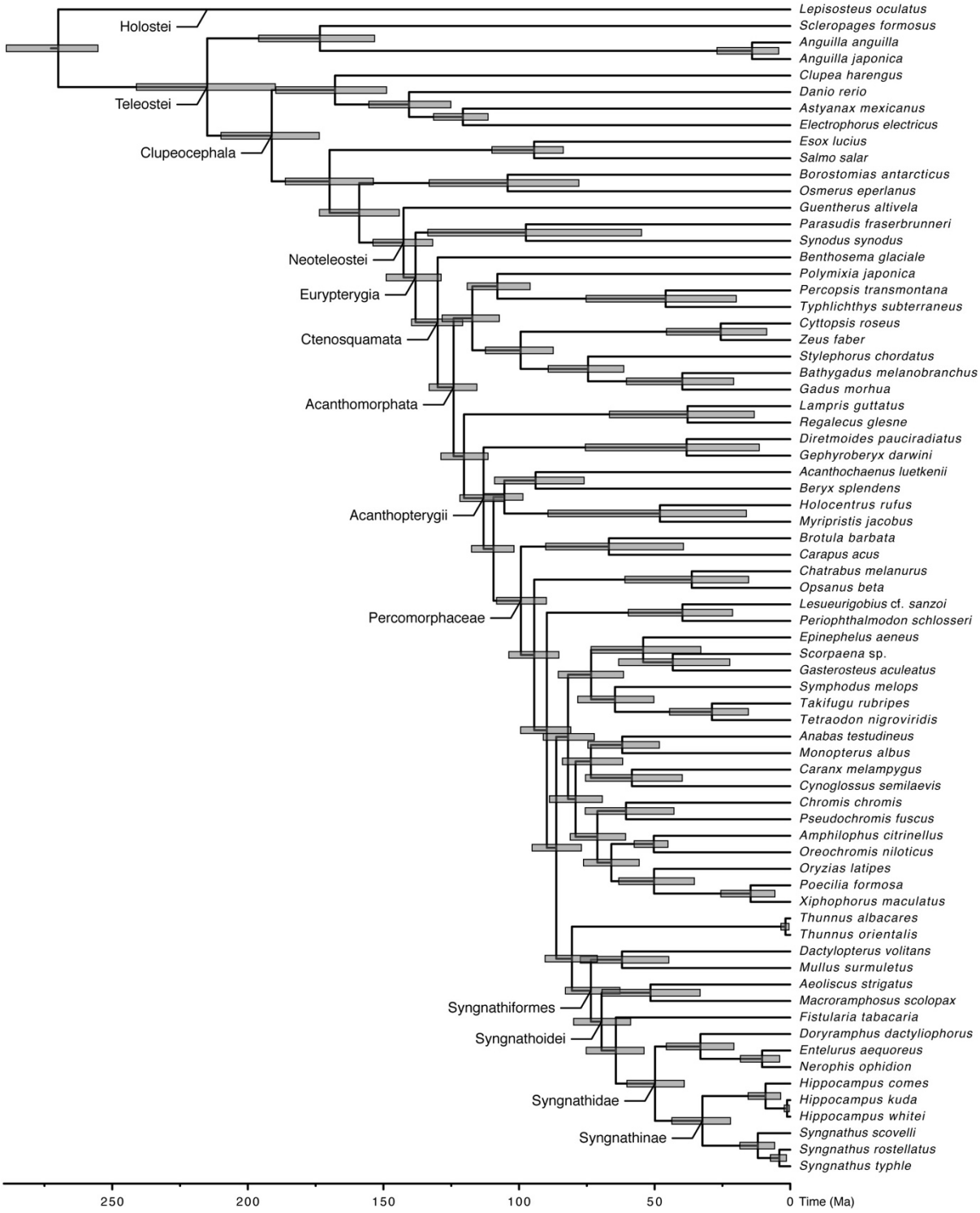
The timeline of species diversification estimated in the second set of our BEAST2 analyses based on 34 fossil constraints is highly congruent with the fossil record. At the root of the tree, the estimated time of divergence between Teleostei and Holostei predates the earliest record of Holostei (†*Acentrophorus varians*, 259.8-254.1 Ma; see above) by no more than 16 myr (269.9 Ma; 95% HPD 289.1-255.3 Ma). Acanthomorph fishes began to diversify around 124.1 Ma (95% HPD 133.2-115.6 Ma), and Syngnathiformes diverged from Scombriformes around 80.5 Ma (95% HPD 90.4-71.2 Ma) according to our estimated timeline. Thus, the age estimate for Acanthomophata predates the earliest acanthomorph crown fossils (e.g., † *Xenyllion zonensis*, 113.0-97.5 Ma; see above) by ≤27 myr, and the estimate for the origin of Syngnathiformes overlaps with the age of the earliest record assigned to Syngnathiformes (†*Gasterorhamphosus zuppichinii*, 89.8-82.0 Ma), which was not used for calibration as morphological synapomorphies are not known for the order(31). The divergence of the syngnathiform suborders Syngnathoidei, Dactylopteroidei, and Mulloidei is estimated between 73.5 Ma (95% HPD 82.9-62.8 Ma) and 62.0 Ma (95% HPD 77.4-44.8 Ma). Within Syngnatoidei, the families Centriscidae, Fistulariidae, and Syngnathidae diverged between 69.6 Ma (95% HPD 79.9-58.9 Ma) and 64.3 Ma (95% HPD 75.2-53.9 Ma), during a time that more generally seemed to promote explosive diversification of marine fishes(34). Finally, the two syngnathid subfamilies Syngnathinae and Nerophinae diverged around 49.9 Ma (95% HPD 60.2-39.1 Ma), predating the earliest fossil records of both groups (†*Syngnathus incompletus* and †*Hipposyngnathus neriticus*, 32.9-29.7 Ma) by appoximatley 20 myr.

SI Appendix Table S3: Genome assemblies used for phylogenomic inference. Unless specified otherwise, NCBI accessions are given. Taxonomic classification follows ref(31)

Species	Family	Order	Group	Accession / Link
Regalecus glesne	Regalecidae	Lampriformes	Division Lampripterygii	GCA_900302585.1
Lampris guttatus	Lampridae	Lampriformes	Division Lampripterygii	GCA_900302545.1
Polymixia japonica	Polymixiidae	Polymixiiformes	Division Polymixiipterygii	GCA_900302305.1
Guentherus altivela	Ateleopodidae	Ateleopodiformes	Infracohort Ateleopodia	GCA_900312595.1
Thunnus albacares	Scombridae	Scombriformes	Series Pelagiaria	GCA_900302625.1
Parasudis fraserbrunneri	Chlorophthalmidae	Aulopiformes	Section Cyclosquamata	GCA_900302295.1
Chatrabus melanurus	Batrachoididae	Batrachoidiformes	Series Batrachoidaria	GCA_900302635.1
Lesueurigobius cf. sanzoi	Gobiidae	Gobiiformes	Series Gobiaria	GCA_900303255.1
Brotula barbata	Ophidiidae	Ophidiiformes	Series Ophidiaria	GCA_900303265.1
Carapus acus	Ophidiidae	Ophidiiformes	Series Ophidiaria	GCA_900312935.1
Symphodus melops	Labridae	Labriformes	Series Eupercaria	GCA_900323315.1
Percopsis transmontana	Percopsidae	Percopsiformes	Series Percopsaria	GCA_900302285.1
Typhlichthys subterraneus	Amblyopsidae	Percopsiformes	Series Percopsaria	GCA_900302405.1
Bathygadus melanobranchus	Gadidae	Gadiformes	Subseries Gadariae	GCA_900302375.1
Zeus faber	Zeidae	Zeiformes	Subseries Zeiariae	GCA_900323335.1
Cyttopsis roseus	Parazenidae	Zeiformes	Subseries Zeiariae	GCA_900302355.1
Stylephorus chordatus	Stylephoridae	Stylephoriformes	Subseries Gadariae	GCA_900312615.1
Osmerus eperlanus	Osmeridae	Osmeriformes	Subcohort Stomiati	GCA_900302275.1
Borostomias antarcticus	Stomiidae	Stomiatiformes	Subcohort Stomiati	GCA_900323325.1
Beryx splendens	Berycidae	Beryciformes	Subdivision Berycimorphaceae	GCA_900312565.1
Acanthochaenus luetkenii	Stephanoberycidae	Beryciformes	Subdivision Berycimorphaceae	GCA_900312575.1
Myripristis jacobus	Holocentridae	Holocentriformes	Subdivision Holocentrimorphaceae	GCA_900302555.1
Holocentrus rufus	Holocentridae	Holocentriformes	Subdivision Holocentrimorphaceae	GCA_900302615.1
Benthosema glaciale	Myctophidae	Myctophiformes	Subsection Myctophata	GCA_900323375.1
Anabas testudineus	Anabantidae	Anabantiformes	Series Anabantaria	GCA_900302665.1
Chromis chromis	Pomacentridae	*incertae sedis*	Series Ovalentaria	GCA_900302755.1
Pseudochromis fuscus	Pseudochromidae	*incertae sedis*	Series Ovalentaria	GCA_900323345.1
Lepisosteus oculatus	Lepisosteidae	Lepisosteiformes	Infraclass Holostei	LepOcu1 (ENSEMBL 87)
Takifugu rubripes	Tetraodontidae	Tetraodontiformes	Series Eupercaria	FUGU4 (ENSEMBL 87)

Astyanax mexicanus	Characidae	Characiformes	Cohort Otomorpha	AstMex102 (ENSEMBL 87)
Danio rerio	Danionidae	Cypriniformes	Cohort Otomorpha	GRCz10 (ENSEMBL 87)
Oryzias latipes	Adrianichthyidae	Beloniformes	Series Ovalentaria	MEDAKA1 (ENSEMBL 87)
Poecilia formosa	Poeciliidae	Cyprinodontiformes	Series Ovalentaria	PoeFor_5.1.2 (ENSEMBL 87)
Xiphophorus maculatus	Poeciliidae	Cyprinodontiformes	Series Ovalentaria	Xipmac4.4.2 (ENSEMBL 87)
Gasterosteus aculeatus	Gasterosteidae	Perciformes	Series Eupercaria	BROADS1 (ENSEMBL 87)
Tetraodon nigroviridis	Tetraodontidae	Tetraodontiformes	Series Eupercaria	TETRAODON8 (ENSEMBL 87)
Anguilla anguilla	Anguillidae	Anguilliformes	Cohort Elopomorpha	GCA_000695075.1
Anguilla japonica	Anguillidae	Anguilliformes	Cohort Elopomorpha	GCA_000470695.1
Thunnus orientalis	Scombridae	Scombriformes	Series Pelagiaria	GCA_000418415.1
Periophthalmodon schlosseri	Oxudercidae	Gobiiformes	Series Gobiaria	GCA_000787095.1
Esox lucius	Esocidae	Esociformes	Subcohort Protacanthopterygii	GCA_000721915.3
Cynoglossus semilaevis	Cynoglossidae	Pleuronectiformes	Series Carangaria	GCA_000523025.1
Amphilophus citrinellus	Cichlidae	Cichliformes	Series Ovalentaria	GCA_000751415.1
Scleropages formosus	Osteoglossidae	Osteoglossiformes	Supercohort Osteoglossomorpha	GCA_001624265.1
Oreochromis niloticus	Cichlidae	Cichliformes	Series Ovalentaria	GCA_001858045.2
Salmo salar	Salmonidae	Salmoniformes	Subcohort Protacanthopterygii	GCA_000233375.4
Hippocampus comes	Syngnathidae	Syngnathiformes	Series Syngnatharia	GCA_001891065.1
Gadus morhua	Gadidae	Gadiformes	Subseries Gadariae	GCA_900302565.1
Scorpaena sp.	Scorpaenidae	Perciformes	Series Eupercaria	this study
Syngnathus scovelli	Syngnathidae	Syngnathiformes	Series Syngnatharia	https://creskolab.uoregon.edu/pipefish/resources/ssc_2016_12_20_chromlevel.fa.gz
Fistularia tabacaria	Fistulariidae	Syngnathiformes	Series Syngnatharia	this study
Macroramphorus scolopax	Centriscidae	Syngnathiformes	Series Syngnatharia	this study
Mullus surmuletus	Mullidae	Syngnathiformes	Series Syngnatharia	this study
Nerophis ophidion	Syngnathidae	Syngnathiformes	Series Syngnatharia	this study
Dactylopterus volitans	Dactylopteridae	Syngnathiformes	Series Syngnatharia	this study
Doryramphus dactylophorus	Syngnathidae	Syngnathiformes	Series Syngnatharia	this study
Hippocampus kuda	Syngnathidae	Syngnathiformes	Series Syngnatharia	this study
Hippocampus whitei	Syngnathidae	Syngnathiformes	Series Syngnatharia	this study
Aeoliscus strigatus	Centriscidae	Syngnathiformes	Series Syngnatharia	this study
Syngnathus rostellatus	Syngnathidae	Syngnathiformes	Series Syngnatharia	this study

Syngnathus typhle	Syngnathidae	Syngnathiformes	Series Syngnatharia	this study
Entelurus aequoreus	Syngnathidae	Syngnathiformes	Series Syngnatharia	this study
Synodus synodus	Synodontidae	Aulopiformes	Section Cyclosquamata	ref(24)
Opsanus beta	Batrachoididae	Batrachoidiformes	Series Batrachoidaria	ref(24)
Diretmoides pauciradiatus	Diretmidae	Trachichthyiformes	Subdivision Berycimorphaceae	ref(24)
Gephyroberyx darwini	Trachichthyidae	Trachichthyiformes	Subdivision Berycimorphaceae	ref(24)
Epinephelus aeneus	Serranidae	Perciformes	Series Eupercaria	(24)
Clupea harengus	Clupeidae	Clupeiformes	Cohort Otomorpha	(24)
Electrophorus electricus	Gymnotidae	Gymnotiformes	Cohort Otomorpha	ref(24)
Monopterus albus	Synbranchidae	Synbranchiformes	Series Anabantaria	ref(24)
Caranx melampygus	Carangidae	Carangiformes	Series Carangaria	ref(24)



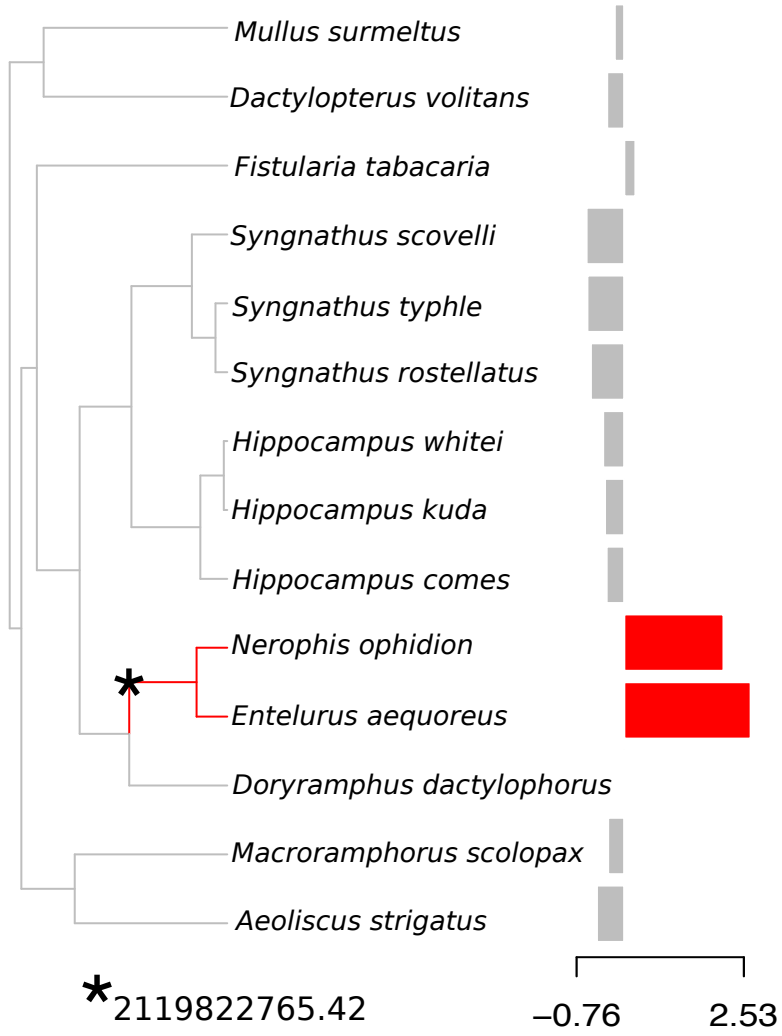
SI Appendix Figure S1: Time-calibrated phylogeny based on genome assemblies of 70 teleost species and *Lepisosteus oculatus*. The phylogeny shown represents the MCC summary tree of a BEAST analysis calibrated with 34 fossil constraints. Taxonomic classification follows Betancur-R. et al. (2017). All nodes received full Bayesian support (BPP 1.0) except the monophyly of Eurypterygia, which was supported by BPP 0.85. Node bars indicate 95% HPD intervals.

4. Genome size evolution

4.1. Optimal genome size evolution

By searching for shifts in the optima of genome size in the different lineages, we aimed to identify a genome size pattern in the evolution of the Syngnathiformes species. We applied the Ornstein–Uhlenbeck process as implemented in the R package(36) using the Syngnathiformes phylogeny and the genome sizes of the species.

While genome sizes of most Syngnathiformes are in the range of 340 to 765 Mbp, both *Nerophis ophidion* and *Entelurus aequoreus* have significantly larger estimated genome sizes (*E. aequoreus*: 1830 Mbp, *N. ophidion*: 1581 Mbp) than all other Syngnathiformes (including their closest relative, *Doryramphus dactylophorus* (650 Mbp)) (SI Appendix Figure S2). As indicated by the BUSCO scores, the expansion of genome size in *N. ophidion* and *E. aequoreus* does not indicate a recent genome tetraploidisation, as the number of duplicated genes is not higher than in all other Syngnathiformes (i.e. between 18 and 127).



SI Appendix Figure S2: Ancestor reconstruction of genome size. A shift in genome size has happened in the evolution of Syngnathiformes, i.e. *Nerophis ophidion* and *Entelurus aequoreus* have larger genomes than all other Syngnathiformes (as indicated by red bars).

4.2. Transposable elements

Transposable elements are known to be the major contributor to genome size evolution(37, 38). To investigate potential reasons for differences in genome size, a library of repeated elements was created as described in ref(39). In brief, RepeatModeler v1.0.8, LTRharvest (part of genometools v1.5.7)(ref(40)) and TransposonPSI were used in combination to create a set of putative repeats. Elements with only a match against a UniProtKB/SwissProt database and not against the database of known repeated elements included in RepeatMasker were removed. The remaining elements were classified and combined with known repeat elements from RepBase v20150807 (ref(41)).

The fraction of transposable elements correlated well with the genome size of the species investigated here (SI Appendix Table S4). This indicate that during 50 million years,

transposable elements seem to have expanded in *Nerophis ophidion* and *Entelurus aequoreus* (which had potentially already started before the split of *Nerophis* and *Entelurus* from *Doryrhamphus* (SI Appendix Figure S1). Further, a possible contraction of transposable elements evolved in the *Hippocampus* and *Syngnathus* species. This quite rapid change suggests that this clade would be a good model for studies in genome size evolution.

SI Appendix Table S4: Overview of genome sizes and the repetitive elements of the Syngnathiformes species.

Species	Assembly size (Mbp)	Estimated genome size (Mbp)	SINEs	LINEs	LTR elements	DNA elements	Unclassified	Total interspersed	Simple repeats
Aeoliscus strigatus	381	403	0.24	1.95	0.82	3.60	5.97	12.58	3.25
Macroramphorus scolopax	418	507	0.24	1.00	0.53	2.59	4.47	8.83	2.67
Mullus surmuletus	469	569	0.07	0.91	0.67	5.08	6.56	13.29	3.64
Dactylopterus volitans	577	499	0.29	3.06	1.23	6.73	8.94	20.25	3.75
Fistularia tabacaria	593	762	1.70	7.58	1.43	4.23	3.98	18.92	4.34
Doryramphus dactylophorus	618	651	0.04	5.10	1.45	16.06	25.56	48.21	1.73
Syngnathus scovelli	307	307	0.31	2.17	0.95	4.20	5.34	12.97	1.73
Syngnathus rostellatus	283	347	0.27	2.12	0.78	4.97	6.46	14.60	1.89
Syngnathus typhle	315	315	0.18	2.60	1.14	4.66	5.47	14.05	1.64
Hippocampus comes	494	494	0.04	2.27	0.69	9.06	12.03	24.09	2.29
Hippocampus kuda	445	478	0.05	2.77	0.74	9.63	12.02	25.21	2.55
Hippocampus whitei	433	461	0.26	2.64	0.65	9.65	12.69	25.89	2.59
Nerophis ophidion	976	1,581	3.09	8.81	1.66	21.27	30.85	65.68	2.78
Entelurus aequoreus	557	1,834	1.62	6.62	1.63	16.69	30.63	57.19	2.82

5. Gene mining and annotation

To investigate how male pregnancy has coevolved with the immune system, we had chosen to analyse a subset of immune, pregnancy and hemoglobin candidate genes in more details (SI Appendix Table S5). For those, query protein sequences were obtained from Ensembl (release 90) or UniProt (2017) by name search with representatives from *Homo sapiens*, *Mus musculus* and at least three different fish species.

SI Appendix Table S5: Immune, pregnancy and hemoglobin genes that were aligned among all sequenced Syngnathiformes species and subjected to an analysis of positive selection upon a gene-wise phylogeny.

Full gene name	**Abbreviation**
Major histocompatibility complex class I and II	*MHC I* and *MHC II*
Activation-induced cytidine deaminase	*AICDA* alias *AID*
Recombination activating 1 and 2	*RAG1* and *RAG2*
Autoimmune regulator	*AIRE*
Transporter 1 and 2 ATP binding cassette subfamily B member	*TAP1* and *TAP2*
T-cell surface glycoprotein Cluster of differentiation 4	*CD4*
T-cell surface glycoprotein Cluster of differentiation 8 alpha and beta	*CD8a* and *CD8b*
Invariant chain	*CD74* alias *Ii* and MHC II gamma chain
Class II Major histocompatibility complex transactivator	*CIITA*
Beta-2-microglobulin	*B2M*
Immunoglobulin heavy and light chains	*IgH* and *IgL*
Apolipoprotein A1	*APOA1*
Activating Transcription Factor 3	*ATF3*
Carbonic Anhydrase 4	*Ca4*
CCAAT Enhancer Binding Protein Beta	*CEBPB*
Distal-Less Homeobox 3	*DLX3b*
Endothelial PAS Domain Protein 1	*EPAS1*
Fatty Acid Binding Protein 1	*FABP1*
Fatty Acid Binding Protein 3	*FABP3*
Fatty Acid Binding Protein 6	*FABP6*
Hydroxysteroid 11-Beta Dehydrogenase 2	*HSD11B2*
JunB Proto-Oncogene, AP-1 Transcription Factor Subunit	*JUNB*
Protein Kinase C Delta	*PRKCD*
Prostaglandin F2 Receptor Inhibitor	*PTGFRN*
Rh Associated Glycoprotein	*RHAG*
Transferrin	*TF*
Hemoglobin	alpha and beta

5.1. Ortholog search within the Syngnathiformes

For immune, pregnancy and hemoglobin genes, translated query sequences were aligned in MEGA7(42) using the MUSCLE alignment algorithm to ensure query protein sequence homology. The protein queries, as whole sequences and for MHC and hemoglobin also split into individual exons, were used as input in a TBLASTN search (blast+-2.2.29) towards the scaffolds from the assembled draft genomes as well as the genomes from

Syngnathus scovelli and *Hippocampus comes*. Initially, the TBLASTN search was run with default parameters meaning an E-value cutoff of 10 and reporting results both with alignments and in tabular format. In some cases, searches towards the scaffolds did not lead to any likely homologous sequence promoting a new TBLASTN search, with the same search parameters as described above, towards the unitigs of the assemblies generated in this project. Unitigs were used for *MHC I* and *MHC II* searches due to large copy numbers and the risk of collapsed sequences with the exception of *Hippocampus comes* and *Syngnathus scovelli* where unitigs were not available.

5.2. Gene alignments

The hit regions with the lowest E-value and best scores were extracted from the assembled draft genomes using bedtools(43). For multicopy genes such as *MHC I* and *MHC II* all putative hits were extracted and subjected to a reciprocal BLASTX search towards the UniProt database reporting the top hit only to select the proper genomic regions for downstream analyses. When extracting genomic regions for downstream analyzes, an additional 2000 to 10000 bp were extracted both up- and downstream of the target region to capture possible smaller leading and trailing exons. Frequently, these were low-complexity stretches coding for hydrophobic tracts, not detected by the TBLASTN search. The resulting nucleotide fasta sequences were imported into MEGA7 together with the codon sequence from the queries. The nucleotide sequences were translated to protein and then aligned using MUSCLE(44). Gaps were manually inserted to adjust the reading frame whenever needed. The hits from the TBLASTN search using individual exons were used as alignment support. All alignments were manually curated before being cleaned and subjected to different selection analyses.

In the analysis of *MHC II* genes we payed particular attention on the beta sequences since these amino acid sequences were nearly complete in more Syngnathiforme species in contrast to those of the alpha gene. For phylogenetic and structural comparison, full-length sequences of other fish species were retrieved from Genbank. For the phylogenetic analysis, alignments of protein translations were done with MUSCLE(44) and subsequently manually edited. To maximize completeness of the alignments across taxa, they had to be split into one covering more the first portion (thus exons 1-3; SI Appendix Data Set 1: 30MHCII_beta_1-86aa) and another one covering more the second portion of the gene (thus exon 3-5, SI Appendix Data Set 1: 31MHCII_beta_91-184aa). We also examined whether or not critical di-sulfide (cystein) bridges within exon 2 and 3 were present, which are required for the appropriate tertiary structure of the *MHC II*

molecules(45, 46). For secondary and tertiary structure prediction we used the neural network algorithm developed by Rost & Sander(47). We applied DISULFIND(48) to identify structurally important di-sulfide bridges among cystein residues; both were implemented in in the online package predictprotein (available at https://www.predictprotein.org/). As "positive" controls we used several full *MHC II beta* sequences of well-curated species obtained from Genbank (stickleback AY713945, seabass DQ821110.1, pikeperch AAO19848.1).

5.3. Gene trees

Trees were generated with RAxML (v. 8.2.10)(49) using the JT T model for MHC I and the VT model for MHC II as determined by the RAxML "PROTGAMMAAUTO" method. 100 rapid bootstraps were run. For *hemoglobin* phylogenetic gene trees for alpha and beta sequences, respectively, were constructed using maximum likelihood in MEGA7(42) using the GTR+G+I model for alpha-sequences and TN93+G model for beta-sequences (these models had the lowest AIC score), with 1000 bootstrap runs (alignments of α and beta sequences are in SI Appendix Data Set 1). Sequences were translated into amino acids and aligned using ClustalW(50) as implemented in MEGA7(42) with default settings for all species (alignments of alpha and beta sequences are in SI Appendix Data Set 1: 1hemoglobin_alpha & 2hemoglobin_beta). The alpha and beta orthologous genes were then mapped on a phylogenomic species tree.

5.4. Local gene synteny

For the additionally lost genes related to the MHC II pathway in *Syngnathus* (*AID* (*AICDA*), *CD4*, *CIITA*) local gene synteny was explored. Syntenic genes were found using the Ensembl genome browser and corresponding peptide sequences of flanking genes directly up- and down-stream of the gene in question were downloaded from several fish species. These peptide queries were used in a TBLASTN as described in the main text (blast+ v 2.6.0, e-value cutoff 1-e10 and tabular output). The output was sorted for genomic contig/scaffold/linkage group and then for start position within each of these. The corresponding hit region for the syntenic genes were subjected to a reciprocal BLASTX using first bedtools for sequence extraction and then the NCBI BLASTX server with default parameters.

In the case of *MHC II*, local gene synteny was drawn as follows: *MHC II* protein sequences (as used in the overall immune gene characterization) were used in a TBLASTN search against the Ensembl genome assemblies of zebrafish, fugu and stickleback using the Ensembl web TBLASTN function with default parameters. BLAST hit regions were

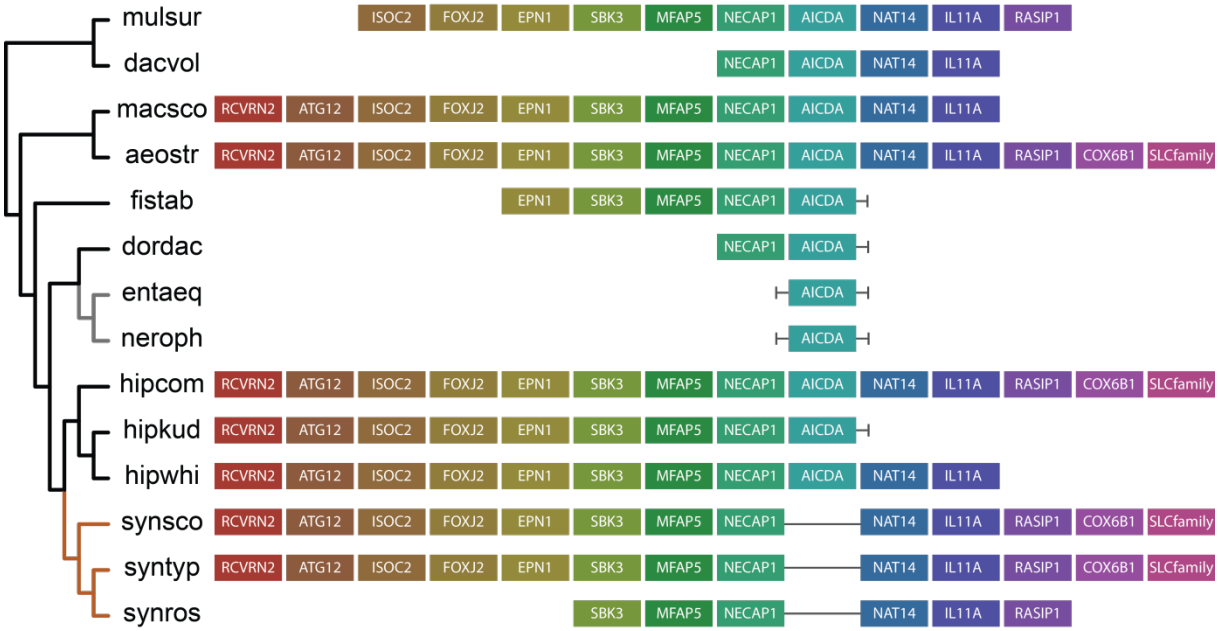
recursively subjected to a BLASTX search using the NCBI blast server with default parameters. If the recursive hits were *MHC II* sequences, we extracted peptide sequences from predicted gene models located up- and downstream from the Ensembl query. These peptide sequences were subjected to a BLASTP, again using the NCBI web BLAST tool, to either verify or obtain annotation. In the cases of no recursive hit, or where the hit was an unknown protein, the query peptide sequence was given "ORF" as annotation. All peptide sequences obtained from Ensembl, together with previously extracted *MHC II* sequences from *Hippocampus comes* as well as the MHC II peptide sequences used for *MHC II* gene mining, were used in a TBLASTN search against the scaffolds of the Syngnatiformes species as well as the downloaded assemblies from *Hippocampus comes* and *Syngnathus scovelli*. Unitigs were not used as they were generally too short for synteny. The regions surrounding *MHC II* hits in *Hippocampus comes* were extracted and subjected to a BLASTX using the NCBI BLAST tool with default parameters to obtain information about surrounding genes. The same was also done for our two *Hippocampus* draft genomes, *Hippocampus kuda* and *Hippocampus whitei*. In some cases, these regions contained different genes than the syntenic genes already downloaded from Ensembl and representative peptide sequences of these were obtained from GenBank and added to the original query list. With these included, the TBLASTN search was rerun (as described above). Synteny was drawn for zebrafish, stickleback and fugu as well as the *Hippocampus comes*, *Hippocampus kuda* and *Hippocampus whitei* based on the tabular BLAST output files. For the *Syngnathus* species we sorted the output, highlighting the best hit for each query and sorted those by genomic region to get preliminary overview of potential syntenic regions compared to zebrafish, stickleback, fugu and the *Hippocampus* species. Having pinpointed potential genomic regions, the entire tabular BLAST output from the *Syngnathus* species was sorted on genomic region and then on query starting point. The gene order of the putative *MHC II* loss regions was drawn into the local synteny accordingly.

5.5. Results of for candidate gene mining and annotation

5.5.1. Activation-induced cytidine deaminase (*AID/AICDA*) - loss of *AICDA* in *Syngnathus*

With its responsibility for the unique receptor diversity of the antibodies, *AICDA* is a key gene of the MHC II pathway. A well conserved copy of *AID* (AICDA) was found in all investigated species except in the three *Syngnathus* species. For *Doryrhamphus dactyliophorus* the gene was divided between two scaffolds. An *AICDA* fragment was found

for *Nerophis ophidion* and *Entelurus aequoreus*. The loss of *AID* (*AICDA*) was further explored by characterizing the local gene synteny between all the investigated species (SI Appendix Figure S3). Genes located directly up- and downstream of *AICDA* in non-*Syngnathus* species were co-localized on the same linkage group or scaffold in the three *Syngnathus* species supporting a loss of *AICDA* in these species (SI Appendix Figure S3, SI Appendix Table S6, SI Appendix Data Set 1: 3AID_clean & 4AID_raw).



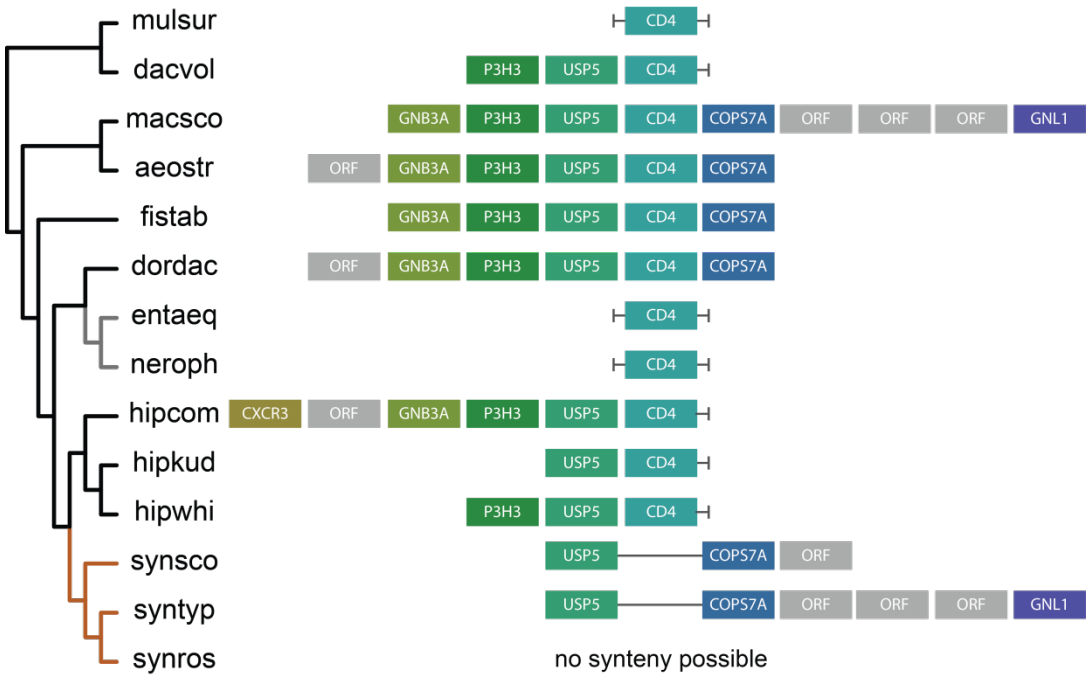
SI Appendix Figure S3: Local gene synteny of the *AICDA* region in all species contrasted to the putative *AICDA* regions of the *Syngnathus* species.

SI Appendix Table S6: Genome assembly localization of *AICDA* (for *Syngnathus*, putative *AICDA* region).

Species	**Genomic location of *AICDA* syntenic region**
Mullus surmuletus	scf7180003351178
Dactylopterus volitans	scf7180002994818
Macroramphorus scolopax	scf7180002135645
Aeoliscus strigatus	scf7180001813954
Fistularia tabacaria	scf7180005990769
Doryramphus dactylophorus	scf7180001777170 scf7180001777885
Entelurus aequoreus	utg7180017948567
Nerophis ophidion	scf7180014164377
Hippocampus comes	KV879925.1
Hippocampus kuda	scf7180002363415
Hippocampus whitei	scf7180002671084
Syngnathus scovelli	LG3
Syngnathus typhle	scaffold00035
Syngnathus rostellatus	scf7180002437117

5.5.2. T-cell surface glycoprotein Cluster of differentiation 4 (*CD4*) - loss of *CD4* in *Syngnathus*

CD4 mediates successful receptor binding and activation of CD4+ T lymphocytes upon MHC II- mediated antigen presentation. A copy of *CD4* with all previously described exons was found in all investigated Syngnathiformes species with the exception of *Syngnathus* – a loss that we further explored using local gene synteny similarly to *AICDA*. In the *CD4*-containing genomes, the gene was often located on a short assembly scaffold with few or no flanking genes. No synteny was possible for *Mullus surmuletus*, *Nerophis ophidion* and *Entelurus aequoreus*. Sparse synteny was obtained for the *Hippocampus* species. The most common pattern of *CD4* flanking genes was used to screen possible scaffolds and linkage groups in the *Syngnathus* species. A possible syntenic scaffold was found in *Syngnathus scovelli* and *Syngnathus typhle* further supporting the genetic loss of *CD4*. For *Syngnathus rostellatus* we found several possible scaffolds containing syntenic genes, but without enough flanking sequence to support the loss of *CD4* in this particular species. However, based on the lack of a *CD4*-like sequence in the remaining *Syngnathus,* we assume a loss of *CD4* in *Syngnathus rostellatus* as well (SI Appendix Figure S4, SI Appendix Table S7, SI Appendix Data Set 1: 5CD4_clean_conserved_block_no_frag & 6CD4_clean_conserved block)).



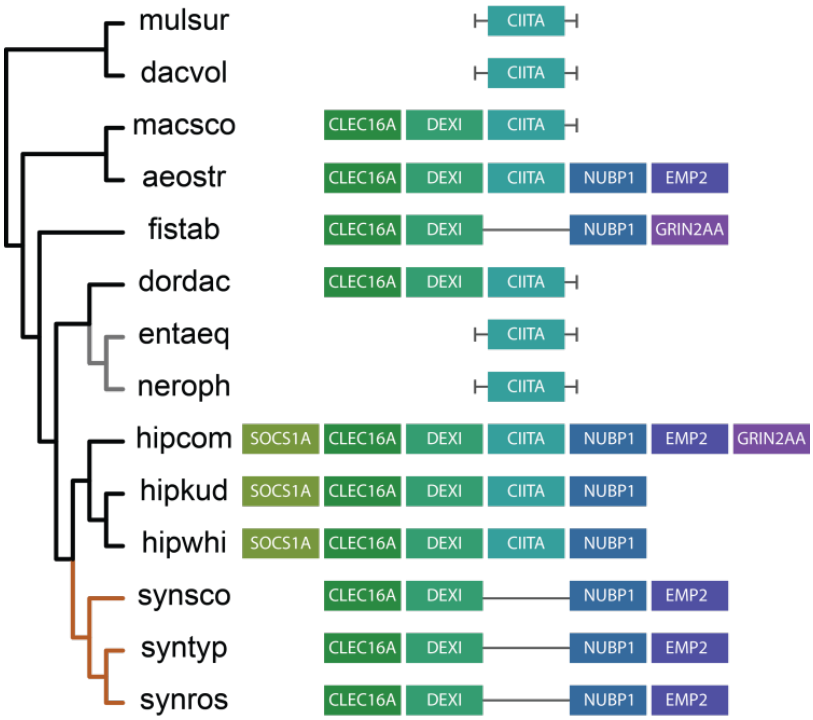
SI Appendix Figure S4: Local gene synteny of the *CD4* region in all species contrasted to the putative *CD4* regions of *Syngnathus* species. Synteny was not possible for *Syngnathus rostellatus* due to the fragmented nature of the genome in this region.

SI Appendix Table S7: Localization of *CD4* (for *Syngnathus*, putative *CD4* region) in the respective genome assembly.

Species	**Genomic location of *CD4* syntenic region**
Mullus surmuletus	scf7180003339999
Dactylopterus volitans	scf7180002998269
Macroramphorus scolopax	scf7180002137464
Aeoliscus strigatus	scf7180001813266
Fistularia tabacaria	scf7180005996797
Doryramphus dactylophorus	scf7180001789914
Entelurus aequoreus	scf7180039925332
Nerophis ophidion	scf7180014150266
Hippocampus comes	KV879925.1
Hippocampus kuda	scf7180002375734
Hippocampus whitei	scf7180002656493
Syngnathus scovelli	scaffold_803
S *Syngnathus typhle*	scaffold00169
Syngnathus rostellatus	No synteny possible

5.5.3. Class II Major histocompatibility complex transactivator (*CIITA*) - lost in *Syngnathus* but also possible loss in *Fistularia tabacaria*

CIITA controls the expression of MHC II genes in antigen presenting cells (APCs). A copy of *CIITA* with all 19 exons was found in all species with the exception of the *Syngnathus* species, but also surprisingly in *Fistularia tabacaria. CIITA*, being a member of the NOD-like receptor family (NLR), displayed significant sequence variation in the first 10 exons in all investigated species (compared to the human exon-intron structure) to such an extent that these exons are difficult or impossible to detect based on sequence similarity alone without the support of RNA data. For this reason, the leading exons are not included in downstream analyses. The remaining part of the gene is well conserved and was easily detected in the no-loss species assemblies. Where *CIITA* was absent, we further explored the loss using local gene synteny. We found good putative *CIITA* containing scaffolds and linkage groups in the *Syngnathus* species and in *Fistularia tabacaria* based on local gene synteny towards the other species, but could not detect any fragments of *CIITA* between the first flanking gene directly up- and downstream. Collectively, this indicates a loss of *CIITA* in all three *Syngnathus* and in *Fistularia tabacaria* (SI Appendix Figure S5, SI Appendix Table S8, SI Appendix Data Set 1: 7CIITA_raw & 8CIITA_temp_section_with_ensembl & 9CIITA_temp_section).



