nature communications

Article

https://doi.org/10.1038/s41467-022-35338-7

Immunogenetic losses co-occurred with seahorse male pregnancy and mutation in *tlx1* accompanied functional asplenia

Received: 9 May 2022

Accepted: 29 November 2022

Published online: 09 December 2022

Check for updates

Yali Liu[1,2,3,8], Meng Qu[1,2,8], Han Jiang[1,3,8], Ralf Schneider[4,8], Geng Qin[1,2], Wei Luo[1], Haiyan Yu[1], Bo Zhang[1], Xin Wang[1,2], Yanhong Zhang[1,2], Huixian Zhang[1,2], Zhixin Zhang[1,5], Yongli Wu[1], Yingyi Zhang[1,3], Jianping Yin[1,2], Si Zhang[1,2], Byrappa Venkatesh[6], Olivia Roth[4] ✉, Axel Meyer[7] ✉ & Qiang Lin[1,2,3] ✉

In the highly derived syngnathid fishes (pipefishes, seadragons & seahorses), the evolution of sex-role reversed brooding behavior culminated in the seahorse lineage's male pregnancy, whose males feature a specialized brood pouch into which females deposit eggs during mating. Then, eggs are intimately engulfed by a placenta-like tissue that facilitates gas and nutrient exchange. As fathers immunologically tolerate allogenic embryos, it was suggested that male pregnancy co-evolved with specific immunological adaptations. Indeed, here we show that a specific amino-acid replacement in the *tlx1* transcription factor is associated with seahorses' asplenia (loss of spleen, an organ central in the immune system), as confirmed by a CRISPR-Cas9 experiment using zebrafish. Comparative genomics across the syngnathid phylogeny revealed that the complexity of the immune system gene repertoire decreases as parental care intensity increases. The synchronous evolution of immunogenetic alterations and male pregnancy supports the notion that male pregnancy co-evolved with the immunological tolerance of the embryo.

The evolutionary diversification of animals went hand-in-hand with an increasing complexity of the immune system[1,2]. As a hallmark of vertebrate evolution, the MHC/B-cell receptor/T-cell receptor system, an essential arm of the adaptive immune system, first appeared in jawed vertebrates and accompanied both the radiation of sharks and rays, and later the radiation of bony fishes[3,4]. The emergence of specialised cells and molecules, as well as the lymphatic system and the spleen as an important vertebrate secondary lymphoid organ permitted complex immunological reorganisation and modifications during the evolution of the vertebrate adaptive immune system[2]. As a rare exception, the spleen is absent in seahorses (Family Syngnathidae)[5,6]. This raises the question of how they cope without this vital immune-organ and what ultimately selected for the evolutionary loss of the spleen. Seahorses are famous for their iconic morphology and their highly unusual life history, which includes sex-role reversed brooding behavior via "male pregnancy"[7,8]. While females in more basal lineages of syngnathids simply glue their eggs to brooding patches on the ventral side of males, the males' seahorse brood pouch represents a

[1]CAS Key Laboratory of Tropical Marine Bio-Resources and Ecology, South China Sea Institute of Oceanology, Chinese Academy of Sciences, 510301 Guangzhou, China. [2]Guangdong Provincial Key Laboratory of Applied Marine Biology, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, PR China. [3]University of Chinese Academy of Sciences, 100101 Beijing, China. [4]Marine Evolutionary Ecology, Zoological Institute, Kiel University, 24118 Kiel, Germany. [5]Graduate School of Marine Science and Technology, Tokyo University of Marine Science and Technology, Minato, Tokyo, Japan. [6]Institute of Molecular and Cell Biology, A*STAR, 138673 Singapore, Singapore. [7]Department of Biology, University of Konstanz, 78464 Konstanz, Germany. [8]These authors contributed equally: Yali Liu, Meng Qu, Han Jiang, Ralf Schneider. ✉e-mail: oroth@zoologie.uni-kiel.de; axel.meyer@uni-konstanz.de; linqiang@scsio.ac.cn

more derived organ for paternal care and is the most complex structure in their family to protect and nourish embryos[9]. Female seahorses transfer eggs during mating into the males' brood pouches where embryos are implanted and nourished by a "pseudoplacenta". Its function is analogous to a mammalian maternal placenta and provides nutrients and oxygen to the developing embryos that hatch inside the males' pouch[10–12] (Fig. 1a).

In vertebrates, viviparity in females—with the unique exception of the sex-role reversed seahorses' male pregnancy—has evolved over 150 times independently[13,14]. While pregnancy provides advantages to the developing offspring, allowing them to be better protected from early-life predation and to be released at an advanced life-history stage, it poses an immunological challenge for the pregnant parent: How are the semi-allogenic embryos immunologically tolerated? As a solution to this immunological challenge mammalian embryos reduce the diversity of MHCI molecules expressed on trophoblasts, which constitute the cell layer in direct contact with maternal tissue[4,15]. In contrast, in seahorses, the co-evolution of the immune system with male pregnancy remains largely unknown. In addition to the spleen, some other important parts of the adaptive immune system's genetic repertoire are absent in seahorses, and these secondary losses have been hypothesized to be linked to the evolutionary novelty that is male pregnancy[5,6].

In an effort to study the immune-related changes during the evolution of seahorse male pregnancy, we comparatively analyze the genomes of two de novo sequenced seahorses together with other teleost genomes that had been previously published, focusing on the