SI Appendix Figure S5: Local gene synteny of the *CIITA* region in all species contrasted to the putative *CIITA* regions of *Syngnathus* species.

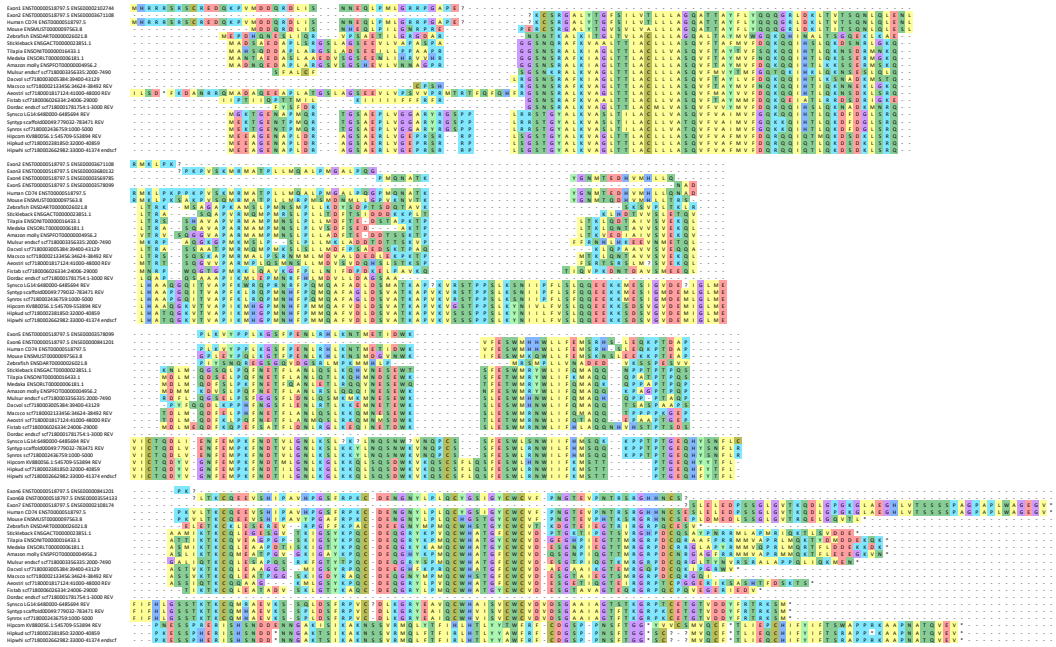
SI Appendix Table S8: Genome assembly localization of *CIITA* (for *Syngnathus*, putative *CIITA* region).

Species	**Genomic location of *CIITA* syntenic region**
Mullus surmuletus	scf7180003352461
Dactylopterus volitans	scf7180003014039
Macroramphorus scolopax	scf7180002139468
Aeoliscus strigatus	scf7180001817874
Fistularia tabacaria	scf7180005993608
Doryramphus dactylophorus	scf7180001773600
Entelurus aequoreus	scf7180039813037
Nerophis ophidion	scf7180014049719
Hippocampus comes	KV879887.1
Hippocampus kuda	scf7180002375048
Hippocampus whitei	scf7180002661233
Syngnathus scovelli	LG16
Syngnathus rostellatus	scf7180002437310
Syngnathus typhle	scaffold00004

5.5.4. MHC II invariant chain (*CD74*) – sequence variation and exon loss in *Hippocampus* and *Syngnathus*

CD74 prevents premature peptide binding of *MHC II*. A copy of *CD74* was found in all investigated species (SI Appendix Data Set 1) (a fragment from *Doryrhamphus dactyliophorus* was, however, not included in the positive selection alignment). However, the *CD74* sequence for *Hippocampus* and *Syngnathus* species diverged from the remaining

species, as well as from the reference sequences of well characterized species obtained from Ensembl. Based on the human exon-intron structure, the coding sequence of exon 1 was very different to the reference species. An existing start codon, however, suggests that *CD74* could be functional in *Hippocampus* and *Syngnathus*, its length and sequence were further conserved between *Hippocampus* and *Syngnathus*. Directly downstream a well conserved exon 2 is followed by a diverged exon 3 (CLIP) in *Hippocampus* and *Syngnathus*. There are some conserved amino acids supporting the placement of exon 3. Exon 4 is very short, but well enough conserved followed by the somewhat longer, and again well conserved exon 5. Exon 6 is well conserved across all teleost species. The exon 6b is in *Syngnathus* divergent from other teleost species and replaced in *Hippocampus*. The final exon 7 is again not only divergent from Ensembl species but the sequence is also different between *Syngnathus* and *Hippocampus* (Figure 3, SI Appendix Figure S6; SI Appendix Data Set 1: 10CD74_clean & 11CD74_exon_loss_determination_clean).



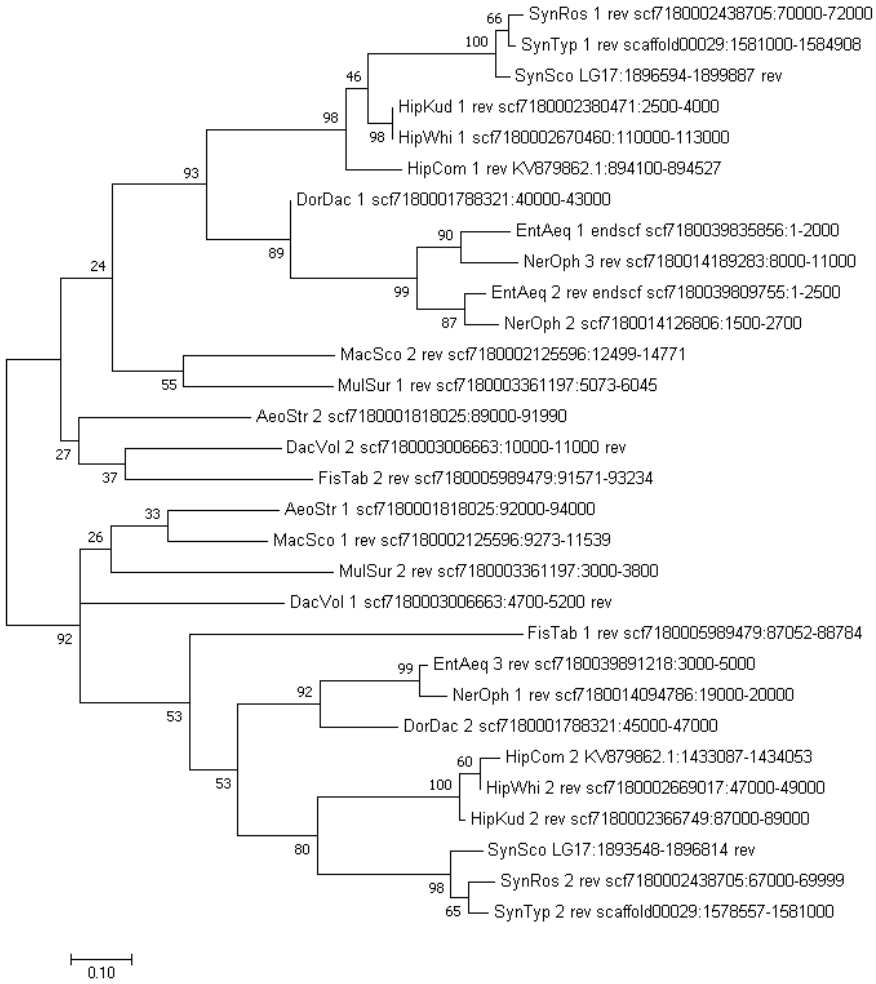
SI Appendix Figure S6: Alignment of *CD74* in all species contrasted with sequences from Ensembl: human, mouse, zebrafish, stickleback, medaka, tilapia and Amazon molly. The exon sequences from the human *CD74* gene is included. CLIP is located in the human exon 3. Exon 7 is not found in the *Hippocampus* species. Exon 8 is highly variable.

5.5.5. Autoimmune regulator *(51)* - found in all species. Increased sequence diversity in *Syngnathus* and *Hippocampus*.

AIRE drives negative selection on self-recognizing T lymphocytes. A copy of *AIRE* was found in all investigated species with fragments in *Nerophis ophidion* and *Entelurus aequoreus*. In *Hippocampus comes*, the sequence contained a misassembly and thus only a fragment is included in the alignment (SI Appendix Data Set 1: 16AIRE_redone_cean_with_HipComfrag). All investigated species displayed well conserved *AIRE* regions determining gene presence. Compared to the human exon-intron structure, exons 7 and 10 were not alignable to the fish sequences. Furthermore, exon 8 appears ultra-conserved across all species. *Hippocampus* and *Syngnathus* species sequences displayed some additional sequence diversity compared to the other teleost species, in particular in exons 3-6 and 12. Thus, the final alignment only contains regions of *AIRE* with confident alignment towards other teleost species (SI Appendix Data Set 1: 18AIRE_redone_with_ensembl).

5.5.6. Beta-2-microglobulin (*B2M*)– present in all species, copy number variation

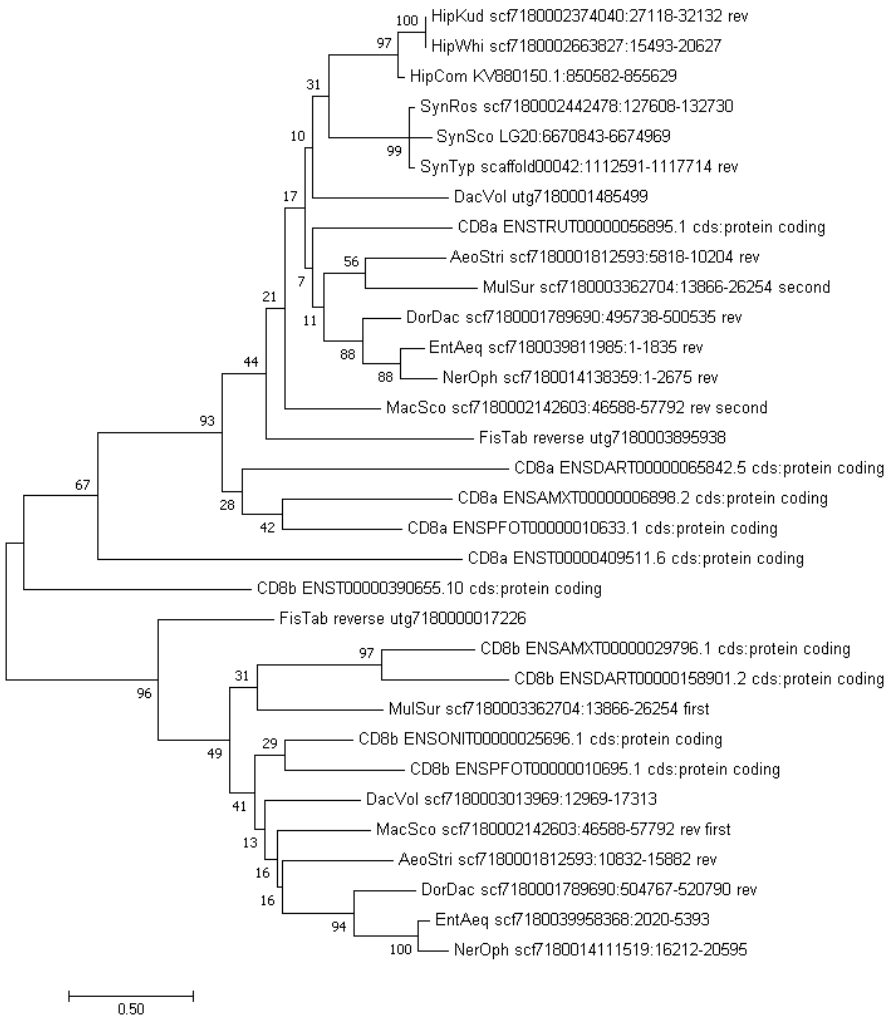
B2M was found in all investigated species with either 2 or 3 copies and appears well conserved. The copies cluster into two well-defined clades and within each clade cluster according to the species phylogeny (SI Appendix Figure S7, SI Appendix Data Set 1: 20B2M_clean_trimmed_nucl & 21B2M_clean & 22B2M_raw).



SI Appendix Figure S7: Maximum likelihood tree of the *B2M* gene. The phylogeny is based on the JTT model, 500 bootstraps, derived from a protein alignment of all detected *B2M* sequences.

5.5.7. *CD8A/B* – *Syngnathus* and *Hippocampus* only have *CD8A*

We found *CD8A* and *CD8B* in all species except in in *Hippocampus* and *Syngnathus* where only *CD8A* was found. For *Fistularia tabacaria*, both *CD8A* and *CD8B* was found in the unitigs and not in the scaffolds. Similarly, for *Dactylopterus volitans*, *CD8A* was found in the unitigs. All obtained *CD8* sequences cluster into two well-defined clades in the phylogenetic tree, either *CD8A* or *CD8B* (SI Appendix Figure S8, SI Appendix Data Set 1: 23CD8_clean_with_ensembl & 24CD8_clean & 25CD8_clean_compactsection).



SI Appendix Figure S8: Maximum likelihood tree of the *CD8A/B* gene, JTT model, 500 bootstrap of the *CD8A/B* protein alignment including reference sequences from Ensembl: human (ENST), zebrafish (ENSDAR), cave fish (ENSAMX), tilapia (ENSONI) and Amazon molly (ENSPFO).

5.5.8. Major Histocompatibility Complex I (*MHC I*)

The expansion of *MHC I* diversity in cod was suggested as a functional compensation for the loss of *MHC II*(52). We aimed to identify whether in Syngnathiformes MHC I coevolved with male pregnancy and the loss of a functional MHC II pathway. Number of detected *MHC I* exon fragments were highly variable within and among species, in line with previous PCR based reports in many fish species. *A. strigatus* and *F. tabacaria* had very low numbers of both *MHC I* exon 2 and exon 3, in contrast to all syngnathids but also *M. surmuletus* and *M. scolopax*. *M. surmuletus* and *N. ophidion* had particularly high numbers of exon 2 and exon 3 fragments. While in the hypervariable exon 2 and exon 3, no phylogenetic pattern of *MHC I* evolution in Syngnathiformes could be detected, using the most conservative exon 4 of the *MHC I* gene, we found that the copy number of *MHC I* was high in all Syngnathiformes species with male pregnancy (the *Nerophinae* with external

male pregnancy (27-42 copies), *Hippocampus* (20-36 copies) and *Syngnathus* (24–44 copies, *Syngnathus scovelli* excluded) with internal male pregnancy) compared to the species without male pregnancy (5-10 copies) (Figure 2; SI Appendix Table S9). The number of *MHC I* fragments in *Syngnathus scovelli* has to be treated with caution, as from this published genome(3) unitigs, from which the search for the other genomes was done, were not available.

We examined the relatedness of Syngnathiformes *MHC I* sequences to homologous sequences in other, well-curated fish species (SI Appendix Data 1: 26MHCI_exon2 & 27MHCI_exon3 &28MHCI_exon4). The resulting tree topologies of two phylogenies obtained by either a long portion of the exon 4 (SI Appendix Figure S9) or the second half of the exon 4 (SI Appendix Figure S10), respectively, revealed that the *MHC I* variants group into two very divergent clades, which are partially distinct from all other *MHC I* types distributed among other fish families. While these sequences are part of the *MHC I* U lineage clade in the gene tree(53), the distinct cluster of syngnathid *MHC I* sequences supports a potential coevolution of *MHC I* with male pregnancy (SI Appendix Figure S9-S10). These lineage-specific *MHC I* variants likely increase the ligand repertoire and suggest a possible function within the cross-presentation pathway, similar to Atlantic cod.

SI Appendix Table S9: Number of detected *MHC I* fragments in the different assemblies. Fragment number does not directly reflect *MHC I* gene number as fragments may overlap. Fragment number is given per exon. Further analyses were done with exon 4. * previously published genome where unitigs were not available, results have to be treated with caution. ** large fragmented genomes.

Species	Exon 2	Exon 3	Exon 4
Mullus surmuletus	72	81	9
Dactylopterus volitans	12	13	11
Macroramphorus scolopax	47	30	5
Aeoliscus strigatus	5	8	5
Fistularia tabacaria	6	8	10
Doryramphus dactylophorus	33	23	27
*Entelurus aequoreus***	88	140	42
*Nerophis ophidion***	50	36	37
*Hippocampus comes**	32	27	36
Hippocampus kuda	34	30	20
Hippocampus whitei	37	29	29
*Syngnathus scovelli **	11	11	10
Syngnathus typhle	28	26	24
Syngnathus rostellatus	53	66	44

5.5.9. Major Histocompatibility Complex II (*MHC II*)

We were unable to identify *MHC II alpha* and *beta* genes in all three *Syngnathus* species in contrast to all other investigated species. In *M. scolopax* we detected an extreme amount

of *beta* fragments compared to the other investigated species. *M. surmuletus* had a similar trend, but not as significant. We also observed that the *MHC II* fragments found in *Hippocampus* looked different compared to the other investigated species. We examined the relatedness of Syngnathiformes *MHC II beta* sequences to homologous sequences in other, well-curated fish species (SI Appendix Data Set 1: 29MHCII_beta_complete & 30MHCII_beta_1-86aa & 31MHCII_beta_91-184aa). The resulting tree topologies of two phylogenies obtained by the first and second portion of the gene, respectively, revealed that the *MHC II beta* variants group into two very divergent clades, which are partially distinct from all other *MHC II beta*-types distributed among other major fish families such as salmonids, flatfishes, carp-like fishes and perches (SI Appendix Figure S11-S12). In particular, the phylogeny based on the more conserved second section of the gene (exon 3, 4) supports the distinct status of those clades from all other Actinopterygii by high bootstrap values (>99%). In contrast, the phylogeny on the first section of the gene under strong positive selection shows two distinct types only, one of which shows several syngnathid species intermingled with other Actinopterygii species. Note that in the more informative second-section phylogeny, aside from genus *Hippocampus*, only *Dactylopterus volitans* as a sister species to the syngnathids possess one gene variant belonging to both distinct sequence groups. Interestingly, only the sequence lineage without two 2 bp deletions within peptide binding exon 2 (SI Appendix Data Set 1: 29MHCII_beta_complete & 30MHCII_beta_1-86aa & 31MHCII_beta_91-184aa) is populated by most of the other sequenced syngnathid taxa including those with external egg carrying (i.e. *Entelurus aequoreus*, *Nerophis ophidion, Doryrhamphus dactylophorus*).

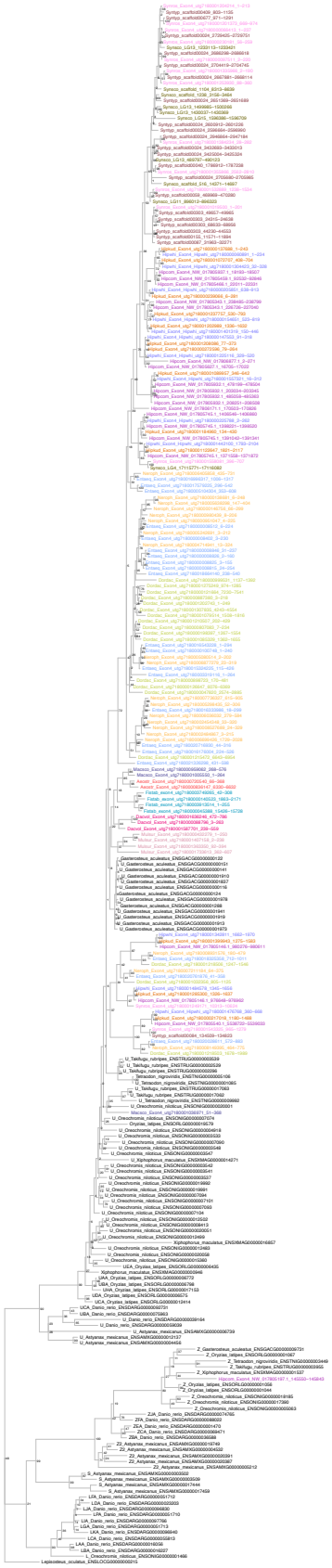
5.5.10. MHC class IIB analysis in Hippocampus

While the genus *Syngnathus* has no detectable genes or pseudogenized remains of both *MHC II* genes (*alpha* and *beta*) in their genomes, the genus *Hippocampus* as second genus with full pregnancy was reported to possess *MHC II alpha* and *beta* sequences(54), which we can confirm in this study. *Hippocampus*-sequences in one of the 2 distinct sequence clades are characterized by two 2-aa indels within the alpha 2 domain (=exon 2, peptide binding region, SI Appendix Data Set 1: 29MHCII_beta_complete & 30MHCII_beta_1-86aa & 31MHCII_beta_91-184aa) with unknown but presumably critical effects on functionality. In *Hippocampus*, in two sequences where the completeness of the *MHC II beta* sequence permits a test for disulfide bridges (one each for *H. comes* and *H. rano*) we found that critical di-sulfide (cystein) bridges within exon 2 and 3 are absent. These are

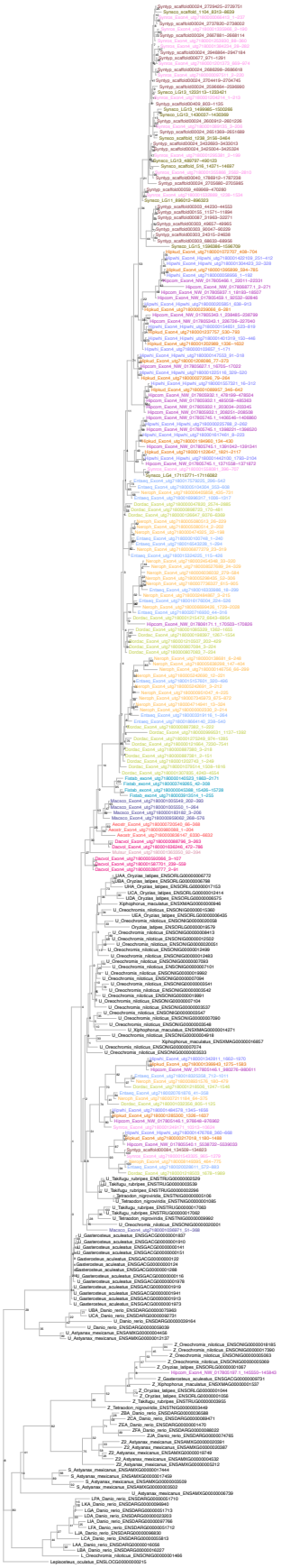
required for the appropriate tertiary structure of the *MHC II* molecules to form the peptide-binding grove(45, 46). None of the *Hippocampus comes MHC II beta* sequences tested is predicted to form the required disulfide bonds between cystein-residues located in exons 2 and 3, respectively, which stabilize the peptide-binding pocket(45, 46). Since the cystein residues themselves are still present (see aa positions 29 - 94; and 132 - 188), this is probably due to major re-arrangements in the secondary structure of the *MHC II beta* molecule. In contrast, with high likelihood, DISULFIND identified two di-sulfide bridges in the *MHC II beta* molecules of all three "control" species. Along with the finding that both major *MHC II beta* sequence variants in *Hippocampus* spp. were very divergent to other homologous Actinopterygii genes (SI Appendix Figure S11-S12), it is doubtful whether *Hippocampus MHC II* molecules can form and be functional, leaving aside other genes required for the MHC class II pathway.

SI Appendix Table S10: *MHC II* fragments detected in the genome assemblies of the different Syngnathiformes species. Fragment number does not directly reflect *MHC II* gene number as fragments may overlap. * deep sequenced genomes. ** large fragmented genomes.

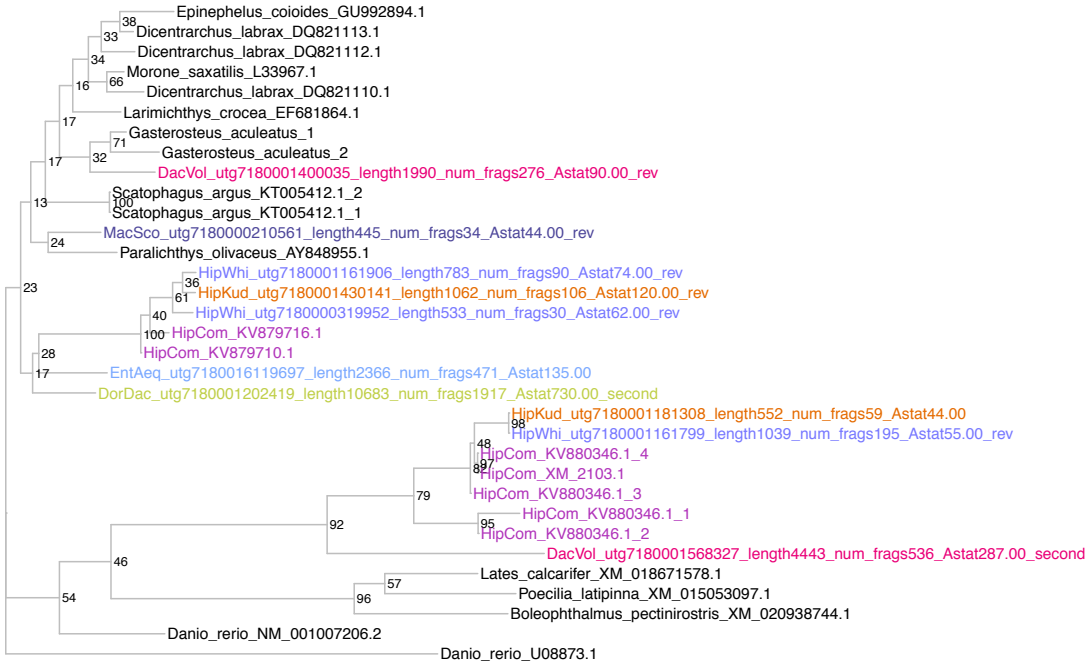
Species	MHC class II Alpha fragments	MHC class II Beta fragments
Mullus surmuletus	9	15
Dactylopterus volitans	4	4
Macroramphorus scolopax	4	43
Aeoliscus strigatus	3	4
Fistularia tabacaria	2	3
Doryramphus dactylophorus	3	4
Entelurus aequoreus **	1	1
Nerophis ophidion **	1	1
Hippocampus comes *	6	6
Hippocampus kuda	5	9
Hippocampus whitei	5	7
Syngnathus scovelli	0	0
Syngnathus typhle *	0	0
Syngnathus rostellatus	0	0



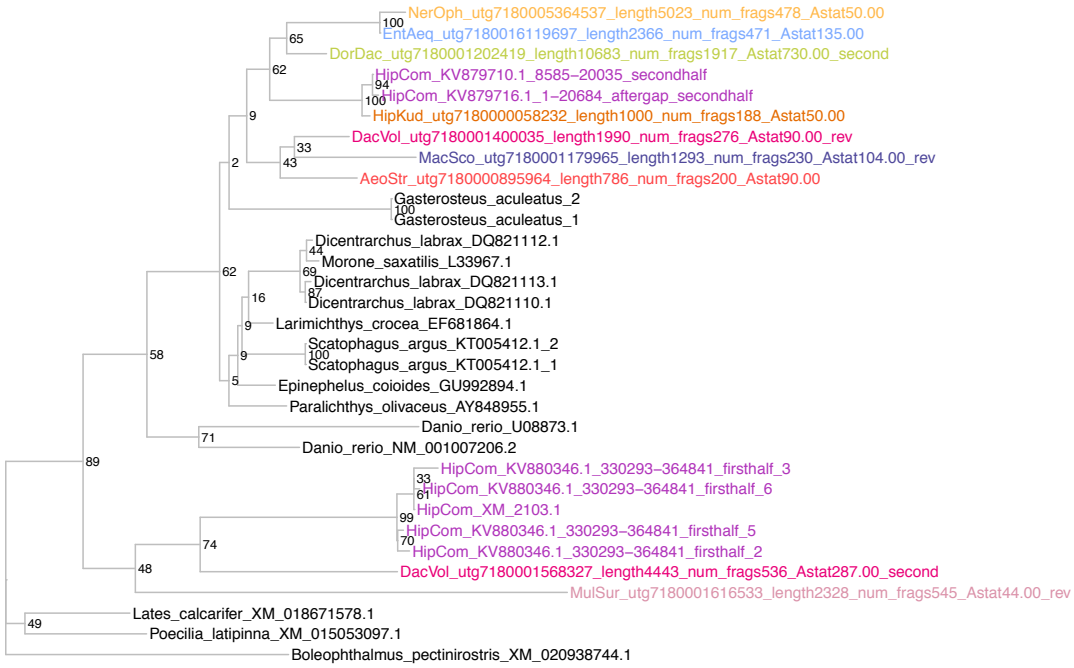
SI Appendix Figure S9: Maximum likelihood phylogeny of exon 4 of the *MHC I* gene in the Syngnathiformes and other fish species using sequences spanning the entire exon. Alignment was performed with MUSCLE and subsequent manual editing.



SI Appendix Figure S10: Maximum likelihood phylogeny of exon 4 of the *MHC I* gene in the Syngnathiformes and other fish species using sequences spanning the second half of the exon. Alignment was performed with MUSCLE and subsequent manual editing.



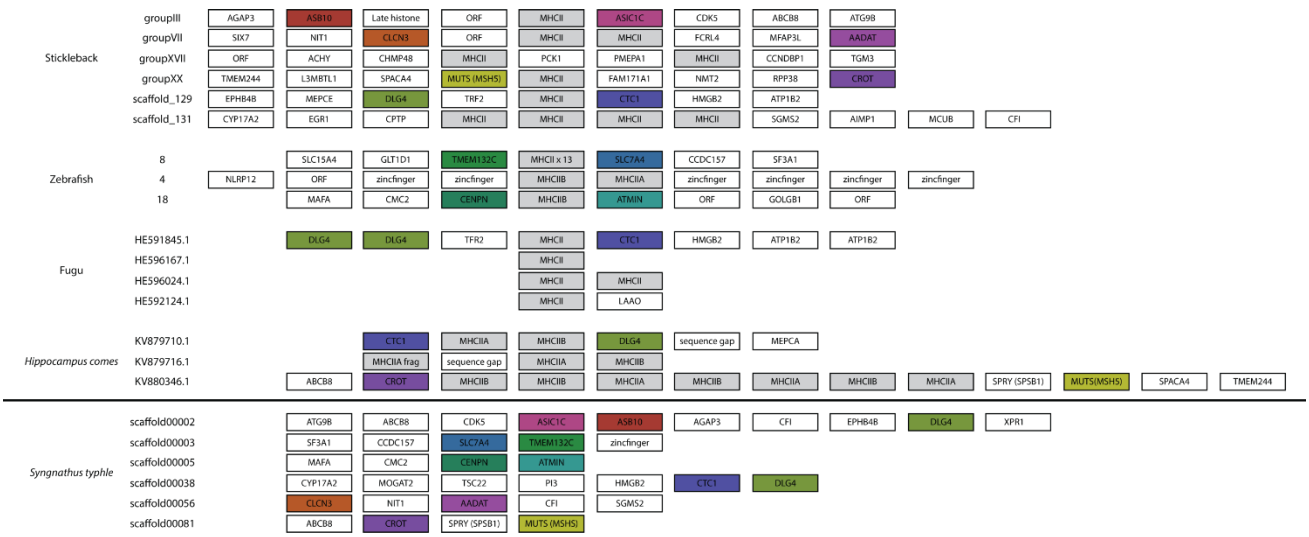
SI Appendix Figure S11: Maximum likelihood phylogeny of exon 4 of the *MHC II* gene in the Syngnathiformes and other fish species using sequences spanning the first half of the exon. Alignment was performed manually.



SI Appendix Figure S12: Maximum likelihood phylogeny of exon 4 of the *MHC II* gene in the Syngnathiformes and other fish species using sequences spanning the second half of the exon. Alignment was performed manually.

5.5.11. MHC II synteny

To obtain safer conclusions upon the loss of the *MHC II*, synteny was characterized. *MHC II* regions are only available in sufficient quality for three reference species; Stickleback, zebrafish and fugu from the Ensembl database. It is exceedingly difficult to assemble *MHC* regions due to recent gene duplications in combination with rampant and ongoing gene conversion(55). The regions were detected using the Ensembl TBLASTN tool with the *MHC II* query sequences listed above. All hit regions were subjected to a reciprocal blast using the NCBI blast tool with default parameters. Regions determined to contain *MHC II* were noted and protein sequences of flanking genes downloaded from Ensembl. These flanking gene protein sequences were used in a TBLASTN towards all Syngnathiformes assemblies together with *MHC II* sequences. In each of the Ensembl reference species, Stickleback, zebrafish and fugu, we detected 6, 3 and 4 *MHC II* regions, respectively. The regions appear to be poorly assembled, but contain some common flanking genes. Synteny was mainly investigated in the high-coverage genomes of *Syngnathus typhle* and *Hippocampus comes*. In *Hippocampus comes* we found 3 putative *MHC II* regions and in *Syngnathus typhle* we investigated local gene synteny in 6 putative *MHC II*-loss regions (SI Appendix Figure S13).



SI Appendix Figure S13: *MHC II* synteny

5.5.12. Recombination activating gene 1 and 2 (*RAG1* and *RAG2*)

RAG1 and *RAG2* facilitate V(D)J recombination. Both *RAG1* and *RAG2* genes appear well conserved in all investigated species, *RAG2* more so than *RAG1*. For *RAG1* there is some

length variation in the first 120 aa, more specifically between aa 46 and aa 92. AA 12 – 310 is likely to interact with importin alpha (SI Appendix Data Set 1: 32RAG1 &33RAG2).

5.5.13. Transporter 1 and 2 ATP binding cassette subfamily B member (*TAP1* and *TAP2*)

TAP1 and *TAP2* function as heterodimers in the transport of antigens(56). We found either two or three *TAP*-like genes in all investigated species with the exception of 5 fragments in *Entelurus aequoreus* collectively likely representing 3 different genes, and 4 in *Dactylopterus volitans* of which two were identical. The genes from 3 well supported phylogenetic clades, all *TAP1* genes together and two clades with *TAP2* (SI Appendix Data Set 1: 34TAPS_clean).

6. Detection of positive selection across Syngnathiformes

6.1. Assess positive selection in pregnancy and immune genes

The large changes in genetic regulation that the evolution of pregnancy requires in developmental and immunological processes can be expected to reflect signs of positive selection on genes as they are recruited to perform novel functions in pregnancy and changed functions in immunity. To assess this, we aimed to detect positive selection in the candidate gene alignments across the Syngnathiformes.

Orthologues of coding sequences for the candidate genes were curated manually as described above and aligned at the codon level. Due to fragmented draft genome sequences, some alignments do not represent the full gene. The bases analyzed are given as a percentage of the zebrafish reference gene in SI Appendix Table S11. For genes that are present in most genomes in two copies, the copies have also been analyzed separately and denoted by numbers, see also gene trees in SI Appendix Data Set 3.

Site-specific, positive selection was assessed using the HYPHY package(57). Gene wide selection was tested using the BUSTED method, a likelihood ratio test that compares a model that allows selection in a chosen foreground branch against a null model without selection. We tested for positive selection across the whole tree and additionally with the syngnathids as the foreground branches, for the latter the Datamonkey web service was employed due to the convenience of foreground branch selection(58). Selection in specific branches was assessed using the adaptive branch-site random effects likelihood (aBSREL) method(59). The gene tree for use in the aBSREL analysis was calculated using

RAxML (ver. 8.2.10) with the GTRCAT model. Selection on specific codons was evaluated with the Mixed Effects Model of Evolution (MEME) method(60). We used a p-value cutoff of 0.05 for all methods.

SI Appendix Table S11: Summary of the results of the tests for positive selection. Shown are the p-values for the BUSTED test over all sequences, the BUSTED test using the syngnathid sequences as foreground branches, and the sites identified by the MEME test. Bold p-values are significant. In cases where the gene alignment had to be truncated due to missing data the percentage of the gene analyzed is shown for the Zebrafish ortholog as reference, the EnsEBML gene ID for this gene is given.

gene	function	p-val BUSTED	p-val LRT BUSTED syn foreground	aBSREL significant branches	MEME significant sites	% analyzed zebrafish gene	zebrafish gene ID
B2M	MHC I pathway	0.0003	0.01	Node13	1; 44; 68; 72; 85; 91	80,34	ENSDARG00000053136
B2M copy 1	MHC I pathway	0.007	0.041	-	-	80,34	ENSDARG00000053136
B2M copy 2	MHC I pathway	0.2511	0.155	Node14	19; 42; 64; 73; 92	80,34	ENSDARG00000053136
CD8	MHC I pathway	0	0	MacSco2, NerOph1, Node26, FisTab1	32; 34; 40; 65; 70; 90; 92; 98; 103; 108; 112; 117; 124; 133; 135; 144; 164; 201; 207; 257; 260; 265; 308; 325; 330; 355; 365	100,00	ENSDARG00000044797
CD8 alpha	MHC I pathway	0	0	Node5, Node4	16; 18; 28; 32; 36; 43; 54; 56; 114; 116; 122; 123; 125; 129; 142		ENSDARG00000044798
TAPs	MHC I pathway	0	0	HipCom1, AeoStr3	16; 141; 171; 208; 277; 325; 422; 452	70,21	ENSDARG00000036787
TAPs copy 1	MHC I pathway	0	0	HipCom	22; 105; 141; 175; 242; 277; 439	70,21	ENSDARG00000036787
TAPs copy 2	MHC I pathway	0.014	0.098	aeostr_1_1	14; 49; 85; 96; 101; 124; 130; 164; 189; 208; 283; 325; 412; 432; 443; 499	70,21	ENSDARG00000036787
AICDA (AID)	MHC II pathway	0.4763	0.841	DorDac	31; 194	97,16	ENSDARG00000015734
AIRE	MHC II pathway	0	0	Node10, HipCom, Node20, Node22, Node1	5; 76; 109; 119; 128; 137; 139; 150; 158; 161; 277; 305; 336	67,32	ENSDARG00000056784
CD4	MHC II pathway	0	0	Node10	2; 39; 46; 60; 66; 99; 105; 110; 145; 160; 161; 210; 211; 222; 254; 301; 345; 372; 388; 428; 445; 457	97,19	ENSDARG00000115309
CD74	MHC II pathway	0	0.071	-	66; 69; 103; 104; 121	67,65	ENSDARG00000009087
CIITA	MHC II pathway	0.0093	0.011	Node12, HipWhi	4; 14; 26; 28; 37; 59; 111; 116; 177; 312; 385; 389; 397; 400; 409; 489; 537; 543; 585; 602; 609; 615; 668; 723; 739	75,39	ENSDARG00000090851
RAG1	MHC II pathway	0.0005	0	Node16, Node20	3; 38; 64; 71; 74; 88; 95; 149; 150; 160; 190; 229; 239; 430; 499; 630; 662; 766; 1084; 1088	100,00	ENSDARG00000052122
RAG2	MHC II pathway	0.2885	0.952	-	213; 455; 476	97,17	ENSDARG00000052121
APOA1	Pregnancy	0	0	Node19	64; 69; 75; 76; 143; 151; 170; 192; 218; 239; 251	100,00	ENSDARG00000012076
ATF3	Pregnancy	0.0054	0.212	-	93; 99; 125	100,00	ENSDARG00000007823
CA4	Pregnancy	0.657	0.247	-	-	93,79	ENSDARG00000043589
CEBPB	Pregnancy	0.1835	0.682	-	53; 109; 140	100,00	ENSDARG00000042725
DLX3b cluster1	Pregnancy	0.4015	1	-	-	100,00	ENSDARG00000014626
DLX3b cluster2	Pregnancy	0	0	Node10	15; 41; 75; 105; 240; 241; 242; 245; 247; 248; 255; 263; 268; 271; 276	100,00	ENSDARG00000014626
JUNB	Pregnancy	1	0.927	-	204	100,00	ENSDARG00000074378
JUNB copy 1	Pregnancy	0.0301	0.372	-	128; 200	100,00	ENSDARG00000074378
JUNB copy 2	Pregnancy	1	1	-	201	100,00	ENSDARG00000074378
PRKCD	Pregnancy	0.0444	0.008	Node31, MulSur1	167; 173; 184; 186; 196; 203; 305; 312; 328; 438	81,46	ENSDARG00000070651

PRKCD copy 1	Pregnancy	0.0317	0.026	-	53; 179; 185; 212; 218; 219; 225	81,46	ENSDARG00000070651
PRKCD copy 2	Pregnancy	0.5162	0.749	-	13; 498	81,46	ENSDARG00000070651
PTGFRN	Pregnancy	0	0	SynRos	101; 111; 273; 403; 405; 416; 449	98,15	ENSDARG00000075505
RHAG	Pregnancy	1	0.932	-	47; 48	45,87	ENSDARG00000019253
TF	Pregnancy	0	0	Node5, HipCom, AeoStr, FisTab	5; 8; 15; 76; 95; 124; 155; 158; 162; 165; 203; 205; 261; 284; 332; 339; 366; 470; 488; 531; 538; 580; 628; 629; 633; 637	96,14	ENSDARG00000016771
hemoglobin beta	Hemoglobin	0	0	synsco	30; 91	100,00	ENSDARG00000109554
hemoglobin alpha 1	Hemoglobin	0.0266	1	-	20; 30; 78	57,75	ENSDARG00000110250
hemoglobin alpha 2	Hemoglobin	0	0	dacvol_1_1_1, dacvol_1_1	3; 6; 20; 21; 85	100,00	ENSDARG00000079305
hemoglobin alpha 3	Hemoglobin	0	0.003	-	5; 20; 45; 53; 73	41,03	ENSDARG00000079078

6.2. Signs of positive selection on immune system genes in syngnathids

We have identified major re-arrangements in the repertoire of immune genes that are considered standard for gnathostome immunity and now examined whether or not this is also associated with accelerated evolution in some of the important target genes. To begin with, *AICDA* (*AID*), responsible for the unique receptor diversity in B-cells, was lost in *Syngnathus*, but is retained in the *Hippocampus* species. We find no signs of positive selection with the BUSTED method with *Hippocampus* species as the foreground branch, however, aBSREL identifies possible positive selection in *Doryrhamphus dactylophorus* belonging to the *Nerophinae*. The *AIRE* gene, involved in negative selection of T-cells, is present in all Syngnathiformes and shows signs of positive selection in syngnathids (BUSTED likelihood ratio test). The aBSREL method indicates selection in multiple branches of the gene tree, specifically in *Hippocampus*, however, not in *Syngnathus*. This was expected given that *AIRE* is not expressed in *S. typhle*. *CD4*, absent in *Syngnathus*, shows significant positive selection overall, and aBSREL suggests that selection occurs in the branch leading to two seahorse species, *H. kuda*, and *H. comes*. Also absent in *Syngnathus*, the MHC II activator *CIITA* is positively selected in *Hippocampus*, as shown by BUSTED with the three *Hippocampus* species as the foreground branch, and aBSREL. This indicates that upon the loss of exons in *CD74* and the potentially resulting non-functional MHC II response in *Hippocampus* and *Syngnathus*, *CIITA* and *CD4* may have acquired a novel function in the seahorses. The products of the recombination-activating genes *RAG1* and *RAG2* facilitate V(D)J recombination in vertebrates. BUSTED detected selection in *RAG1*, but not *RAG2*. This is localized by aBSREL to *H. kuda* and *H. whitey*, and also *N. ophidion* and *E. aequoreus*, but not *Syngnathus*. *CD74* is lacking functional domains in *Syngnathus* and *Hippocampus*. When testing the whole gene tree with BUSTED, the LRT test shows significant selection, but not so when only the syngnathids were used as the foreground branch.