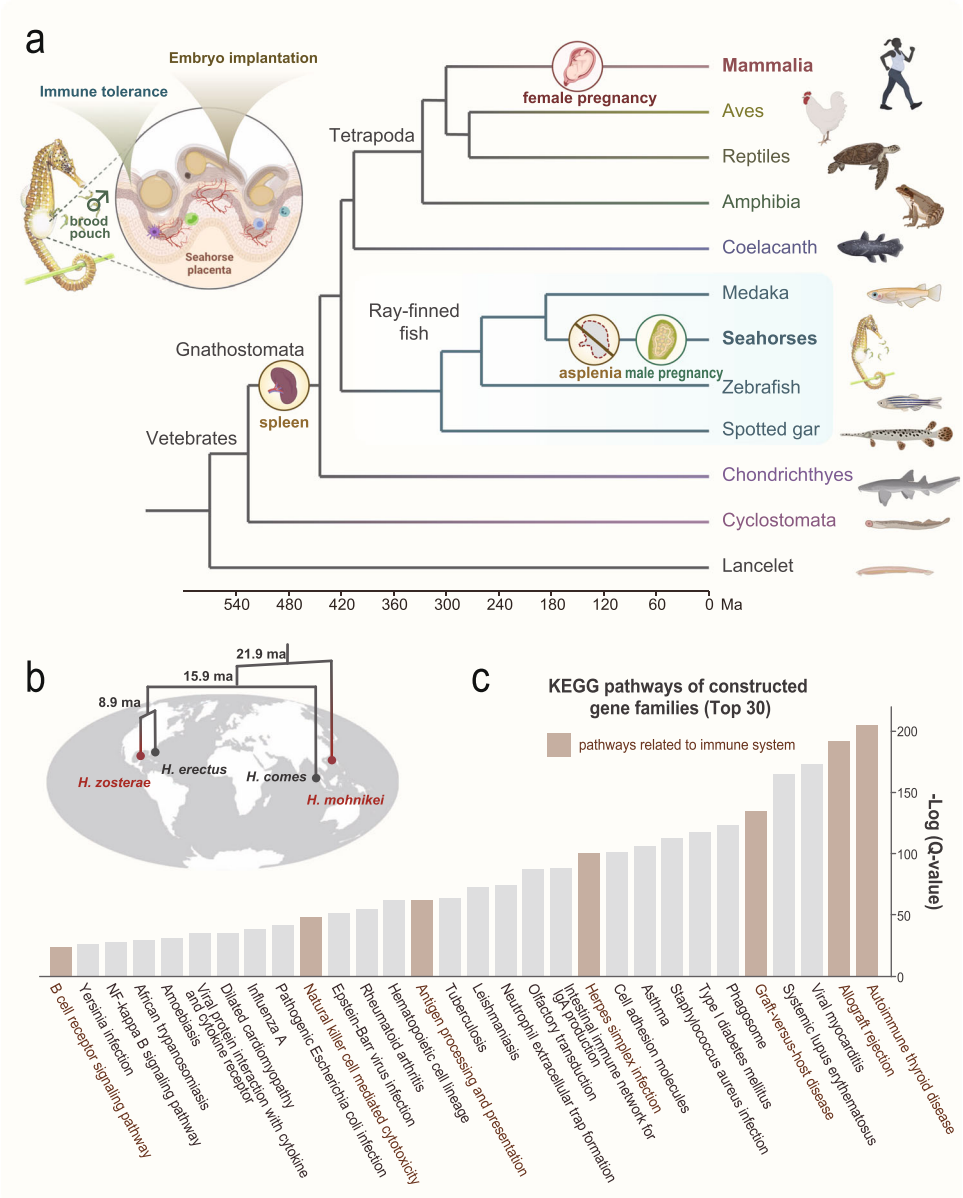

Fig. 1 | Unique features of immunity and reproduction in seahorses. a A species tree showing the evolution of specialized male pregnancy and asplenia traits in the seahorse. During pregnancy, embryos implanted in the "pseudoplacenta" of male seahorses are recognized by the paternal immune system. **b** Four seahorse species, from the major lineages of seahorses were included in the comparative genomic analyses. Newly sequenced species are indicated in red. Ma, million years ago. **c** The top 30 KEGG pathways of contracted gene families in seahorses. Categories involved in immunity are colored in brown. The figures were created with BioRender.com. The used map was downloaded from a free world map website (https://www.freeworldmaps.net/outline/maps.html). Source data are provided as a Source Data file.

https://doi.org/10.1038/s41467-022-35338-7

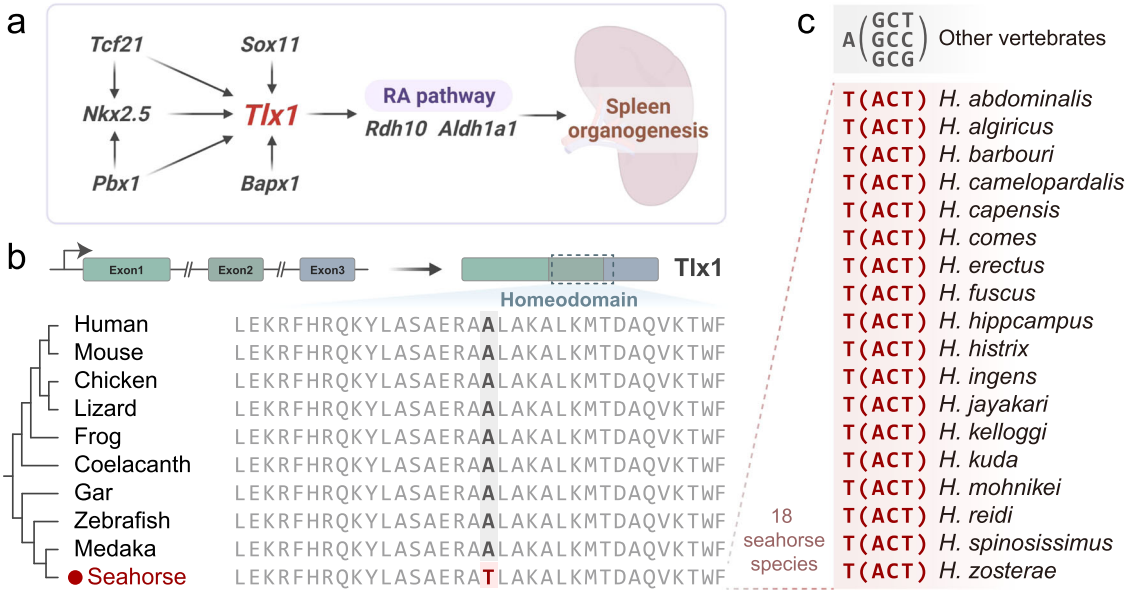

Fig. 2 | Seahorse-specific mutation of the *tlx1* gene. a Schematic diagram of the involved genes during spleen growth and morphogenesis (modified from previous studies[18]). Red color Tlx1, the sequence-specific variation in seahorses; RA, retinoic acid. **b** Gene structure (upper panel) and multiple amino acid alignments (lower panel) of Tlx1. Seahorse harbors a missense mutation of A (Alanine, Ala) to T (Threonine, Thr) in the homeodomain. **c** Extended Data examination of the identified *tlx1* mutation in 18 seahorse species. All detected seahorses exhibit T (ACT) instead of A (GCT or GCC or GCG), unlike other vertebrates. The figure was created with BioRender.com.

evolutionary innovation of the immune-related genes (Fig. 1b and Supplementary Table 1). We discover that a single amino acid mutation in the *tlx1* transcription factor is most likely causal for the most drastic change, the loss of the spleen (termed "asplenia") in seahorses. This we verify by knockout and missense mutations of *tlx1* in zebrafish. Novel insights into the modified immune system, including the complement system and immunoglobulins in this lineage support its previously hypothesized link to the evolutionary origin of male pregnancy.