The gene encoding β2-Microglobulin, *B2M*, exists in two or three paralogs in the genomes analysed, named *B2M-1* and *B2M-2* in SI Appendix Table S11. It is a component of MHC I, and the BUSTED test finds evidence for selection only in *B2M*-1, however, the aBSREL test does not. Conversely, the BUSTED p-value is not significant for *B2M*-2, but the aBSREL methods shows selection in the syngnathids (Node 12, SI Appendix Table S11). For the *CD8* gene, *Syngnathus* and *Hippocampus* have lost the *beta* chain, however, the *alpha* chain

shows traces of positive selection in the syngnathids according to the BUSTED test. *TAP1* and *TAP2* function as a heterodimer in transport of antigens. The BUSTED test indicates positive selection in both genes, also with the syngnathids as foreground. The aBSREL test, however, identified selection *in H. comes* for *TAP1* only, and in A. *strigosus* for *TAP2*, this test has less power than BUSTED.

Generally, we find strong signs of positive selection in the candidate genes involved in immunity that were selected for analysis in contrast to randomly selected genes (SI Appendix Table S11). This suggests that the immune system is coevolving with male pregnancy in the Syngnathiformes and supports our results of major immune system rearrangement in the syngnathids.

6.3. Pregnancy genes and signs of positive selection

To identify genes with a function in male pregnancy, candidate genes known to be involved in female pregnancy that are also differentially expressed during pipefish pregnancy, were analysed for signs of positive selection across the Syngnathiformes but also specifically within syngnathids.

Apolipoprotein A-I, produced by the *APOA1* gene, is a cholesterol transporter involved in hormone production in the corpus luteum. BUSTED indicates positive selection in the species we tested, and also specifically in the syngnathids. The aBSREL test, however, only detected selection in the branches leading to species without male pregnancy.

The *ATF3* gene produces a transcription factor that is involved in corpus luteum development and regression. While the BUSTED test is significant for positive selection overall, the results for BUSTED with the syngnathids as foreground shows that the selection impacted species other than the syngnathids.

Carbonic anhydrase 4 (*CA4*) is an enzyme involved in a large number of metabolic processes, among them putatively the paternal-embryonal transport of CO_2 and ammonia. We find no positive selection in the species tested using the BUSTED method.

CEBPB produces a transcription factor important in regulating immune genes, especially in early embryogenesis. We find no indication of positive selection on this gene in our species set.

The transcription factor coding gene *JUNB* is involved in several pathways, including the estrogen pathway. It is present in two copies in most of the genomes analysed here, positive selection was detected in neither copy.

The protein kinase encoded by *PRKCD* is present in two copies and has been shown to be involved in the sperm acrosome and embryo implantation. The first *PRKCD* copy is detected as having undergone positive selection in the syngnathids by the BUSTED test, but there is no significant result for aBSREL.

PTGFRN codes for a protein involved in the prostaglandin pathway, and is positively selected in the syngnathids as shown with the BUSTED test, aBSREL identifies the *S. rostellatus* branch as having experienced positive selection, specifically.

The gene Rh associated glycoprotein *RHAG* gene, responsible for ammonium transport in mammals, does not seem to be under positive selection from our tests.

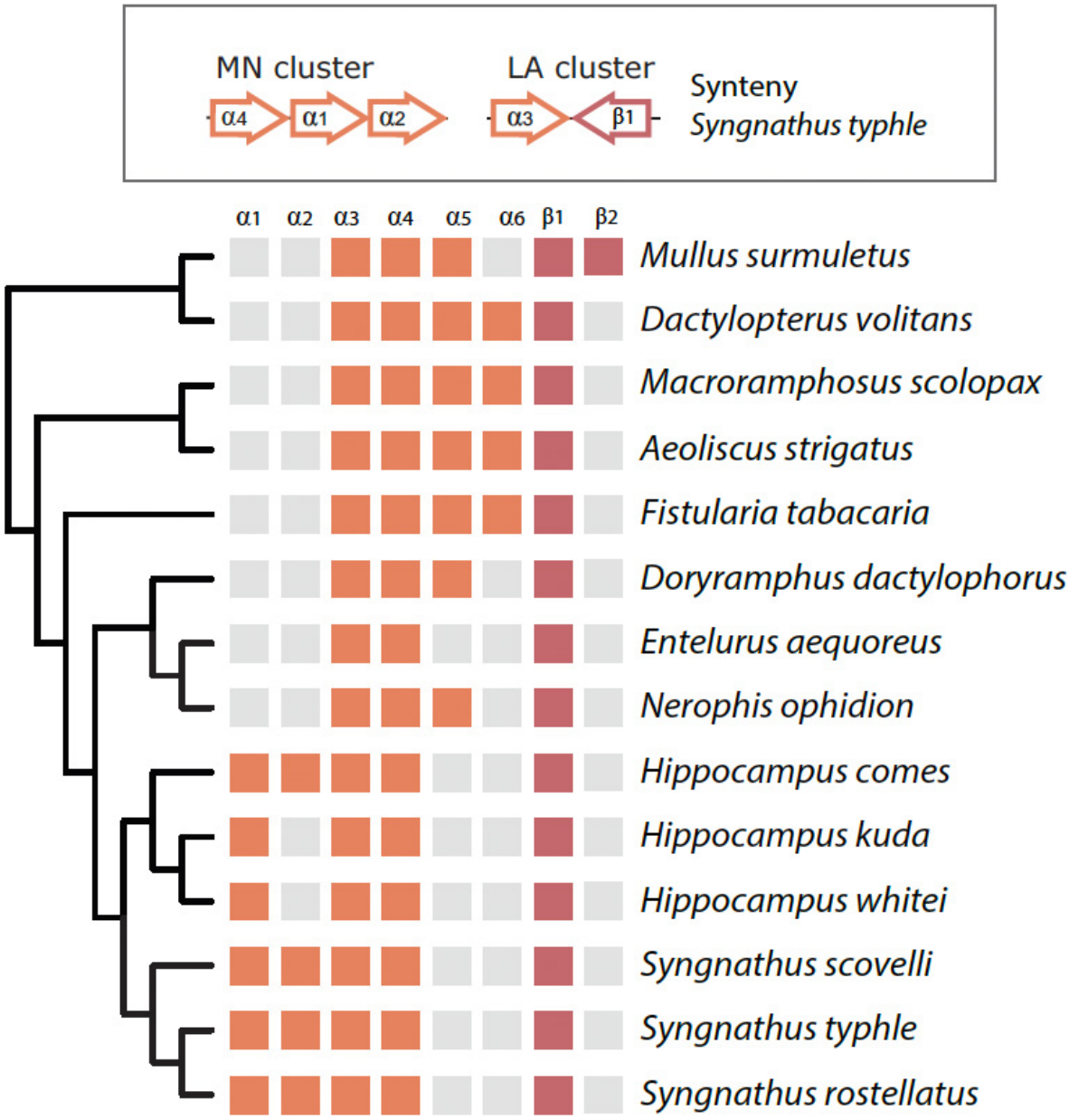
Only in *PRKCD* and *PTGFRN* seem to be under positive selection within the syngnathids. We could thus only find limited support that selection acts on these genes that are potentially involved in syngnathid male pregnancy.

6.4. Oxygen transport and Hemoglobin genes (*Hb*)

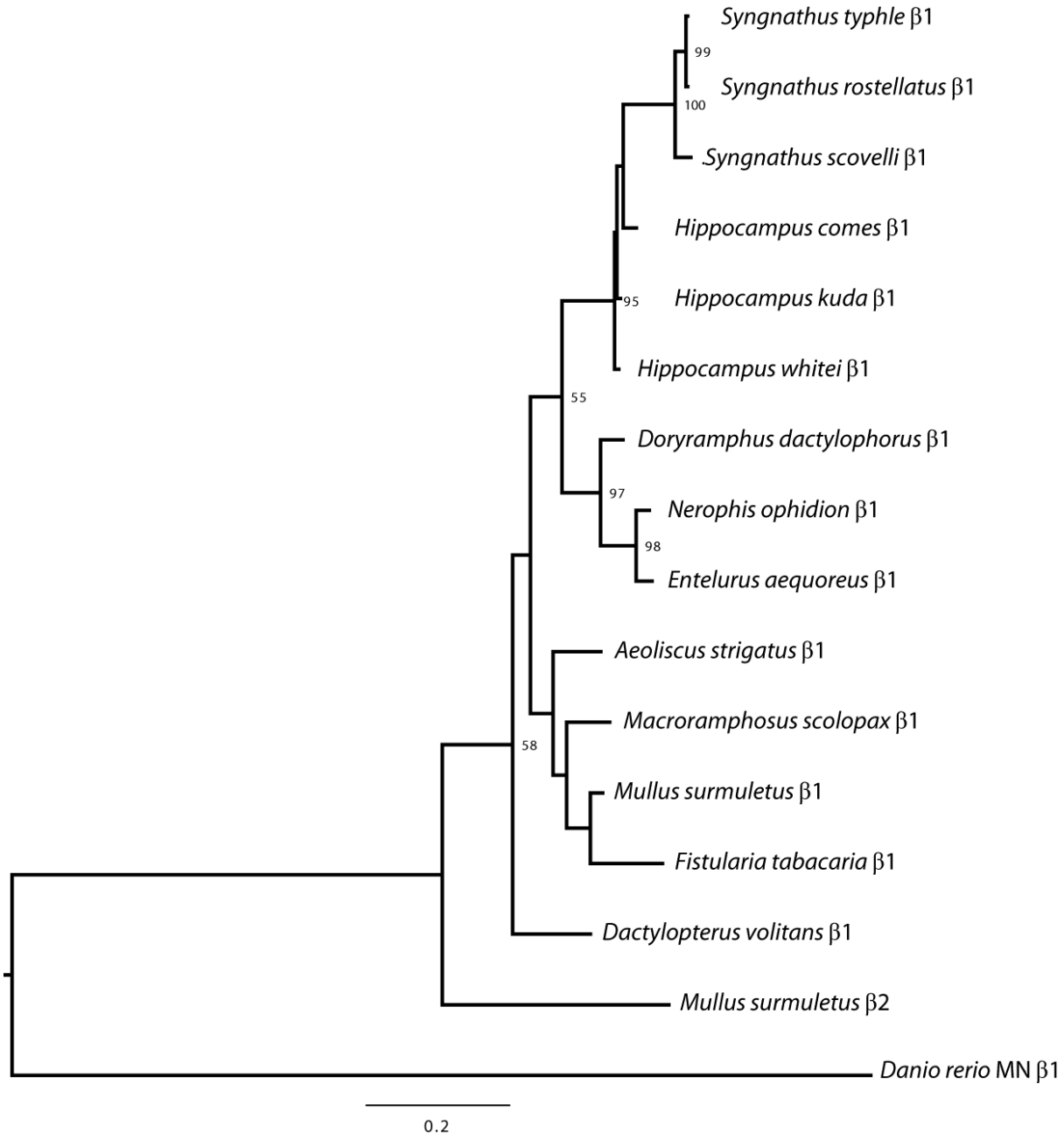
Oxygen transport from parent to the embryo is a key requirement for a successful pregnancy. To assess evolution of oxygen transport with male pregnancy in syngnathids, hemoglobin genes were analysed. While teleosts usually have a wide diversity of hemoglobins genes (*Hbs*) to cope with different environments and different life stages(61-63), Syngnathiformes feature only one *beta*-globin gene, except *M. surmuletus* having two copies (SI Appendix Figure S14-16). Syngnathiformes, such as most teleost species, possess a swimbladder filled with oxygen due to a highly specialized *Hb* isoform that can deliver oxygen against a concentration gradient, a phenomenon known as the Root effect(64, 65). Although the molecular mechanisms responsible for the Root effect are still not completely understood, it has been hypothesized that substitutions on one copy of the *beta*-globin is responsible, while the other *beta*-globin(s) have a normal oxygen affinity to allow efficient gas exchange in the remainder of the body(65). However, as most of the Syngnathiformes have only one *beta*-globin gene, this points to a different gene being responsible for the Root effect, possibly the *alpha*-globin gene. All of the Syngnathiformes species investigated had several copies of *alpha*-globin, with *alpha 3* and *alpha 4* being shared between all species (SI Appendix Figure S14 and S16). Further analyses into the function of these genes could shed light on the evolution of the Root effect in teleosts.

We also discovered a notable difference in the *Hb* repertoire of Syngnathiformes species with an inverted (*Syngnathus*) or sealed (*Hippocampus*) brood pouch being the only

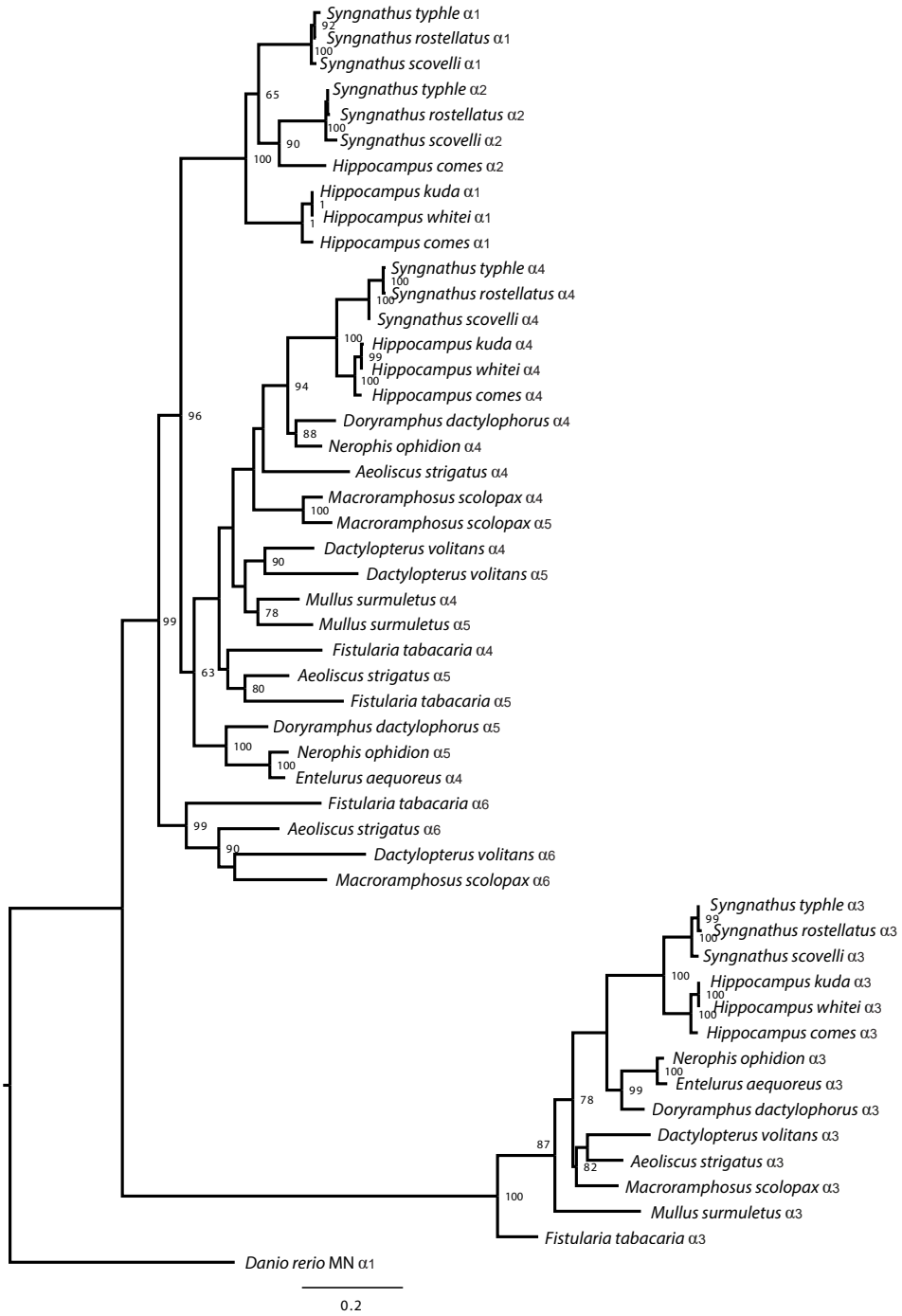
species to have *alpha 1* and *alpha 2*, with an apparent loss of *alpha 2* in *H. kuda* and *H. whitei* (SI Appendix Figure S14). Furthermore, *Hippocampus* and *Syngnathus* species do not have the *Hb* variants *alpha 5* and *alpha 6*, suggesting that species with full male pregnancy have a unique *Hb* repertoire. We can only speculate on the functional implications. In mammals, *Hb* isoforms with a higher oxygen affinity due to specific *beta*-globin genes are expressed at the embryonic and fetal life stages(66). Examining the gene expression results from various life stages in *S. typhle* shows that only *alpha 4* is expressed in embryos and larvae, whereas adults express several types including *alpha 4* and *alpha 2* at high levels and some *alpha 1* and *alpha 3.* At this stage, we lack sufficient experimental replication to further analyse expression differences at various life stages. However, it is important to note that in Syngnathiformes pregnancy adaptation might involve switching between *alpha*-globins. This is in contrast to mammals where a switch occurs between *beta*-globins as an adaptation to pregnancy.



SI Appendix Figure S14: Overview of hemoglobin repertoire in Syngnathiformes. Top panel shows synteny for the *Hb* gene clusters MN and LA *in S. typhle*. The lower panel shows *Hb* genes mapped on a phlyogenetic species tree, with *alpha*-globins in orange and *beta*-globins in red and grey indicating absence.



SI Appendix Figure S15: Phylogenetic tree of *beta-globin* genes constructed using ML with 1000 bootstrap replicates. Numbers at nodes show percentage bootstrap values >50.



SI Appendix Figure S16: Phylogenetic tree of *alpha-globin* genes constructed using ML with 1000 bootstrap replicates. Numbers at nodes show percentage bootstrap values >50.

7. Differential gene expression (transcriptome sequencing) in *S. typhle*

7.1. Sampling and experimental design, RNA extraction, Sequencing

In order to assess which genes are differentially expressed during male pregnancy and during the development of the brood pouch tissue, we sequenced the transcriptome (RNAseq) of the pipefish *Syngnathus typhle* in the following tissue types (i) brood pouch tissue of males with an undeveloped brood pouch, (27) a developed pouch, (iii) the pouch tissue at early pregnancy and (iv) the pouch tissue at late pregnancy. For this purpose, *S. typhle* males were kept in our laboratories in groups of 10 individuals. Fish with undeveloped brood pouches, developed pouches, those in early pregnancy stage and those in late pregnancy stage were dissected upon killing fish in a water bath with an overdose of MS22. From all fishes, pouch tissue was dissected and stored in RNA later upon further processing. All samples were taken at the same day, five replicates per developmental stage. These libraries were also used as reference transcriptomes for the *S. typhle* genome annotation.

Total RNA was extracted using the Machery-Nagel NucleoSpin 96 RNA kit following the manufacturer's protocol, and RNA quality was checked with an Experion Automated Electrophoresis system (Bio-Rad). Samples were prepared for sequencing using the Illumina TruSeq RNA sample preparation Kit, according to the standard protocol. Quality control and quantification of libraries was examined by using the LapChip GX (Caliper) and the HT DNA High Sensitivity Kit. For sequencing, indexed libraries were diluted to 2nmol/ L and pooled. To control quality of the sequencing run, a 1% PhiX control library (PhiX Control Kit v3, Illumina) was added to each lane. cDNA libraries were paired-end sequenced on the Illumina HiScan SQ platform for 2*101 bp. We sequenced 20 indexed samples (as part of a 96 sample setup) distributed over one flowcell. Raw image data were transformed and de-multiplexed using CASAVA 1.8 software.

7.2. Differential gene expression analysis Syngnathus typhle

For the expression analysis the “Tuxedo” suite of tools were used(67). Raw RNA reads were mapped against the genome assembly using TopHat2 (v2.0.12) (ref(68)), with an average mapping rate of 74% across samples. The mapping data was used by Cufflinks (v2.2.1) to calculate splice sites and generate transcripts for each sample separately. The transcripts were then merged into a unified transcriptome by Cuffmerge. Expression was

calculated using Cuffdiff (v2.2.1) with a FDR-corrected p-value of < 0.05 (ref(67)), plots and lists of differentially expressed genes were generated with the CummRbund R package (v2.24.0) (ref(69)).

Differential gene expression was calculated pairwise against the undeveloped pouch as reference tissue. All differentially expressed genes were searched for potential functions via homology, using reported functions in female pregnancy of mammals, in the squamate reptile *Chalcides ocellatus* (70) and in male pregnancy of *S. scovelli* (3) and *H. abdominalis* (71). These homologous genes are in the following displayed in ***bold italics***. Gene expression data were compared in more detail to RNAseq data collected during male pregnancy of the pipefish *Syngnathus scovelli* (3).

7.3. Convergent evolution of female and male pregnancy

A total of 73M reads were sequenced, with an average of 3.6M reads per sample. In total we identified 116 genes that were differentially expressed in male pregnancy in *S. typhle*, which also are reported to display a function in female pregnancy in reptiles and mammals (SI Appendix Figure S12).

In mammals, trophoblast interferon (*IFNT*) is key in initiating early pregnancy. *IFNT* disrupts the uterine release of prostaglandin that inhibits luteolysis and maintains a critical progesterone/estrogen level(72) required for the survival of the developing embryo. An inflammation response was suggested to be important for successful implantation in mammalian pregnancy (73). The key genes mediating this specific inflammation at implantation in mammals, ***IL6R (M/P)***, ***TNF (M)*** and ***PTGS2 (M)*** (73, 74), were upregulated during pouch development in pipefish. The implantation of the eggs further takes place via activation of the mTOR pathway. To initiate the maternal pregnancy, *IFNT* induces downstream signal transducers and further transcription activators such as *STAT1, STAT2, IRF1*(75, 76). The enhanced *IRF1* transcription, in particular, has substantial immune-modulating effects mediating embryo tolerance and preventing its rejection. *IRF1* induces the expression of *IL-4*, reduces the lymphocyte proliferation in response to *IL-2*(77) and changes *MHC I* expression on trophoblasts by silencing classical *MHC I* expression and enhancing the non-classical *MHC I* expression on trophoblasts over the JAK /STAT pathway via *IRF2* and *CASP3*. The *MHC II* expression is silenced on trophoblasts via the *PSMB* gene. In addition, mammalian pregnancy involves a shift from Th1 (proinflammatory, tissue damage: *IFN-gamma*) towards a Th2 immunological environment (IgE promotion, anti-inflammation and eosinophilic

responses: *IL-4, IL-5, IL-13, IL-10*). This shift was shown to reduce inflammation (cytotoxic T-cells and proinflammatory interleukins (*IL-1a, IL-2, IL-6*)) in favour of B-cell activation (78).

7.3.1. Genes of the immune system

During the development of the brood pouch and during male pregnancy several genes of *Syngnathus typhle* were differentially expressed that modulate the maternal immune system during mammalian pregnancy. The critical shift from Th1 towards Th2 immune response during mammalian pregnancy, i.e. the suppression of inflammatory responses, was also identified in pipefish male pregnancy. ***Epx*** with its dual function in mammalian placenta development and in eosinophil activation was upregulated during pouch development in *S. typhle* and remained so during pregnancy(79). The same pattern was apparent for ***TNF***, involved in cytokine production (Th1) but also in uterine receptivity and placental development(80). Moreover, ***IL-6*** was upregulated during pouch development but not during pregnancy, in mammals the ***IL-6*** downregulation during pregnancy is essential for placental functioning and the immunological tolerance of the embryo(81). ***CLCF1***, ***KLF4*** and ***CEBPB*** that are all involved in inflammation response were upregulated during pipefish pregnancy(82). ***CEBPB*** that represses *MHC* expression but facilitates Th2 immune responses was consistently upregulated in both pipefish species and the seahorse(71), which is in line with the shift towards Th2 type immune response. In mammals ***KLF4*** is key for the maintenance of gestation(83), and ***CEBPB*** for the regulation of decidual gene expression(84). Consistent with mammalian pregnancy, the proinflammatory interleukins ***Il1*** and ***IL2*** were downregulated during pouch development, and the latter gene also during pregnancy. In mammals ***IL1*** induces ***CLCF1*** expression that mediates the onset of labours at term, a mechanism that resembles an inflammation(82). Also ***RPL18A***, responsible for T-cell proliferation (Th1) and known to be downregulated in lizard pregnancy(70) is downregulated during pipefish pregnancy. Three genes involved in Th2 immune response are downregulated in pipefish tissues, ***MEF2C*** (maturation of lymphocytes and important in mammalian embryo development by transcribing Bisphenol A)(85) is downregulated during pouch development, in the seahorse ***MEF2C*** was upregulated after pregnancy during parturition(71). Expression of Interleukin secretion genes ***S100A13*** and ***IL20RB*** is lowered during pipefish pouch development (***S100A13***) and pregnancy (***S100A13***, ***IL20RB***), which is consistent with seahorse pregnancy(71) but in contrast to reptile pregnancy(70). The transcriptional repressor ***PRDM1*** (lymphocyte maturation and proliferation (Th2)) defines in mammals

a lineage-restricted progenitor cell population contributing to placental growth and morphogenesis(86) and is upregulated in pipefish brood pouch development, while ***MEF2C*** that enhanced lymphocyte maturation (85) was downregulated. Gelsolin (***GSN***) that is present in midtrimester amniotic fluid, that binds to LPS and inhibits the induction of ***TNF*** (87), was upregulated during pipefish pouch development.

Similar to mammals, during pipefish pregnancy ***CASP3*** was upregulated, while we find a downregulation of ***STAT1*** and ***IRF1***. In mammals, binding of trophoblast interferon (*IFNT*) activates ***STAT*1** that via ***IRF1***, *IRF2* and ***CASP3*** modifies the MHC class I pathway(88, 89) to maintain immunological tolerance**.** ***CASP3*** is a progesterone responsible for the uterine quiescence necessary during pregnancy. At the onset of labors, ***CASP3*** levels decrease in mammals(90). In pipefish a series of putative homologous genes to mammals involved in antigen recognition, presentation and processing are differentially expressed during pregnancy such as the *Class I histocompatibility antigen* ***F10***, ***TAP1***, ***H2-K1*** *(two orthologues)*, ***IRF1***, ***IRF8***, ***MR1*** (4 orthologues). In mammals during early pregnancy, ***TAP1*** is increasingly expressed on placenta-specific trophoblasts, and plays an important role in the prevention of the embryos from maternal immune attacks(91). As a barrier between the mother and the embryo trophoblasts build a layer of tissue in the placenta. They only express non-classical *MHC I* that does not induce a non-self-reaction against the embryo. Most other genes involved in adaptive immune defence (*Class I histocompatibility antigen* ***F10***, ***H2-K1*** *(two orthologues)*, ***IRF1***, ***IRF8***, ***MR1*** *(4 orthologues*), ***FUT 9***) are silenced during mammalian pregnancy to protect the embryonal allograft(92, 93). Also in pipefish the expression of those genes is downregulated during early pregnancy, which also holds true for other genes with a function in antigen processing (***Ig kappa chain V***, ***Ig mu chain C***, ***IGLC1***, ***FUT9***). All those genes are involved in the MHC I pathway. The only gene with an original function in antigen processing that was upregulated during pipefish pregnancy is the MHC II invariant chain ***CD74***. However, due to the ambiguous sequence of CLIP, ***CD74*** cannot function as invariant chain of MHC II**.** In humans, the silencing of ***CD74*** during pregnancy is key for maintaining the acceptance of the semi-allogenic embryo(94).

7.3.2. Expression patterns of prostaglandin, progesterone, and the estrogen pathway

A shift in the expression of genes involved in the biosynthesis of prostaglandin (PTG) is well established in mammalian menstrual cycle and pregnancy. The activity of these genes prevents the degradation of the corpus luteum at the onset of pregnancy. These genes were also differentially regulated in the pipefish pouch development and pregnancy. ***PTGFRN***, ***PTGIS*** and ***PTGS2*** that are all involved in the biosynthesis of prostaglandin with a role in mammalian pregnancy are upregulated during pouch development. In mammals ***PTGFRN*** is upregulated during early pregnancy(95), similar to ***PTGIS*** that converts prostaglandin to a functional form during pregnancy(96), in contrast to ***PTGS2*** that is increased at luteolysis, but suppressed during pregnancy(97). ***PTGFR*** is downregulated during pipefish pouch development, but upregulated during early pregnancy. Also ***PTGS1*** is upregulated during early pipefish pregnancy, while ***PTGDR2*** and ***PTGER4*** are downregulated during pipefish late pregnancy. In mammals ***PTGFR*** is high during luteolysis but suppressed later in pregnancy(97), while ***PTGS1*** is also upregulated during mammalian pregnancy(97).

APOA1 is a key gene for the corpus luteum development. ***APOA1*** expression decreases over the corpus luteum regression in menstrual cycle but is highly expressed in the early corpus luteum at the time of embryo implantation and post-implantation(98). In contrast, ***ATF3*** is only involved in the regression of the corpus luteum(99). In line with these expectations from mammals, in all three syngnathids examined (*Syngnathus typhle, Syngnathus scovelli* and *Hippocampus abdominalis*), ***APOA1*** and ***ATF3*** (the latter only in *S. typhle* and *S. scovelli*) expression was upregulated during early and late pregnancy(3, 71).

An upward shift in progesterone level is critical for the survival of the developing embryo. Two key transcription factors of the estrogen pathway (***JUN*** and ***JUNB***)(100) were upregulated during *S. typhle* pregnancy, the latter one (***JUNB***) was also upregulated during *S. scovelli* pregnancy(3).Various genes regulated by progesterone are suggested to be involved in mammalian pregnancy. ***DKK2*** as an antagonist of WNT signaling, mediating reproductive events, is downregulated during pouch development and early pregnancy(97). ***NFATC4*** that can block pregnancy-induced cardiac hypertrophy in humans(101) was downregulated both during pouch development and early pregnancy in pipefish. ***HAVCR1*** expression can induce preeclampsia and increase inflammation(102)

in mammals, in pipefish it was downregulated during pouch development and pregnancy. Triggered by the shift of mammalian estrogen levels, ***RCN3*** expression is induced in the endometrium in mammals(103), consistent with a constant upregulation during pouch development and pregnancy in pipefish. The action of the insulin-like growth factors (IGF) important for tissue homeostastis, regulation of cell proliferation, differentiation and survival during embryonic development(104) are mediated by six secreted high-affinity binding proteins (*IGFBP*s) that, however, also have an independent function(105). Two *IGFBP*s were upregulated during pouch development (***IGFBP6***) or pregnancy (***IGFBP1***). ***HSD11B1*** and ***HSD11VB2*** convert cortisol into cortisone and are inhibited by progesterone(106). ***HSD11B1*** was upregulated during pouch development, while ***HSD11B2*** was downregulated during late pregnancy, the opposite pattern was observed in seahorse pregnancy(71).

7.3.3. Parent – fetus transport

For the development of the embryo, the transport of nutrients and oxygen from the pregnant parent to offspring is essential. The two transport genes ***SLC2A1*** and ***SLC4A1*** are susceptible to a shift in progesterone level. In mammalian pregnancy they have a dual function, mediating the maternal-fetal transport (***SLC2A1***: glucose(107); ***SLC4A1***: CO_2 transport and bicarbonate balance(108)), as well as activating targets of the mTOR pathway that stimulates proliferation, migration, gene expression and mRNA translation by the conceptus(109, 110). While ***SLC2A1*** was upregulated during pouch development, ***SLC4A1*** was downregulated during pouch development and pregnancy.

Two putative homologs to mammalian transport genes were upregulated during pouch development and early pipefish pregnancy: ***HK2***, involved in glycolytic and glycerol lipid biosynthesis(111) and ***AQP***, responsible for water homeostasis in reproductive cells, an impairment of its function can result in attenuated male and female fertility(112). Other genes with a putative function in transport were already previously described to being differentially expressed during seahorse pregnancy: ***GLT8D*** that could be involved in membrane transport, ***MAT2A*** as well as ***CALR,*** suggested to play a role in paternal-embryonal transport, were significantly upregulated during pipefish pouch development and both pipefish and seahorse pregnancy, similar to ***CA4*** that is involved in CO_2 and ammonia transport(71). Key candidates representing convergent pregnancy evolution are the fatty acid binding proteins encoded by *FABP* genes. In mammals they are supposed

to be involved in fatty acid delivery to the fetus, and potentially play a role in fetal protection from hypoxia and are under control of the hypoxia inducible factor (HIF)(113). In pregnant reptiles and seahorses, they are differentially regulated as well. In *S. typhle* ***FABP3*** is upregulated during pouch development and early pregnancy, ***FABP1*** is upregulated during pregnancy, while ***FABP6*** is downregulated during late pregnancy.
During pipefish pregnancy, ***TF***, a gene that is known to be involved in transport to the fetus in mammals(114) and ***RHAG*** involved in ammonium transport in mammals were upregulated. ***CLDN5*** belonging to human tight junctions factors is downregulated during pipefish late pregnancy(115).

7.3.4. mTor pathway and implantation

In mammals, the mTOR pathway regulates the expression of genes responsible for implantation. Three genes homologous to the mammalian counterparts of the mTOR pathway were downregulated either during pouch development (***DDIT4***) or during pipefish pregnancy (***TSC1***). In mammals, ***DDIT4*** downregulation activates the mTOR signaling pathway, its downregulation during pregnancy enhances fetal growth(116). ***TSC1*** was suggested to mediate conceptus growth over the mTOR pathway(117). Already during pouch development, we observed a downregulation of ***LGALS3***. As a member of the galectins, ***LGALS3*** was in mammals suggested to play a role in uterine receptivity and implantation by modulating the maternal immune response and to be receptive to progesterone(118). All other genes with a known function in implantation of mammals were upregulated in pouch development, such as ***PRKCD*** that is in male sperm important to penetrate the zona pellucida, in females it supports the embryo implantation(119). Also ***DAG1*** was upregulated during pouch development. ***DAG1*** is a target for evolutionary host-pathogen interactions at the maternal-fetal interface. As such, it is important for implantation but also serves as a receptor for viral recognition that can cause placental infections(120). Further, ***RAC1*** & ***RHOA*** (the invasion of the maternal decidua by trophoblasts requires those two genes for the establishment of a normal placenta), ***COL5A1*** & ***COL6A3*** (involved in the implantation of the egg and the muscle development of the embryo(121, 122)) and ***IST1*** (affects cell cycle and differentiation during pre-implantation phase(123)) were upregulated during pouch development. During pipefish early pregnancy, however, ***PLAU*** (involved in signal transduction(124)) and ***RHOA*** expression decrease, while ***COL5A1*** and ***COL6A3*** remain upregulated, as well as ***COL12A1*** (involved in progesterone pathway that maintains the myometrical quiescence and

blocks myometrical contractility(125)). During late pregnancy, ***PRKCD*** and ***DAG1*** were again upregulated, as well as ***RAC1*** and ***COL5A1***.

7.3.5. Endometrium and placenta development

Many of the hypothetically homologous genes with a role in placenta and endometrium development in mammals were differentially expressed in the development of the pouch and during male pregnancy of *S. typhle*. Several placenta-specific genes were upregulated during pouch development, such as ***AHNAK***(126), ***FURIN*** (important for the invasion and migration of trophoblast cells during early mammalian pregnancy(127)), ***FN1*** (involved in proliferation of trophoblastic cells and the tissue organization of the placenta(128)) and ***TNFRSF21*** (role in placental immune defence(129)). The latter remains upregulated during early pregnancy in *S. typhle*, as well as ***DLX3B***, a regulator of mammalian placental growth and ***EPAS1*** (involved in placental blood flow), which is also differentially expressed during reptile pregnancy and seahorse male pregnancy(70, 71, 130). In late pipefish pregnancy, in addition to ***DLXB*** and ***TNFRSF21***, also ***PRDM1*** (a transcriptional repressor that defines a lineage-restricted progenitor cell population contributing to placental growth and morphogenesis(86)), and ***FRRS1*** (involved in potential placental abruption(131)) were upregulated, while ***BNIP3*** was downregulated.

7.4. Gene co-option for male pregnancy in *Syngnathus typhle*

We compared differential gene expression during pouch development and male pregnancy (early and late pregnancy) from RNA isolated in brood pouch tissue against tissue of undeveloped pouches in order to study which genes were co-opted for male pregnancy. We only considered genes with a clear expression change (2fold) in at least one of the possible 3 comparisons, in total 220 genes (SI Appendix Table S13). Genes in *italics* have not been shown to have a function also in female pregnancy. Genes in ***bold italics*** have known functions in female pregnancy and have already been discussed in the section 7.3.

All differentially expressed genes were assigned to categories. Immune system (29), developmental processes (59), apoptosis (7), cell differentiation (14), gene expression (8), membrane (132), metabolic process (12), protein modification (132), stress regulation (16), signaling (5), transport (7), sensory reaction (27), or another/ unknown function (24).

7.4.1. Genes of the immune system

We identified 14 innate immune system genes that were differentially expressed during pouch development or male pregnancy. During pouch development, we find a downregulation of *ADSSL1C* (involved in antimicrobial peptide synthesis) and the antimicrobial peptide *PLE3*, the pro-inflammatory cytokine *MIF*, *FHL2* (inflammation response), *S100A13* (responsible for stress-induced export of IL1a), *JUND* (involved in LPS response) and *Gal lectin* that mobilizes neutrophils. *PlA2G4A* and *IL17REL* (both inflammation response), ***EPX*** that is involved in bacterial fragmentation and lysing over eosinophil activity, *TNF* & *GREM2* (both cytokines), and *TNFAIP8L1* are upregulated during pouch development. *PLE3* and *S100A13* remained downregulated throughout pregnancy while *FHL2* shifted towards downpregulation during pregnancy. *C6* (complement component C6) and *JUND* were upregulated during early and ***EPX*** and ***TNF*** during late pregnancy.

15 genes with putative function in the adaptive immune system were differentially expressed during pouch development or pregnancy. Three genes involved in antigen binding were downregulated during pregnancy (*IGLC1*, *IGkappa*, *IgMu*), the latter two also during pouch development. In contrast, ***FUT9*** and ***CD74*** both also involved in antigen processing are upregulated during pouch development. The invariable chain of the MHC II ***CD74*** remained consistently upregulated during pregnancy. However, as stated above, ***CD74*** cannot be functional due to losses of exons and very divergent sequences. During pouch development, one other gene of the adaptive immune system, *CHIA* was downregulated, lowering T-cell proliferation. Three genes involved in lymphocyte response were upregulated during pouch development: ***PRDM1*** (B-cell maturation), ***CLCF1*** (B-cell stimulation) and *GIMAP4* (lymphocyte apoptosis), as well as ***TNFRSF21*** with a function in T-cell proliferation. One gene enhancing B-cell development, ***FCRL5*** was lowered during early pregnancy, while ***CLCF1*** remained upregulated over pregnancy. During late pregnancy also *GPR97* and *MFNG* (both responsible for B-cell differentiation) were downregulated as well as ***NFATC4*** and ***HAVCR1*** (T-cell development). This suggests that both antigen processing and lymphocyte responses were affected by male pregnancy.

7.4.2. Genes involved in developmental processes

The majority of differentially expressed genes during pregnancy are involved in developmental processes (89 genes). Among those, genes can be divided into sub-categories, corresponding to putative functions in fertilization and pregnancy (18 genes), digestive system (2 genes), bone development (8 genes), brain development (5 genes), embryo/ organism development (20), muscle development (5 genes), cardiovascular system (5 genes) and muscle contraction (21 genes).

Among the genes involved in pregnancy and fertilization ***NOXO1*** (angiotensin induced enlargement of the uterus), ***ADAMTS1*** (female fertilization, ureteropelvic junction), *RAPGEF3* (sperm development), ***ADM2*** (trophoblast invasion and migration), *MKK* (male gonad development), ***PRELP*** (gonad development), ***SRC*** (regulation of increased heart rate during pregnancy), ***PTGS2*** (in mammals prevents corpus luteum degradation over prostaglandin pathway), ***HSD11B1L*** (uterine receptivity), ***CUZD1*** and ***IGFBP6*** (both important for initiation of labors) were upregulated during pouch development compared to undeveloped pouch tissue. ***CUZD1*** and ***ADAMTS1*** remained upregulated also during pregnancy. Induced gene expression was also observed in ***KLF9*** and ***FAM64A***, two genes important for endometrium development, and ***APOA1***, which prevents the corpus luteum from degradation at the onset of a pregnancy.
Only ***DKK2*** (anterior-posterior axis development, an antagonist of WNT signalling, mediating reproductive events) and two copies of ***Tropomyosin alpha-1 chain*** (important for in utero development) were downregulated during pouch development. ***DKK2*** remained downregulated also during pregnancy, as well as ***CLDN5*** (responsible for tight junction). ***PLAC8*** (mammalian placenta development), ***PTGFR*** (in mammals prevents corpus luteum degradation over prostaglandin pathway) and ***Klf9*** mediating inflammation during pregnancy.

Among the 8 genes playing a role in bone formation, 4 were downregulated during pregnancy (*PDLIM7*, *PDLIM3*, *SMAD9*, *KLHL41*), and 2 were upregulated (*COL11A1*, *COL2A1* (all collagen alpha chain)). During pregnancy *SOST*, *COL11A1* and *COL10A1* were upregulated. This most likely suggests an important function of those genes during pregnancy rather than during pouch development.

Genes mediating heart development were consistently downregulated either during pouch development (*HEY1*, *SMYD1*, *SMYD2B)* or during pregnancy (*HEY1*, *ANKRD1*) suggesting an important function for the development of the cardiovascular system during pouch development and pregnancy. The only gene involved in digestive system development was in one copy highly downregulated during pouch development but then in the other copy upregulated during pregnancy (*BHMT*).

Among genes with a putative function in embryonal development, *HCEA* (egg hatching), *GPX4* (embryogenesis), *ARG8*, *CKM* (tissue energy transduction), *EPGN* (EPGF response) were downregulated during pouch development. In contrast, *ALDH1A2*, *TGFA*, *STX2*, *TMEM79*, *MSCX*, *ENAH*, *SEMA3G*, *SOX7* and *ZNF513* were upregulated in the same treatment. During early pregnancy *ALDH1A2*, *CRABP2* and *BGLAP* were upregulated and not a single embryogenesis gene was downregulated, while during late pregnancy the only downregulated embryogenesis gene was *TTC36* (otolith morphogenesis). During late pregnancy *ALDH1A2*, *MSXS* (ear morphogenesis), *NMRK2* (myoblast differentiation), *MEF2D* (transcription factor). *MYL2* and *BGLAP* were upregulated.

REG1B, *KCNN2* and ***CDKN3*** were the three genes involved in brain development that had a higher expression during pouch development. In contrast, *DCHS1* (neuronal cell differentiation) was downregulated. During pregnancy *SERPINI1* (nervous system organization) was highly upregulated and *PLDX1* (spinal cord development) was downregulated. The two genes with a function in keratinization were upregulated either during pouch development (*CNFN-B*) or during pregnancy (*ALOX8*).

All but one gene (*SMTNL2*) involved in muscle development showed a decreased expression during pouch development (*JPH2*, *TNNI2*, *ENO3*, *MURC*), while during pregnancy no change in expression could be observed.
Genes mediating muscle contraction were highly lowered in their expression during pouch development (*MYLPF*, *ACTC1*, *TTN*, *MYOM2, TRDN*, *MYOSIN H*, *MYOM2*, *ACTA1*, *TNNC, ACTC, CASQ1, MYLPF, MYOLCHS, MYOLCHC*). During early pregnancy *CASQ2, MYLPF* and *ACTC* were still downregulated, while during late pregnancy *MYOLCHC was* upregulated. This suggests a decreased muscle activity during pouch development.

SI Appendix Table S12: Genes with a function in mammalian (M), reptile (R) or seahorse (S) pregnancy that were differentially expressed during pipefish brood pouch development or pipefish pregnancy. Comparison was: developed pouch against undeveloped brood pouch tissue, early pregnant pouch tissue against undeveloped brood pouch tissue and late pregnant pouch tissue against undeveloped pouch tissue. Positive values indicate upregulation in undeveloped pouch tissues.

Gene name	full gene name	Function category	homologous	id	devel_vs_undev	e-preg_vs_undev	l-preg_vs_undev	ref
CD74	H-2 class histocompatibility antigen gamma chain	immune	M/ R	maker-scaffold00049-augustus-gene-8.11-	-2,0334	-2,7864	-3,2362	(94)
CEBPB	CCAAT/ enhancer-binding protein beta	immune	M/ S	snap_masked-scaffold00001-processed-gene-82.40	0,0000	-1,1242	-0,9878	(84)
F3	Tissue Factor	immune	M/ S	maker-scaffold00002-augustus-gene-85.35	0,0000	0,0000	1,7166	(133)
GSN	Gelsolin	immune	M/ R	genemark-scaffold00022-processed-gene-1.11-	-0,9516	0,0000	0,0000	(87)
H2-K1	H-2 class I histocompatibility antigen, K-Q alpha chain	immune	M	genemark-scaffold00040-processed-gene-18.19-	0,0000	0,9503	0,0000	(92)
H2-K1	H-2 class I histocompatibility antigen, K-Q alpha chain	immune	M	genemark-scaffold00303-processed-gene-0.11-	0,0000	0,8641	0,0000	(92)
I hist anti	Class I histocompatibillity antigen	immune	M	maker-scaffold00303-augustus-gene-0.46-	0,0000	1,0500	0,0000	(134)
IL1B	Interleukin-1 beta	immune	M	maker-scaffold00023-augustus-gene-34.52-	1,1526	0,0000	0,0000	(75)
IL20RB	Interleukin-20 receptor subunit beta	immune	R/ S	maker-scaffold00002-snap-gene-44.66	0,0000	0,8893	0,0000	(71)
IL2RG	Cytokine receptor common subunit gamma	immune	M	maker-scaffold00008-augustus-gene-47.57	1,4007	1,6468	1,3735	(75)
IL6R	Interleukin-6 receptor subunit alpha	immune	M	maker-scaffold00081-augustus-gene-3.57	-0,9301	0,0000	0,0000	(75)
Klf4	Krueppel-like factor 4	immune / embryonal development	M/ S	snap_masked-scaffold00048-processed-gene-11.20	0,0000	-2,0034	-2,4307	(83)
MEF2C	Myocyte-specific enhancer factor 2C	**immune**	M/ S	maker-scaffold00050-augustus-gene-14.90	1,3723	0,0000	0,0000	(85)
MEF2D	Myocyte-specific enhancer factor 2D	immune	M	maker-scaffold00058-snap-gene-5.52-	-1,7967	0,0000	-2,1529	(135)
MR1	Major histocompatibility complex class I-related gene protein	immune	M	maker-scaffold00024-augustus-gene-22.54-	0,9626	1,4606	1,9281	(93)
MR1	Major histocompatibility complex class I-related gene protein	immune	M	maker-scaffold00040-augustus-gene-18.60-	-1,1531	0,0000	0,0000	(93)
MR1	Major histocompatibility complex class I-related gene protein	immune	M	maker-scaffold00303-augustus-gene-0.47-	0,0000	0,8694	0,0000	(93)
MR1	Major histocompatibility complex class I-related gene protein	immune	M	maker-scaffold00303-augustus-gene-0.48-	0,0000	1,2369	0,0000	(93)
NFATC4	Nucllear factor of activate T cells, cytoplasmic 4	immune	M	maker-scaffold00002-augustus-gene-71.51-	1,0427	1,4050	2,2371	(101)
RPL18A	60S ribosomal protein L18a	immune	R	maker-scaffold00064-augustus-gene-12.94-	0,0000	0,0000	0,8577	(70)
TAP1	Antigen peptie transporter 1	immune	M	maker-scaffold00058-snap-gene-12.59-	0,0000	0,9055	0,0000	
TNFRSF21	Tumor necrosis factor receptor superfamily member 21	immune / placenta	M	maker-scaffold00058-augustus-gene-13.61-	-2,8935	0,0000	-3,0368	(136)
STAT1	Signal transducer and activator of transcription 1-alpha/beta	immune / in utero development	M	genemark-scaffold00127-processed-gene-3.25-	0,0000	0,9079	0,0000	(88)
HAVCR1	Hepatitis A virus cellular receptor 1 homolog	immune / progesterone	M	genemark-scaffold00008-processed-gene-22.3-	1,2544	1,9961	2,3375	(102)

IRF1	Interferon regulatory factor 1	immune/ estrogen	M	maker-scaffold00008-augustus-gene-12.49-	0,0000	0,9705	0,0000	(88)
IRF8	Interferon regulatory factor 8	immune/ estrogen	M	maker-scaffold00005-augustus-gene-14.34-	0,0000	1,0972	1,5468	(56)
FUT9	Alpha-(1,3)-fucosyltransferase 9	immune/ implantation	M	maker-scaffold00010-snap-gene-27.42-	-3,0834	0,0000	0,0000	(137)
STAT1	Signal transducer and activator of transcription 1-alpha/beta	immune/ in utero development	M/ M	maker-scaffold00067-augustus-gene-4.51-	0,0000	1,0695	1,3682	(88)
CLCF1	Cardiotrophin-like cytokine factor 1	immune/ labor	M	maker-scaffold00013-snap-gene-49.40-	-2,8488	-3,2282	-3,3566	(82)
EPX	Eosinophil peroxidase	immune/ placenta	M	maker-scaffold00022-augustus-gene-39.76-	-4,0782	-1,5275	-2,1218	(Knudsen, 1997
BTG1	B-cell translocation gene 1 protein	conceptus growth	M	maker-scaffold00112-augustus-gene-2.23-	-0,9321	0,0000	0,0000	(138)
DUSP5	Dual specificity protein phosphatase 5	conceptus growth	S	maker-scaffold00033-augustus-gene-1.29	0,0000	-1,2488	0,0000	(139)
MAPK3	Mitogen-activated protein kinase 3	conceptus growth	M	maker-scaffold00004-snap-gene-59.75-	-1,3088	0,0000	-1,1125	(140)
ODC1	Ornithine decarboxylase	conceptus growth	M	genemark-scaffold00010-processed-gene-6.8-	0,0000	0,0000	-0,8851	(141)
TSC1	Hamartin	conceptus growth	M	genemark-scaffold00027-processed-gene-1.45-	1,1008	0,8962	1,3468	(117)
TSPAN1	Tetraspanin-1	conceptus growth	M	maker-scaffold00045-augustus-gene-11.30-	0,0000	0,0000	-4,9276	(142)
CDKN3	Cyclin-dependent kinase inhibitor 3	embryonal development	M	augustus_masked-scaffold00010-processed-gene-17.38-	-2,1323	0,0000	0,0000	(143)
ENPP1	Ectonucleotide pyrophosphatase	embryonal development	M/ S	maker-scaffold00060-snap-gene-13.81	0,0000	0,0000	0,8806	(144)
FAM20C	Family with sequence similarity 20	embryonal development	M/ S	maker-scaffold00004-snap-gene-8.38	0,0000	0,0000	0,9938	(145)
FKBP10	Peptidyl-prolyl cis-trans isomerase FKBP1A	embryonal development	M	maker-scaffold00051-augustus-gene-10.63-	-0,8530	0,0000	0,0000	(146)
FAM64A	Family with sequence similarity 64	endometrium	M	maker-scaffold00008-snap-gene-21.86-	0,0000	-2,4650	0,0000	(147)
KLF9	Krueppel-like factor 9	endometrium	M	maker-scaffold00027-augustus-gene-13.35-	0,0000	2,3370	1,6691	(148)
RCN3	Reticulocalbin-3	endometrium / progesterone	M	augustus_masked-scaffold00047-processed-gene-6.11-	-1,9337	-2,0191	-2,1853	(103)
DKK2	Dickkopf-related protein 2	endometrium/ progesterone	M	maker-scaffold00034-augustus-gene-5.34-	3,2396	1,2866	2,8916	(97)
JUN	Transcription factor AP-1	estrogen	M	augustus_masked-scaffold00011-processed-gene-21.36-	0,0000	-1,3252	0,0000	(100)
JUNB	Transcription factor jun-B	estrogen	M/ S	augustus_masked-scaffold00053-processed-gene-9.6	0,0000	-1,7124	-1,4201	(100)
ADAMTS 1	A disintegrin and metalloproteinase with thrombospondin motifs 1	female fertilization	M	maker-scaffold00039-snap-gene-14.31-	-2,1682	0,0000	-3,0538	(149)
PCK2	Phosphoenolpyruvate carboxykinase (GTP)	gene expression / metabolic process	M/ R	genemark-scaffold00018-processed-gene-38.2-	-2,8818	-1,8954	-2,6371	(150)

RPLP2	60S acidic ribosomal protein P2	gene expression	R	maker-scaffold00102-augustus-gene-2.35-	0,0000	0,0000	0,9219	(70)
PRELP	Prolargin	gonad development	M	maker-scaffold00042-snap-gene-17.52-	-2,2768	0,0000	0,0000	(151)
SRC	Proto-oncogene tyrosine-protein kinase Src	heart rate regulation	M	maker-scaffold00025-snap-gene-8.40-	-2,7005	0,0000	0,0000	(152)
FOS	Proto-oncogene c-FOS	implantation	M	genemark-scaffold00077-processed-gene-7.22-	0,0000	-3,0759	-2,4711	(153)
LGALS3	Galectin-3	implantation	M	maker-scaffold00205-snap-gene-0.72-	0,8557	0,0000	0,0000	(118)
PLAU	Urokinase-type plasminogen activator	implantation	M/S	maker-scaffold00015-augustus-gene-48.136	0,0000	1,0030	0,0000	(154)
PRKCD	Protein kinase C delta type	implantation	M/ S	genemark-scaffold00062-processed-gene-2.8-	-1,6277	0,0000	-1,2271	(119)
RAC1	Ras-related protein Rac1	implantation	M	maker-scaffold00015-augustus-gene-6.53-	-1,8460	0,0000	-1,4744	(117)
RHOA	Transforming protein RhoA	implantation	M	maker-scaffold00020-augustus-gene-5.50-	0,0000	0,7965	1,0854	(117)
RHOA	Transforming protein RhoA	implantation	M	maker-scaffold00123-augustus-gene-3.70-	-0,8305	0,0000	0,0000	(117)
DAG1	Dystroglycan	implantation / immune	M/ S	maker-scaffold00001-augustus-gene-94.19	-1,7515	0,0000	-1,3516	(120)
IST1	IST1 homolog	implantation/ cell cycle	M/ S	maker-scaffold00102-augustus-gene-2.32	-1,8704	0,0000	0,0000	(123)
COL12A1	Collagen alpha-1 (XII) chain	implantation/ muscle development	M/ R	maker-scaffold00046-augustus-gene-18.81-	0,0000	-1,0186	0,0000	(125)
COL5A1	Collagen alpha-1 (V) chain	implantation/ muscle development	M	maker-scaffold00050-snap-gene-3.39-	-1,2435	-1,2225	-0,9425	(121)
COL6A3	Collagen alpha-3 (VI) chain	implantation/ muscle development	M/ R	maker-scaffold00014-snap-gene-12.69-	-1,2294	-1,0016	0,0000	(122)
BTF3	Transcription factor BTF3	in utero development	M/ R	maker-scaffold00050-augustus-gene-7.45-	0,0000	0,0000	0,9253	(155)
DDIT4	DNA damage-inducible transcript 4 protein	in utero development	M/ R/ S	maker-scaffold00015-augustus-gene-12.55	1,6882	0,0000	0,0000	
PGR	Progesterone receptor	in utero development	M/ R	genemark-scaffold00009-processed-gene-58.22-	0,0000	-1,0004	0,0000	(117)
APOA1	Apolipoprotein A-I	in utero development / progesterone	M/ R/ S	genemark-scaffold00009-processed-gene-42.17-	0,0000	-4,6493	-5,4239	(98)
ATF3	Cyclic AMP-dependent transcription factor ATF-3	in utero development / progesterone	M/ S	maker-scaffold00046-augustus-gene-16.69	0,0000	-1,5039	-1,6576	(99)
CASP3	Caspase-3	in utero development / progesterone	M/ S	maker-scaffold00012-snap-gene-37.32	-1,1179	0,0000	0,0000	(90)
HSD11B1 L	Hydroxysteroid 11-beta-dehydrogenase 1-like protein	in utero development / progesterone	M	maker-scaffold00011-augustus-gene-7.29-	-2,4510	0,0000	0,0000	(106)
HSD11B2	Corticosterod 11-beta-dehydrogenase isozyme 2	in utero development / progesterone	M/ S	maker-scaffold00036-snap-gene-1.71	0,0000	0,0000	1,1837	(106)

IGFBP1	Insulin-like growth factor-binding protein 1	in utero development / progesterone	M	maker-scaffold00090-augustus-gene-1.37-	0,0000	-1,1028	-0,9644	(156)
IGFBP6	Insulin-like growth factor-binding protein 6	in utero development / progesterone	M	maker-scaffold00001-augustus-gene-86.42-	-3,5975	0,0000	0,0000	(157)
SH3D19	SH3 domain-containing protein 19	labor	M/ S	maker-scaffold00038-augustus-gene-18.55	0,0000	0,0000	1,2053	(158)
TGFB3	Transforming growth factor beta-3 proprotein	labor	M	maker-scaffold00010-augustus-gene-54.34-	-1,7547	0,0000	-1,2334	(157)
PDE4B	Phosphodiesterase	labor / uterus relaxant	S	maker-scaffold00002-snap-gene-49.81	-1,3441	0,0000	-1,1591	(159)
ACTA	Actin	placenta	M/ R/ S	augustus_masked-scaffold00199-processed-gene-0.56	0,0000	-1,8485	0,0000	(130)
ADM2	Adrenomedullin	placenta	M	maker-scaffold00029-augustus-gene-25.43-	-3,0992	-1,8960	-2,5006	(160)
AHNAK	Neoroblast differentiation-associated protein AHNAK	placenta	M/ R	maker-scaffold00013-snap-gene-47.40-	-1,9331	0,0000	0,0000	(126)
BNIP3	BCL2-interacting protein 3	placenta	M	maker-scaffold00043-augustus-gene-19.80-	1,6160	2,0576	2,5098	(161)
CREBL2	cAMP-responsive element-binding protein-like 2	placenta	M	maker-scaffold00064-augustus-gene-13.135-	0,9927	1,1813	1,0872	(162)
DLX3B	Homeobox protein Dlx3b	placenta	M/ S	maker-scaffold00015-augustus-gene-4.41	0,0000	-0,7628	-0,8589	(71)
FRRS1	Ferric-chelate reductase 1	placenta	M/ S	genemark-scaffold00002-processed-gene-68.36	-2,3810	0,0000	-1,9142	(131)
FURIN	Furin	placenta	M	maker-scaffold00005-augustus-gene-17.32-	-0,9000	0,0000	0,0000	(127)
GHR	Growth hormone receptor	placenta	M	maker-scaffold00063-snap-gene-13.80-	0,0000	0,9393	0,0000	(163)
GPX4	Phospholipid hydroperoxide glutathione peroxidase	placenta	M/ R	maker-scaffold00011-augustus-gene-14.50-	2,2665	0,0000	0,0000	(164)
PLAC8	Placenta-specific gene 8 protein	placenta	M	maker-scaffold00050-augustus-gene-0.51-	0,0000	1,8074	2,1520	(165)
PRDM1	PR domain zinc finger protein 1	immune/ placenta	M	augustus_masked-scaffold00058-processed-gene-3.7-	-3,2592	0,0000	-2,1695	(86)
FN1	Fibronectin	placenta	M/ S	snap_masked-scaffold00014-processed-gene-18.32-	-1,2274	0,0000	0,0000	(128)
TNFRSF21	Tumor necrosis factor receptor superfamily member 21	immune / placenta	M	genemark-scaffold00135-processed-gene-1.14-	-2,2489	-1,2337	-1,5501	(129)
SLC9A1	Sodium/ hydrogen exchanger 1	regulation	M/ S	genemark-scaffold00035-processed-gene-13.8	-0,9568	0,0000	0,0000	(166)
NR4A3	Nuclear receptor subfamily 4 group A member 3	sperm	M/ S	maker-scaffold00081-augustus-gene-6.41	-3,4252	0,0000	0,0000	(167)
CALCOCO1	Calcium-binding and coiled-coil domain-containing protein 1	transcriptional activator	S	maker-scaffold00067-snap-gene-9.48	0,0000	2,4180	0,0000	(71)
AQP3	Aquaporin-3	transport	M/ S	maker-scaffold00027-augustus-gene-21.29	-0,9670	0,0000	0,0000	(112)
CA4	Carbonic anhydrase 4	transport	S	maker-scaffold00008-augustus-gene-25.70	0,0000	-1,6192	-1,9324	(71)
CALR	Calreticulin	transport	S	maker-scaffold00159-augustus-gene-1.75	-1,0409	0,0000	-1,5459	(71)
CLDN5	Claudin-5	transport	M/ R	augustus_masked-scaffold00027-processed-gene-3.10-	0,0000	0,0000	2,1372	(115)
FABP1	Fatty acid-binding protein, liver	transport	M/ R/S	maker-scaffold00030-augustus-gene-13.52	0,0000	-2,0520	-3,1418	(113)
FABP3	Fatty acid-binding protein, heart	transport	M/ R/ S	augustus_masked-scaffold00165-processed-gene-0.29	-1,6310	-1,3859	0,0000	(113)
FABP6	Fatty acid-binding protein, ileal (gastrotropin)	transport	M/ R/ S	maker-scaffold00012-snap-gene-35.32-	0,0000	0,0000	1,1277	(113)
GLT8D2	Glycosyltransferase 8 domain-containing protein 2	transport	S	maker-scaffold00029-augustus-gene-25.49	-1,5780	0,0000	0,0000	(71)
HK2	Hexokinase-2	transport	M	maker-scaffold00003-augustus-gene-56.20-	-0,8645	0,0000	0,0000	(111)
MAT2A	S-adenosylmethionine synthase isoform type-2	transport	S	maker-scaffold00022-augustus-gene-39.77	-0,9045	0,0000	0,0000	(71)
RHAG	Ammonium transporter Rh type A	transport	M/ S	maker-scaffold00046-augustus-gene-3.64	0,0000	0,0000	-1,3216	(168)
SLC2A1	Solute carrier family 2	transport	M	maker-scaffold00025-augustus-gene-15.30-	-0,9338	0,0000	0,0000	(107)
SLC4A1	Anion exchange protein	transport	M/ S	genemark-scaffold00015-processed-gene-35.30	1,3207	0,9607	1,3026	(169)
TF	Serotransferrin	transport	M/ S	maker-scaffold00096-augustus-gene-3.52	0,0000	-2,5512	-1,7603	(114)

NOXO1	NADPH oxidase 1	uterine receptivity	M	genemark-scaffold00004-processed-gene-7.16-	-1,0101	0,0000	0,0000	(170)
NOXO1	NADPH oxidase 1	uterine receptivity	M	maker-scaffold00051-augustus-gene-13.52-	-2,5288	0,0000	0,0000	(170)
PTGDR2	Prostaglandin D2 receptor 2	uterine receptivity / prostaglandin	M	genemark-scaffold00055-processed-gene-8.7-	0,0000	0,0000	1,3894	(97)
PTGER4	Prostaglandin E2 receptor EP4 subtype	uterine receptivity / prostaglandin	M	maker-scaffold00050-augustus-gene-11.64-	0,0000	0,0000	0,9639	(97)
PTGFR	Prostaglandin F2-alpha receptor	uterine receptivity / prostaglandin	M	maker-scaffold00011-augustus-gene-20.61-	3,0461	-1,1214	0,0000	(97)
PTGFRN	Prostaglandin F2 receptor inhibitor	uterine receptivity / prostaglandin	M	maker-scaffold00079-snap-gene-5.50-	-1,6674	0,0000	0,0000	(97)
PTGIS	Prostacyclin synthase	uterine receptivity / prostaglandin	M/ R	maker-scaffold00020-augustus-gene-1.27-	-1,5772	0,0000	0,0000	(97)
PTGS1	Prostaglandin G/H synthase 1	uterine receptivity / prostaglandin	M	maker-scaffold00022-augustus-gene-21.56-	0,0000	-0,8775	0,0000	(97)
PTGS2	Prostaglandin G/H synthase 2	uterine receptivity / prostaglandin	M	maker-scaffold00011-augustus-gene-46.41-	-2,2945	0,0000	0,0000	(97)

SI Appendix Table S13: The genes with most expression changes during pipefish brood pouch development or pipefish pregnancy. Comparison was: developed pouch against undeveloped brood pouch tissue, early pregnant pouch tissue against undeveloped brood pouch tissue and late pregnant pouch tissue against undeveloped pouch tissue. Positive values indicate upregulation in undeveloped pouch tissues.

Gene name	full gene name	process	function	id	dev_vs_undev	e-preg_vs_undev	l-preg_vs_undev
CD74	H-2 class II histocompatibility antigen gamma chain	adaptive immune defence	antigen processing	maker-scaffold00049-augustus-gene-8.11-	-2,0334	-2,78642	-3,23619
FUT9	Alpha-(1,3)-fucosyltransferase 9	adaptive immune defence	antigen processing	maker-scaffold00010-snap-gene-27.42-	-3,08337	0	0
IG KAPPA CHAIN V	Ig kappa chain V-V region MOPC 21	adaptive immune defence	antigen processing	augustus_masked-scaffold00089-processed-gene-7.36-	2,24574	2,36928	0
IGLC1	Ig lambda-1 chain C regions	adaptive immune defence	antigen processing	augustus_masked-scaffold00089-processed-gene-7.39-	0	2,90306	0
IGMU	Ig mu chain C region membrane-bound form	adaptive immune defence	antigen processing	genemark-scaffold00036-processed-gene-9.27-	3,02107	2,77106	3,54336
CLCF1	Cardiotrophin-like cytokine factor 1	adaptive immune defence	B cell maturation	maker-scaffold00013-snap-gene-49.40-	-2,84877	-3,22815	-3,35656
FCRL5	Fc receptor-like protein 5	adaptive immune defence	B cell maturation	maker-scaffold00021-augustus-gene-18.49-	0	2,3117	2,52554
GPR97	Probable G-protein coupled receptor 97	adaptive immune defence	B cell maturation	maker-scaffold00036-augustus-gene-9.58-	1,71716	1,44545	2,23253
MFNG	Similar to mfng: Beta-1,3-N-acetylglucosaminyltransferase manic fringe	adaptive immune defence	B cell maturation	maker-scaffold00091-augustus-gene-4.55-	1,84666	0	2,04404
PRDM1	PR domain zinc finger protein 1	adaptive immune defence	B cell maturation	augustus_masked-scaffold00058-processed-gene-3.7-	-3,25921	0	-2,1695
GIMAP4	GTPase IMAP family member 4	adaptive immune defence	T / B cell maturation	snap_masked-scaffold00861-processed-gene-0.1-	-2,74912	0	-1,90482
CHIA	Acidic mammalian chitinase	adaptive immune defence	T cell maturation	maker-scaffold00135-snap-gene-0.29-	3,30331	0	0
HAVCR1	Hepatitis A virus cellular receptor 1 homolog	adaptive immune defence	T cell maturation	genemark-scaffold00008-processed-gene-22.3-	1,25439	1,9961	2,33745
NFATC4	Nuclear factor of activated T-cells, cytoplasmic 4	adaptive immune defence	T cell maturation	maker-scaffold00002-augustus-gene-71.51-	1,04269	1,40499	2,23711
TNFRSF21	Tumor necrosis factor receptor superfamily member 21	adaptive immune defence	T cell maturation	genemark-scaffold00135-processed-gene-1.14-	-2,24893	-1,23367	-1,55011
IL17REL	Putative interleukin-17 receptor E-like	innate immune defence	inflammation	maker-scaffold00029-augustus-gene-28.42-	-2,0652	0	-1,86166
NLRP6	NACHT, LRR and PYD domains-containing protein 6	innate immune defence	inflammation	maker-scaffold00064-snap-gene-8.117-	0	2,16319	2,88737
PLA2G4A	Cytosolic phospholipase A2	innate immune defence	inflammation	maker-scaffold00011-augustus-gene-46.33-	-2,80696	0	0
ADSSL1C	Adenylosuccinate synthetase isozyme 1 C	innate immune defence	innate immune defence	maker-scaffold00044-augustus-gene-0.31-	3,80359	0	0
C6	Complement component C6	innate immune defence	innate immune defence	maker-scaffold00063-snap-gene-13.81-	0	-2,30338	0
EPX	Eosinophil peroxidase	innate immune defence	innate immune defence	maker-scaffold00022-augustus-gene-39.76-	-4,07819	-1,52751	-2,12184
GAL LECTIN	Galactose-specific lectin nattectin	innate immune defence	innate immune defence	maker-scaffold00399-augustus-gene-0.5-	3,41083	0	0

GREM2	Gremlin-2	innate immune defence	innate immune defence	genemark-scaffold00019-processed-gene-12.2-	-3,02402	0	-1,50035
JUND	Transcription factor jun-D	innate immune defence	innate immune defence	augustus_masked-scaffold00088-processed-gene-6.31-	0	-2,47438	-1,86318
MIF	Macrophage migration inhibitory factor	innate immune defence	innate immune defence	maker-scaffold00219-augustus-gene-0.52-	2,13754	0	0
PLE3	Pleurocidin-like peptide WF3	innate immune defence	innate immune defence	maker-scaffold00046-augustus-gene-11.64-	4,52164	3,51947	2,43555
S100A13	Protein S100-A13	innate immune defence	innate immune defence	augustus_masked-scaffold00024-processed-gene-19.8-	4,06137	2,78926	1,93793
TNF	Tumor necrosis factor	innate immune defence	innate immune defence	maker-scaffold00058-augustus-gene-13.61-	-2,89354	0	-3,03681
TNFAIP8L1	Tumor necrosis factor alpha-induced protein 8-like protein 1	innate immune defence	innate immune defence	augustus_masked-scaffold00037-processed-gene-23.37-	-2,12959	0	-1,32696
RAPGEF3	Rap guanine nucleotide exchange factor 3	developmental process	angiogenesis	maker-scaffold00001-augustus-gene-28.24-	-2,07274	0	0
COL10A1	Collagen alpha-1(X) chain	developmental process	bone development	maker-scaffold00010-augustus-gene-27.38-	-1,24964	0	-2,5573
COL11A1	Collagen alpha-1(XI) chain	developmental process	bone development	genemark-scaffold00002-processed-gene-59.6-	-3,32668	-2,88118	-3,34878
COL2A1	Collagen alpha-1(27) chain	developmental process	bone development	maker-scaffold00003-snap-gene-14.44-	-2,31012	-1,32123	0
KLHL41	Kelch-like protein 41	developmental process	bone development	maker-scaffold00066-augustus-gene-7.29-	2,06188	0	0
PDLIM3	PDZ and LIM domain protein 3	developmental process	bone development	maker-scaffold00034-augustus-gene-21.35-	4,66674	0	0
PDLIM7	PDZ and LIM domain protein 7	developmental process	bone development	maker-scaffold00027-snap-gene-28.44-	2,84204	0	0
SMAD9	Mothers against decapentaplegic homolog 9	developmental process	bone development	maker-scaffold00008-augustus-gene-12.43-	3,08407	0	0
SOST	Sclerostin	developmental process	bone development	maker-scaffold00015-augustus-gene-14.39-	0	-2,08699	0
ANKRD1	Ankyrin repeat domain-containing protein 1	developmental process	cardiovascular development	maker-scaffold00145-augustus-gene-0.39-	0	3,6418	3,69306
HEY1	Hairy/enhancer-of-split related with YRPW motif protein 1	developmental process	cardiovascular development	maker-scaffold00026-augustus-gene-31.86-	2,07784	1,43095	2,95744
SMYD!	Histone-lysine N-methyltransferase Smyd1	developmental process	cardiovascular development	maker-scaffold00054-augustus-gene-2.37-	3,12441	0	0
SMYD1	Histone-lysine N-methyltransferase SMYD1	developmental process	cardiovascular development	maker-scaffold00030-augustus-gene-13.62-	3,45469	1,28781	0
SMYD2B	N-lysine methyltransferase SMYD2-B	developmental process	cardiovascular development	augustus_masked-scaffold00077-processed-gene-3.32-	3,45326	0	0
BHMT	Betaine-homocysteine S-methyltransferase 1	developmental process	digestive system	maker-scaffold00050-augustus-gene-5.58-	5,10341	0	0
BHMT	Betaine-homocysteine S-methyltransferase 1	developmental process	digestive system	maker-scaffold00030-augustus-gene-10.33-	0	-2,66341	-3,19067
ALDH1A2	Retinal dehydrogenase 2	developmental process	embryonal development	maker-scaffold00005-augustus-gene-21.23-	-3,17177	-2,18398	-2,42251
ARG 8	[Arg8]-vasotocin receptor	developmental process	embryonal development	augustus_masked-scaffold00017-processed-gene-0.2-	2,04437	0	0
CRABP2	Cellular retinoic acid-binding protein 2	developmental process	embryonal development	maker-scaffold00101-snap-gene-6.79-	-1,96219	-2,13417	-1,92426

HCEA	High choriolytic enzyme 1	developmental process	embryonal development	maker-scaffold00102-augustus-gene-1.33-	2,76964	0	0
KLF4	Krueppel-like factor 4	developmental process	embryonal development	snap_masked-scaffold00048-processed-gene-11.20-	0	-2,00339	-2,43069
KLF9	Krueppel-like factor 9	developmental process	endometrium	maker-scaffold00027-augustus-gene-13.35-	0	2,33702	1,66913
DKK2	Dickkopf-related protein 2	developmental process	endometrium/ progesterone	maker-scaffold00034-augustus-gene-5.34-	3,23958	1,28659	2,89159
CKM	Creatine kinase M-type	developmental process	energy transduction	maker-scaffold00074-snap-gene-2.47-	4,43949	0	0
EPGN	Epigen	developmental process	EPGF response	maker-scaffold00027-augustus-gene-2.67-	2,12098	0	1,29576
TGFA	Protransforming growth factor alpha (Fragment)	developmental process	EPGF response	maker-scaffold00020-augustus-gene-9.47-	-3,723	0	0
STX2	Syntaxin-2	developmental process	epithelial morphogenesis	maker-scaffold00003-augustus-gene-38.26-	-5,23102	0	0
TMEM79	Transmembrane protein 79	developmental process	epithelial morphogenesis	genemark-scaffold00167-processed-gene-0.4-	-2,35773	0	0
ADAMTS1	A disintegrin and metalloproteinase with thrombospondin motifs 1	developmental process	female fertilization	maker-scaffold00039-snap-gene-14.31-	-2,16821	0	-3,05377
CUZD1	CUB and zona pellucida-like domain-containing protein 1	developmental process	in utero development	maker-scaffold00126-snap-gene-3.81-	-4,82703	0	-2,87368
TPM1	Tropomyosin alpha-1 chain	developmental process	in utero development	maker-scaffold00017-snap-gene-28.74-	4,60837	0	0
TPM1	Tropomyosin alpha-1 chain	developmental process	in utero development	maker-scaffold00094-snap-gene-4.43-	2,15154	0	-1,51268
ALOX9	Arachidonate 8S-lipoxygenase	developmental process	keratinization	snap_masked-scaffold00034-processed-gene-6.31-	0	-2,02663	0
CNFN-B	Cornifelin homolog B	developmental process	keratinization	maker-scaffold00074-augustus-gene-7.58-	-2,53543	-1,75197	-1,8148
ENO3	Beta-enolase	developmental process	muscle development	genemark-scaffold00038-processed-gene-4.5-	4,76423	0	0
JPH2	Junctophilin-2	developmental process	muscle development	maker-scaffold00001-augustus-gene-92.25-	3,79542	0	0
MURC	Muscle-related coiled-coil protein	developmental process	muscle development	augustus_masked-scaffold00065-processed-gene-8.5-	2,89224	0	0
SMTNL2	Smoothelin-like protein 2	developmental process	muscle development	genemark-scaffold00009-processed-gene-29.24-	-2,19152	0	0
TNNI2	Troponin I, fast skeletal muscle	developmental process	muscle development	augustus_masked-scaffold00017-processed-gene-40.18-	4,58781	0	0
DCHS1	Protocadherin-16	developmental process	neuronal development	maker-scaffold00034-snap-gene-20.67-	2,07693	0	0
KCNN2	Small conductance calcium-activated potassium channel protein 2	developmental process	neuronal development	augustus_masked-scaffold00026-processed-gene-22.6-	-3,62704	0	0
REG1B	Lithostathine-1-beta	developmental process	neuronal development	maker-scaffold00006-augustus-gene-61.44-	-3,03354	0	0
SERPINI1	Neuroserpin	developmental process	neuronal development	maker-scaffold00116-augustus-gene-2.25-	0	-4,97898	0
BGLAP	Osteocalcin	developmental process	organism tissue development	maker-scaffold00004-augustus-gene-64.40-	0	-3,04357	-2,03638

ENAH	Protein enabled homolog	developmental process	organism tissue development	maker-scaffold00061-augustus-gene-14.38-	-2,6377	0	0
MSXC	Homeobox protein MSH-C	developmental process	organism tissue development	maker-scaffold00023-augustus-gene-21.40-	-4,62086	0	-3,76791
MYL2	Myosin regulatory light chain 2, ventricular/cardiac muscle isoform	developmental process	organism tissue development	maker-scaffold00042-augustus-gene-7.44-	0	0	-2,25899
SEMA3G	Semaphorin-3G	developmental process	organism tissue development	maker-scaffold00001-augustus-gene-22.27-	-2,3677	0	0
SOX7	Transcription factor Sox-7	developmental process	organism tissue development	maker-scaffold00077-snap-gene-5.92-	-2,10665	-0,966524	-1,5667
TTC36	Tetratricopeptide repeat protein 36	developmental process	organism tissue development	maker-scaffold00021-augustus-gene-40.55-	1,7581	1,8393	2,57374
Znf513	Zinc finger protein 513	developmental process	organism tissue development	maker-scaffold00079-augustus-gene-1.28-	-2,03022	0	0
Adm2	Adrenomedullin	developmental process	placenta	maker-scaffold00029-augustus-gene-25.43-	-3,09917	-1,89595	-2,5006
GPX4	Phospholipid hydroperoxide glutathione peroxidase 2C mitochondrial	developmental process	placenta	maker-scaffold00011-augustus-gene-14.50-	2,26651	0	0
PLAC8	Placenta-specific gene 8 protein	developmental process	placenta	maker-scaffold00050-augustus-gene-0.51-	0	1,80736	2,15204
NR4A3	Nuclear receptor subfamily 4 group A member 3	developmental process	sperm	maker-scaffold00081-augustus-gene-6.41-	-3,42519	0	0
PLDXC1	Plexin domain-containing protein 1	developmental process	spinal cord development	maker-scaffold00015-augustus-gene-15.54-	2,37789	1,81104	3,0739
FABP6	Gastrotropin	developmental process	transport / apoptosis	augustus_masked-scaffold00012-processed-gene-35.1-	-2,11932	-2,36517	-2,07423
CLDN5	Claudin-5	developmental process	transport / tight junctions	augustus_masked-scaffold00027-processed-gene-3.10-	0	0	2,13715
NOXO1	NADPH oxidase organizer 1	developmental process	uterine receptivity	maker-scaffold00051-augustus-gene-13.52-	-2,52875	0	0
BNIP3	BCL2/adenovirus E1B 19 kDa protein-interacting protein 3 (Homo sapiens)	apoptosis	apoptosis	maker-scaffold00043-augustus-gene-19.80-	1,61604	2,05762	2,50975
FHL2	Four and a half LIM domains protein 2	apoptosis	apoptosis	augustus_masked-scaffold00014-processed-gene-24.1-	0	-2,57293	-2,67605
FHL2	Four and a half LIM domains protein 2	apoptosis	apoptosis	maker-scaffold00010-augustus-gene-27.35-	3,2775	0	-2,21539
PIM3	Serine/threonine-protein kinase pim-3	apoptosis	apoptosis	maker-scaffold00029-augustus-gene-27.40-	-1,34117	-1,66014	-2,3264
SIX1B	Homeobox protein six1b	apoptosis	apoptosis	maker-scaffold00093-augustus-gene-3.28-	2,44617	0	0
SLC25A4	ADP/ATP translocase 1	apoptosis	apoptosis	augustus_masked-scaffold00045-processed-gene-3.42-	3,00264	0	0
VOPP1	Vesicular, overexpressed in cancer, prosurvival protein 1	apoptosis	apoptosis	genemark-scaffold00018-processed-gene-19.12-	0	2,34141	0
CAMK1G	Calcium/calmodulin-dependent protein kinase type 1G	binding	binding	maker-scaffold00001-augustus-gene-87.45-	-2,74046	0	0
RCN3	Reticulocalbin-3	binding	binding	augustus_masked-scaffold00047-processed-gene-6.11-	-1,9337	-2,01913	-2,18533
TMOD4	Tropomodulin-4	blood	blood	maker-scaffold00058-snap-gene-2.65-	3,85016	0	0
B3GNT7	N-acetyllactosaminide beta-13-N-acetylglucosaminyltransferase 7	cell differentiation	cell adhesion / proliferation	augustus_masked-scaffold00082-processed-gene-0.1-	-2,82901	0	-1,86002
BCR	Breakpoint cluster region protein	cell differentiation	cell adhesion / proliferation	genemark-scaffold00003-processed-gene-60.11-	-2,824	0	0