Results

Contraction and loss of immune-related gene families in seahorse genomes

We generated chromosome-level reference genomes of *Hippocampus zosterae* (2n = 44) and *H. mohnikei* spanning 458 Mb and 527 Mb, respectively. With contig N50 of 7.3 Mb and 18.1 Mb and nearly complete BUSCO genes (92.95% and 93.46%), we succeeded in generating high-quality assemblies (Supplementary Table 5, 6, Supplementary Note 1). By including the already published genomes of *H. comes*[16] and *H. erectus*[17], we could comprehensively conduct comparative genomics analyses of four seahorse species (subfamily Syngnathinae) that evolved male pregnancy (Fig. 1b) by comparison with nine other teleost genomes from different orders (Supplementary Note 2). We show that 1260 gene families have undergone significant contraction in the seahorse lineage (Supplementary Fig. 4a and Supplementary Data 1–2). KEGG pathway enrichment analysis of the contracted gene families revealed signatures related to immune response pathways, such as the autoimmune thyroid disease, allograft rejection, and antigen processing and presentation (Fig. 1c and Supplementary Table 13), which was supported by a gene loss analysis (Supplementary Fig. 4b and Supplementary Data 1). The contraction or loss of immune-related gene families have substantially contributed to the modifications identified in the seahorse immune system and are presumably linked to the unique evolution of male pregnancy and its brood pouch structure (Supplementary Note 3). As to none of these genes, a role in spleen development can be assigned, their loss cannot be the molecular basis for asplenia in seahorses.

Lineage-specific missense mutation of *tlx1* in *Hippocampus* and *Syngnathus*

Our comparative genomic analysis identified genes undergoing positive selection, rapid evolution, or containing lineage-specific mutations in seahorses (Supplementary Data 4–8). Spleen development is mediated by the precise regulation of several transcription factors, including *Pbx1*, *Tlx1*, *Nkx3-2*, and *Nkx2-5*[18–20] (Fig. 2a). In the homeobox domain of *tlx1* (initially mistaken for *Hox11*[19,21]) a seahorse-specific mutation was identified (Fig. 2a and Supplementary Figs. 8–10). *Tlx1*, which contains three exons and a homeobox domain (Fig. 2b), has been reported to control splenic primordia cell fate specification and organ expansion. It is the only known gene whose pseudofunctionalization results in spleen loss without causing other developmental abnormalities[19]. To validate the specificity of the mutation, we extended our analyses to a larger phylogenetic range (34 species), including mammals, birds, reptiles, amphibians, and additional fishes. The amino acid sequences of *tlx1's* homeobox were identical in all vertebrates except for a seahorse-specific mutation, where a hydrophilic threonine (T, Thr) has been replaced by the hydrophobic alanine (A, Ala) (Fig. 2b and Supplementary Fig. 11a). Further validation of *tlx1* exon2 in 18 seahorse species confirmed this mutation to be a common feature of all seahorses (Fig. 2c, Supplementary Figs. 12–13 and Supplementary Note 4). The family *Syngnathidae* is a large (>350 species) and diverse clade of morphologically unique teleosts. They can be divided into two subfamilies: the *Nerophinae* and the *Syngnathinae*[22,23]. When assessing the mutation's locus also in all other available members of the subfamily Syngnathinae, we detected that all investigated members of the genus *Syngnathus* (all with quite derived, closed brooding organs and closely related to seahorses[8]) share the described mutation. In contrast, more distantly related members with less complex brooding organs, such as the seadragon and the alligator pipefish – where the eggs are simply attached to the ventral side of the males' tail[8] – have retained the ancestral Alanine in this position. As for the subfamily *Nerophinae*, we found the TLX1 sequences exhibit amino acid substitutions of A to L and A to I in *Oostethus manadensis* and *Nerophis ophidion*, respectively (Supplementary Fig. 11b). We also

provided the splenic phenotype for a number of syngnathid species using morphological, histological, and Micro CT methods. Dissections showed that species of the genera *Hippocampus* and *Syngnathus* (both belonging to Syngnathinae) have evolutionarily lost an unambiguous spleen, but not *Syngnathoides biaculeatus* (belongs to the subfamily *Syngnathinae*) nor *Nerophis ophidion* (that belong to the subfamily *Nerophinae*) (Supplementary Fig. 14). As mutations in exons of protein-coding genes can lead to substantial phenotypic changes, we hypothesized that the identified mutation in the conserved homeobox domain of *tlx1* might be associated with the loss of the spleen in the *Hippocampus* and *Syngnathus* species.

Missense mutation of *tlx1* leads to an asplenia phenotype

To test this hypothesis, we conducted CRISPR/Cas9-mediated genome editing to generate two zebrafish lineages: the full gene knockout line (*tlx1*$^{▲}$) and the specific point-mutation line (*tlx1*^{A208T}) mimicking the seahorse missense mutation (Fig. 3a, Supplementary Fig. 15 and Supplementary Note 5). As shown in Fig. 3b and c, both *tlx1*$^{▲}$ and *tlx1*^{A208T} zebrafish were found to lose their spleen in all examined individuals, suggesting that the seahorse-specific mutation is causally linked to a functional alteration of *tlx1*, which leads to functional asplenia. Whether this mutation originally caused functional asplenia in this syngnathid lineage, or whether another mutation was causal, removed stabilizing selection from genes involved in spleen development, and *tlx1* mutated only then, cannot be resolved using our data-set. Previous studies in both mammals and zebrafish have shown that *tlx1* knockouts result in congenital asplenia[19,24], indicating a crucial and conserved vertebrate developmental pathway. In an additional zebrafish lineage, another A to T point mutation was produced at the site adjacent to the seahorse-specific mutation as a control point mutation. Correspondingly, the *tlx1*^{A207T} line had a normal spleen, similar to that in the wild-type zebrafish, lending further support to the hypothesis that the A208T mutation in seahorses might be causal for the loss of the spleen (Fig. 3b, c and Supplementary Figs. 16–18).

The transcription factor *tlx1* contributes not only to spleen organogenesis but also has a crucial role in brain development and function[25]. In situ experiments of seahorse *tlx1* showed expression in the hindbrain and pharyngeal arches in embryos, similar to that in mice and zebrafish during embryogenesis (Supplementary Fig. 19a and Supplementary Note 6)[19,21,24]. To investigate the effects of seahorse-specific mutants and compare them with the full gene knockout line, transcriptomic-wide gene expression analyses (RNAseq) of the brain and three immune-related organs (liver, kidney, and intestine) were performed (Supplementary Note 7). Compared to the *tlx1*$^{▲}$ line, the *tlx1*^{A208T} line always shows transcriptomic patterns more similar to those of the wild type, with a lower number of differentially expressed genes (DEGs) in all tested tissues (Fig. 3d, Supplementary Fig. 19b–e and Supplementary Data 11). Compared with the wild-type zebrafish, the asplenia zebrafish (*tlx1*^{A208T}) exhibited differential gene expression profiles of genes involved in the MHC and complement pathways, e.g., *mhc1zka* and *ciita* downregulated, while *chia*, *C3a.2*, and *C9* were upregulated specifically in the kidney (Supplementary Fig. 19i–k).

In addition, most DEGs were specific for either *tlx1*$^{▲}$ or *tlx1*^{A208T}, except for a small number shared by both lines (75 genes, brain), indicating that more severe effects are caused by the full gene knockout compared to the point mutation (Fig. 3e, Supplementary Fig. 19f–h and Supplementary Data 11). GO enrichment analysis of DEGs in the brain between *tlx1*$^{▲}$ and *tlx1*^{A208T} lines revealed signatures in the membrane component, oxidase enzyme activity, and biosynthetic processes (Fig. 3f, Supplementary Fig. 20 and Supplementary Data 12), which are vital for brain development and function[26,27]. Our results suggest that although *tlx1*$^{▲}$ and *tlx1*^{A208T} zebrafish had the same asplenia phenotype, genome editing led to different effects on other organs.

Evolutionary immunogenomic modification of the unique male pregnancy

In an effort to test the hypothesis that regressive evolution of the immune system is linked to the evolution of unique male-pregnancy, we scanned genomes for modifications in immune-related genes involved in pregnancy informed by previous publications[15,28–32] and recorded their complete loss, reduced copy number or other noticeable sequence variation (Fig. 4a and Supplementary Note 8). In this study, we observed ubiquitous presence of MHCII molecules (three copies of MHC IIa and three to four copies of MHC IIb) and a conspicuous loss of not only *cd8b*[6], but also *batf3* in seahorses (Fig. 4a, Supplementary Figs. 21a, 22). Meanwhile, *il12a/b* and *ifng* genes, which encode the cytokines that are important for the alternative Batf3-independent pathway[33], are present in seahorses and pipefishes (Fig. 4a).

The spleen is essential for producing B-1a B lymphocytes, which are responsible for the production of natural antibodies in mammals[34]. Importantly, also *cd5*, the gene homologous to mammalian CD5, which encodes the surface molecule of B-1a B cells, was lost in seahorses (Fig. 4, Supplementary Figs. 21b, 23). In addition, we found that one of the necessary immunoreceptor tyrosine-based activation motifs (ITAMs) in *cd79a* is absent in seahorses (Fig. 4, Supplementary Figs. 24–25). As a vital element of the B cell receptor complex, the structural variation of *cd79a* might change the signaling ability of B cell receptors[35]. The hallmark of the vertebrate adaptive immune system is the somatic diversification of the immunoglobulin family (VDJ-rearrangement) as an advancement of the innate immune system[36]. Herein, we found a significantly reduced number of V domains in the heavy chain coding genes (*ighv)* in seahorses and other syngnathid fishes (only 2-20) compared to other teleosts, including tilapia, platyfish, medaka, stickleback, fugu, zebrafish and spotted gar (≥35) (Fig. 4 and Supplementary Fig. 26). In addition, antibodies can induce rejection of non-self through activation of the classical complement system, especially the complement component 4d (C4d)[37]. In mammals, several pregnancy complications are associated with excessive or misdirected activation of the complement system[38]. In this study, we found only two copies of the *C3* gene in seahorses - fewer than in all other teleosts (three to eight copies) - while *C4* is entirely absent (Fig. 4, Supplementary Figs. 21c, 28a). Finally, in mammals, immune tolerance is enhanced by increasing Treg cells during pregnancy[39]. Our study revealed that *foxp3* - a crucial regulator in the establishment and maintenance of Treg phenotypes[40] - was also lost in seahorses (Fig. 4 and Supplementary Fig. 28b). Moreover, several vital genes involved in pregnancy including *cd8a*, *T-bet*, *C3*, *il12a*, *il12b* and *cd79a* also showed different expression patterns during seahorses' pregnancy, further indicating the complex and unique immune features of male pregnancy (Supplementary Fig. 29). In addition, an expanded comparative approach revealed that asplenia and male pregnancy characters evolved independently but cooccurred on the same branches (*Hippocampus* and *Syngnathus*) (Supplementary Fig. 30).