CDKN3	Cyclin-dependent kinase inhibitor 3	cell differentiation	cell adhesion / proliferation	augustus_masked-scaffold00010-processed-gene-17.38-	-2,13229	0	0
FAM64A	Protein FAM64A	cell differentiation	cell adhesion / proliferation	maker-scaffold00008-snap-gene-21.86-	0	-2,46495	0
IGFBP6	Insulin-like growth factor-binding protein 6	cell differentiation	cell adhesion / proliferation	maker-scaffold00001-augustus-gene-86.42-	-3,59747	0	0
MKX	Homeobox protein Mohawk	cell differentiation	cell adhesion / proliferation	maker-scaffold00176-augustus-gene-1.62-	-2,39185	0	0
MSLNL	Mesothelin-like protein	cell differentiation	cell adhesion / proliferation	maker-scaffold00034-augustus-gene-1.45-	-2,06612	0	-1,53551
NMRK2	Nicotinamide riboside kinase 2	cell differentiation	cell adhesion / proliferation	maker-scaffold00037-augustus-gene-23.60-	0	0	-2,45289
PRELP	Prolargin	cell differentiation	cell adhesion / proliferation	maker-scaffold00042-snap-gene-17.52-	-2,27678	0	0
SRC	Proto-oncogene tyrosine-protein kinase Src	cell differentiation	cell adhesion / proliferation	maker-scaffold00025-snap-gene-8.40-	-2,70049	0	0
TUBA1A	Tubulin alpha-1A chain	cell differentiation	cell adhesion / proliferation	maker-scaffold00001-augustus-gene-30.30-	0	-3,04922	-2,754
TUBA1B	Tubulin alpha-1B chain	cell differentiation	cell adhesion / proliferation	augustus_masked-scaffold00001-processed-gene-30.4-	0	-3,96079	-3,65919
TUBA1C	Tubulin alpha-1C chain	cell differentiation	cell adhesion / proliferation	augustus_masked-scaffold00001-processed-gene-47.16-	0	-2,18976	-1,79009
TSPAN1	Tetraspanin-1	cell differentiation	conceptus growth / cell adhesion /	maker-scaffold00045-augustus-gene-11.30-	0	0	-4,92762
APOBEC2	Probable C->U-editing enzyme APOBEC-2	gene expression	epigenetic / gene expression	maker-scaffold00162-augustus-gene-1.96-	3,0827	0	0
DXO	Decapping and exoribonuclease protein	gene expression	epigenetic / gene expression	maker-scaffold00004-snap-gene-3.83-	0	2,88047	2,8908
DXO	Decapping and exoribonuclease protein	gene expression	epigenetic / gene expression	maker-scaffold00004-snap-gene-3.82-	0	3,12669	2,37762
ERCC4	DNA repair endonuclease XPF	gene expression	epigenetic / gene expression	maker-scaffold00051-snap-gene-11.27-	-2,09033	0	0
KLHL31	Kelch-like protein 31	gene expression	epigenetic / gene expression	maker-scaffold00024-augustus-gene-25.46-	3,11654	0	0
MEF2D	Myocyte-specific enhancer factor 2D homolog	gene expression	epigenetic / gene expression	maker-scaffold00058-snap-gene-5.52-	-1,79673	0	-2,15293
PABPC4	Polyadenylate-binding protein 4	gene expression	epigenetic / gene expression	maker-scaffold00044-snap-gene-14.43-	2,07771	0	0
TRIM55	Tripartite motif-containing protein 55	gene expression	epigenetic / gene expression	maker-scaffold00002-augustus-gene-43.71-	2,90674	1,32087	0
APOA1	Apolipoprotein A-I	hormone	steroid / hormone	genemark-scaffold00009-processed-gene-42.17-	0	-4,64929	-5,42389
CALCOCO1	Calcium-binding and coiled-coil domain-containing protein 1	hormone	steroid / hormone	maker-scaffold00067-snap-gene-9.48-	0	2,41804	0
HSD11B1	Hydroxysteroid 11-beta-dehydrogenase 1-like protein	hormone	steroid / hormone	maker-scaffold00011-augustus-gene-7.29-	-2,45103	0	0
PTGFR	Prostaglandin F2-alpha receptor	hormone	steroid / hormone	maker-scaffold00011-augustus-gene-20.61-	3,04614	-1,12135	0

PTGS2	Prostaglandin G/H synthase 2	hormone	steroid / hormone	maker-scaffold00011-augustus-gene-46.41-	-2,29452	0	0
FAM69A	Protein FAM69A	membrane	membrane	maker-scaffold00002-snap-gene-19.50-	-2,20669	0	0
GANZ	Guanine nucleotide-binding protein G(z) subunit alpha	membrane	membrane	maker-scaffold00022-augustus-gene-38.32-	0	-2,17574	0
GPR155	Integral membrane protein GPR155	membrane	membrane	maker-scaffold00007-augustus-gene-25.36-	-2,89823	0	-1,59949
SGCG	Gamma-sarcoglycan	membrane	membrane	augustus_masked-scaffold00009-processed-gene-3.1-	2,535	0	0
TMEM238	Transmembrane protein 238	membrane	membrane	maker-scaffold00015-augustus-gene-42.47-	-2,20134	0	0
TMEM38A	Trimeric intracellular cation channel type A	membrane	membrane	maker-scaffold00011-augustus-gene-41.50-	3,95238	0	0
A1AT	Alpha-1-antitrypsin homolog	metabolic process	metabolic process	maker-scaffold00075-augustus-gene-9.25-	0	-4,61005	-3,91652
ALD	Fructose-bisphosphate aldolase A	metabolic process	metabolic process	maker-scaffold00015-augustus-gene-17.64-	4,6383	0	0
CYP24A1	1,25-dihydroxyvitamin D(3) 24-hydroxylase, mitochondrial	metabolic process	metabolic process	maker-scaffold00025-augustus-gene-8.38-	2,22375	0	1,70951
FAM213A	Redox-regulatory protein FAM213A	metabolic process	metabolic process	maker-scaffold00015-augustus-gene-15.56-	-2,16579	-0,899201	-0,867901
HGD	Homogentisate 1,2-dioxygenase	metabolic process	metabolic process	maker-scaffold00018-augustus-gene-23.52-	1,36034	0	2,10152
HOMER2	Homer protein homolog 2	metabolic process	metabolic process	maker-scaffold00005-augustus-gene-45.56-	-2,2788	0	0
LPL	Lipoprotein lipase	metabolic process	metabolic process	maker-scaffold00011-augustus-gene-3.53-	1,25086	1,87975	2,22114
PCK2	Phosphoenolpyruvate carboxykinase	metabolic process	metabolic process	genemark-scaffold00018-processed-gene-38.2-	-2,88181	-1,89538	-2,63708
PFKM	ATP-dependent 6-phosphofructokinase, muscle type	metabolic process	metabolic process	maker-scaffold00031-augustus-gene-2.23-	0	2,98092	2,8252
PGM1	Phosphoglucomutase-1	metabolic process	metabolic process	maker-scaffold00011-snap-gene-50.32-	2,83088	0	0
TPI1B	Triosephosphate isomerase B	metabolic process	metabolic process	genemark-scaffold00035-processed-gene-2.14-	2,6974	0	0
FABP1	Fatty acid-binding protein%2C liver-type	metabolic process	metabolic process / transport	maker-scaffold00030-augustus-gene-13.52-	0	-2,05204	-3,14184
SH3BGR	SH3 domain-binding glutamic acid-rich protein	mitochondrium	mitochondrium	maker-scaffold00009-augustus-gene-2.28-	5,26001	0	0
MARCH7	E3 ubiquitin-protein ligase MARCH7	protein	protein modification	augustus_masked-scaffold00005-processed-gene-49.1-	-2,10917	0	0
ANKRD28	Serine/threonine-protein phosphatase 6 regulatory ankyrin repeat subunit A	protein	protein modification	maker-scaffold00028-snap-gene-28.47-	-2,38827	0	0
GFPT2	Glutamine--fructose-6-phosphate aminotransferase [isomerizing] 2	protein	protein modification	maker-scaffold00049-augustus-gene-11.41-	-2,26512	0	0
KLHL38	Kelch-like protein 38	protein	protein modification	maker-scaffold00073-augustus-gene-1.58-	2,7606	3,77597	3,18103
MYOZ2	Myozenin-2	protein	protein modification	maker-scaffold00038-augustus-gene-14.29-	3,19367	1,49046	1,63145
RNF144AA	Probable E3 ubiquitin-protein ligase RNF144A-A	protein	protein modification	maker-scaffold00010-snap-gene-47.34-	0	1,63057	2,02624
FAM177A1	Protein FAM177A1	protein	unknown	maker-scaffold00075-snap-gene-7.13-	-2,33533	0	0
ACTA1	Actin, alpha skeletal muscle	sensory reaction	muscle contraction	augustus_masked-scaffold00005-processed-gene-1.34-	4,1379	0	0
ACTC	Actin, alpha cardiac	sensory reaction	muscle contraction	maker-scaffold00010-augustus-gene-52.35-	4,02399	0	0
ACTC	Actin, alpha cardiac	sensory reaction	muscle contraction	maker-scaffold00107-augustus-gene-2.53-	4,24573	3,87088	0
ACTC1	Actin, alpha cardiac muscle 1	sensory reaction	muscle contraction	augustus_masked-scaffold00010-processed-gene-29.2-	2,69478	1,94864	0
CASQ1	Calsequestrin-1 (	sensory reaction	muscle contraction	maker-scaffold00038-snap-gene-2.65-	4,26665	0	0

CASQ2	Calsequestrin-2	sensory reaction	muscle contraction	genemark-scaffold00007-processed-gene-59.17-	1,43488	2,55707	1,09615
M-PROT	M-protein, striated muscle	sensory reaction	muscle contraction	maker-scaffold00125-snap-gene-3.146-	4,84647	0	0
MYLPF	Myosin regulatory light chain, skeletal muscle isoform type 2	sensory reaction	muscle contraction	maker-scaffold00047-augustus-gene-13.99-	4,41798	0	0
MYLPF	Myosin regulatory light chain 2, skeletal muscle isoform	sensory reaction	muscle contraction	maker-scaffold00015-augustus-gene-36.73-	2,09656	2,65305	0
MYOHCH	Myosin heavy chain, fast skeletal muscle	sensory reaction	muscle contraction	maker-scaffold00020-snap-gene-8.51-	0	-1,47376	-2,2916
MYOHCH	Myosin heavy chain, fast skeletal muscle	sensory reaction	muscle contraction	maker-scaffold00004-augustus-gene-31.43-	0	0	-3,25145
MYOHCH	Myosin heavy chain, fast skeletal muscle	sensory reaction	muscle contraction	genemark-scaffold00020-processed-gene-8.10-	3,89969	0	0
MYOLCHC	Myosin light chain 1, cardiac muscle	sensory reaction	muscle contraction	augustus_masked-scaffold00027-processed-gene-1.14-	4,50642	2,90846	0
MYOLCHS	Myosin light chain 3, skeletal muscle isoform	sensory reaction	muscle contraction	maker-scaffold00070-augustus-gene-6.38-	4,59611	0	0
MYOM2	Myomesin-2	sensory reaction	muscle contraction	genemark-scaffold00034-processed-gene-26.31-	3,07832	0	0
MYOM2	MYOM2: Myomesin-2	sensory reaction	muscle contraction	genemark-scaffold00002-processed-gene-18.10-	3,56919	0	0
MYOM2	MYOM2: Myomesin-2	sensory reaction	muscle contraction	maker-scaffold00005-augustus-gene-7.28-	3,96532	0	0
PVALBB2	Parvalbumin beta-2	sensory reaction	muscle contraction	augustus_masked-scaffold00004-processed-gene-9.39-	6,09956	0	0
TNNC	Troponin C, skeletal muscle	sensory reaction	muscle contraction	augustus_masked-scaffold00001-processed-gene-23.9-	4,20496	0	0
TRDN	Triadin	sensory reaction	muscle contraction	maker-scaffold00077-augustus-gene-1.98-	3,57133	1,6103	0
TTN	Titin	sensory reaction	muscle contraction	snap_masked-scaffold00014-processed-gene-1.9-	2,85435	0	0
FAM134P	Protein FAM134B	sensory reaction	sensory reaction	maker-scaffold00002-augustus-gene-66.64-	0	2,6039	0
FAM222A	Protein FAM222A	sensory reaction	sensory reaction	maker-scaffold00022-augustus-gene-29.43-	-2,50465	0	0
S100P	Protein S100-P	sensory reaction	sensory reaction	maker-scaffold00155-augustus-gene-2.81-	0	0	-2,77289
TMEM150A	Transmembrane protein 150A	sensory reaction	sensory reaction	genemark-scaffold00043-processed-gene-13.31-	-3,02026	0	-1,94394
VEGP	Lipocalin	sensory reaction	sensory reaction	maker-scaffold00018-augustus-gene-27.23-	2,55689	0	0
GMPR	GMP reductase 1	sensory reaction	stress regulation	snap_masked-scaffold00024-processed-gene-32.60-	2,27393	0	0
CIRL	Latrophilin Cirl	signaling	signaling	maker-scaffold00023-augustus-gene-32.77-	0	-1,8218	-2,3406
DHRS7CA	Dehydrogenase/reductase SDR family member 7C-A	signaling	signaling	maker-scaffold00015-augustus-gene-45.34-	4,1668	0	0
FOS	Proto-oncogene c-Fos	signaling	signaling	genemark-scaffold00077-processed-gene-7.22-	0	-3,07587	-2,47111
LGR6	Leucine-rich repeat-containing G-protein coupled receptor 6	signaling	signaling	maker-scaffold00070-augustus-gene-11.61-	0	-3,31032	0
RGS3	Regulator of G-protein signaling 3	signaling	signaling	genemark-scaffold00048-processed-gene-4.2-	-2,37802	0	0
CILP	Cartilage intermediate layer protein 1	stress regulation	insulin pathway	maker-scaffold00121-augustus-gene-4.34-	0	0	-5,25234
PDK2	Pyruvate dehydrogenase (acetyl-transferring) kinase isozyme 2, mitochondrial	stress regulation	insulin pathway	maker-scaffold00004-augustus-gene-49.23-	2,31714	1,1115	1,37197
SIK2	Similar to SIK2: Serine/threonine-protein kinase SIK2 (Gallus gallus)	stress regulation	insulin pathway	maker-scaffold00007-augustus-gene-64.40-	0	0	-2,55981
ALAS2	5-aminolevulinate synthase, erythroid-specific, mitochondrial	stress regulation	stress regulation	augustus_masked-scaffold00031-processed-gene-16.9-	-1,54571	0	-2,01443
ARRDC3	Arrestin domain-containing protein 3	stress regulation	stress regulation	maker-scaffold00050-augustus-gene-14.86-	0	1,74748	2,48304
CAHZ	Similar to cahz: Carbonic anhydrase (Danio rerio)	stress regulation	stress regulation	maker-scaffold00002-augustus-gene-83.39-	-1,39043	0	-2,12137
CILP	Cartilage intermediate layer protein 1	stress regulation	stress regulation	maker-scaffold00121-snap-gene-3.41-	0	0	-7,15441
HSPA8	Heat shock cognate 71 kDa protein	stress regulation	stress regulation	maker-scaffold00026-snap-gene-22.59-	2,93941	0	0

HSPB1	Heat shock protein beta-1	stress regulation	stress regulation	maker-scaffold00008-augustus-gene-41.38-	2,42169	0	0
LS-12	Type-4 ice-structuring protein LS-12	stress regulation	stress regulation	maker-scaffold00024-augustus-gene-17.72-	0	-4,58777	-5,00146
MAP3K7CL	MAP3K7 C-terminal-like protein	stress regulation	stress regulation	maker-scaffold00008-snap-gene-20.49-	2,19517	0	0
METRNL	Meteorin-like protein	stress regulation	stress regulation	genemark-scaffold00053-processed-gene-11.20-	-2,12656	0	0
MYLK	Myosin light chain kinase, smooth muscle (Fragment)	stress regulation	stress regulation	maker-scaffold00051-snap-gene-7.73-	2,27832	0	0
NEOVE	Neoverrucotoxin subunit alpha ´	stress regulation	stress regulation	genemark-scaffold00047-processed-gene-2.8-	-4,09709	-3,76871	-3,92026
PIP4K2A	Phosphatidylinositol 5-phosphate 4-kinase type-2 alpha	stress regulation	stress regulation	genemark-scaffold00018-processed-gene-26.16-	-2,04972	0	0
PLCL1	Inactive phospholipase C-like protein 1	stress regulation	stress regulation	genemark-scaffold00014-processed-gene-50.6-	-2,17933	0	0
FRRS1	Putative ferric-chelate reductase 1	transport	transport	genemark-scaffold00002-processed-gene-68.36-	-2,38096	0	-1,91422
NAGLT1	Sodium-dependent glucose transporter 1	transport	transport	maker-scaffold00135-augustus-gene-3.51-	-2,36216	0	0
SLC43A2	Large neutral amino acids transporter small subunit 4	transport	transport	maker-scaffold00008-augustus-gene-2.26-	3,48693	1,4322	0
SLC4A1	Band 3 anion exchange protein	transport	transport	maker-scaffold00047-snap-gene-15.64-	0	0	-2,65229
SRL	Sarcalumenin	transport	transport	maker-scaffold00141-snap-gene-3.61-	3,58937	0	0
TF	Serotransferrin	transport	transport	maker-scaffold00096-augustus-gene-3.52-	0	-2,55122	-1,76032
TOMM40	Mitochondrial import receptor subunit TOM40 homolog	transport	transport	maker-scaffold00028-snap-gene-31.109-	0	-4,39196	-3,10469
unknown	Protein of unknown function	unknown	unknown	maker-scaffold00003-augustus-gene-27.38	0	0	-2,57625
unknown	Protein of unknown function	unknown	unknown	maker-scaffold00015-snap-gene-31.46-	-3,97258	0	0
unknown	Protein of unknown function	unknown	unknown	maker-scaffold00042-augustus-gene-16.56-	-2,44314	0	0
unknown	Protein of unknown function	unknown	unknown	maker-scaffold00037-augustus-gene-8.38-	-2,11303	0	-1,71089
unknown	Protein of unknown function	unknown	unknown	maker-scaffold00039-augustus-gene-6.26-	-1,86397	-3,2515	0
unknown	Protein of unknown function	unknown	unknown	maker-scaffold00024-augustus-gene-35.150-	0	-4,75471	-4,97039
unknown	Protein of unknown function	unknown	unknown	maker-scaffold00101-augustus-gene-0.61-	0	-4,44566	-5,02445
unknown	Protein of unknown function	unknown	unknown	maker-scaffold00101-augustus-gene-0.59-	0	-4,25239	-3,90736
unknown	Protein of unknown function	unknown	unknown	genemark-scaffold00002-processed-gene-7.20-	0	0	-2,751
unknown	Protein of unknown function	unknown	unknown	genemark-scaffold00032-processed-gene-26.8-	0	0	2,09359
unknown	Protein of unknown function	unknown	unknown	maker-scaffold00049-augustus-gene-11.39-	0	2,62214	3,62875
unknown	Protein of unknown function	unknown	unknown	maker-scaffold00006-augustus-gene-49.46-	1,28609	1,12395	2,02365
unknown	Protein of unknown function	unknown	unknown	maker-scaffold00047-snap-gene-2.41-	2,39669	0	0
UPF0501	Similar to UPF0501 protein KIAA1430 homolog	unknown	unknown	maker-scaffold00045-augustus-gene-3.61-	2,39658	0	0

7.5. Gene expression during pregnancy, comparison *Syngnathus typhle* vs. *Syngnathus scovelli*

To validate our findings with another independent study, we compared the differentially expressed genes during pregnancy of *S. scovelli*(3) and *S. typhle* (undeveloped pouch *vs.* early pregnant). We identified 141 genes that are significantly up- or down regulated during male pregnancy in both species (SI Appendix Table S14).

The direction of expression differences in differentially expressed genes correlated between *S. typhle* and *S. scovelli* (R^2 = 0.767), implying that the direction of up- or downregulation during pregnancy was mostly consistent in both pipefish species. In particular, this was the case for the four genes with the most extreme upregulation during pregnancy (*MYOC*, *HCEA*, *LS-12*, ***APOA1***) and for the two genes that showed the most massive downregulation during pregnancy (*STX2* and *MSXC*).

Several immune genes were overlapping in their expression change in the two species.

CEBPB has in mammals an important role in the early embryogenesis due to its antiproliferative effect on T-cells by repressing MHC expression and facilitating the onset of the anti-inflammatory Th2 response upon the identified upregulation, which is supported by its upregulation in both species during pregnancy. Similar to *SOCS3*, which is also upregulated both in *S. typhle* and in *S. scovelli*, it regulates TH2 immune responses including the maintenance of allergic responses over *IL6*. In line with this, ***IL6R*** was downregulated in both species. *PTX3* (upregulated in *S. typhle,* downregulated in *S. covelli*) has a dual function in the innate immune system and in female fertility, similar to *ID2* that is responsible for the differentiation of leukocytes and also plays a role in embryonal development). The complement system seems activated in both species during pregnancy, as indicated by the upregulation of *C6* and *CFD* (alternative pathway). In line with the predictions for a successful pregnancy, inflammation responses were downregulated in both species (*PLA2G4A*, *STAP2*). Also adaptive immune responses including B-cell maturation were mostly downregulated when they were not simultaneously involved in embryonal development (***PRDM1***, *CLCF1*, however, the latter was upregulated during pregnancy in *S. typhle*).

On the side of the innate immune system *CXCL9*, a cytokine driving the activation of diverse immune cells is downregulated in *S. typhle* but upregulated in *S. scovelli*. *TLR2-1*,

responsible for cytokine secretion and inflammatory response, was upregulated in both species.
APP2 that was upregulated in *S. typhle* but downregulated in *S. scovelli*, is involved in the antimicrobial humoral immune responses. *RAC1* involved in phagocytosis was downregulated during pregnancy in both species. Other innate immune genes were *MAP3K5*, which was slightly upregulated in both species, while ***ADAMTS1*** was upregulated in *S. typhle* but downregulated in *S. scovelli.*

From all genes that were not involved in immune responses, we only discuss those where at least once species had a fold change difference of >2.
Most genes that were differentially expressed both in *S. typhle* and *S. scovelli* during male pregnancy are involved in embryonal development. As such, *STX2* (involved in epithelia morphogenesis), *MSXC* (in humans responsible for the development of the inner ear), *KCNN2* & *ISL3* (neuronal synapses) and *BCR* and *TFGB3* (various functions in embryogenesis) are consistently downregulated during pregnancy in both species.
The opposite pattern could be found for genes involved in osteoblast differentiation and bone formation (*SOST*, *SMAD9*, *MYOC*), egg hatching (*HCEA*, ***APOA1***), they were highly upregulated during pregnancy in both species. For some embryonal development genes, however, we found a downregulation in *Syngnathus typhle* and an upregulation in *S. scovelli* (***DKK2***) or the opposite pattern (*MFAP2*, *ALDH1A2*).

HSD11B1 involved in steroid pathway was downregulated in both species assessed. ***IGFBP6***, known from mammalian pregnancy to inhibit labours, is downregulated in *S. typhle* but upregulated in *S. scovelli.* The stress gene *LS-12* was highly upregulated in both species during pregnancy. Several other genes with functions in cell adhesion and transport were also differentially expressed in both species. The full list of differentially expressed genes during pregnancy of both *S. typhle* and *S. scovelli* can be found in SI Appendix Table S13. This suggests that in both species immune and embryonal developmental genes seem to play major roles during male pregnancy and that differentially expressed genes considerably overlap.

SI Appendix Table S14: Genes with expression changes both during pregnancy of *S. typhle* (1) and *S. scovelli* (SSC)

name	ID STY	*S. typhle* log2FC (undev_epreg)	Gene_ID SSC	*S. scovelli* log2FC	category	function
STX2	maker-scaffold00003-augustus-gene-38.26	-3,80149	SSCG00000004506	-2,585547457	embryonal development	epithelia morphogenesis
MSXC	maker-scaffold00023-augustus-gene-21.40	-3,67628	SSCG00000014649	-3,343167499	embryonal development	ear morphogenesis
KCNN2	augustus_masked-scaffold00026-processed-gene-22.6	-3,02652	SSCG00000008660	-1,066440089	embryonal development	synaptic transmission
CAMK1G	maker-scaffold00001-augustus-gene-87.45	-2,84637	SSCG00000006897	-1,248528878	other	ATP binding
unknown	maker-scaffold00042-augustus-gene-16.56	-2,50322	SSCG00000000884	-1,024166971	unknown	unknown
TMEM150A	genemark-scaffold00043-processed-gene-13.31	-2,41694	SSCG00000000528	-1,158732973	metabolism	catabolism fasting induced
PLA2G4A	maker-scaffold00011-augustus-gene-46.33	-2,36275	SSCG000000011782	-0,728505619	immune	inflammation
PRDM1	augustus_masked-scaffold00058-processed-gene-3.7	-2,32956	SSCG00000002979	-1,981114972	immune	B cell maturation
Mar 07	augustus_masked-scaffold00005-processed-gene-49.1	-2,31625	SSCG00000013551	-1,328816247	protein	protein ubiquination
unknown	maker-scaffold00037-augustus-gene-8.38	-2,2855	SSCG00000005233	-1,056617348	unknown	unknown
HSD11B1L	maker-scaffold00011-augustus-gene-7.29	-2,20912	SSCG00000008093	-1,831251973	hormone	steroid pathway
NAGLT1	maker-scaffold00135-augustus-gene-3.51	-2,19711	SSCG00000000079	-0,931791631	transport	sodium-ion transport
MSLINL	maker-scaffold00034-augustus-gene-1.45	-2,05823	SSCG00000005045	-1,612759244	cell adhesion	cell adhesion
BCR	genemark-scaffold00003-processed-gene-60.11	-2,0315	SSCG00000002754	-1,241047337	embryonal development	embryonal development
FAM69A	maker-scaffold00002-snap-gene-19.50	-2,01196	SSCG00000010175	-1,509702695	other	membrane component
KLHDC1	maker-scaffold00015-augustus-gene-15.62	-1,98957	SSCG00000019315	-0,834554107	other	cytoplasm
TTC36	maker-scaffold00021-augustus-gene-40.55	-1,8393	SSCG00000018122	1,246676237	embryonal development	embryonal development
MSXC	maker-scaffold00008-augustus-gene-9.39	-1,83683	SSCG00000001114	-1,135429194	embryonal development	embryonal development
A33	maker-scaffold00017-augustus-gene-38.36	-1,81953	SSCG00000019560	-0,650342488	unknown	unkonwn
ISLR2	augustus_masked-scaffold00017-processed-gene-34.5	-1,76514	SSCG00000015892	-2,733441317	embryonal development	neural development
ZNF513	maker-scaffold00079-augustus-gene-1.28	-1,72221	SSCG00000003113	-1,004574642	embryonal development	retina development
MKX	maker-scaffold00176-augustus-gene-1.62	-1,67741	SSCG00000008274	-1,34205988	spermatogenesis	male gonad development
ENAH	maker-scaffold00061-augustus-gene-14.38	-1,62648	SSCG00000000062	-0,828343468	embryonal development	filipodia formation
PRKCD	genemark-scaffold00062-processed-gene-2.8	-1,56253	SSCG00000001318	-1,537627077	apoptosis	apoptosis
FGD5	maker-scaffold00062-augustus-gene-2.47	-1,54367	SSCG00000001316	-1,963972199	cell structure	cell shape organisation
TGFB3	maker-scaffold00010-augustus-gene-54.34	-1,53164	SSCG00000015695	-2,197119068	embryonal development	embryogenesis
DGKZ	maker-scaffold00001-augustus-gene-64.58	-1,52896	SSCG00000001663	-0,499073096	blood	blood coagulation
IST1	maker-scaffold00102-augustus-gene-2.32	-1,52575	SSCG00000005977	-0,596301624	cell proliferation	cell cycle/ cell division
TNFAIP8L1	augustus_masked-scaffold00037-processed-gene-23.37	-1,51986	SSCG00000015792	-0,665722321	embryonal development	TOR signaling
RORB	maker-scaffold00011-snap-gene-36.34	-1,51397	SSCG00000016888	-1,731705396	embryonal development	osteoblast differentiation
TTC39A	maker-scaffold00004-augustus-gene-71.33	-1,50919	SSCG00000008919	-0,514041341	unknown	unknown
CHCHD6	maker-scaffold00001-snap-gene-8.27	-1,45905	SSCG00000010556	-1,368986749	unknown	unknown

KIAA0922	genemark-scaffold00034-processed-gene-24.13	-1,45329	SSCG00000016784	-0,842141574	other	membrane component
ATP2A3	maker-scaffold00008-snap-gene-40.40	-1,43516	SSCG00000002189	-0,532445586	transport	calcium ion transport
GRAMD3	maker-scaffold00002-snap-gene-55.34	-1,43461	SSCG00000010479	0,560329335	other	microtubule
HEY1	maker-scaffold00026-augustus-gene-31.86	-1,43095	SSCG00000010880	1,484072479	embryonal development	heart development
RDX	snap_masked-scaffold00014-processed-gene-20.32	-1,41713	SSCG00000014669	0,473036595	cell structure	regulation of cell structure
FMO5	maker-scaffold00011-snap-gene-22.66	-1,41074	SSCG00000015973	0,633769074	other	NADP binding
EHF	maker-scaffold00006-augustus-gene-22.22	-1,36355	SSCG00000011332	-0,941628315	cell proliferation	cell proliferation
RAP1GAP2	maker-scaffold00009-augustus-gene-38.22	-1,34857	SSCG00000009559	-0,690259582	embryonal development	neuronal development
LRIG1	maker-scaffold00062-augustus-gene-5.23	-1,32577	SSCG00000001327	-1,164889806	embryonal development	embryonal development
BTBD7	maker-scaffold00093-snap-gene-5.58	-1,30115	SSCG00000015617	-0,675918772	embryonal development	embryonal development
DKK2	maker-scaffold00034-augustus-gene-5.34	-1,28659	SSCG00000006209	2,706555135	embryonal development	embryonal development
FURIN	maker-scaffold00005-augustus-gene-17.32	-1,2747	SSCG00000013367	-0,831488069	other	membrane component
FAM213A	maker-scaffold00015-augustus-gene-15.56	-1,26659	SSCG00000019326	-1,054638155	stress	redox regulation
IGFBP6	maker-scaffold00001-augustus-gene-86.42	-1,2641	SSCG00000006889	2,071881685	in utero development	uterine function regulation, and inhibited labors
DNMT3b	maker-scaffold00120-snap-gene-2.53	-1,25611	SSCG00000009469	-0,940052021	unknown	unknown
CHST11	maker-scaffold00025-snap-gene-30.44	-1,25114	SSCG00000013736	-1,30119718	embryonal development	embryonal development
ZBTB11	genemark-scaffold00007-processed-gene-60.14	-1,22536	SSCG00000010753	-0,616092414	transcription	transcription
SPAG1	maker-scaffold00041-snap-gene-12.26	-1,2218	SSCG00000017172	-0,506689918	fertilization	fertilization
ALDH2	maker-scaffold00041-augustus-gene-16.41	-1,21399	SSCG00000017205	-0,946325591	hormone	response to steroid hormone
RAB11FIP2	maker-scaffold00069-augustus-gene-9.30	-1,20465	SSCG00000004610	-0,756704951	insulin	insulin pathway
ADM2	maker-scaffold00029-augustus-gene-25.43	-1,20322	SSCG00000007876	-1,230938333	in utero development	trophoblast invasion and migration, angiogenesis (cardiovascular)
PITPNA	maker-scaffold00009-snap-gene-28.27	-1,12423	SSCG00000014483	-0,589303601	transport	transport
MAP4K3	genemark-scaffold00056-processed-gene-9.15	-1,11779	SSCG00000007960	-0,49944484	stress	JUN pathway; stress
NSMAF	genemark-scaffold00018-processed-gene-11.5	-1,10047	SSCG00000014091	-1,21647643	apoptosis	apoptosis
GMDS	maker-scaffold00051-snap-gene-0.45	-1,10005	SSCG00000005931	-0,879010674	transport	iron-ion transport
CXCL9	maker-scaffold00023-augustus-gene-21.37	-1,08909	SSCG00000014646	0,89697591	immune	cytokine; immune response; activation of immune cells
HSH2D	snap_masked-scaffold00011-processed-gene-3.23	-1,0804	SSCG00000008128	-0,971320031	signaling	tyrosin kinase signaling
NPD21	maker-scaffold00022-augustus-gene-18.53	-1,0798	SSCG00000006332	-0,583638368	cell proliferation	neural cell proliferation
DAG1	maker-scaffold00001-augustus-gene-94.19	-1,07588	SSCG00000012110	-0,941749525	cell proliferation	cell survival
RAC1	maker-scaffold00015-augustus-gene-6.53	-1,0643	SSCG00000013113	-0,449310942	immune	immune; phagocytosis
DDX56	maker-scaffold00050-augustus-gene-14.91	-1,04373	SSCG00000019137	0,482487778	other	rRNA processing
THBS2	maker-scaffold00019-snap-gene-33.37	-1,03978	SSCG00000017775	-1,301736445	embryonal development	angiogenesis
OPTN	maker-scaffold00029-augustus-gene-25.39	-1,03675	SSCG00000009097	-0,626096359	apoptosis	TNF-alpha; apoptosis
GSR	maker-scaffold00012-snap-gene-9.46	-1,02414	SSCG00000020143	-0,771293376	stress	cell redox homeostasis
SCGN	maker-scaffold00026-augustus-gene-13.29	-1,01562	SSCG00000012784	-1,139278181	transport	calcium ion binding
unknown	maker-scaffold00035-augustus-gene-8.46	-1,01289	SSCG00000018949	0,785698511	other	

ZNF503	maker-scaffold00080-augustus-gene-1.23	-1,01219	SSCG00000018251	-1,196293656	embryonal development	embryonal development
B4GALNT1	maker-scaffold00057-augustus-gene-2.61	-1,00791	SSCG00000009983	0,646198776	spermatogenesis	spermatogenesis
RAB15	maker-scaffold00010-augustus-gene-40.35	-0,995149	SSCG00000001701	-0,585142975	cell proliferation	protein transport
GRINA	maker-scaffold00026-snap-gene-30.57	-0,967896	SSCG00000010850	-0,482856252	apoptosis	apoptosis
STAP2	maker-scaffold00011-augustus-gene-13.41	-0,9584	SSCG00000015086	-1,6695311	immune	inflammation; STAT 3 activity
unknown	maker-scaffold00005-augustus-gene-59.31	-0,951653	SSCG00000002767	-0,835423	other	unknown
MAL2	maker-scaffold00026-snap-gene-12.18	-0,948849	SSCG00000012789	-0,642548547	other	membrane component
NOB1	maker-scaffold00036-augustus-gene-0.61	-0,948241	SSCG00000018086	0,58421508	other	rRNA processing, visual perception
ZDHHC9	maker-scaffold00013-snap-gene-21.27	-0,947763	SSCG00000005638	-0,59721255	other	membrane component
TSG101	maker-scaffold00006-augustus-gene-25.21	-0,93006	SSCG00000011340	-0,480638358	cell proliferation	cell cycle/ cell division
PTK7	maker-scaffold00019-snap-gene-19.41	-0,905429	SSCG00000004758	-1,034793559	embryonal development	Wnt signaling pathway
IL6R	maker-scaffold00081-augustus-gene-3.57	-0,865669	SSCG00000010964	-0,454217454	immune	receptor for IL6; cytokine receptor; immune
ETV6	maker-scaffold00097-augustus-gene-5.67	-0,852208	SSCG00000006150	0,588871674	embryonal development	transcriptional repressor, hematopoiesis, neurogenesis, vitellogenesis
FAM57A	maker-scaffold00009-augustus-gene-30.35	-0,840976	SSCG00000014478	-0,742397283	unknown	unknown
ATP7A	maker-scaffold00012-snap-gene-28.61	-0,840934	SSCG00000000923	-0,59260857	transport	copper transport
CAP1	augustus_masked-scaffold00024-processed-gene-0.11	-0,823277	SSCG00000018678	-0,642434836	embryonal development	embryonal development
ANKRD13A	maker-scaffold00003-augustus-gene-11.40	-0,8216	SSCG00000005515	-0,919749535	other	plasma membrane
AP3M2	maker-scaffold00003-augustus-gene-55.25	-0,806796	SSCG00000002711	-0,594759908	transport	intracellular protein transport
GNPTAB	maker-scaffold00064-snap-gene-6.84	-0,788319	SSCG00000015315	4,431434331	transport	lysosomal enzyme transport
ARID5B	maker-scaffold00022-augustus-gene-0.45	0,87321	SSCG00000007995	0,985657959	embryonal development	liver development
PTGS1	maker-scaffold00022-augustus-gene-21.56	0,877511	SSCG00000006297	0,604683999	hormone	prostaglandin synthesis, inflammatory response, blood pressure, cell proliferation
SMARCD2	maker-scaffold00015-augustus-gene-35.59	0,88465	SSCG00000013629	-0,476408745	transcription	transcription regulation
CERS4	genemark-scaffold00037-processed-gene-21.1	0,890948	SSCG00000015805	0,858871956	other	sphingolipid production
FAM213A	maker-scaffold00015-augustus-gene-15.56	0,899201	SSCG00000019326	-1,054638155	metabolism	metabolic process
MAP3K5	genemark-scaffold00073-processed-gene-2.8	0,913159	SSCG00000020032	0,709011763	immune	innate immune: MAP kinase signaling
VIT	maker-scaffold00019-augustus-gene-26.35	0,968969	SSCG00000017805	-1,212819013	cell adhesion	promotes matrix assembly and cell adhesiveness - positive regulation of cell-substrate adhesion
ABI3BP	maker-scaffold00001-augustus-gene-72.38	0,969748	SSCG00000017701	-0,773286116	blood	collagen binding, herparin binding
PGR	genemark-scaffold00009-processed-gene-58.22	1,00038	SSCG00000018116	1,319065897	hormone	steroid hormones and receptors involved in eukaryotic gene exrpression. Cellular proliferation and differentiation.
LIMK2	maker-scaffold00003-augustus-gene-13.19	1,00501	SSCG00000007434	0,842158691	spermatogenesis	Spermatogenesis
DLX4A	maker-scaffold00047-augustus-gene-1.30	1,03152	SSCG00000008604	0,62294535	embryonal development	embryonal development