Discussion

Seahorses possess the most advanced form of full internal male pregnancy compared to other syngnathids, and simultaneously, exhibit asplenia. A pregnant male seahorse faces the dilemma to immunologically defend both itself and the embryo against prevailing pathogens, while concurrently the semi-allogenic embryo has to be immunologically tolerated[41]. Pipefishes and seahorses adapted their immune biology encompassing modifications and losses of immune-related genes, e.g. CD8b and MHC II pathway genes, which illustrates the remarkable flexibility of the vertebrate immune system in general[6]. Pregnant males further change the expression of immune genes like *ptgs2*, *pla2g4a*, *chia* and *b2m*[6,42] presumably supporting the immunological tolerance of the embryo. As a vital part of the adaptive immune system, the lack of the MHC class II pathway apparently does not

https://doi.org/10.1038/s41467-022-35338-7

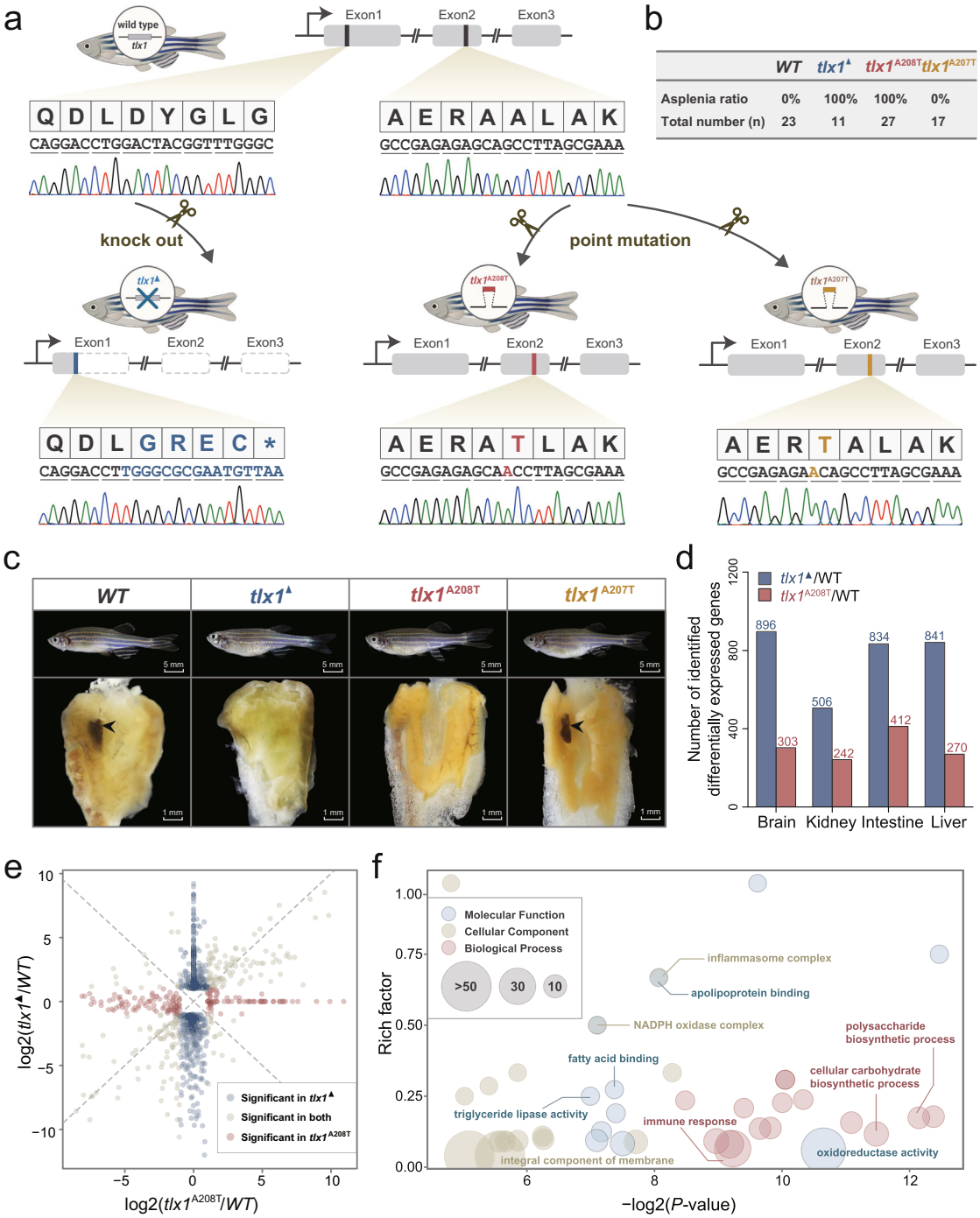

Fig. 3 | Missense mutation of *tlx1* controls asplenia phenotype in zebrafish. **a** Genome editing of three zebrafish lines (*tlx1*▲, *tlx1*^{A208T} and *tlx1*^{A207T}) using CRISPR/Cas9 technology. Procedures are detailed in Supplementary Fig. 15. ▲ represents full knockout of the gene, while ^{A208T} and ^{A207T} indicate the specific point mutations at sites 622 and 619 of zebrafish *tlx1*, respectively. Blue, *tlx1*▲; Red, *tlx1*^{A208T}; Yellow, *tlx1*^{A207T}. The *tlx1* nucleotide sequence of genomic DNA of three zebrafish lines was validated by Sanger sequencing. The figure was created with BioRender.com. **b**, **c** *tlx1*▲ and *tlx1*^{A208T} zebrafish exhibited asplenia while *tlx1*^{A207T} zebrafish had an intact spleen; the asplenia ratio of individuals was calculated and listed ($n = 11$–23). The splenic phenotype of other 74 individuals were listed in Supplementary Figs. 16–18. WT wild type. **d** Differentially expressed gene (DEG) number of *tlx1*▲ and *tlx1*^{A208T} compared to that of wild-type individuals via comparisons of the brain-, kidney-, intestine-, and liver-transcriptomes. **e** Volcano map of the shared and group-specific DEGs in the brain of *tlx1*▲ and *tlx1*^{A208T} compared to that in the wild-type group. Blue, red, and gray represent significantly changed DEGs in *tlx1*▲, *tlx1*^{A208T}, and both groups, respectively. **f** GO enrichment of the DEGs between *tlx1*▲ and *tlx1*^{A208T} illustrate the different in gene expression patterns associated with brain development and function. The enrichment was conducted using the GOseq R package, and corrected $P < 0.05$ indicated significant enrichment. The size of the circle represents the number of genes in each category. Source data are provided as a Source Data file.

always lead to asplenia, since spleens are present in other species that lost MHCII, as is known from Gadiformes (e.g., Atlantic cod) and the non-parasitizing anglerfish *Lophius piscatorius*[43–46]. In addition, as an important secondary lymphoid organ that performs mostly immunological functions, the loss of a spleen - mediated by knockout of *tlx1* - is linked to an immunological deficiency in mice and zebrafish[19,24,47]. Our results strongly suggest that a seahorse-specific mutation in *tlx1* leads to asplenia, as confirmed by CRISPR-Cas9

https://doi.org/10.1038/s41467-022-35338-7

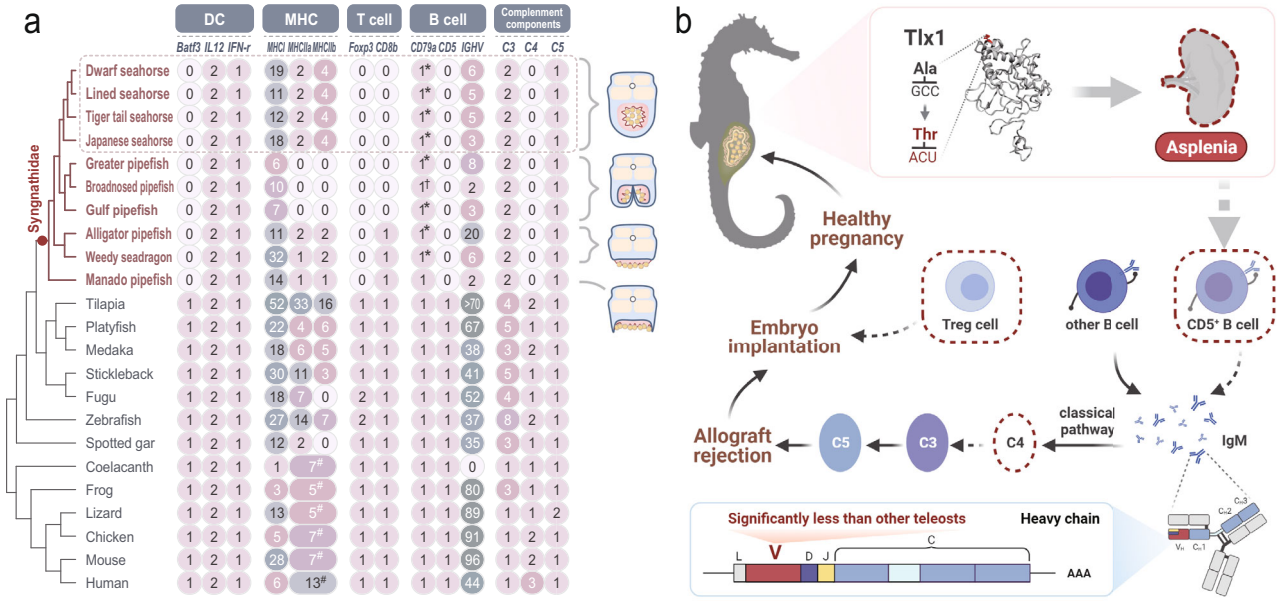

Fig. 4 | Immunogenomic basis of asplenia and male pregnancy in seahorses. **a** Copy numbers of key genes involved in the development and function of DC, MHC, T/B lymphocytes, and complement components 3–4 in seahorses and other vertebrates. DC dendritic cell, MHC major histocompatibility complex. *, variation of ITAM (immunoreceptor tyrosine-based activation motif) region; †, based on partial sequence (low-coverage sequencing and genome assemblies), but featuring the variation of ITAM; #, Total number of MHC II. The brood pouch types are shown on the right. *Hippocampus* and *Syngnathus* species exhibit the closed brood pouches. **b** Schematic map illustrating the hypothesized molecular trade-offs in the male pregnancy of seahorses. Asplenia caused by the *tlx1* mutation could decrease the CD5⁺ B cell populations. In addition to the corresponding reduced diversity in antibodies, the classical complement pathway triggered by antibody-antigen-interactions might also be weakened, affecting the downstream events usually triggering allograft rejection. Loss of *C4* is expected to be associated with decreased abundance of Treg cells, which orchestrate self-immune tolerance. The shaded icon and dotted line indicate gene loss in seahorse genomes. Treg cell, regulatory T cell. The figures were created with BioRender.com.

experiments. This finding confirms that *tlx1* is part of a non-redundant transcription factor network required for the development of the spleen, and also attests to the evolutionary stability of transcription factor networks in regulating organ development between teleosts and mammals.

Immunological adaptations are key in the evolution of extraordinary reproduction strategies. Female pregnancy evolved convergently over 100 times in vertebrates and coevolved by fine-tuning of the immune system. Similarly, in deep-sea angler fishes, the loss of parts of their immune system crucially permitted the integration of the male body into the female body[48]. Male pregnancy has evolved along a gradient of increasing parental care complexity, ranging from simple egg attachment to the males' ventral side (e.g., *Nerophina*) to the complexity seen in *Syngnathus* and *Hippocampus* that evolved brood pouches with intricate placenta-like structures. The loss of the identified immunogenomic features in syngnathid fishes emerged concomitantly with the evolution of increasing parental care complexity, as was previously shown for the loss of MHC II and *cd8b*[6]. Combined with the here identified loss of *batf3*, this suggests a modification of antigen processing and antigen-specific immunity in seahorses. Alternatively, however, seahorse male pregnancy may also rely on alternative dendritic cell regulation via *batf*, *il12* and *ifng* - genes which are present in all syngnathids, as *cd8a* likely compensates for the loss of its paralog[6]. Coincidentally, a previous study on single-cell transcriptomics of immune cells in *Syngnathus typhle* have identified a number of MHC I-related pathway constituents and cytotoxic-related genes, which supports the presence of the CD8 + T-cell subset in pipefish[49]. The reduced copy number of *C3* and the loss of *C4* genes in seahorses further suggest an impeded activation of the complement system, which can reduce antibody-mediated allograft recognition and inflammation. Along the same lines, our current data add to the hypothesis that syngnathid immune biology might evolve synchronously with their unique male pregnancy[16].