VEGFAA	maker-scaffold00060-snap-gene-4.40	1,03408	SSCG00000009790	0,826983399	embryonal development	angiogenesis; cell proliferation, apoptosis inhibition
FSTL3	maker-scaffold00011-augustus-gene-6.39	1,03595	SSCG00000008107	1,051651377	immune; osteoblast; spermatogensis	inhibits BMP2 induced cellular signaling, osteoblast differentiation, male gonad development, spermatogenesis
ADD3	maker-scaffold00033-augustus-gene-0.17	1,04518	SSCG00000003279	-0,641435421	cell adhesion	transmembrane protein, binds to calmodulin
PODN	genemark-scaffold00071-processed-gene-7.16	1,05692	SSCG00000007802	2,052556461	cell proliferation	cell proliferation
MARCKSL1	maker-scaffold00010-augustus-gene-10.30	1,05906	SSCG00000014536	0,682994875	cell proliferation	positive regulation of cell proliferation
GLUL	augustus_masked-scaffold00002-processed-gene-23.8	1,07209	SSCG00000010218	0,618107399	embryonal development	embryonal development
SOX9-B	genemark-scaffold00004-processed-gene-23.8	1,07967	SSCG00000020573	1,153511858	embryonal development	embryonal development: ear and neurones
TLR-1	maker-scaffold00058-augustus-gene-3.48	1,10002	SSCG00000002988	1,183578784	immune	innate immune response to microbial agents; regulation of cytokine secretion
CEBPB	snap_masked-scaffold00001-processed-gene-82.40	1,12416	SSCG00000003201	1,08772866	immune	immune and inflammatory responses; early embryogenesis;antiproliferative effect on T-cells by repressing MHC expression
APP	maker-scaffold00039-augustus-gene-13.22	1,14625	SSCG00000008907	-0,581190911	immune	antimicrobial humal imune response; innate immune response
C6	maker-scaffold00063-snap-gene-13.81	1,17047	SSCG00000016399	2,866862829	immune	complement system membrane attack complex
HEYL	maker-scaffold00024-augustus-gene-11.44	1,19457	SSCG00000011738	1,466891982	embryonal development	embryonal development ; Notch signaling
STK32B	maker-scaffold00008-augustus-gene-3.41	1,2481	SSCG00000007324	-1,561114243	signaling	ATP binding, intracellular signaling
MFAP2	maker-scaffold00020-augustus-gene-4.63	1,25488	SSCG00000000663	-2,017390244	embryonal development	embryo eye morphogenesis
RGCC	maker-scaffold00014-augustus-gene-47.21	1,25563	SSCG00000019682	0,871544096	cell proliferation	cell cycle
CFD	genemark-scaffold00124-processed-gene-1.4	1,28116	SSCG00000003998	0,901691371	immune: complement	complement system: alternative pathway
EPGN	maker-scaffold00027-augustus-gene-2.67	1,29584	SSCG00000003546	1,766589097	embryonal development	EPGF response
FAM53B	maker-scaffold00069-augustus-gene-2.35	1,30305	SSCG00000014325	1,013183687	unknown	unknown
FSTL3	maker-scaffold00011-augustus-gene-6.39	1,34566	SSCG00000008107	1,051651377	immune	immune; osteoblast; spermatogenesis
JUNB	augustus_masked-scaffold00053-processed-gene-9.6	1,41208	SSCG00000008537	1,594511043	hormone	hormone
SOCS3	maker-scaffold00004-augustus-gene-37.31	1,42842	SSCG00000014355	1,803038185	immune	immune defence; regulates onset and maintenance of allergic responses mediated by T-helper

						type 2 cells. regulates IL-6 signaling in vivo. : immune system
ADAMTS	maker-scaffold00059-augustus-gene-2.37	1,43043	SSCG00000011493	-1,017376544	immune	metalloproteinase
NET1	maker-scaffold00029-augustus-gene-26.43	1,4505	SSCG00000007875	1,54851508	apoptosis	apoptosis
PTX3	maker-scaffold00116-augustus-gene-2.31	1,48777	SSCG00000005081	-1,214684839	immune	innate immune system, resistance to pathogens, inflammatory reactions, clearence of self-components and female fertility
TF	maker-scaffold00096-augustus-gene-3.52	1,55518	SSCG00000001812	-0,904967042	transport	transport- iron
ID2	maker-scaffold00010-augustus-gene-46.30	1,61915	SSCG00000015149	1,381549857	immune	immune response (leukocyt differenciation, NK cell) & embryonal development
CA4	maker-scaffold00008-augustus-gene-25.70	1,61922	SSCG00000012480	1,084144244	transport	CO2 and Sodium transport
FOSL2	genemark-scaffold00010-processed-gene-6.6	1,67003	SSCG00000014514	1,654752711	bone	osteoblast survival and size. Activates transcription of LIF and CEBPB together with jun
HEPHL1	maker-scaffold00008-snap-gene-2.32	1,6723	SSCG00000007326	1,644280235	transport	copper transport
SEMASA	maker-scaffold00018-snap-gene-25.33	1,69227	SSCG00000001240	-0,916270573	apoptosis	increases cell proliferation and migration and inhibits apoptosis, which potentially promotes angiogenesis
JUNB	augustus_masked-scaffold00004-processed-gene-67.32	1,69312	SSCG00000020360	1,341703143	hormone	estrogen pathway, uterine response to estrogen, transcription, gene expression
CXCL12	maker-scaffold00023-augustus-gene-21.38	1,69662	SSCG00000014643	1,005308194	unknown	unknown
SOCS3	maker-scaffold00004-augustus-gene-37.31	1,70321	SSCG00000014355	1,803038185	immune	immune
COL18A1	snap_masked-scaffold00014-processed-gene-8.33	1,70373	SSCG00000009070	0,849437876	embryonal development	angiogenesis, visual perception
ID2	maker-scaffold00010-augustus-gene-46.30	1,7072	SSCG00000015149	1,381549857	immune	immune
JUNB	augustus_masked-scaffold00053-processed-gene-9.6	1,71243	SSCG00000008537	1,594511043	hormone	estrogen pathway, uterine response to estrogen, transcription, gene expression
ID1	maker-scaffold00062-augustus-gene-8.24	1,7301	SSCG00000007259	1,088895166	embryonal development	transcriptional regulator; angiogenesis; clock
JUNB	augustus_masked-scaffold00004-processed-gene-67.32	1,75392	SSCG00000020360	1,341703143	hormone	hormone
MYOC	maker-scaffold00011-snap-gene-21.51	1,75578	SSCG00000015977	3,629778253	bone	bone formation, promotes ostoeoblast differentiation
INHBB	maker-scaffold00020-augustus-gene-28.20	1,75731	SSCG00000014188	-0,654325889	hormone	activates follitropin secretion; gonadal hormone secretion; germ cell development and maturation, embryonal development, bone growth
ADM2	maker-scaffold00029-augustus-gene-25.43	1,89595	SSCG00000007876	-1,230938333	pregnancy	pregnancy
HEPHL1	maker-scaffold00008-snap-gene-2.32	2,00396	SSCG00000007326	1,644280235	transport	transport
SOST	maker-scaffold00015-augustus-gene-14.39	2,08699	SSCG00000019330	1,403085855	bone	bone formation

SOST	maker-scaffold00015-augustus-gene-14.39	2,16699	SSCG00000019330	1,403085855	embryonal development	bone formation
ALDH1A2	maker-scaffold00005-augustus-gene-21.23	2,18398	SSCG00000013402	-1,947742459	embryonal development	embryonal development
C6	maker-scaffold00063-snap-gene-13.81	2,30338	SSCG00000016399	2,866862829	immune	complement system
RASD2	augustus_masked-scaffold00045-processed-gene-6.2	2,52683	SSCG00000014747	2,460472039	locomotion	locomotion
TF	maker-scaffold00096-augustus-gene-3.52	2,55122	SSCG00000001812	-0,904967042	transport	iron ion transport
TLR2-1	maker-scaffold00058-augustus-gene-3.48	3,03175	SSCG00000002988	1,183578784	immune	immune
CLCF1	maker-scaffold00013-snap-gene-49.40	3,22815	SSCG00000005729	-1,110167551	immune	adaptive immune defence, B cell maturation
PRR33	maker-scaffold00017-snap-gene-40.105	3,27326	SSCG00000019602	1,038278575	unknown	unknown
SMAD9	maker-scaffold00008-augustus-gene-12.43	3,30442	SSCG00000002651	0,7253261	embryonal development	osteoblast differentiation
MYOC	maker-scaffold00011-snap-gene-21.51	3,49787	SSCG00000015977	3,629778253	embryonal development	bone formation
HCEA	maker-scaffold00102-augustus-gene-1.33	3,53311	SSCG00000005974	4,395061631	egg hatching	egg hatching
LS-12	maker-scaffold00024-augustus-gene-17.72	4,58777	SSCG00000018317	2,610053533	stress	stress regulation
APOA1	genemark-scaffold00009-processed-gene-42.17	4,64929	SSCG00000009542	5,314326365	hormone	adrenal gland development, steroid / hormone pathway
LS-12	maker-scaffold00024-augustus-gene-17.72	5,36141	SSCG00000018317	2,610053533	stress	stress
APOA1	genemark-scaffold00009-processed-gene-42.17	5,97501	SSCG00000009542	5,314326365	embryonal development	adrenal gland development

8. SI Appendix Data Set 1

Alignment files in FASTA format.
See SI Appendix data files at figshare: doi: 10.6084/m9.figshare.11499360

1hemoglobin_alpha.fas
2hemoglobin_beta.fas
3AID_clean.fas
4AID_clean.fas
5A_CD4clean.fas
5CD4_clean_conserved_block_no_frag.fas
6CD4_clean_conserved_block.fas
7CIITA_raw.fas
8CIITA_temp_section_with_ensemble.fas
9CIITA_temp_section.fas
10CD74_clean.fas
11CD74_exon_loss_determination_clean.fas
12AIRE_clean_long.fas
13AIRE_exon_overview.fas
14AIRE_redone_clean_no_Sp180_Sp663_NerOph.fas
15AIRE_redone_clean_with_Hip_syn_frag.fas
16AIRE_redone_clean_with_Hipcomfrag.fas
17AIRE_redone_clean.fas
18AIRE_redone_with_ensembl.fas
19AIRE_redone.fas
20B2M_clean_trimmed_nucl.fas
21B2M_cean.fas
22B2M_raw.fas
23CD8_clean_with_ensembl.fas
24CD8_clean.fas
25CD8_clean_compactsection.fas
26MHCI_exon2.fas
27MHCI_exon3.fas
28MHCI_exon4.fas
29MHCII_beta_complete.fas
30MHCII_beta_1-86aa.fas
31MHCII_beta_91-184aa.fas
32RAG1.fas
33RAG2.fas
34TAPs_clean.fas

9. SI Appendix Data Set 2

Details on the selected taxon set for phylogenetic analyses. Taxonomy follows ref[31] unless indicated otherwise.

Holostei (Infraclass)
Monophyly support: BS 100% (ref(27))
Possible sister groups: Teleostei (Infraclass)(27).
Diagnostic characters: 13 synapomorphies(171).
Earliest record: †*Acentrophorus varians* Kirkby, 1862 (ref(172)).
Formation: Raisby Formation at Fulwell Hill (Sunderland, UK)(173).
Age: Wuchiapingian (259.8-254.1 Ma).
Used as calibration: Yes.
Sampled species: *Lepisosteus oculatus*.

Teleostei (Infraclass)
Monophyly support: BS 100% (27, 174)
Possible sister groups: Holostei (Infraclass)(27).
Diagnostic characters: Eleven synapomorphies(175).
Earliest record: †*Malingichthys* spp. Tintori et al., 2015.
Formation: Zhuganpo Member of the Falang Formation at Xingyi City (Guizhou Province, China).
Age: Middle Late Ladinian (239.5-237.0 Ma)(176).
Used as calibration: Yes.
Sampled species: 70 species.

Elopocephalai (Megacohort)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Osteoglossomorpha (Supercohort), Clupeocephala (Supercohort)(27).
Diagnostic characters: Five synapomorphies of Elopomorpha, the only cohort of megacohort Elopocephalai(175)
Earliest record: †*Anaethalion zapporum* Arratia, 2000(177).
Formation: Rögling Formation at Schamhaupten (Eichstätt, Germany)(177).
Age: Latest Kimmeridgian (154.7-152.1 Ma)(177).
Used as calibration: Yes.
Sampled species: *Anguilla anguilla*, *Anguilla japonica*.

Osteoglossomorpha (Supercohort)
Monophyly support: BS 99%(27).
Possible sister groups: Elopocephalai (Megacohort), Clupeocephala (Supercohort)(27).
Diagnostic characters: Four synapomorphies(175).
Earliest record: †*Lycoptera* spp. (Sauvage, 1880).
Formation: Lower Yixian Formation at Jianshangou (China)(178).
Age: 128.4-121.0 Ma(179, 180).
Used as calibration: Yes.
Sampled species: *Scleropages formosus*.

Clupeocephala (Supercohort)
Monophyly support: BS 100% ref(27).

Possible sister groups: Elopocephalai (Megacohort), Osteoglossomorpha (Supercohort)(27).
Diagnostic characters: Ten synapomorphies(175, 181).
Earliest record: †*Leptolepides sprattiformis* Blainville, 1818 (ref(182)).
Formation: Lithographic limestone of Cerin (Ain, France).
Age: Kimmeridgian (157.3-152.1 Ma)(183).
Used as calibration: Yes.
Sampled species: 67 species.

Otomorpha (Cohort)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Euteleosteomorpha (Cohort)(27).
Diagnostic characters: Five synapomorphies(175).
Earliest record: †*Tischlingerichthys viohli* Arratia, 1997. Otomorpha combine subcohorts Clupei, Alepocephali, and Ostariophysi(27). No stem-group fossils are known for Otomorpha and both Clupei and Alepocephali appear in the fossil record later than Ostariophysi(177, 184). Thus, the earliest fossil record of Otomorpha is †*Tischlingerichthys viohli*, the oldest member of Ostariophysi(177).
Formation: Mörnsheim Formation at Mühlheim (Bavaria, Germany)(177).
Age: 152.1-150.9 Ma(177).
Used as calibration: Yes.
Sampled species: *Clupea harengus*, *Danio rerio*, *Astyanax mexicanus*, *Electrophorus electricus*.

Cypriniphysae (Superorder)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Characiphysae *sensu*(177). (Superorder).
Diagnostic characters: Nine synapomorphies(175).
Earliest record: cf. Cypriniformes indet(185).
Formation: Tremp Formation at Els Nerets (Spain)(185).
Age: C31r magnetochron (71.5-68.3 Ma)(186).
Used as calibration: Yes.
Sampled species: *Danio rerio*.

Characiphysae (Superorder) *sensu* Betancur-R. et al. 2013
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Cypriniphysae (Superorder)(27).
Diagnostic characters: 13 synapomorphies(175).
Earliest record: †*Santanichthys diasii* (Silva Santos, 1958). According to(27), Characiphysae combine the three orders Gymnotiformes, Characiformes, and Siluriformes(27). No stem-group fossils are known for Characiphysae and both Gymnotiformes and Siluriformes appear in the fossil record later than Characiformes. Thus, the earliest fossil record of Characiphysae is †*Santanichthys diasii*, the oldest member of Characiformes(187-189). Note that in the more recent classification of ref(27), Characiphysae include just the single order Characiformes; however, the monophyly of the clade combining Characiformes, Gymnotiformes, and Siluriformes is not debated.
Formation: Romualdo Member of the Santana Formation in the Araripe Basin (Brazil)(187).
Age: Late Aptian to Early Albian (123.0-110.8 Ma)(190).

Used as calibration: Yes.
Sampled species: *Astyanax mexicanus*, *Electrophorus electricus*.

Characiformes (Order)
Monophyly support: BS 100% (ref[25]) GGI 100% (ref(191)).
Possible sister groups: The clade combining Gymnotiformes (Order) and Siluriformes (Order)(191).
Diagnostic characters: Seven synapomorphies(175).
Earliest record: †*Santanichthys diasii* (Silva Santos, 1958).
Formation: Romualdo Member of the Santana Formation in the Araripe Basin (Brazil)(187).
Age: Late Aptian to Early Albian (123.0-110.8 Ma)(190).
Used as calibration: Yes.
Sampled species: *Astyanax mexicanus*.

Euteleosteomorpha (Cohort)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Otomorpha (Cohort)(27).
Diagnostic characters: Three synapomorphies(175).
Earliest record: †*Leptolepides sprattiformis* Blainville, 1818 (ref(182)).
Formation: Lithographic limestone of Cerin (Ain, France).
Age: Kimmeridgian (157.3-152.1 Ma)(183).
Used as calibration: Yes.
Sampled species: 63 species.

Esociformes (Order)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Salmoniformes (Order)(27). The monophyly of a clade combining the two orders Esociformes and Salmoniformes is strongly supported by molecular data(27),(192).
Diagnostic characters: Five synapomorphies(193).
Earliest record: †*Estesesox foxi*(194)
Formation: Milk River Formation (southern Alberta, Canada)(194).
Age: 84.5-83.5 Ma(195).
Used as calibration: Yes.
Sampled species: *Esox lucius*.

Salmoniformes (Order)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Esociformes (Order)(27). The monophyly of the two orders Esociformes and Salmoniformes, to the exclusion of the orders Galaxiiformes and Argentiniformes, is strongly supported by molecular data(27, 192).
Diagnostic characters: Six synapomorphies(175) given for "Salmonoidei".
Earliest record: †*Eosalmo driftwoodensis* Wilson, 1977.
Formation: Driftwood Creek Formation (British Columbia, Canada)(196).
Age: 52.1-51.4 Ma(197, 198).
Used as calibration: Yes.
Sampled species: *Salmo salar*.

Stomiatiformes (Order)

Monophyly support: BS 100% (ref(27)).
Possible sister groups: Osmeriformes (Order)(27).
Diagnostic characters: Ten synapomorphies(175).
Earliest record: †*Telepholis* von der Marck and Schlüter, 1868 (ref(199)).
Formation: Coesfeld Member of the deposits of Sendenhorst in the Münster Basin (Westphalia, Germany)(199).
Age: Campanian (83.6-72.1 Ma)(199).
Used as calibration: Yes.
Sampled species: *Borostomias antarcticus*.

Osmeriformes (Order)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Stomiatiformes (Order)(27).
Diagnostic characters: No synapomorphies are known for the phylogeny-based circumscription of Osmeriformes, combining Osmeridae, Plecoglossidae, Retropinnidae, and Salangidae(27, 31).
Earliest record: †*Enoplophthalmus schlumbergeri* Sauvage, 1880 (ref(200)).
Formation: Formation de Campagne-Calavon near Céreste (France)(200).
Age: Lower Oligocene (33.9-28.1 Ma)(200).
Used as calibration: No.
Sampled species: *Osmerus eperlanus*.

Neoteleostei (Subcohort)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Stomiatii (Subcohort), the clade combining Stomiatii (Subcohort) and Argentiniformes (Order)(192).
Diagnostic characters: No synapomorphies are known for the phylogeny-based circumscription of Neoteleostei, combining Ateleopodia and Eurypterygia(27).
Earliest record: †*Atolvorator longipectoralis* Gallo and Coelho, 2008 (see below, Eurypterygia).
Formation: Morro do Chaves Member of the Coqueiro Seco Formation in the Sergipe-Alagoas Basin (Brazil)(201).
Age: Barremian (129.4-125.0)(201).
Used as calibration: No.
Sampled species: 59 species.

Ateleopodia (Infracohort)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Eurypterygia (Infracohort)(27).
Diagnostic characters: Seven synapomorphies(175).
Earliest record: No skeletal fossils are known of Ateleopodia.
Formation: NA.
Age: NA.
Used as calibration: No.
Sampled species: *Guentherus altivela*.

Eurypterygia (Infracohort)
Monophyly support: BS 96% (ref(27)).
Possible sister groups: Ateleopodia (Infracohort)(27).
Diagnostic characters: Five synapomorphies(175).

Earliest record: †*Atolvorator longipectoralis* Gallo and Coelho, 2008. Eurypterygia combine the two sections Cyclosquamata and Ctenosquamata(27). No stem-group fossils are known for Eurypterygia and Ctenosquamata appear in the fossil record later than Cyclosquamata. Thus, the earliest fossil record of Eurypterygia is †*Atolvorator longipectoralis*, the oldest member of Cyclosquamata(184, 201).
Formation: Morro do Chaves Member of the Coqueiro Seco Formation in the Sergipe-Alagoas Basin (Brazil)(201).
Age: Barremian (129.4-125.0)(201).
Used as calibration: Yes.
Sampled species: 58 species.

Cyclosquamata (Section)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Ctenosquamata (Section)(27).
Diagnostic characters: Seven synapomorphies(175).
Earliest record: †*Atolvorator longipectoralis* Gallo and Coelho, 2008 (ref(184, 201)).
Formation: Morro do Chaves Member of the Coqueiro Seco Formation in the Sergipe-Alagoas Basin (Brazil)(201).
Age: Barremian (129.4-125.0)(201).
Used as calibration: Yes.
Sampled species: *Parasudis fraserbrunneri*, *Synodus synodus*.

Ctenosquamata (Section)
Monophyly support: BS 97% (ref(27)), BPP 1.0 (ref(30)).
Possible sister groups: Cyclosquamata (Section)(27).
Diagnostic characters: Five synapomorphies(175).
Earliest record: †*Xenyllion zonensis* Wilson and Murray, 1996, or †*Plesioberyx maximus* Gayet, 1980, †*Plesioberyx discoides* Gayet, 1980, †*Caproberyx pharsus* Patterson, 1967, †*Stichopteryx lewisi* Davis, 1887, †*Plectocretacicus clarae* Sorbini, 1979, †*Pharmacichthys venenifer* Smith Woodward, 1942, †*Pharmacichthys numismalis* Gayet, 1980a, *Aipichthys minor* (Pictet, 1850), and *Freigichthys elleipsis* Otero, 1997, or †*Pharmacichthys judensis* Gayet, 1980b, or †*Muhichthys cordobai* González-Rodríguez & Fielitz, 2008, †*Handuichthys interopercularis*, †*Pseudomonocentris microspinosus*, and †*Dalgoichthys tropicalis* González-Rodríguez et al. 2013. Ctenosquamata combine the two subsections Myctophata and Acanthomorphata(27). No stem-group fossils are known for Ctenosquamata and Myctophata appear in the fossil record later than Acanthomorphata. Thus, the earliest fossil record of Ctenosquamata is represented by the oldest members of Acanthomorphata.
Formation: Fish Scales Formation in north-western Alberta (Canada)(202), Lithographic Limestone of Haqil (Lebanon)(203, 204), Amminadav Formation at Ein Yabrud (Palestine)(205), and El Doctor Formation in Hidalgo (Mexico)(206, 207).
Age: 113.0-98.0 Ma. Both the Fish Scales Formation and the El Doctor Formation are Albian (113.0-100.5 Ma) to Cenomanian (100.5-93.9 Ma) in age(208) and co-occurrence with *Neogastroplites americanus* gives the minimum age of 97.53 Ma for †*Xenyllion zonensis*(209, 210). The occurrence of *Mantelliceras mantelli* in the Cenomanian Lithographic Limestone of Haqil(177) and the Amminadav Formation(211) indicates an age of at least 98.0 Ma (ref(177)).
Used as calibration: Yes.
Sampled species: 56 species.

Myctophata (Subsection)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Acanthomorphata (Subsection)(27).
Diagnostic characters: Seven synapomorphies(175).
Earliest record: †*Sardinioides* spp. van der Marck, 1858 (ref(184, 199)).
Formation: Coesfeld Member of the deposits of Sendenhorst in the Münster Basin (Westphalia, Germany)(199).
Age: Campanian (83.6-72.1 Ma)(199).
Used as calibration: Yes.
Sampled species: *Benthosema glaciale*.

Acanthomorphata (Subsection)
Monophyly support: BS 97% (ref(27)), BPP 1.0 (ref(30)).
Possible sister groups: Myctophata (Subsection)(27).
Diagnostic characters: Nine synapomorphies(175).
Earliest record: †*Xenyllion zonensis* Wilson and Murray, 1996, or †*Plesioberyx maximus* Gayet, 1980, †*Plesioberyx discoides* Gayet, 1980, †*Caproberyx pharsus* Patterson, 1967, †*Stichopteryx lewisi* Davis, 1887, †*Plectocretacicus clarae* Sorbini, 1979, †*Pharmacichthys venenifer* Smith Woodward, 1942, †*Pharmacichthys numismalis* Gayet, 1980a, †*Aipichthys minor* (Pictet, 1850), and †*Freigichthys elleipsis* Otero, 1997, or †*Pharmacichthys judensis* Gayet, 1980b, or †*Muhichthys cordobai* González-Rodríguez & Fielitz, 2008, †*Handuichthys interopercularis*, †*Pseudomonocentris microspinosus*, and †*Dalgoichthys tropicalis* González-Rodríguez et al. 2013.
Formation: Fish Scales Formation in north-western Alberta (Canada), Lithographic Limestone of Haqil (Lebanon)(203, 204), Amminadav Formation at Ein Yabrud (Palestine)(205), and El Doctor Formation in Hidalgo (Mexico)(206, 207).
Age: 113.0-98.0 Ma (see above, Ctenosquamata).
Used as calibration: Yes.
Sampled species: 55 species.

Polymixiipterygii (Division)
Monophyly support: BS 100% (ref[25]).
Possible sister groups: Euacanthomorphacea (Division), Lampripterygii (Division), Percopsaria (Series), Zeiogadaria (Series)(27, 30).
Diagnostic characters: Two synapomorphies(175).
Earliest record: †*Homonotichthys rotundus* Smith Woodward, 1902 (ref(29)).
Formation: English Chalk of Malling and Brighton (United Kingdom)(212).
Age: Middle to upper Cenomanian (96.5-93.9 Ma)(29). †*Homonotichthys rotundus* is recorded from the zone of †*Holaster subglobosus*, which can be constrained as middle to upper Cenomanian(211).
Used as calibration: Yes.
Sampled species: *Polymixia japonica*.

Percopsaria (Series)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Polymixiipterygii (Division), Zeiogadaria (Series)(27).
Diagnostic characters: Two synapomorphies(175).
Earliest record: †*Mcconichthys longipinnis* Grande, 1988 (ref(213, 214)).
Formation: Tullock Formation in eastern Montana(46).
Age: 66.0-65.0 (ref(215)).

Used as calibration: Yes.
Sampled species: *Percopsis transmontana*, *Typhlichthys subterraneus*.

Zeiogadaria (Series)
Monophyly support: BS 98% (ref[25]). BPP 1.0 (ref(30)).
Possible sister groups: Polymixiipterygii (Division), Percopsaria (Series)(27, 30).
Diagnostic characters: No synapomorphies are known for the phylogeny-based circumscription of Zeiogadaria, combining Zeariae and Gadariae(27).
Earliest record: †*Cretazeus rinaldii* Tyler et al., 2000 (ref(216)).
Formation: Calcare di Melissano at Cavetta near Nardò (Italy)(217).
Age: 89.8-82.0 Ma(218, 219).
Used as calibration: No.
Sampled species: *Cyttopsis rosea*, *Zeus faber*, *Stylephorus chordatus*, *Bathygadus melanobranchus*, *Gadus morhua*.

Zeariae (Subseries)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Gadariae (Subseries)(27, 30).
Diagnostic characters: Twelve synapomorphies(175).
Earliest record: †*Cretazeus rinaldii* Tyler et al., 2000 (ref(216)).
Formation: Calcare di Mellissano near Nardò (Italy)(217).
Age: 89.8-82.0 Ma(218, 219).
Used as calibration: Yes.
Sampled species: *Cyttopsis rosea*, *Zeus faber*.

Gadariae (Subseries)
Monophyly support: BS 100% (ref(27)), BPP 1.0 (ref(30))
Possible sister groups: Zeariae (Subseries)(27, 30).
Diagnostic characters: No synapomorphies are known for the phylogeny-based circumscription of Gadariae, combining Stylephoriformes and Gadiformes[25].
Earliest record: †"*Protocodus*" sp. (see below, Gadiformes).
Formation: Eqalulik Formation (Greenland)(30, 220).
Age: 62.8-59.7 Ma(30, 221)
Used as calibration: No.
Sampled species: *Stylephorus chordatus*, *Bathygadus melanobranchus*, *Gadus morhua*.

Stylephoriformes (Order)
Monophyly support: Monotypic.
Possible sister groups: Gadiformes (Order)(27, 30).
Diagnostic characters: All diagnostic characters of *Stylephorus chordatus*.
Earliest record: No fossils are known of Stylephoriformes.
Formation: NA.
Age: NA.
Used as calibration: No.
Sampled species: *Stylephorus chordatus*.

Gadiformes (Order)
Monophyly support: BS 100% (ref(27, 30)).
Possible sister groups: Stylephoriformes (Order)(27, 30).

Diagnostic characters: Five synapomorphies(175).
Earliest record: †"*Protocodus*" sp. While a formal description has never been published, a nearly complete specimen assigned to †"*Protocodus*" at the Mineralogical and Geological Museum of the University of Copenhagen has often been considered the earliest skeletal fossil record of Gadiformes(220, 222-226).
Formation: Eqalulik Formation (Greenland)(30, 220).
Age: 62.8-59.7 Ma(30, 221).
Used as calibration: Yes.
Sampled species: *Bathygadus melanobranchus*, *Gadus morhus*.

Lampripterygii (Division)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Euacanthomorphacea (Division), Polymixiacea (Division), Percopsaria (Series), Zeiogadaria (Series)(27, 30).
Diagnostic characters: Five synapomorhpies(175).
Earliest record: †*Pharmacichthys venenifer* Smith Woodward, 1942, †*Pharmacichthys numismalis* Gayet, 1980a, †*Pharmacichthys judensis* Gayet, 1980b, †*Aipichthys minor* (Pictet, 1850), †*Freigichthys elleipsis* Otero, 1997. Specimens of these species are recorded from the deposits of Haqil and Ein Yabrud, which predate other occurrences of Lampripterygii(205).
Formation: Lithographic Limestone of Haqil (Lebanon), and Amminadav Formation at Ein Yabrud (Palestine)(204, 205).
Age: 100.5-98.0 Ma.
Used as calibration: Yes.
Sampled species: *Regalecus glesne*, *Lampris guttatus*.

Acanthopterygii (Division)
Monophyly support: BS 99% (ref(27)), BPP 1.0 (ref(30)).
Possible sister groups: Lampripterygii (Division), Polymixiacea (Division), Percopsaria (Series), Zeiogadaria (Series)(27, 30).
Diagnostic characters: Three synapomorphies(227).
Earliest record: †*Plesioberyx maximus* Gayet, 1980, †*Plesioberyx discoides* Gayet, 1980, †*Caproberyx pharsus* Patterson, 1967, †*Stichopteryx lewisi* Davis, 1887, and †*Plectocretacicus clarae* Sorbini, 1979 (ref(29)), or †*Handuichthys interopercularis* and †*Pseudomonocentris microspinosus*(206).
Formation: Lithographic Limestone of Haqil (Lebanon)(203, 204, 228) and El Doctor Formation in Hidalgo (Mexico)(206, 207).
Age: 113.0-98.0 Ma (see above, Ctenosquamata).
Used as calibration: Yes.
Sampled species: 45 species.

Trachichthyiformes (Order)
Monophyly support: BS 100% (ref(27, 229)), BPP 1.0 (ref(24)).
Possible sister groups: Holocentrimorphaceae (Subdivision), Beryciformes (Order), Percomorphaceae (Subdivision)(27).
Diagnostic characters: Three synapmorphies(230).
Earliest record: †*Stichopteryx lewisi* Davis, 1887 (ref(231)).
Formation: Lithographic Limestone of Haqil (Lebanon)(203, 232).
Age: 100.5-98.0 Ma(177).
Used as calibration: Yes.

Sampled species: *Diretmoides pauciradiatus*, *Gephyroberyx darwinii*.

Beryciformes (Order)
Monophyly support: The monophyly of a clade combining Barbourisiidae, Berycidae, Cetomimidae, Megalomycteridae, Melamphaidae, Mirapinnidae, and Rondeletiidae is supported by BS 84 (ref(27)) and corroborated by BPP 1.0 (ref(24)) for the monophyly of the Berycidae, Cetomimidae, Rondeletiidae (other members of the clade were not included). The addition of Stephanoberycidae to this clade is supported by BPP 1.0 (ref(24, 30)).
The inclusion of Gibberichthyidae, Megalomycteridae, Mirapinnidae and Hispidoberycidae in this clade is assumed based on the taxonomic placement of Gibberichthyidae, Megalomycteridae, and Mirapinnidae within superfamily Cetomimoidea(233) and the placement of Hispidoberycidae in superfamily Stephanoberycoidea(233).
Possible sister groups: Trachichthyiformes (Order), Holocentrimorphaceae (Subdivision), Percomorphaceae (Subdivision)(27).
Diagnostic characters: No synapomorphies are known for the phylogeny-based circumscription of Beryciformes.
Earliest record: †*Berycomorus firdoussii* Arambourg, 1967 (ref(234)).
Formation: Elam Formation (Iran)(235).
Age: Rupelian (33.9-28.1 Ma).
Used as calibration: No.
Sampled species: *Acanthochaenus luetkenii*, *Beryx splendens*.

Holocentrimorphaceae (Subdivision)
Monophyly support: BS 100% (27), BPP 1.0 (ref(24)).
Possible sister groups: Trachichthyiformes (Order), Beryciformes (Order), Percomorphaceae (Subdivision)(27).
Diagnostic characters: Nine synapomorphies(236).
Earliest record: †*Caproberyx pharsus* Patterson, 1967 (ref(29)).
Formation: Lithographic Limestone of Haqil (Lebanon)(203).
Age: 100.5-98.0 Ma(177).
Used as calibration: Yes.
Sampled species: *Holocentrus rufus*, *Myripristis jacobus*.

Percomorphaceae (Subdivision)
Monophyly support: BS 99% (ref(27)), BPP 1.0.
Possible sister groups: Trachichthyiformes (Order), Beryciformes, Holocentrimorphaceae (Subdivision)(27, 30) (24).
Diagnostic characters: Eight synapomorphies(175) given for "Percomorphacea".
Earliest record: †*Plectocretacicus clarae* Sorbini, 1979, or †*Cretatriacanthus guidottii* Tyler and Sorbini, 1996. †*Plectocretacicus clarae* has often been considered to be the earliest record of Tetraodontiformes; however, evidence for the assignment of †*Plectocretacicus clarae* to Tetraodontiformes was based on an assumed close relationship between Tetraodontiformes and Zeiformes(237), which is now known to be incorrect. Similarities between †*Plectocretacicus clarae* and recently discovered Cretaceous "monocentrid-like" acanthomorphs of Mexico(177, 206) suggest that †*Plectocretacicus clarae* may not even be a member of Percomorphaceae. In this case, †*Cretatriacanthus guidottii* would be the oldest reliable record of Percomorphaceae (Santini, priv. comm.).

Formation: Lithographic Limestone of Haqil (Lebanon)(228) and Calcare di Melissano at Canale near Nardò (Italy)(177).
Age: 100.5-98.0 Ma(177) or 89.8-82.0 Ma(218, 219).
Used as calibration: No. Due to the uncertain earliest record of Percomorphaceae, neither †*Plectocretacicus clarae* nor †*Cretatriacanthus guidottii* are used to calibrate the age of Percomorphaceae.
Sampled species: 39 species.

Ophidiaria (Series)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Batrachoidaria (Series), Gobiaria (Series), Scombrimorpharia *sensu* Betancur-R. et al. 2013, Eupercaria (Series), Carangimorpharia sensu Betancur-R. et al. 2013 (ref(27)).
Diagnostic characters: Two probable synapomorphies(238) (239).
Earliest record: †*Pastorius methenyi* Carnevale and Johnson, 2015.
Formation: Liburnica Formation near Trebiciano (Italy)(239).
Age: Late Campanian to early Maastrichtian (76.4-69.8 Ma)(239).
Used as calibration: No.
Sampled species: *Brotula barbata*, *Carapus acus*.

Batrachoidaria (Series)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Ophidiaria (Series), Gobiaria (Series), Scombrimorpharia *sensu* Betancur-R. et al. 2013, Eupercaria (Series), Carangimorpharia *sensu* Betancur-R. et al. 2013 (ref(27)).
Diagnostic characters: Nine synapomorphies(175).
Earliest record: †*Zappaichthys harzhauseri* Carnevale and Collette, 2014.
Formation: Leitha Limestone of St. Margarethen in Burgenland (Austria)(240).
Age: Late Badenian (13.8-12.8 Ma)(240, 241).
Used as calibration: Yes.
Sampled species: *Chatrabus melanurus*, *Opsanus beta*.

Gobiaria (Series)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Ophidiaria (Series), Batrachoidaria (Series), Scombrimorpharia *sensu* Betancur-R. et al. 2013, Eupercaria (Series), Carangimorpharia *sensu* Betancur-R. et al. 2013 (ref(27)).
Diagnostic characters: One synapomorphy(27, 242).
Earliest record: †*Carlomonnius quasigobius* Bannikov and Carnevale, 2016, †*Eosphaeramia margaritae* Sorbini, 1983, †*Eoapogon fraseri*, †*Bolcapogon johnsoni*, and †*Apogoniscus pauciradiatus* Bannikov, 2005. Gobiaria combine the two orders Kurtiformes and Gobiiformes. No stem-group fossils are known for Gobiaria. Kurtiformes and Gobiformes appear simultaneously in the fossil record.
Formation: "Calcari nummulitici" at Monte Bolca near Verona (Italy)(243).
Age: 49.1-48.9. The age of the Monte Bolca deposits can be constrained to the overlap of nannoplankton zone NP 14 and and shallow benthic zone SBZ 11.
Used as calibration: Yes.
Sampled species: *Lesueurigobius* cf. *sanzi*, *Periophthalmodon schlosseri*.

Scombrimorpharia *sensu* Betancur-R. et al. 2013

Monophyly support: BS 97% (ref(27)), BP 1.0 (ref(24)).
Possible sister groups: Ophidiaria (Series), Batrachoidaria (Series), Gobiaria (Series), Eupercaria (Series), Carangimorpharia *sensu* Betancur-R. et al. 2013 (ref(27)).
Diagnostic characters: No synapomorphies are known for the phylogeny-based circumscription of Scombrimorpharia *sensu* Betancur-R. et al. 2013, combining Scombriformes and Syngnathiformes(27).
Earliest record: †*Gasterorhamphosus zuppichinii* Sorbini, 1981. No stem-group fossils are known for Scombrimorpharia *sensu* Betancur-R. et al. 2013 and Scombriformes appear in the fossil record later than Syngnathiformes(244). Thus, the earliest fossil record of Scombrimorpharia *sensu* Betancur-R. et al. 2013 is †*Gasterorhamphosus zuppichinii*, the oldest member of Syngnathiformes(244).
Formation: Calcare di Melissano at Cavetta near Nardò (Italy)(219).
Age: 89.8-82.0 Ma(218, 219).
Used as calibration: No.
Sampled species: 16 species.

Pelagiaria (Series)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Syngnatharia (Series)(27).
Diagnostic characters: Six synapomorphies(175).
Earliest record: †*Eutrichiurides opiensis* (Leriche 1906) Casier, 1944, †*Sphyraenodus multidentatus* Dartevelle and Casier, 1959, †*Ardiodus mariotti* White, 1931 [a detailed discussion of the earliest scombriform records is given in the SI Appendix Material of (ref(29)).
Formation: "Montian Phosphates, Morocco"(245), Landana Cliffs of the Cabinda enclave (Angola), and Oldhaven Beds near Upnor in Kent (United Kingdom)(29).
Age: 66.0-56.5 Ma(29).
Used as calibration: Yes.
Sampled species: *Thunnus albacares*, *Thunnus orientalis*.

Syngnatharia (Series)
Monophyly support: BS 97% (ref(27)).
Possible sister groups: Pelagiaria (Series)(27).
Diagnostic characters: No synapomorphies are known for the phylogeny-based circumscription of Syngnathiformes, combining Aulostomidae, Callionymidae, Centriscidae, Creediidae, Dactylopteridae, Draconettidae, Fistulariidae, Mullidae, Leptoscopidae, Pegasidae, Solenostomidae, and Syngnathidae(27, 246).
Earliest record: †*Gasterorhamphosus zuppichinii* Sorbini, 1981.
Formation: Calcare di Melissano at Cavetta near Nardò (Italy)(219).
Age: 89.8-82.0 Ma(218, 219).
Used as calibration: No.
Sampled species: *Dactylopterus volitans*, *Mullus surmuletus*, *Aeoliscus strigatus*, *Macroramphosus scolopax*, *Fistularia tabacaria*, *Doryrhamphus dactyliophorus*, *Entelurus aequoreus*, *Nerophis ophidion*, *Hippocampus comes*, *Hippocampus cuda*, *Hippocampus whitei*, *Syngnathus scovelli*, *Syngnathus rostellatus*, *Syngnathus typhle*.

Syngnathinae (Subfamily) *sensu* Hamilton et al. (2017)
Monophyly support: BS 95%, BPP 1.0 (ref(35)).
Possible sister groups: Nerophinae (Subfamily)(35).
Diagnostic characters: Tail-bearing male brood pouch(247).

Earliest record: †*Syngnathus incompletus* Cosmovici, 1887 (ref(247)).
Formation: IPM2 zone of the Menilite Formation in the outer Carpathian basin (Poland).
Age: 32.9-29.7 Ma. The IPM2 zone of the Menilite Formation corresponds to nannoplankton zones NP 22 to NP 23 (ref(248)).
Used as calibration: Yes.
Sampled species: *Hippocampus comes*, *Hippocampus cuda*, *Hippocampus whitei*, *Syngnathus scovelli*, *Syngnathus rostellatus*, *Syngnathus typhle*.

Nerophinae (Subfamily) *sensu* Hamilton et al. (2017)
Monophyly support: BS 99%, BPP 1.0 (ref(35)).
Possible sister groups: Syngnathinae (Subfamily)(35).
Diagnostic characters: Trunk-bearing male brood pouch(247).
Earliest record: †*Hipposyngnathus neriticus* Jerzmańska, 1968 (ref(247)).
Formation: Jamna Dolma Member of the Menilite Formation in the outer Carpathian basin (Poland).
Age: 32.9-32.0 Ma. The Jalmna Dolma Member of the Menilite Formation is synchronous with nannoplankton zone NP 22 (ref(248)).
Used as calibration: Yes.
Sampled species: *Doryrhamphus dactyliophorus*, *Entelurus aequoreus*, *Nerophis ophidion*.

Eupercaria (Series)
Monophyly support: BS 99% (ref(27)).
Possible sister groups: Ophidiaria (Series), Batrachoidaria (Series), Gobiaria (Series), Scombrimorpharia *sensu* Betancur-R. et al. 2013, Carangimorpharia *sensu* Betancur-R. et al. 2013 (ref(27)).
Diagnostic characters: No synapomorphies are known for the phylogeny-based circumscription of Eupercaria, combining Acanthuriformes, Centrarchiformes, Cirrhitiformes, Ephippiformes, Labriformes, Lophiiformes, Perciformes, Pempheriformes, Spariformes, Tetraodontiformes, and Uranoscopiformes, as well as 39 *incertae sedis* families(27).
Earliest record: †*Plectocretacicus clarae* Sorbini, 1979, or †*Cretatriacanthus guidottii* Tyler and Sorbini, 1996 (see above, Percomorphaceae).
Formation: Lithographic Limestone of Haqil (Lebanon)(228) and Calcare di Melissano at Canale near Nardò (Italy)(177).
Age: 100.5-98.0 Ma(177) or 89.8-82.0 Ma(218, 219).
Used as calibration: No.
Sampled species: *Epinephelus aeneus*, *Scorpaena* sp., *Gasterosteus aculeatus*, *Symphodus melops*, *Takifugu rubripes*, *Dichotomyctere nigroviridis*.

Anabantaria (Series)
Monophyly support: BS 94% (ref(27)).
Possible sister groups: Carangaria (Series), Ovalentaria (Series)(27, 30).
Diagnostic characters: No synapomorphies are known for the phylogeny-based circumscription of Anabantomorphariae, combining Synbranchiformes and Anabantiformes(27).
Earliest record: †*Eochanna chorlakkiensis* Roe, 1991 (see below, Anabantiformes).
Formation: Kuldana Formation near Chorlakki in the Kohat District (Pakistan).
Age: 48.8-43.6 Ma.
Used as calibration: No.
Sampled species: *Anabas testudineus*, *Monopterus albus*.