The main responsibility for the transient T cell tolerance specific for paternal antigens during pregnancy was assigned to regulatory T cells (Tregs), as these enhance immune tolerance during pregnancy[39]. Asplenia might also be directly or indirectly linked to the loss of *foxp3*, as splenectomy patients exhibit a decrease of blood FOXP3⁺ Treg cells[40], implying a potential link between the spleen and the abundance of FOXP3⁺ Treg cells. In addition, sex-dependent immunological differences affect immune responses to both self and foreign antigens, and Treg numbers are associated not only with sex steroid levels but also with sex[50]. The loss of *foxp3* indicates that seahorses seem to have adopted a unique evolutionary strategy for their unparalleled sex-role reversal brooding and nourishing paternal care. Indeed, as a conserved organ, the loss of the spleen may not only be accompanied by the evolution of male pregnancy but may also bring about changes in some other traits, such as the observed immune response changes (including the lost/contracted genes like *batf3*, *C4*, *C3*, *CD5*, and *foxp3*). Hence, our speculation that the synchronized evolution of asplenia and male pregnancy may not be accidental but instead an evolutionary link exists. In addition, as demonstrated here, this reduction in the immune repertoire is shared across the family *Syngnathidae* and is not unique to seahorses; thus, further analyses will be necessary to evaluate the immune response strategy in species with different levels of pregnancy complexity in the future.

Seahorses represent a single lineage within the Syngnathidae, a family with >350 species of which many show less derived forms of male pregnancy - potentially facing pregnancy-related immune challenges that seahorses seemingly have overcome. Moreover, much of the syngnathids' diversity still remains unexplored to date and the variations in morphological and physiological complexity of immune and brooding organs have only been described superficially[9]. Functional redundancy in the immune system makes it hard to predict the consequences that gene loss might have on an organism. However, it is also this redundancy that provides the evolutionary opportunity to

organisms to tinker with the regulatory pathways orchestrating an effective immune response while accommodating for physiological challenges, such as pregnancy, resulting in an unexpected genomic flexibility of vertebrate immune systems. For now our understanding of the intricate immune challenges syngnathids face at different stages of male pregnancy is still incomplete, our study sheds light on the genetic basis of a major adaptive immunological novelty putatively associated with the evolution of sex-role reversal and male pregnancy in seahorses. Considering the entire syngnathid phylogeny, there is still deep divergence between the *Hippocampus* and *Syngnathus* lineages[23], and the splenic phenotype and *tlx1* gene sequence information of lineages more closely related to *Hippocampus*, which are missing in our study, would improve the resolution of the dataset and thus allow us to formulate better informed conclusions. Therefore, we encourage further studies filling these gaps in genomic knowledge.

Methods

Animal use ethics

All animal experiments were conducted per the guidelines and approval of the respective Animal Research and Ethics Committees of the South China Sea Institute of Oceanology, Chinese Academy of Sciences.

Hippocampus spp. specimens and nucleic acid sample preparation

Two individuals, one adult male Japanese seahorse *H. mohnikei* and one adult male dwarf seahorse *H. zosterae*, were used for whole-genome sequencing in this study. *H. mohnikei* was collected from local markets in Rizhao (Shandong, P.R. China) in 2017. *H. zosterae* was donated by aquariums in Guangzhou to the Marine Biodiversity Collections of South China Sea, Chinese Academy of Science in 2015 (Supplementary Table 1). Genomic DNA was extracted from each sample using a standard phenol-chloroform protocol.

Genome sequencing and Hi-C sequencing

Two paired-end libraries with 350 bp insert size were constructed for seahorses using the Illumina HiSeq 2500 system (San Diego, USA). The quality-filtered reads were used for genome size estimation via the *K-mer* method[51]. The genome size was then estimated as follows: Genome Size = K_{num}/K_{depth}. We also calculated and plotted the 19-mer depth distribution. In addition, Nanopore sequencing was performed for *H. mohnikei*. Briefly, libraries were constructed and sequenced on R9.4 FlowCells using the MinION sequencer (ONT, UK). All sequencing was performed by BioMarker Technologies Company (Beijing, China). PacBio sequencing was performed for *H. zosterae*. The genomic DNA of the sample was sheared to an average size of 20 Kb using a g-TUBE device (Covaris, Woburn, MA, USA). The sheared DNA was purified and end-repaired using polishing enzymes, followed by blunt end ligation reaction and exonuclease treatment to create a SMRTbell template according to the PacBio 20-kb template preparation protocol. A BluePippin device (Sage Science, Beverly, USA) was used to size-select the SMRTbell template and enrich large (>10 Kbp) fragments. Single-molecule sequencing was then conducted on a PacBio Sequel platform to generate long-read data.

A blood sample of the male *H. zosterae* was also used for Hi-C sequencing by the Illumina HiSeq 2500 platform, using paired-end of 150-bp reads. DNA was isolated from the sample and the fixed chromatin was digested with the restriction enzyme DpnII overnight. Subsequently, the DNA was sheared by sonication to the mean size of 350 bp. Hi-C libraries were generated using NEBNext Ultra enzymes and Illumina-compatible adaptors. Biotin-containing fragments were isolated using streptavidin beads. All libraries were quantified by Qubit2.0, and the insert size was checked using an Agilent 2100. The libraries were then quantified by qPCR. The Hi-C data were mapped to PacBio-based contigs using BWA (version 0.7.10-r789; mapping method: aln). Uniquely mapped data were used for chromosome-level scaffolding. HiC-Pro (version 2.8.1)[52] was used for duplicate removal and quality controls, and the remaining reads were valid interaction pairs for further assembly.

Genome assembly

PacBio subreads were corrected and trimmed using Canu (version 1.5, available at https://github.com/marbl/canu)[53]. wtdbg generated a draft assembly with the command 'wtdbg -i pbreads.fasta -t 40 -H -k 21 -S 1.02 -e 3 -o wtdbg' using error-corrected reads from Canu. A consensus assembly was obtained with the command 'wtdbg-cns -t 40 -i wtdbg.ctg.lay -o wtdbg.ctg.lay.fa -k 15'. Next, Illumina paired-end reads were also aligned for consensus assembly using BWA (version 0.7.10-r789; mapping method: MEM) and the polishing step was performed with the Pilon (version 1.22) software, with the following parameters: --mindepth 10 --changes --threads 4 --fix bases. The polishing step was iterated two times. The assembly statistics are shown in Supplementary Table 5. Benchmarking Universal Single-Copy Orthologs (BUSCOs) assessment showed that our assembly captured 93.46% and 92.95% of complete BUSCOs of *H. mohnikei* and *H. zosterae*, respectively.