Anabantiformes (Order)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Synbranchiformes (Order)(27).
Diagnostic characters: Seven synapomorphies(175).
Earliest record: †*Eochanna chorlakkiensis* Roe, 1991 (ref(182, 249)).
Formation: Kuldana Formation near Chorlakki in the Kohat District (Pakistan).
Age: 48.8-43.6 Ma. The Kuldana Formation can be constrained to shallow benthic zones SB 12 to SB 13 (ref(250)).
Used as calibration: Yes.
Sampled species: *Anabas testudineus*.

Synbranchiformes (Order)
Monophyly support: BS 98% (ref(27)).
Possible sister groups: Anabantiformes (Order)(27).
Diagnostic characters: No synapomorphies are known for the phylogeny-based circumscription of Synbranchiformes, combining Chaudhuriidae, Indostomidae, Mastacembelidae, and Synbranchidae(27).
Earliest record: No fossils are known of Synbranchiformes.
Formation: NA.
Age: NA.
Used as calibration: No.
Sampled species: *Monopterus albus*.

Carangaria (Series)
Monophyly support: BS 100% (ref(27)).
Possible sister groups: Anabantaria (Series), Ovalentaria (Series)(27).
Diagnostic characters: No synapomorphies are known for the phylogeny-based circumscription of Carangaria, combining Istiophoriformes, Carangiformes, and Pleuronectiformes(27).
Earliest record: †*Hemingwaya sarissa* Sytchevskaya and Prokofiev, 2002 (ref(251)).
Formation: Danata Formation near Uylya-Kushlyuk (Turkmenistan).
Age: Thanetian (59.2-56.0 Ma)(252)
Used as calibration: No.
Sampled species: *Caranx melampygus*, *Cynoglossus semilaevis*.

Ovalentaria (Series)
Monophyly support: BS 100% (ref(27, 253)).
Possible sister groups: Anabantaria (Series), Carangaria (Series)(27).
Diagnostic characters: Demersal, adhesive eggs with chorionic filaments(253).
Earliest record: †*Palaeopomacentrus orphae* Bellwood and Sorbini, 1996, †*Lorenzichthys olihan* Bellwood, 1999, †?*Oncolepis isseli* Bassani, 1898, and †*Rhamphexocoetus volans* Bannikov et al., 1985. Older records based on otoliths and "Poeciliidae indet." are not considered here [see discussion in the SI Appendix Material of (29).
Formation: "Calcari nummulitici" at Monte Bolca near Verona (Italy)(243).
Age: 49.1-48.9.
Used as calibration: No. As the single known synapomorphy of Ovalentaria would likely not be recognized in the fossil record, direct assignment of fossils to this clade would probably not be possible.
Sampled species: *Chromis chromis*, *Pseudochromis fuscus*, *Amphilophus citrinellus*,

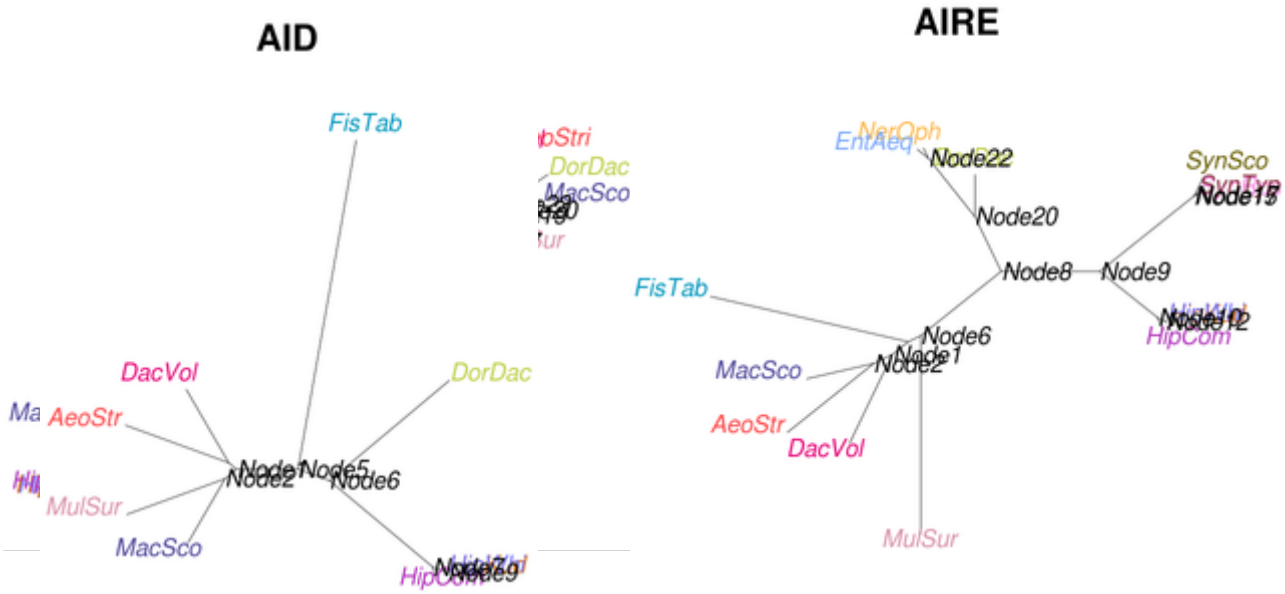
Oreochromis niloticus, *Oryzias latipes*, *Poecilia formosa*, *Xiphophorus maculatus*.

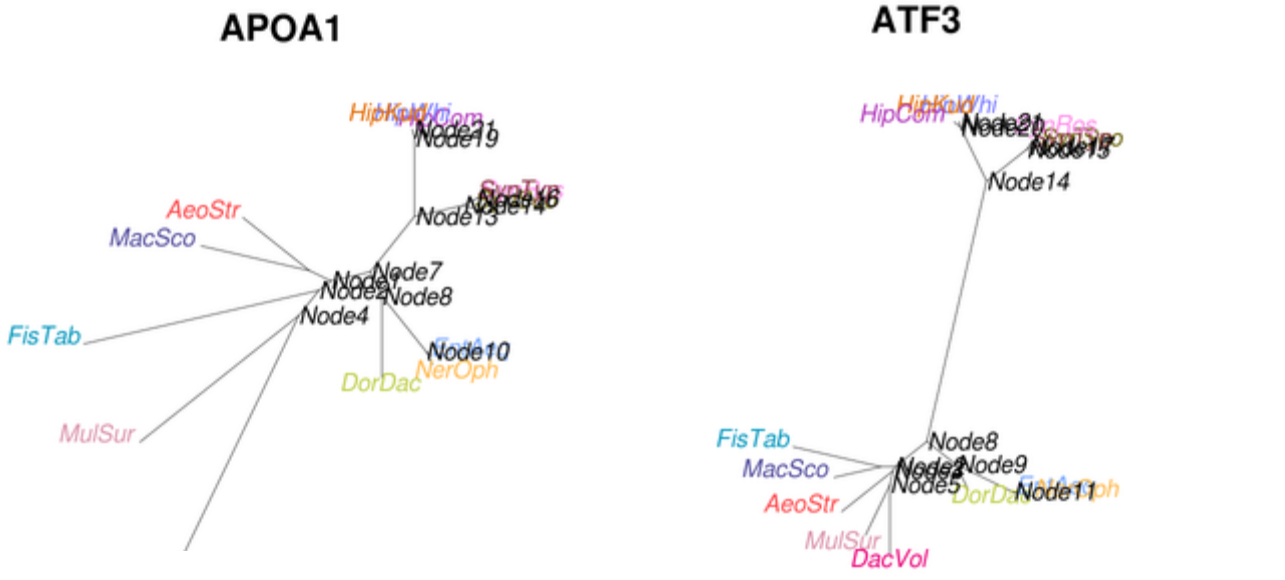
Pseudocrenilabrinae (Subfamily)
Monophyly support: BS 100% (ref(29)).
Possible sister groups: Cichlinae (Subfamily)(29).
Diagnostic characters: No synapomorphies are known for Pseudocrenilabrinae including *Heterochromis*(254). However, as all members of Pseudocrenilabrinae are geographically separated from those of other cichlid subfamilies, fossils can be assigned directly to Pseudocrenilabrinae, not only through synapomorphies with subclades.
Earliest record: †*Mahengechromis* spp. Murray, 2000 (ref(255)).
Formation: Mahenge paleolake in the Singida District (Tanzania)(255).
Age: 46.0-45.0 Ma(256).
Used as calibration: Yes.
Sampled species: *Oreochromis niloticus*.

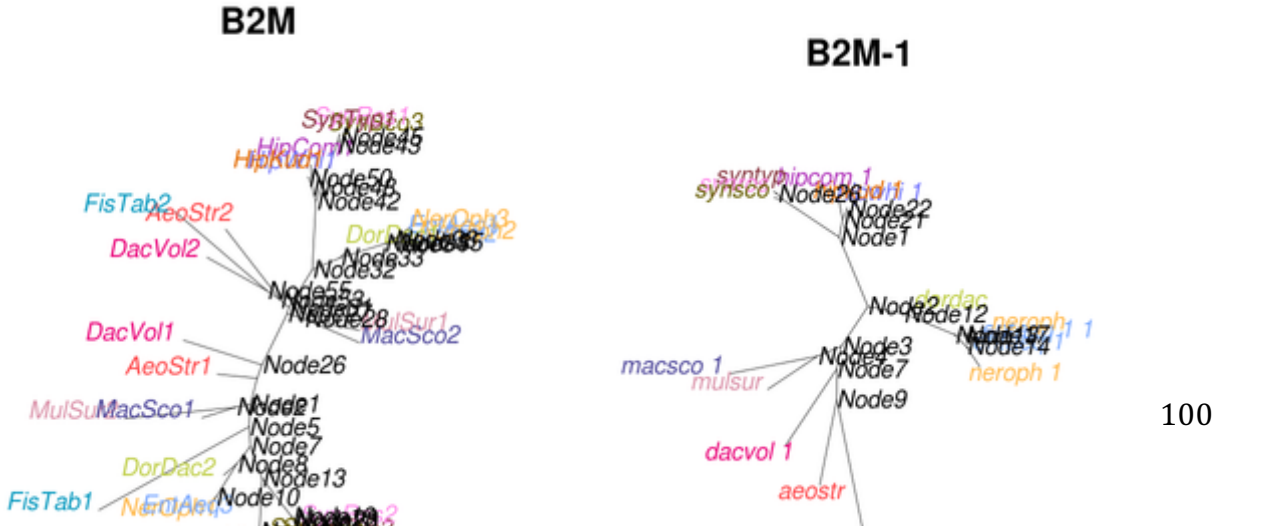
Cichlinae (Subfamily)
Monophyly support: BS 98% (ref(29)).
Possible sister groups: Pseudocrenilabrinae (Subfamily)(29).
Diagnostic characters: No synapomorphies are known for Cichlinae excluding *Heterochromis* (254). However, as all members of Cichlinae are geographically separated from those of other cichlid subfamilies, fossils can be assigned directly to Cichlinae, not only through synapomorphies with subclades.
Earliest record: *Gymnogeophagus* †*eocenicus* Malabarba, Malabarba, and del Papa, 2010, †*Plesioheros chauliodus* Alano Perez, Malabarba, and del Papa, 2010, and †*Proterocara argentina* Malabarba, Zuleta, and del Papa, 2006 (ref(257-259)).
Formation: Lumbrera Formation in the Salta Province (Argentina).
Age: 45.4-39.9 Ma(29).
Used as calibration: Yes.
Sampled species: *Amphilophus citrinellus*.

10. SI Appendix Data Set 3

Gene trees used for the aBSREL test for positive selection. Node labels refer to those listed in SI Appendix Table S11.



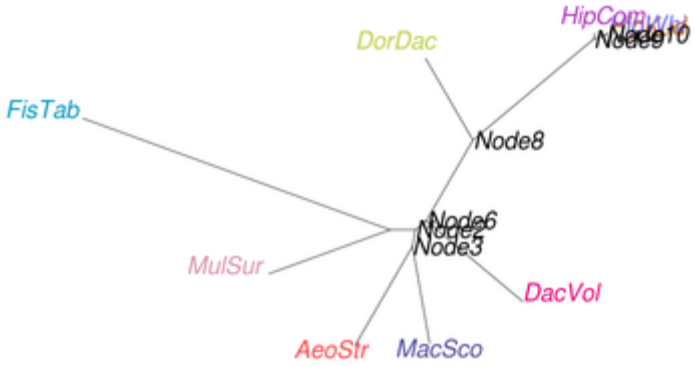




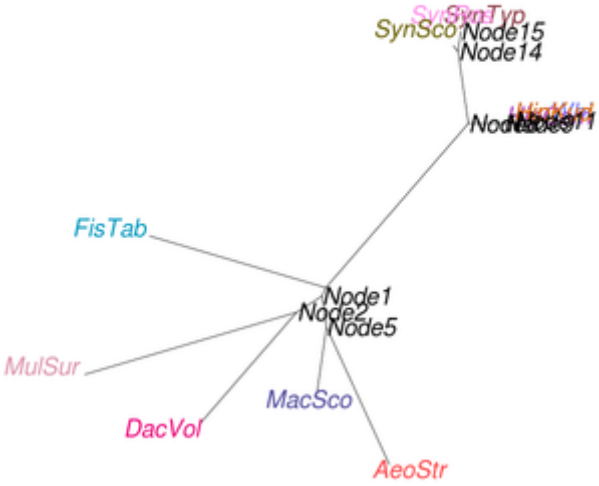
B2M-2

Ca4

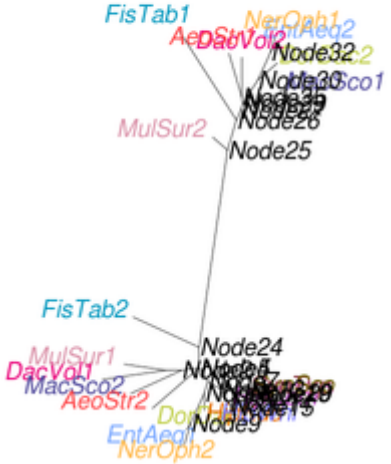
CD4



CD74

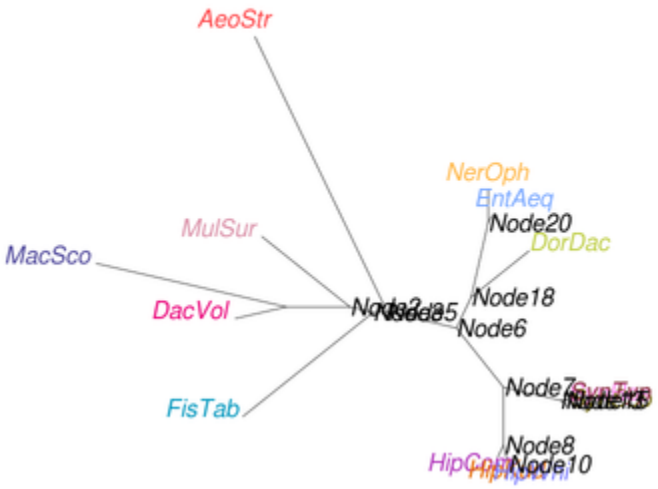


CD8

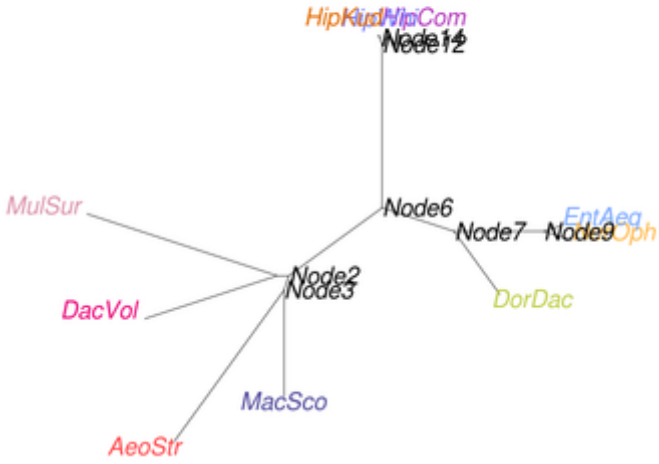


CD8-alpha

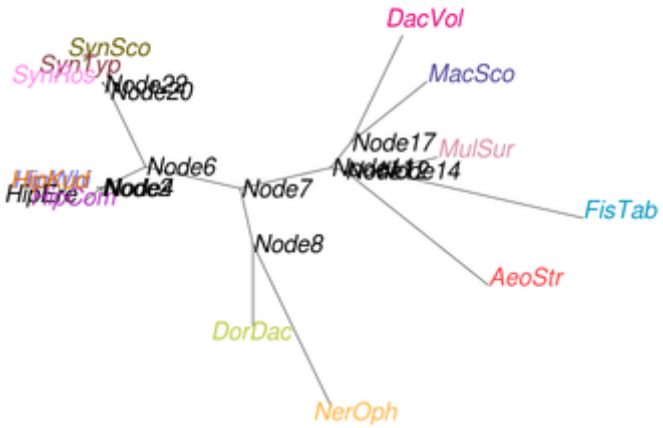
CEBPB



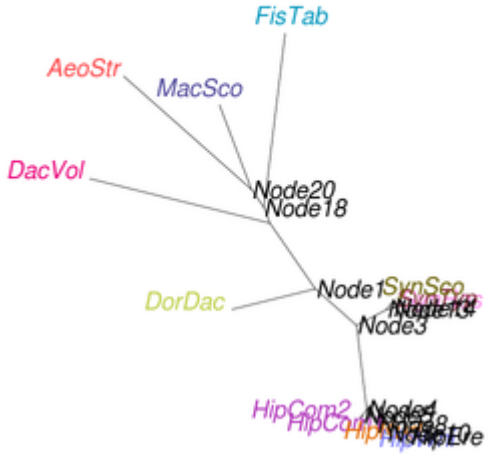
CIITA



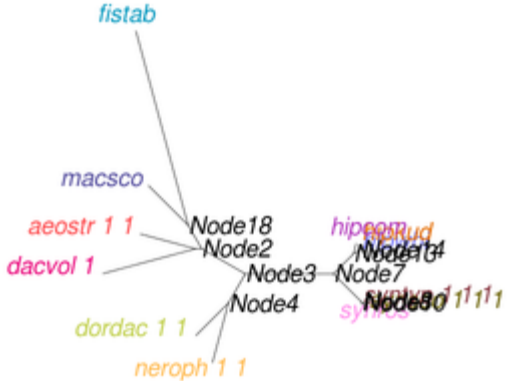
DLX3b-cluster1



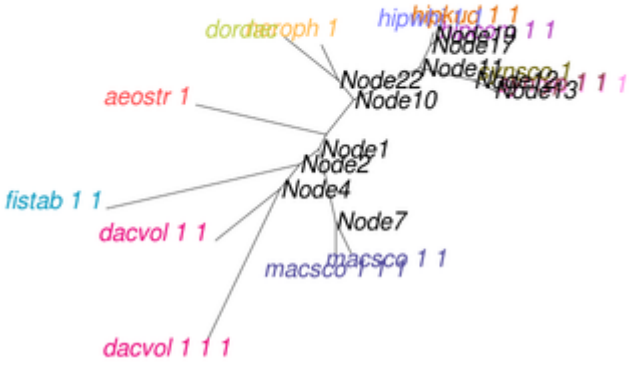
DLX3b-cluster2



hemoglobin-Alpha-1

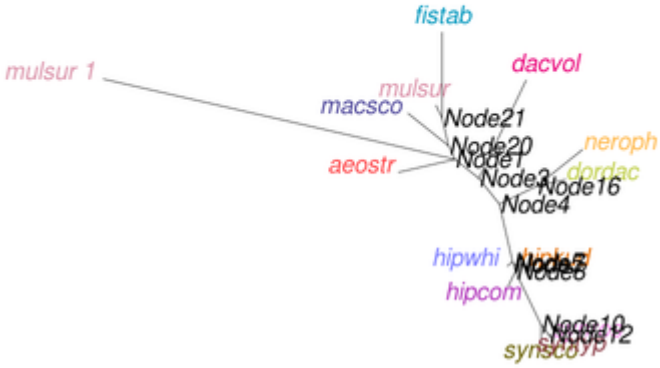


hemoglobin-Alpha-2

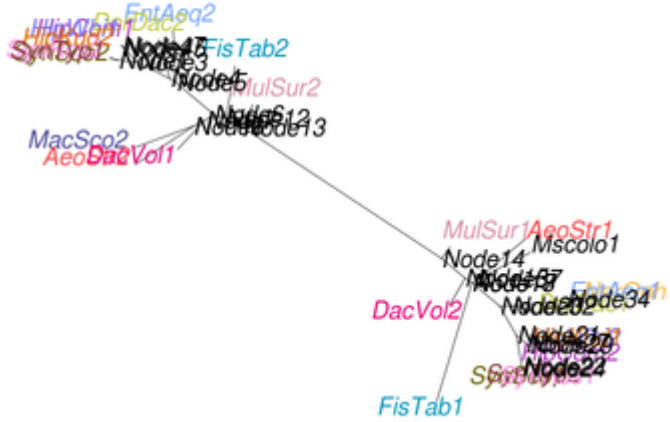


hemoglobin-Alpha-3

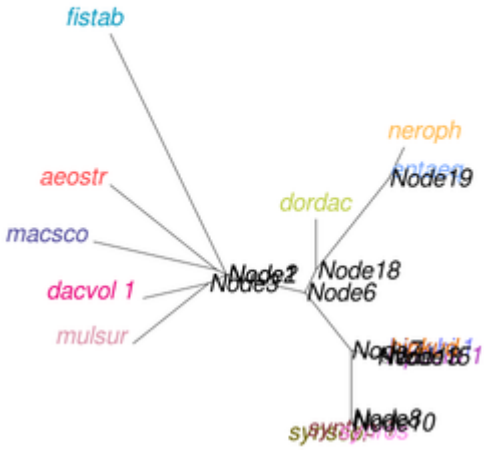
hemoglobin-Beta



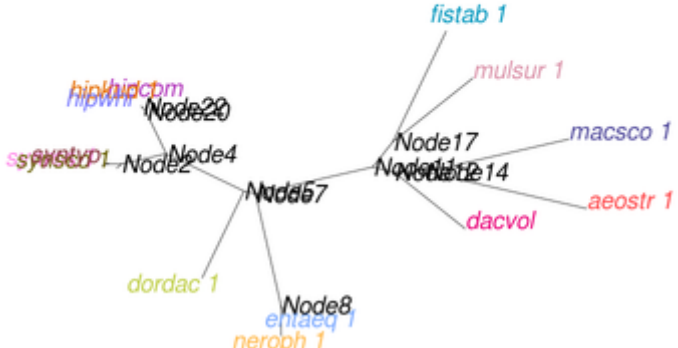
JUNB



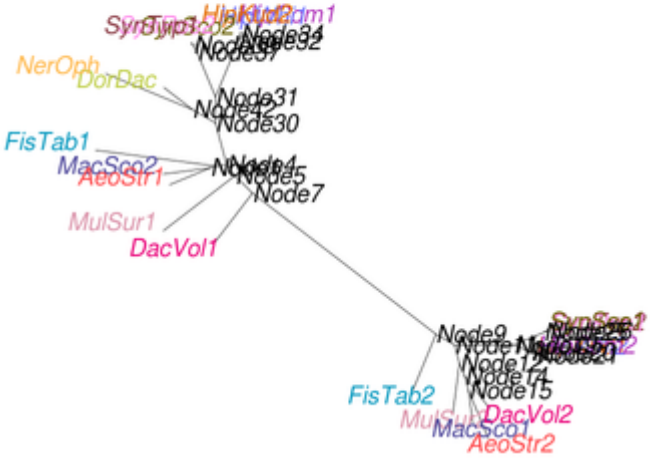
JUNB-1



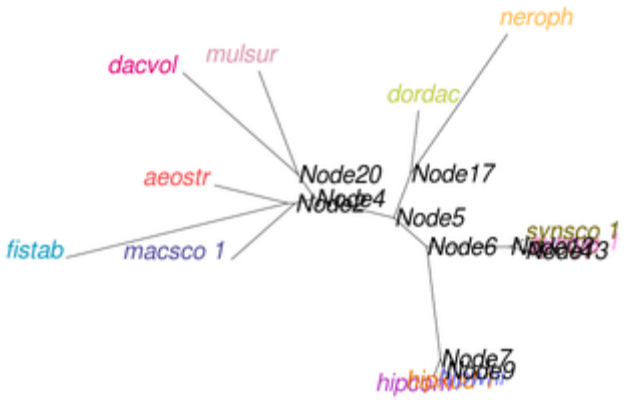
JUNB-2



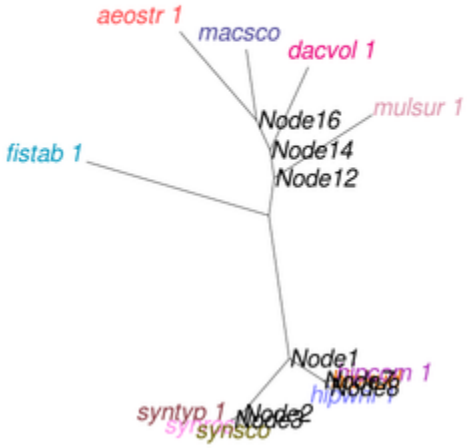
PRKCD



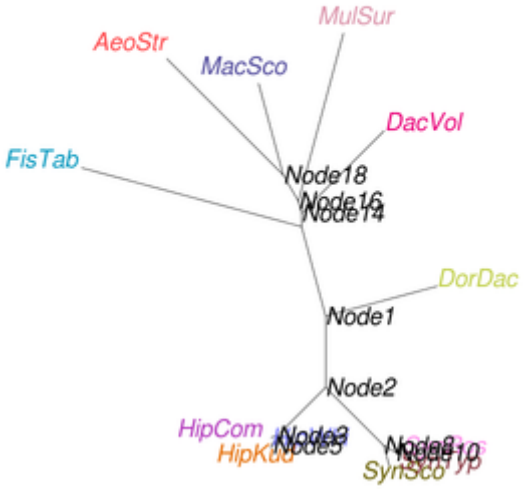
PRKCD-1



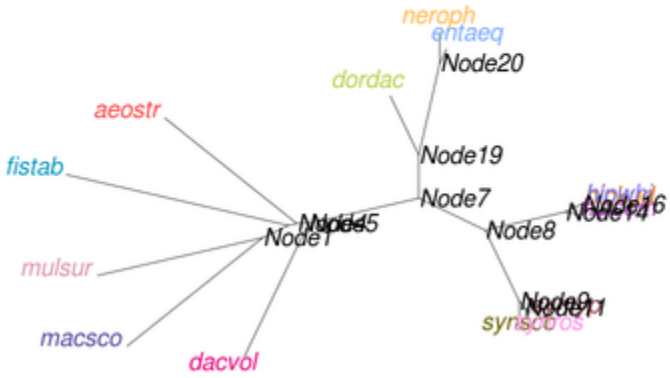
PRKCD-2



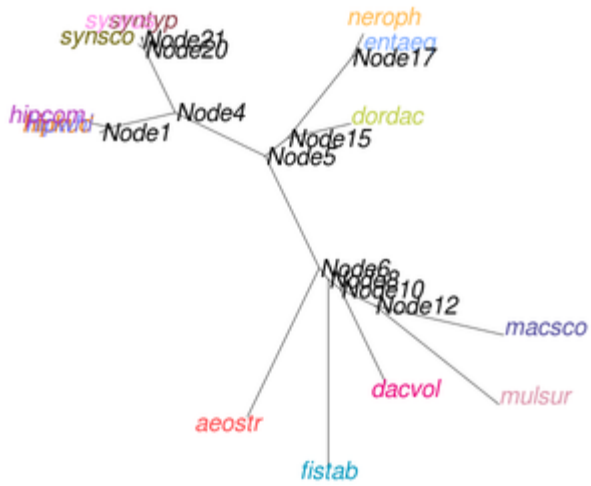
PTGFRN



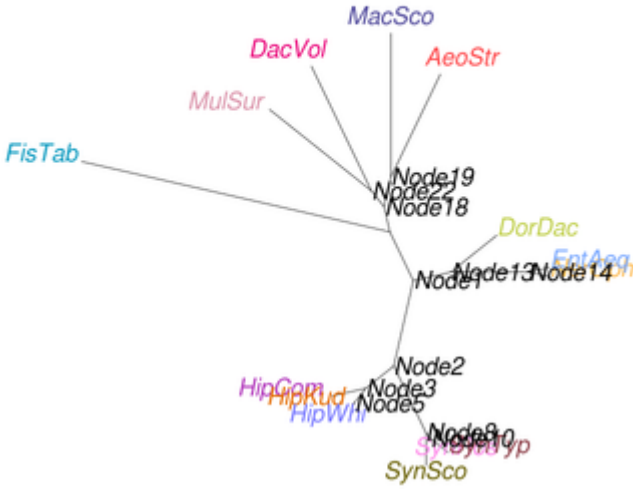
RAG1



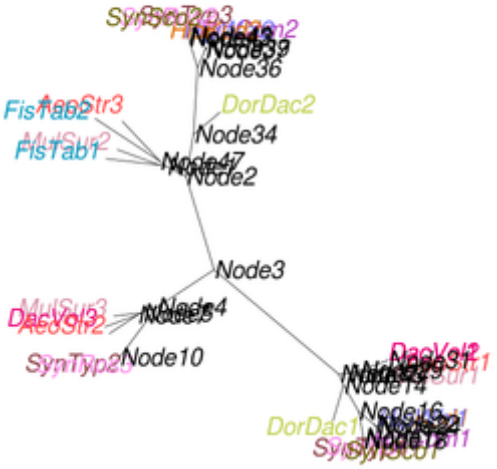
RAG2



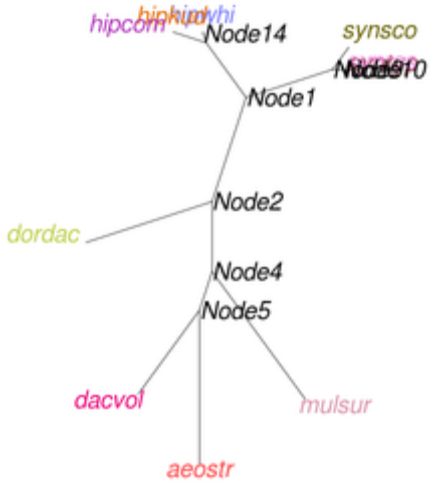
RHAG



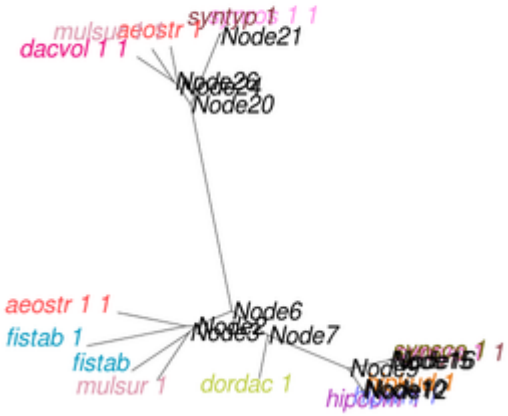
TAPs



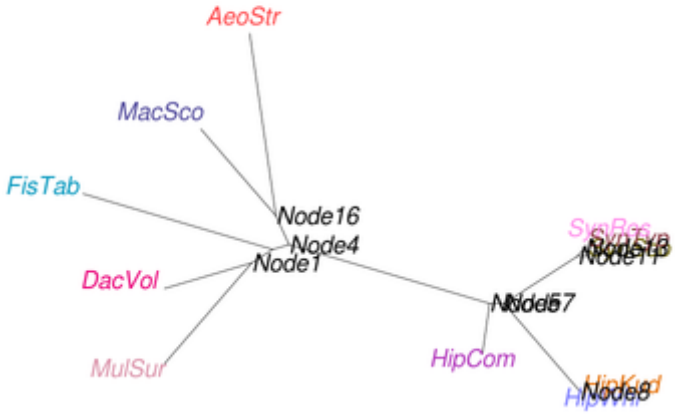
TAPs-1



TAPs-2



TF



References

1. M. A. Nowak, K. T. Tarczy-Hornoch, J. M. Austyn, The optimal number of major histocompatibility complex molecules in an individual. *Proc Natl Acad Sci USA* **89**, 10896 - 10899 (1992).
2. Q. Lin *et al.*, The seahorse genome and the evolution of its specialized morphology. *Nature* **540**, 395-399 (2016).
3. C. M. Small *et al.*, The genome of the Gulf pipefish enables understanding of evolutionary innovations. *Genome Biology* **17**, 258 (2016).
4. M. Malmstrøm *et al.*, Evolution of the immune system influences speciation rates in teleost fishes. *Nature Genetics* **48**, 1204-1210 (2016).
5. M. Malmstrøm, M. Matschiner, O. K. Torresen, K. S. Jakobsen, S. Jentoft, Whole genome sequencing data and de novo draft assemblies for 66 teleost species. *Scientific data* **4**, 160132 (2017).
6. S. Gnerre *et al.*, High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proceedings of the National Academy of Sciences* **108**, 1513-1518 (2011).
7. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (2011).
8. J. R. Miller *et al.*, Aggressive assembly of pyrosequencing reads with mates. *Bioinformatics* **24**, 2818-2824 (2008).
9. J. T. Simpson, Exploring genome characteristics and sequence quality without a reference. *Bioinformatics* **30**, 1228-1235 (2014).
10. G. Parra, K. Bradnam, Z. Ning, T. Keane, I. Korf, Assessing the gene space in draft genomes. *Nucleid Acids Research* **37**, 289-297 (2009).
11. G. Parra, K. Bradnam, I. Korf, CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* **23**, 1061-1067 (2007).
12. F. A. Simao, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseava, E. M. Zdobnov, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210-3212 (2015).
13. K. R. Bradnam *et al.*, Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species. *GigaScience* **2**, 10 (2013).
14. M. G. Grabherr *et al.*, Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology* **29**, 644-652 (2011).
15. A. Lomsadze, V. Ter-Hovhannisyan, Y. O. Chernoff, M. Borodovsky, Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic acids research* **33**, 6494-6506 (2005).
16. I. Korf, Gene finding in novel genomes. *BMC bioinformatics* **5**, 59 (2004).
17. C. Holt, M. Yandell, MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC bioinformatics* **12**, 491 (2011).
18. M. S. Campbell *et al.*, MAKER-P: a tool kit for the rapid creation, management, and quality control of plant genome annotations. *Plant physiology* **164**, 513-524 (2014).
19. W. Bao, K. K. Kojima, O. Kohany, Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA* **6**, 11 (2015).
20. U. Consortium, UniProt: the universal protein knowledgebase. *Nucleic acids research* **45**, D158-D169 (2017).
21. M. S. Campbell, C. Holt, B. Moore, M. Yandell, Genome annotation and curation using MAKER and MAKER-P. *Current Protocols in Bioinformatics*, 4.11. 11-14.11. 39 (2014).

22. M. Stanke, M. Diekhans, R. Baertsch, D. Haussler, Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* **24**, 637-644 (2008).
23. P. Jones *et al.*, InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236-1240 (2014).
24. Z. Musilová *et al.*, Vision using multiple distinct rod opsins in deep-sea fishes. *bioRxiv* (2018).
25. B. L. Aken *et al.*, Ensembl 2017. *Nucleic acids research* **45**, D635-D642 (2016).
26. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *Journal of molecular biology* **215**, 403-410 (1990).
27. R. Betancur-R *et al.*, The tree of life and a new classification of bony fishes. *PLoS currents* **5** (2013).
28. R. Bouckaert *et al.*, BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS computational biology* **10**, e1003537 (2014).
29. M. Matschiner *et al.*, Bayesian phylogenetic estimation of clade ages supports trans-Atlantic dispersal of cichlid fishes. *Systematic biology* **66**, 3-22 (2017).
30. M. Malmstrøm *et al.*, Evolution of the immune system influences speciation rates in teleost fishes. *Nature Genetics* **48**, 1204 (2016).
31. R. Betancur-R *et al.*, Phylogenetic classification of bony fishes. *BMC evolutionary biology* **17**, 162 (2017).
32. T. J. Near *et al.*, Resolution of ray-finned fish phylogeny and timing of diversification. *Proceedings of the National Academy of Sciences* **109**, 13698-13703 (2012).
33. D. L. Rabosky *et al.*, An inverse latitudinal gradient in speciation rate for marine fishes. *Nature* **559**, 392 (2018).
34. M. E. Alfaro *et al.*, Explosive diversification of marine fishes at the Cretaceous–Palaeogene boundary. *Nature ecology & evolution*, 1 (2018).
35. H. Hamilton *et al.*, Molecular phylogeny and patterns of diversification in Syngnathid fishes. *Molecular phylogenetics and Evolution* **107**, 388-403 (2017).
36. M. Khabbazian, R. Kriebel, K. Rohe, C. Ané, Fast and accurate detection of evolutionary shifts in Ornstein–Uhlenbeck models. *Methods in Ecology and Evolution* **7**, 811-824 (2016).
37. D. Chalopin, M. Naville, F. Plard, D. Galiana, J.-N. Volff, Comparative analysis of transposable elements highlights mobilome diversity and evolution in vertebrates. *Genome biology and evolution* **7**, 567-580 (2015).
38. A. Canapa, M. Barucca, M. A. Biscotti, M. Forconi, E. Olmo, Transposons, genome size, and evolutionary insights in animals. *Cytogenetic and genome research* **147**, 217-239 (2015).
39. O. K. Tørresen *et al.*, An improved genome assembly uncovers prolific tandem repeats in Atlantic cod. *BMC genomics* **18**, 95 (2017).
40. D. Ellinghaus, S. Kurtz, U. Willhoeft, LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC bioinformatics* **9**, 18 (2008).
41. J. Jurka *et al.*, Repbase Update, a database of eukaryotic repetitive elements. *Cytogenetic and genome research* **110**, 462-467 (2005).
42. S. Kumar, G. Stecher, K. Tamura, MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular biology and evolution* **33**, 1870-1874 (2016).
43. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010).
44. R. C. Edgar, MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC bioinformatics* **5**, 113 (2004).
45. J. H. Brown *et al.*, Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* **364**, 33 (1993).

46. F. Buonocore *et al.*, Molecular cloning, differential expression and 3D structural analysis of the MHC class-II β chain from sea bass (Dicentrarchus labrax L.). *Fish & shellfish immunology* **23**, 853-866 (2007).
47. B. Rost, C. Sander, Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins: Structure, Function, and Bioinformatics* **19**, 55-72 (1994).
48. A. Ceroni, A. Passerini, A. Vullo, P. Frasconi, DISULFIND: a disulfide bonding state and cysteine connectivity prediction server. *Nucleic acids research* **34**, W177-W181 (2006).
49. A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313 (2014).
50. J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin, D. G. Higgins, The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic acids research* **25**, 4876-4882 (1997).
51. M. Tine *et al.*, European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. *Nature communications* **5**, 5770 (2014).
52. B. Star *et al.*, The genome sequence of Atlantic cod reveals a unique immune system. *Nature* **477**, 207-210 (2011).
53. U. Grimholt *et al.*, A comprehensive analysis of teleost MHC class I sequences. *BMC evolutionary biology* **15**, 32 (2015).
54. A. Bahr, A. B. Wilson, The evolution of MHC diversity: evidence of intralocus gene conversion and recombination in a single-locus system. *Gene* **497**, 52-57 (2012).
55. K. Wegner *et al.*, Genetic variation in MHC class II expression and interactions with MHC sequence polymorphism in three-spined sticklebacks. *Molecular Ecology* **15**, 1153-1164 (2006).
56. R. Cerri *et al.*, Effects of lactation and pregnancy on gene expression of endometrium of Holstein cows at day 17 of the estrous cycle or pregnancy. *Journal of dairy science* **95**, 5657-5675 (2012).
57. S. L. K. Pond, S. V. Muse, "HyPhy: hypothesis testing using phylogenies" in Statistical methods in molecular evolution. (Springer, 2005), pp. 125-181.
58. S. Weaver *et al.*, Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. *Molecular biology and evolution* **35**, 773-777 (2018).
59. M. D. Smith *et al.*, Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Molecular biology and evolution* **32**, 1342-1353 (2015).
60. B. Murrell *et al.*, Detecting individual sites subject to episodic diversifying selection. *PLoS genetics* **8**, e1002764 (2012).
61. T. Borza, C. Stone, A. K. Gamperl, S. Bowman, Atlantic cod (Gadus morhua) hemoglobin genes: multiplicity and polymorphism. *BMC genetics* **10**, 51 (2009).
62. J. C. Opazo, G. T. Butts, M. F. Nery, J. F. Storz, F. G. Hoffmann, Whole-genome duplication and the functional diversification of teleost fish hemoglobins. *Molecular biology and evolution* **30**, 140-153 (2012).
63. J. Feng *et al.*, Channel catfish hemoglobin genes: identification, phylogenetic and syntenic analysis, and specific induction in response to heat stress. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* **9**, 11-22 (2014).
64. M. Berenbrink, Historical reconstructions of evolving physiological complexity: O2 secretion in the eye and swimbladder of fishes. *Journal of Experimental Biology* **210**, 1641-1652 (2007).
65. T. Brittain, Root effect hemoglobins. *Journal of Inorganic Biochemistry* **99**, 120-129 (2005).

66. R. C. Hardison, Evolution of hemoglobin and its genes. *Cold Spring Harbor perspectives in medicine* **2**, a011627 (2012).
67. C. Trapnell *et al.*, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols* **7**, 562-578 (2012).
68. D. Kim *et al.*, TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology* **14**, R36 (2013).
69. L. Goff, C. Trapnell, D. Kelley, cummeRbund: Analysis, exploration, manipulation, and visualization of Cufflinks high-throughput sequencing data. *R package version* **2** (2013).
70. M. C. Brandley, R. L. Young, D. L. Warren, M. B. Thompson, G. P. Wagner, Uterine gene expression in the live-bearing lizard, Chalcides ocellatus, reveals convergence of squamate reptile and mammalian pregnancy mechanisms. *Gen Biol Evol* **4**, 394-411 (2012).
71. C. M. Whittington, O. W. Griffith, W. Qi, M. B. Thompson, A. B. Wilson, Seahorse brood pouch transcriptome reveals common genes associated with vertebrate pregnancy. *Mol Biol Evol* **32**, 3114-3131 (2015).
72. T. E. Spencer, T. R. Hansen, *Implantation and establishment of pregnancy in ruminants*, Regulation of implantation and establishment of pregnancy in mammals (Springer International Publishing, 2015).
73. A. R. Chavan, O. W. Griffith, G. P. Wagner, The inflammation paradox in the evolution of mammalian pregnancy: turning a foe into a friend. *Current opinion in genetics & development* **47**, 24-32 (2017).
74. R. Geisert, A. Fazleabas, M. Lucy, D. Mathew, Interaction of the conceptus and endometrium to establish pregnancy in mammals: role of interleukin 1β. *Cell and tissue research* **349**, 825-838 (2012).
75. T. R. Hansen *et al.*, Mechanism of action on interferon-tau in the uterus during early pregnancy. *Journal of reproduction and fertility* **54**, 329-339 (1999).
76. G. E. Mann, G. E. Lamming, R. S. Robinson, D. C. Wathes, The regulation of interferon-tau production and uterine hormone receptors during early pregnancy. *Journal of reproduction and fertility* **54**, 317-328 (1999).
77. Y. Niwano *et al.*, Suppression of T-Lymphocyte Blastogenesis by Ovine Trophoblast Protein-1 and Human Interferon-α May Be Independent of Interleukin-2 Production. *American Journal of Reproductive Immunology* **20**, 21-26 (1989).
78. T. G. Wegmann, H. Lin, L. Guilbert, T. R. Mosmann, Bidirectional cytokine interactions in the maternal-fetal relationship: is successful pregnancy a TH2 phenomenon? *Immunology today* **14**, 353-356 (1993).
79. U. B. Knudsen, N. Uldbjerg, T. Rechberger, K. Fredens, Eosinophils in human cervical ripening. *Eur J Obstretics Gynecol Reprod Biol* **72**, 165-168 (1997
).
80. K. Chwalisz *et al.*, Pregnancy: Cervical ripening with the cytokines interleukin 8, interleukin 1β and tumour necrosis factor α in guinea-pigs. *Human Reproduction* **9**, 2173-2181 (1994).
81. Z. Zhang *et al.*, Alternations of IL-6, IL-6R and gp130 in early and late onset severe preeclampsia. *Hypertension in pregnancy* **32**, 270-280 (2013).
82. S. Khanjani *et al.*, NF-κB regulates a cassette of immune/inflammatory genes in human pregnant myometrium at term. *Journal of Cellular and Molecular Medicine* **15**, 809-824 (2011).
83. L. Blanchon *et al.*, Activation of the human pregnancy-specific glycoprotein PSG-5 promotor by KLF4 and Sp1. *Biochem Biophys Res Comm* **343**, 745-753 (2006).

84. G. P. Wagner, K. Kin, L. Muglia, M. Pavlicev, Evolution of mammalian pregnancy and the origin of the decidual stromal cells. *International Journal of Developmental Biology* **58**, 117-126 (2014).
85. S. Tait, R. Tassinari, F. Maranghi, A. Montovani, Bisphenol A affects placental layers morphology and angiogenesis during early pregnancy phase in mice. *J Appl Toxicol* **35**, 1278-1291 (2015).
86. A. Mould, M. A. Morgan, L. Li, E. K. Bikoff, E. J. Robertston, Blimp1/Prdm1 governs terminal differentiation of endovascular trophoblast giant cells and defines multipotent progenitors in the developing placenta. *Genes & development* **26**, 2063-2074 (2012).
87. D. Sezen *et al.*, Gelsolin down-regulated lipopolysaccharide-induced intraamniotic tumor necrosis factor-a production in the midtrimester of pregnancy. *American J Obstetrics Gynecol* **200**, 191.e191-191.e194 (2009).
88. M. M. Joyce *et al.*, Pig Conceptuses Increase Uterine Interferon-Regulatory Factor 1 (IRF1), but Restrict Expression to Stroma Through Estrogen-Induced IRF2 in Luminal Epithelium. *Biology of Reproduction* **77**, 292-302 (2007).
89. S. P. Murphy *et al.*, Interferon gamma in successful pregnancies. *Biology of reproduction* **80**, 848-859 (2009).
90. P. Jeyasuria, J. Wetzel, M. Bradley, K. Subedi, J. C. Condon, Progesterone-regulated caspase 3 action in the mouse may play a role in uterine quiescence during pregnancy through fragmentation of uterine myocyte contractile proteins. *Biol Reprod* **928-934** (2009).
91. L. M. Clover, I. L. Sargent, A. Townsend, R. Tampe, C. W. G. Redman, Expression of TAP1 by human trophoblast. *Eur J Immunol* **25**, 543-548 (1995).
92. X. Zhang *et al.*, Ovarian stimulation retards postimplantation development and alters global gene expression profile of blastocytsts in mouse . *Fertility and Sterility* **93**, 2770-2773 (2010).
93. D. R. Bainbridge, S. A. Ellis, I. L. Sargent, The short forms of HLA-G are unlikely to play a role in pregnancy because they are not expressed at the cell surface. *Journal of reproductive immunology* **47**, 1-16 (2000).
94. L. Przybyl *et al.*, Cd74-dysregulation of Macrophage-trophoblastic Interactions in the Preeclamptic Placenta. *Hypertension* **66**, A042 (2015).
95. A. Franczak, B. Wojciechowicz, G. Kotwica, Transcriptomic analysis of the porcine endometrium during early pregnancy and the estrous cycle. *Reproductive Biology* **13**, 229-237 (2013).
96. S. E. Ulbrich *et al.*, Quantitative characterization of prostaglandins in the uterus of early pregnant cattle. *Reproduction* **138**, 371-382 (2009).
97. M. O. Atli, A. Guzeloglu, D. Ali Dinc, Expression of wingless type (WNT) genes and their antagonists at mRNA levels in equine endometrium during the estrous cycle and early pregnancy. *Animal Reprod Sci* **125**, 94-102 (2011).
98. M. Arianmanesh, R. H. McIntosh, R. G. Lea, P. A. Fowler, K. H. Al-Gubory, Ovine corpus luteum proteins, with functions including oxidative stress and lipid metabolism, show complex alterations during implantation. *J Endocrinol* **210**, 47-58 (2011).
99. D. Mao *et al.*, ATF3 expression in the corpus luteum: possible role in luteal regression. *Molecular Endocrinology* **27**, 2066-2079 (2013).
100. S. A. Khan, G. M. Stancel, *Protooncogenes and growth factors in stroid hormoneinduced grwoth and differentiation* (CRC press London, 1993).
101. E. Chung, F. Yeung, L. A. Leinwand, Calcineurin activity is required for cardiac remodelling in pregnancy. *Cardiovasc Res* **100**, 402-410 (2013).
102. K. Shirasuna *et al.*, NLRP3 deficiency improves angiotensin II-induced hypertension but not fetal growth restriction during pregnancy. *Endocrinol* **156**, 4281-4292 (2015).

103. T. Shimizu *et al.*, Actions and interactions of progesterone and estrogen on transcriptome profiles of the bovine endometrium. *Physiological genomics* **42**, 290-300 (2010).
104. D. Le Roith, The Insulin-Like Growth Factor System. *Experimental Diabesity Research* **4**, 205-212 (2003).
105. S. M. Firth, R. C. Baxter, Cellular Actions of the Insulin-Like Growth Factor Binding Proteins. *Endocrine Reviews* **23**, 824-854 (2002).
106. K. Sun, K. Yang, J. R. Challis, Regulation of 11β-hydroxysteroid dehydrogenase type 2 by progesterone, estrogen, and the cyclic adenosine 5′-monophosphate pathway in cultured human placental and chorionic trophoblasts. *Biology of Reproduction* **58**, 1379-1384 (1998).
107. P. M. Coan *et al.*, Adaptations in placental phenotype support fetal growth during undernutrition of pregnant mice. *The Journal of Physiology* **588**, 527-538 (2010).
108. M. F. Romero, w. C. M. Fulton, W. F. Boron, The SLC4 family of HCO3-transporters. *Pflügers Arch* **447**, 495-509 (2004).
109. E. D. Watson, J. C. Cross, Development of structures and transport functions in the mouse placenta. *Physiology* **20**, 180-193 (2005).
110. A. Moffett, C. Loke, Immunology of placentation in eutherian mammals. *Nature Reviews Immunology* **6**, 584-594 (2006).
111. J. Sonanez-Organis, M. Hernandez-Palomares, A. Virgen-Ortiz, J. Rosas-Rodriguez, R. Ortiz, Pregnancy up-regulates genes involved in glucose and lipid metabolism in the left ventricle of rat. *FASEB J* **29**, LB576 (2015).
112. D. Zhang, Y.-J. Tan, F. Qu, J. k.-Z. Sheng, H. F. Huang, Functions of water channels in male and female reproductive systems. *Water Channel Prot*, 676-690 (2012).
113. P. Cunningham, L. McDermott, Long chain PUFA transport in human term placenta. *The Journal of Nutrition* **139**, 636-639 (2009).
114. D. Leger, B. Campoin, J. P. Decottignies, J. Montreuil, G. Spik, Physiological significanc of the marked increased branching of the glycans of human serotransferring during pregnancy. *Biochem J* **257**, 231-238 (1989).
115. Y. Zhang, N. Zhao, H. Zhao, Y. Cui, J. Liu, Expression of tight junction factors in human placental tissues derived from assisted reproductive technology and natural pregnancy. *Europe PMC* **49**, 125-129 (2014).
116. O. Vaughan, A. Sferruzzi-Perri, A. Fowden, Maternal corticosterone regulates amino acid allocation to fetal growth in mice. *Endocrine Abstracts* **31**, 289 (2013).
117. F. W. Bazer *et al.*, Select nutrients in the uterine lumen of sheep and pigs affect conceptus development. *J Reprod Dev* **58**, 180-188 (2012).
118. L. A. Okumu *et al.*, Endometrial expression of progesterone-induced blocking factor and galectins-1,-3,-9, and 3-binding protein in the luteal phase and early pregnancy in cattle. *Physiological genomics* **43**, 903-910 (2011).
119. W. Ma, C. Baumann, M. M. Viveiros, Lack of protein kinase C-delta (PKCδ) disrupts fertilization and embryonic development. *Molecular reproduction and development* **82**, 797-808 (2015).
120. M. Santhanakrishnan, K. Ray, K. Oppenheimer, E. A. Bonney, Dynamic regulation of alpha-dystroglycan in mouse placenta. *Placenta* **29**, 932-936 (2008).
121. J. M. Stoler *et al.*, Separation of amniotic membranes after amniocentesis in an individual with the classic form of EDS and haploinsufficiency for COL5A1 expression. *American journal of medical genetics* **101**, 174-177 (2001).
122. C. G. Bönnemann, The collagen VI-related myopathies: muscle meets its matrix. *Nature Reviews Neurology* **7**, 379 (2011).

123. C. Agca, Y. Agca, Molecular and ultrastuctural changes of rat pre-implantation embryos during two-cell developmental arrest. *Journal of assisted reproduction and genetics* **31**, 767-780 (2014).
124. A. El-Sayet *et al.*, Large-scale transcriptional analysis of bovine embryo biopsies in relation to pregnancy successafter transfer to recipients. *Physiological genomics* **28**, 84-96 (2006).
125. Y. Lee *et al.*, Interactions between inflammatory signals and the progesterone receptor in regulating gene expression in pregnant human uterine myocytes. *J Cell Mol Med* **16**, 1487-2503 (2012).
126. K. M. Downs, J. McHugh, A. J. Copp, E. Shtivelman, Multiple developmental roles of Ahnak are suggested by localization to sites of placentation and neural plate fusion in the mouse conceptus. *Mechanisms of development* **119**, S31-S38 (2002).
127. Z. Zhou *et al.*, The proprotein convertase furin in human trophoblast: possible role in promoting trophoblast cell migraton and invasion. *Placenta* **30**, 929-938 (2009).
128. Y. Yamaguchi *et al.*, Changes in the distribution of fibronectin in the placenta during normal human pregnancy. *American journal of obstetrics and gynecology* **152**, 715-718 (1975).
129. J. S. Hunt, J. L. Pace, R. M. Gill, Immunoregulatory molecules in human placentas: potential for diverse roles in pregnancy. *Journal of Developmental Biology* **54**, 457-467 (2009).
130. S. Bauersachs, K. Mitko, S. E. Ulbrich, H. Blum, E. Wolf, Transcriptome studies of bovine endometrium reveal molecular profiles characteristic for specific stages of estrous cycle and early pregnancy. *Exp Clin Endocrinol Diabetes* **116**, 1-14 (2008).
131. T. Workalemahu *et al.*, Genome-wide and candidate gene association studies of placental abruption. *International journal of molecular epidemiology and genetics* **4**, 128 (2013).
132. J. P. Boisseau, Prolactin et l'incubation chez l'hippocampe. *Comp Rend Acad Sci Paris* **177**, 205-215 (1969).
133. S. Renaud, P. Gautheron, Role of platelet factor 3 in the hypercoagulability induced by pregnancy and oral contraceptives. *Thrombosis and Haemostasis* **30**, 299-306 (1973).
134. G. T. Colbern, M. H. Chiang, E. K. Main, Expression of the nonclassic histocompatibility antigen HLA-G by preeclamptic placenta. *American journal of obstetrics and gynecology* **170**, 1244-1250 (1994).
135. E. Crosley, M. Elliot, J. Christians, B. Crespi, Placental invasion, preeclampsia risk and adaptive molecular evolution at the origin of the great apes: evidence from genome-wide analyses. *Placenta* **34**, 127-132 (2013).
136. I. J. Elenkov *et al.*, IL-12, TNF-α, and hormonal changes during late pregnancy and early postpartum: implications for autoimmune disease activity during these times. *The Journal of Clinical Endocrinology & Metabolism* **86**, 4933-4938 (2001).
137. J. D. Aplin, C. J. Jones, Fucose, placental evolution and the glycocode. *Glycobiology* **22**, 470-478 (2011).
138. F. Li, J. Liu, E.-S. Park, M. Jo, T. E. Curry Jr, The B cell translocation gene (BTG) family in the rat ovary: hormonal induction, regulation, and impact on cell cycle kinetics. *Endocrinology* **150**, 3894-3902 (2009).
139. M. Breuiller-Fouche, G. Charpigny, G. Germain (2007) Functional genomics of the pregnant uterus: from expectations to reality, a compilation of studies in the myometrium. in *BMC pregnancy and childbirth* (BioMed Central), p S4.
140. Q. E. Yang, M. I. Giassetti, A. D. Ealy, Fibroblast growth factors activate mitogen-activated protein kinase pathways to promote migration in ovine trophoblast cells. *Reproduction*, REP-10-0541 (2011).

141. X. Wang *et al.*, Functional role of arginine during the peri-implantation period of pregnancy. I. Consequences of loss of function of arginine transporter SLC7A1 mRNA in ovine conceptus trophectoderm. *The FASEB Journal* **28**, 2852-2863 (2014).
142. S. Betsha *et al.*, Transcriptome profile of bovine elongated conceptus obtained from SCNT and IVP pregnancies. *Molecular reproduction and development* **80**, 315-333 (2013).
143. S. H. Zeisel, Importance of methyl donors during reproduction–. *The American journal of clinical nutrition* **89**, 673S-677S (2008).
144. D. S. Iravathy Goud Kalal, A. Panda, Y. Nitschke, F. Rutsch, Molecular diagnosis of generalized arterial calcification of infancy (GACI). *Journal of cardiovascular disease research* **3**, 150 (2012).
145. O. Kohannim, J. Peredo, K. M. Dipple, F. Quintero-Rivera, Clinical Findings Associated with a De Novo Partial Trisomy 10p11. 22p15. 3 and Monosomy 7p22. 3 Detected by Chromosomal Microarray Analysis. *Case reports in genetics* **2011** (2011).
146. S. Bauersachs, E. Wolf, Transcriptome analyses of bovine, porcine and equine endometrium during the pre-implantation phase. *Animal reproduction science* **134**, 84-94 (2012).
147. S. M. Waters, G. S. Coyne, D. A. Kenny, D. G. Morris, Effect of dietary n-3 polyunsaturated fatty acids on transcription factor regulation in the bovine endometrium. *Molecular biology reports* **41**, 2745-2755 (2014).
148. J. M. P. Pabona, Z. Zeng, F. A. Simmen, R. C. Simmen, Functional differentiation of uterine stromal cells involves cross-regulation between bone morphogenetic protein 2 and Kruppel-like factor (KLF) family members KLF9 and KLF13. *Endocrinology* **151**, 3396-3406 (2010).
149. T. Shindo *et al.*, ADAMTS-1: a metalloproteinase-disintegrin essential for normal growth, fertility, and organ morphology and function. *The Journal of clinical investigation* **105**, 1345-1352 (2000).
150. M. Ostrowska, B. Ż. Górka, K. Słoniewski, L. Z. Kowalski, Expression of PC, PCK1, PCK2, LDHB, FBP1 and G6PC genes in the liver of cows in the transition from pregnancy to lactation. *Animal Science Papers and Reports* **31**, 281-290 (2013).
151. V. Maillo *et al.*, Oviduct-embryo interactions in cattle: two-way traffic or a one-way street? *Biology of reproduction* **92**, 144, 141-148 (2015).
152. M. Eghbali *et al.*, Molecular and functional signature of heart hypertrophy during pregnancy. *Circulation research* **96**, 1208-1216 (2005).
153. J. Zheng, M. L. Johnson, D. A. Redmer, L. P. Reynolds, Estrogen and progesterone receptors, cell proliferation, and c-fos expression in the ovine uterus during early pregnancy. *Endocrinology* **137**, 340-348 (1996).
154. A. El-Sayed *et al.*, Large-scale transcriptional analysis of bovine embryo biopsies in relation to pregnancy success after transfer to recipients. *Physiological genomics* (2017).
155. M. C. Rudolph, J. L. McManaman, L. Hunter, T. Phang, M. C. Neville, Functional development of the mammary gland: use of expression profiling and trajectory clustering to reveal changes in gene expression during pregnancy, lactation, and involution. *Journal of mammary gland biology and neoplasia* **8**, 287-307 (2003).
156. A. Larsson, M. Palm, S. Basu, O. Axelsson, Insulin-like growth factor binding protein-1 (IGFBP-1) during normal pregnancy. *Gynecological Endocrinology* **29**, 129-132 (2013).
157. O. Shynlova, P. Tsui, A. Dorogin, B. L. Langille, S. J. Lye, The expression of transforming growth factor β in pregnant rat myometrium is hormone and stretch dependent. *Reproduction* **134**, 503-511 (2007).

158. Y. W. Chan, H. A. Berg, J. D. Moore, S. Quenby, A. M. Blanks, Assessment of myometrial transcriptome changes associated with spontaneous human labour by high-throughput RNA-seq. *Experimental physiology* **99**, 510-524 (2014).
159. C. Méhats *et al.*, Pregnancy induces a modulation of the cAMP phosphodiesterase 4-conformers ratio in human myometrium: consequences for the utero-relaxant effect of PDE4-selective inhibitors. *Journal of Pharmacology and Experimental Therapeutics* **292**, 817-823 (2000).
160. M. Chauhan, U. Yallampalli, Y.-L. Dong, G. D. Hankins, C. Yallampalli, Expression of adrenomedullin 2 (ADM2)/intermedin (IMD) in human placenta: role in trophoblast invasion and migration. *Biology of reproduction* **81**, 777-783 (2009).
161. H. Stepan, C. Leo, S. Purz, M. Höckel, L.-C. Horn, Placental localization and expression of the cell death factors BNip3 and Nix in preeclampsia, intrauterine growth retardation and HELLP syndrome. *European Journal of Obstetrics & Gynecology and Reproductive Biology* **122**, 172-176 (2005).
162. D. A. Enquobahrie, C. Qiu, S. Y. Muhie, M. A. Williams, Maternal peripheral blood gene expression in early pregnancy and preeclampsia. *International journal of molecular epidemiology and genetics* **2**, 78 (2011).
163. M. Waters, P. Kaye, The role of growth hormone in fetal development. *Growth Hormone & IGF Research* **12**, 137-146 (2002).
164. H. D. Mistry *et al.*, Differential expression and distribution of placental glutathione peroxidases 1, 3 and 4 in normal and preeclamptic pregnancy. *Placenta* **31**, 401-408 (2010).
165. N. Mansouri-Attia *et al.*, Endometrium as an early sensor of in vitro embryo manipulation technologies. *Proceedings of the National Academy of Sciences* **106**, 5687-5692 (2009).
166. T. Gu *et al.*, Endometrial gene expression profiling in pregnant Meishan and Yorkshire pigs on day 12 of gestation. *BMC genomics* **15**, 156 (2014).
167. S. García-Herrero *et al.*, The transcriptome of spermatozoa used in homologous intrauterine insemination varies considerably between samples that achieve pregnancy and those that do not. *Fertility and sterility* **94**, 1360-1373 (2010).
168. C. Le Van Kim, Y. Colin, J.-P. Cartron, Rh proteins: key structural and functional components of the red cell membrane. *Blood reviews* **20**, 93-110 (2006).
169. G. A. Johnson *et al.* (2010) Glucose, Amino Acids, and Their Transporters in Pig Uteri, Conceptuses, and Placentae. (Oxford University Press).
170. X.-L. Cui, B. Chang, L. Myatt, Expression and distribution of NADPH oxidase isoforms in human myometrium—role in angiotensin II-induced hypertrophy. *Biology of reproduction* **82**, 305-312 (2010).
171. L. Grande, An empirical synthetic pattern study of gars (Lepisosteiformes) and closely related species, based mostly on skeletal anatomy. The resurrection of Holostei. *Copeia* **10**, 1 (2010).
172. R. E. Broughton, R. Betancur-R, C. Li, G. Arratia, G. Ortí, Multi-locus phylogenetic analysis reveals the pattern and tempo of bony fish evolution. *PLoS currents* **5** (2013).
173. A. S. Woodward, *Catalogue of the fossil fishes in the British Museum (Natural History)* (British Museum, 1895), vol. 3.
174. C. C. Wendling, A. Piecyk, D. Refardt, O. Roth, Tripartite species interaction: phage susceptible bacterial phenotypes show higher virulence against eukaryotic hosts. *In review* (2015).
175. E. Wiley, G. D. Johnson, A teleost classification based on monophyletic groups. *Origin and phylogenetic interrelationships of teleosts* (2010).
176. X. Zou *et al.*, Ammonoids from the Zhuganpo Member of the Falang Formation at Nimaigu and their relevance for dating the Xingyi fossil-lagerstätte (late Ladinian,

Guizhou, China). *Rivista Italiana di Paleontologia e Stratigrafia (Research In Paleontology and Stratigraphy)* **121** (2015).
177. M. J. Benton *et al.*, Constraints on the timescale of animal evolutionary history. *Palaeontologia Electronica* **18**, 1-106 (2015).
178. X. Wang, Y. Wang, F. Jin, X. Xu, Y. Wang (1999) Vertebrate assemblages of the Jehol biota in western Liaoning, China. in *Seventh Annual Meeting of the Chinese Society of Vertebrate Paleontology*, pp 1-12.
179. J. E. O'Reilly, M. dos Reis, P. C. Donoghue, Dating tips for divergence-time estimation. *Trends in Genetics* **31**, 637-650 (2015).
180. Z. Zhonghe, Evolutionary radiation of the Jehol Biota: chronological and ecological perspectives. *Geological Journal* **41**, 377-393 (2006).
181. G. Arratia, The Clupeocephala re-visited: analysis of characters and homologies. *Rev Biol Mar Oceanog* **45**, 635-657 (2010).
182. F. Santini, L. J. Harmon, G. Carnevale, M. E. Alfaro, Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. *BMC evolutionary biology* **9**, 194 (2009).
183. S. Wenz *et al.*, L'ichthyofaune des calcaires lithographiques du Kimméridgien supérieur de Cerin (Ain, France). *Geobios* **27**, 61-70 (1994).
184. M. P. Davis, C. Fielitz, Estimating divergence times of lizardfishes and their allies (Euteleostei: Aulopiformes) and the timing of deep-sea adaptations. *Molecular Phylogenetics and Evolution* **57**, 1194-1208 (2010).
185. A. Blanco, M. Szabó, À. Blanco-Lapaz, J. Marmi, Late Cretaceous (Maastrichtian) Chondrichthyes and Osteichthyes from northeastern Iberia. *Palaeogeography, Palaeoclimatology, Palaeoecology* **465**, 278-294 (2017).
186. V. Fondevilla, J. Dinarès-Turell, O. Oms, The chronostratigraphic framework of the South-Pyrenean Maastrichtian succession reappraised: Implications for basin development and end-Cretaceous dinosaur faunal turnover. *Sedimentary geology* **337**, 55-68 (2016).
187. A. Filleul, J. G. Maisey, Redescription of Santanichthys diasii (Otophysi, Characiformes) from the Albian of the Santana Formation and comments on its implications for otophysan relationships. *American Museum Novitates*, 1-21 (2004).
188. O. Otero, X. Valentin, G. Garcia, Cretaceous characiform fishes (Teleostei: Ostariophysi) from Northern Tethys: description of new material from the Maastrichtian of Provence (Southern France) and palaeobiogeographical implications. *Geological Society, London, Special Publications* **295**, 155-164 (2008).
189. M. E. Alfaro *et al.*, Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences* **106**, 13410-13414 (2009).
190. U. Heimhofer *et al.*, Isotope and elemental geochemistry of black shale-hosted fossiliferous concretions from the Cretaceous Santana Formation fossil Lagerstätte (Brazil). *Sedimentology* **64**, 150-167 (2017).
191. D. Arcila *et al.*, Genome-wide interrogation advances resolution of recalcitrant groups in the tree of life. *Nature Ecology & Evolution* **1**, 0020 (2017).
192. M. A. Campbell, M. E. Alfaro, M. Belasco, J. A. López, Early-branching euteleost relationships: areas of congruence between concatenation and coalescent model inferences. *PeerJ* **5**, e3548 (2017).
193. G. D. Johnson, C. Patterson, Relationships of lower euteleostean fishes. *Interrelationships of fishes* (1996).
194. M. V. Wilson, D. B. Brinkman, A. G. Neuman, Cretaceous Esocoidei (Teleostei): early radiation of the pikes in North American fresh waters. *Journal of Paleontology* **66**, 839-846 (1992).

195. T. H. Payenberg, D. R. Braman, D. W. Davis, A. D. Miall, Litho-and chronostratigraphic relationships of the Santonian Campanian Milk River Formation in southern Alberta and Eagle Formation in Montana utilising stratigraphy, U Pb geochronology, and palynology. *Canadian Journal of Earth Sciences* **39**, 1553-1577 (2002).
196. M. V. WILSON, G.-Q. Li, Osteology and systematic position of the Eocene salmonid† Eosalmo driftwoodensis Wilson from western North America. *Zoological Journal of the Linnean Society* **125**, 279-311 (1999).
197. P. T. Moss, D. R. Greenwood, S. B. Archibald, Regional and local vegetation community dynamics of the Eocene Okanagan Highlands (British Columbia Washington State) from palynology. *Canadian Journal of Earth Sciences* **42**, 187-204 (2005).
198. J. J. Eberle, N. Rybczynski, D. R. Greenwood, Early Eocene mammals from the Driftwood Creek beds, Driftwood Canyon Provincial Park, northern British Columbia. *Journal of Vertebrate Paleontology* **34**, 739-746 (2014).
199. K. Dietze, Morphology and phylogenetic relationships of certain neoteleostean fishes from the Upper Cretaceous of Sendenhorst, Germany. *Cretaceous Research* **30**, 559-574 (2009).
200. J. Gaudant, Présence d'un Osmeridae: Enoplophthalmus schlumbergeri Sauvage, 1880 dans l'Oligocène inférieur des environs de Céreste (Alpes-de-Haute-Provence, France). *Geodiversitas* **35**, 345-358 (2013).
201. V. Gallo, P. M. Coelho, First occurrence of an aulopiform fish in the Barremian of the Sergipe-Alagoas Basin, northeastern Brazil. *Mesozoic fishes* **4**, 351-371 (2008).
202. M. V. Wilson, A. M. Murray, Early Cenomanian acanthomorph teleost in the Cretaceous Fish Scale Zone, Albian/Cenomanian boundary, Alberta, Canada. *Mesozoic fishes–systematics and paleoecology*, 369-382 (1996).
203. P. L. Forey, L. Yi, C. Patterson, C. E. Davies, Fossil fishes from the Cenomanian (Upper Cretaceous) of Namoura, Lebanon. *Journal of Systematic palaeontology* **1**, 227 (2003).
204. M. Gayet, P. Abi Saad, O. Gaudant, The Fossils of Lebanon. *éditions désiris.* (2012).
205. D. Davesne *et al.*, Early fossils illuminate character evolution and interrelationships of Lampridiformes (Teleostei, Acanthomorpha). *Zoological Journal of the Linnean Society* **172**, 475-498 (2014).
206. K. A. González-Rodríguez, H.-P. Schultze, G. Arratia, M. Wilson, Miniature armored acanthomorph teleosts from the Albian/Cenomanian (Cretaceous) of Mexico. *Mesozoic fishes* **5**, 457e487 (2013).
207. K. A. González-Rodríguez, C. Fielitz, V. M. Bravo-Cuevas, R. E. Baños-Rodríguez, Cretaceous osteichthyan fish assemblages from Mexico. *New Mexico Museum of Natural History and Science Bulletin* **71**, 1-14 (2016).
208. M. G. Newbrey, A. M. Murray, M. V. Wilson, D. B. Brinkman, A. G. Neuman, A new species of the paracanthopterygian Xenyllion (Sphenocephaliformes) from the Mowry Formation (Cenomanian) of Utah, USA. *Mesozoic fishes*, 363-384 (2013).
209. J. Stewart, Cretaceous acanthomorphs of North America. *Mesozoic Fishes–Systematics and Paleoecology*, 383-394 (1996).
210. R. W. Scott, F. Oboh-Ikuenobe, D. G. Benson, J. M. Holbrook, Numerical age calibration of the Albian/Cenomanian boundary. (2009).
211. L. Jacobs, K. Ferguson, M. Polcyn, C. Rennison, Cretaceous δ 13 C stratigraphy and the age of dolichosaurs and early mosasaurs. *Netherlands Journal of Geosciences* **84**, 257-268 (2005).
212. C. Patterson, An overview of the early fossil record of acanthomorphs. *Bulletin of Marine Science* **52**, 29-59 (1993).

213. A. M. Murray, M. V. Wilson, Contributions of fossils to the phylogenetic relationships of the percopsiform fishes (Teleostei: Paracanthopterygii): order restored. *Mesozoic fishes* **2**, 397-411 (1999).
214. A. M. Murray, A new Paleocene genus and species of percopsid,† Massamorichthys wilsoni (Paracanthopterygii) from Joffre Bridge, Alberta, Canada. *Journal of Vertebrate Paleontology* **16**, 642-652 (1996).
215. L. J. Noorbergen *et al.* (2015) Integrated stratigraphy of Paleocene lignite seams of the fluvial Tullock Formation, Montana (USA). in *EGU General Assembly Conference Abstracts*.
216. D. Davesne, G. Carnevale, M. Friedman, Bajaichthys elegans from the Eocene of Bolca (Italy) and the overlooked morphological diversity of Zeiformes (Teleostei, Acanthomorpha). *Palaeontology* **60**, 255-268 (2017).
217. J. C. Tyler, F. Santini, A phylogeny of the fossil and extant zeiform-like fishes, Upper Cretaceous to Recent, with comments on the putative zeomorph clade (Acanthomorpha). *Zoologica Scripta* **34**, 157-175 (2005).
218. M. Schlüter, T. Steuber, M. Parente, Chronostratigraphy of Campanian–Maastrichtian platform carbonates and rudist associations of Salento (Apulia, Italy). *Cretaceous Research* **29**, 100-114 (2008).
219. W.-J. Chen *et al.*, New insights on early evolution of spiny-rayed fishes (Teleostei: Acanthomorpha). *Frontiers in Marine Science* **1**, 53 (2014).
220. A. Rosenkrantz, *Marine Upper Cretaceous and lowermost Tertiary deposits in west Greenland: Investigations before and since 1938* (Fr. Bagges Kgl. Hofbogtrykkeri, 1970).
221. G. Dam *et al.*, Lithostratigraphy of the Cretaceous–Paleocene Nuussuaq Group, Nuussuaq Basin, West Greenland. *Geological Survey of Denmark and Greenland Bulletin* **19**, 171 (2009).
222. D. E. Rosen, C. Patterson, *The structure and relationships of the paracanthopterygian fishes* (1969).
223. D. Cohen (1984) Gadiformes: Overview. in (HG Moser, WJ Richards ed.) Ontogeny and systematics of fishes. Spec. Publ. No 1, American Society of Ichthyologists and Herpetologists. (and Allen Press, Lawrence, Kansas).
224. J. T. Eastman, L. Grande, Late Eocene gadiform (Teleostei) skull from Seymour Island, Antarctic Peninsula. *Antarctic Science* **3**, 87-95 (1991).
225. H. Endo, Phylogeny of the order gadiformes (Teleostei, Paracanthopterygii). *Memoirs of the Graduate School of Fisheries Sciences, Hokkaido University* **49**, 75-149 (2002).
226. J. Kriwet, T. Hecht, A review of early gadiform evolution and diversification: first record of a rattail fish skull (Gadiformes, Macrouridae) from the Eocene of Antarctica, with otoliths preserved in situ. *Naturwissenschaften* **95**, 899-907 (2008).
227. D. G. Johnson, C. Patterson, Percomorph phylogeny: a survey of acanthomorphs and a new proposal. *Bulletin of Marine Science* **52**, 554-626 (1993).
228. J. C. Tyler, L. Sorbini, New superfamily and three new families of tetraodontiform fishes from the Upper Cretaceous: the earliest and most morphologically primitive plectognaths. (1996).
229. M. Miya, T. P. Satoh, M. Nishida, The phylogenetic position of toadfishes (order Batrachoidiformes) in the higher ray-finned fish as inferred from partitioned Bayesian analysis of 102 whole mitochondrial genome sequences. *Biological Journal of the Linnean Society* **85**, 289-306 (2005).
230. J. A. Moore, Phylogeny of the Trachichthyiformes (Teleostei: Percomorpha). *Bulletin of Marine Science* **52**, 114-136 (1993).
231. J. Alvarado-Ortega, B. A. Than-Marchese, The first record of a North American Cenomanian Trachichthyidae fish (acanthomorpha, acanthopterygii), Pepemkay maya,

gen. et sp. nov., from El Chango quarry (Sierra Madre Formation), Chiapas, Mexico. *Journal of Vertebrate Paleontology* **33**, 48-57 (2013).
232. A. M. Murray, Mid-Cretaceous acanthomorph fishes with the description of a new species from the Turonian of Lac des Bois, Northwest Territories, Canada. *Vertebrate Anatomy Morphology Palaeontology* **1**, 101-115 (2016).
233. J. Nelson, Fishes of the world. Hoboken, New Jersey: John Wiley and Sons, Inc., 601 p. (2006).
234. M. Friedman, Ecomorphological selectivity among marine teleost fishes during the end-Cretaceous extinction. *Proceedings of the National Academy of Sciences* **106**, 5218-5223 (2009).
235. C. Aramburg, Résultats scientifiques de la mission C. Arambourg en Syrie et en Iran (1938-1939). II. Les poissons Oligocènes de l'Iran. Notes et Mémoires sur le Moyen-Orient. (1967).
236. A. Dornburg *et al.*, Molecular phylogenetics of squirrelfishes and soldierfishes (Teleostei: Beryciformes: Holocentridae): Reconciling more than 100 years of taxonomic confusion. *Molecular phylogenetics and Evolution* **65**, 727-738 (2012).
237. F. Santini, J. C. Tyler, A phylogeny of the families of fossil and extant tetraodontiform fishes (Acanthomorpha, Tetraodontiformes), Upper Cretaceous to Recent. *Zoological Journal of the Linnean Society* **139**, 565-617 (2003).
238. G. Howes, Notes on the anatomy and classification of ophidiiform fishes with particular reference to the abyssal genus Acanthonus Gunther, 1878. *Bull Br Mus Nat Hist (Zool)* **52**, 95-131 (1992).
239. G. Carnevale, G. D. Johnson, A Cretaceous cusk-eel (Teleostei, Ophidiiformes) from Italy and the Mesozoic diversification of percomorph fishes. *Copeia* **103**, 771-791 (2015).
240. G. Carnevale, B. B. Collette, † Zappaichthys harzhauseri, gen. et sp. nov., a new Miocene toadfish (Teleostei, Batrachoidiformes) from the Paratethys (St. Margarethen in Burgenland, Austria), with comments on the fossil record of batrachoidiform fishes. *Journal of Vertebrate Paleontology* **34**, 1005-1017 (2014).
241. J. Hohenegger, S. Ćorić, M. Wagreich, Timing of the middle miocene Badenian stage of the central Paratethys. *Geologica Carpathica* **65**, 55-66 (2014).
242. C. E. Thacker, Phylogeny of Gobioidei and placement within Acanthomorpha, with a new classification and investigation of diversification and character evolution. *Copeia* **2009**, 93-104 (2009).
243. A. F. Bannikov, G. Carnevale, † Carlomonnius quasigobius gen. et sp. nov.: the first gobioid fish from the Eocene of Monte Bolca, Italy. *Bulletin of Geosciences* **91**, 13-22 (2016).
244. M. Miya *et al.*, Evolutionary origin of the Scombridae (tunas and mackerels): members of a Paleogene adaptive radiation with 14 other pelagic fish families. *PloS one* **8**, e73535 (2013).
245. C. Patterson, Osteichthyes: teleostei. *The fossil record* **2**, 621-656 (1993).
246. H. Y. Song *et al.*, Mitogenomic circumscription of a novel percomorph fish clade mainly comprising "Syngnathoidei"(Teleostei). *Gene* **542**, 146-155 (2014).
247. A. Wilson, J. Orr, The evolutionary origins of Syngnathidae: pipefishes and seahorses. *Journal of Fish Biology* **78**, 1603-1623 (2011).
248. J. Kotlarczyk, A. Jerzmańska, E. Świdnicka, T. Wiszniowska (2006) A framework of ichthyofaunal ecostratigraphy of the Oligocene-Early Miocene strata of the Polish Outer Carpathian basin. in *Annales Societatis Geologorum Poloniae*, pp 1-111.
249. L. J. Roe, Phylogenetic and ecological significance of Channidae (Osteichthyes Teleostei) from the Early Eocene Kuldana Formation of Kohat, Pakistan. (1991).

250. P. D. Gingerich, Stratigraphic and micropaleontological constraints on the middle Eocene age of the mammal-bearing Kuldana Formation of Pakistan. *Journal of Vertebrate Paleontology* **23**, 643-651 (2003).
251. H. L. Fierstine, Fossil history of billfishes (Xiphioidei). *Bulletin of Marine Science* **79**, 433-453 (2006).
252. N. G. Muzylev, *Anoksicheskie sobytiya paleotsena–srednego eotsena (Anoxic events of the Paleocene-Middle Eocene).* . A. Y. Rozanov, M. A. Semikhatov, Eds., Ekosistemnye perestroiki i evolyutsiya biosfery (Ecosystem restructuring and the evolution of the biosphere) (Nedra, Moscow, Russia, 1994).
253. P. C. Wainwright *et al.*, The evolution of pharyngognathy: a phylogenetic and functional appraisal of the pharyngeal jaw key innovation in labroid fishes and beyond. *Systematic Biology* **61**, 1001-1027 (2012).
254. J. S. Sparks, W. L. Smith, Phylogeny and biogeography of cichlid fishes (Teleostei: Perciformes: Cichlidae). *Cladistics* **20**, 501-517 (2004).
255. A. M. Murray, The oldest fossil cichlids (Teleostei: Perciformes): indication of a 45 million-year-old species flock. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **268**, 679-684 (2001).
256. T. Harrison *et al.*, "Paleontological investigations at the Eocene locality of Mahenge in north-central Tanzania, East Africa" in Eocene Biodiversity. (Springer, 2001), pp. 39-74.
257. W. Leo Smith, P. Chakrabarty, J. S. Sparks, Phylogeny, taxonomy, and evolution of Neotropical cichlids (Teleostei: Cichlidae: Cichlinae). *Cladistics* **24**, 625-641 (2008).
258. P. A. Perez, M. C. Malabarba, C. del Papa, A new genus and species of Heroini (Perciformes: Cichlidae) from the early Eocene of southern South America. *Neotropical Ichthyology* **8**, 631-642 (2010).
259. M. C. Malabarba, L. R. Malabarba, C. D. Papa, Gymnogeophagus eocenicus, n. sp.(Perciformes: Cichlidae), an Eocene cichlid from the Lumbrera Formation in Argentina. *Journal of Vertebrate Paleontology* **30**, 341-350 (2010).