Using contigs assembled from the PacBio data, Hi-C data were used to correct misjoins in contigs, order, and orient contigs. Pre-assembly was performed for contig correction by splitting the contigs into segments with an average length of 300 kb. The segments were then preassembled with Hi-C data. Misassembled points were defined and broken when split segments could not be placed to the original position. Next, the corrected contigs were assembled using LACHESIS with parameters CLUSTER_MIN_RE_SITES = 225, CLUSTER_MAX_LINK_DENSITY = 2, ORDER_MIN_N_RES_IN_TRUN = 105, and ORDER_MIN_N_RES_IN_SHREDS = 105 with Hi-C valid pairs. Gaps between ordered contigs were filled with 100 'N's. The sequence interaction matrices are shown in Supplementary Fig. 2, and the statistical analysis results of chromosome assemblies are summarized in Supplementary Table 6.

Genome annotation

De novo identification of repeats and transposable elements were performed using the PILER-DF[54] and the RepeatScout[55] under default parameters. RepeatMasker and RepeatProteinMask (version 3.3.0) were employed to identify transposable elements (TEs) based on homology searches against the Repbase library (release 16.03) using the parameters "-nolow -no_is -norna -parallel 1" and "-noLowSimple –pvalue 1e-4". Ab initio gene prediction was performed using three programs, namely Augustus[56], GlimmerHMM[57], and SNAP[58]. The GeMoMa[59] program was run for homology-based prediction by aligning the assembled genome against *O. latipes*, *S. salar*, *H. comes*, and *H. erectus*. Next, the transcriptome data of *H. erectus* from our previous work[16] were downloaded and mapped onto the genome, and gene prediction was performed by TransDecoder (http://transdecoder.github.io) and GeneMarkS-T. PASA[60] was used to predict the Unigene sequences without reference assembly based on transcriptome data. We then combined the results from these three methods with EvidenceModeler[61]. Gene functions were further annotated by searching publicly available databases, including the NR, KOG, KEGG, GO, and TrEMBL databases.

Phylogenetic analysis

Protein datasets were obtained from Ensembl-FTP release-96 [*T. rubripes*, *G. aculeatus*, *O. latipes*, *O. niloticus*, *P. magnuspinnatus*, *D. rerio*, *X. maculatus*, and *L. oculatus*] or other sources [*G. morhua*[62], *H. comes*[16], *H erectus*[17], *H. zosterae*, and *H. mohnikei* (present study)]. One-to-one orthologues were identified using OrthoFinder version 2.2.7 at default settings from the 13 ray-finned fish species. Protein sequences for these one-to-one orthologues were then extracted into their

respective orthogroups using an in-house Perl script. Multiple alignments were generated for each of the orthogroups using MAFFT (version 7.475). The individual protein alignments were concatenated together using an in-house Perl script. A trimmed concatenated alignment was generated using Gblocks 0.91b with the 'allowed gap positions' set to "With Half". ModelFinder[63] was used to deduce the best-suited substitution model for the trimmed alignment (JTT + F + I + G4 model). For the maximum likelihood analysis, we employed the best fit substitution model as deduced by ModelFinder and 1000 replicates using the ultrafast bootstrap approximation approach[64] as implemented in IQ-TREE version 1.6.10.

Expansion and contraction of gene families

Expansion and contraction in gene families were calculated by the CAFÉ program (v 3.1) based on the birth-and-death model[65]. The parameters "-p 0.01, -r 10000, -s" were set to search the birth and death parameter (λ) of genes based on a Monte Carlo resampling procedure, and birth and death parameters in gene families with a *P*-value ≤ 0.01 have been reported. The gene families without homology in the SWISS-PROT database were filtered out to reduce the potential false positive expansions or contractions caused by gene prediction. Go and KEGG terms of all proteins used in the comparative analysis were annotated with eggnog 5.0[66]. The gene family with more than 90% of its members sharing the same annotations was considered as the single functional family and its weighting was set to 1. For gene families containing sequences having multiple functional annotations, different weighting values were assigned to each functional annotations according to the ratio of the annotated times of each term to the total annotated times of all members. The total weighting value were 1 for each family.

Gene loss

When a gene has no homologs within the seahorse clade, but the homologs of that gene are present in the closest sister lineage of the seahorse clade, we consider that the gene was lost in seahorses[67]. One-to-one orthologous genes were extracted from each species, and multiple sequence alignments were generated accordingly. The gene loss analysis was performed using an in-house script in *R*.

Positive selection of genes (PSGs)

We tested for PSGs in the seahorse lineages compared to all other background species. Orthologous genes were extracted from the same species selected for the gene family expansion/contraction analysis. Multiple sequence alignments were generated using the MUSCLE software (v3.8.31). Positive selection analyses were conducted with the branch-site model using PAML[68]. We compared model A (allows sites to be under positive selection; fix_omega = 0) with the null model A1 (sites may evolve neutrally or under purifying selection; fix_omega = 1 and omega = 1) via likelihood ratio test using the Codeml program in PAML. The significance of the compared likelihood ratios was evaluated by χ2 tests from PAML. Then the p.adjust function embedded in the R language was used to adjust the *P* value using all the original *p*-values. Only the genes with an FDR-corrected *P* value smaller than 0.05 were considered positively selected.

Rapidly evolving genes (REGs)

The same orthologous genes and tree topology as those for PSGs were used in the analysis. The branch model in PAML was used, with the free-ratio model (model = 1) allowing the value to vary on each branch. The dN values of all nodes and tips on each tree were tested using the t.test function in R programming and then the *p*-values were adusted using FDR. Genes with an FDR-corrected *P* value smaller than 0.05 and the dN of the seahorse lineage higher than that of the sister lineage of seahorse were considered as rapidly evolved on seahorse lineage. GO and KEGG enrichment analyses of REGs were then Performed.

Lineage-specific mutated genes (LSGs)

To identify LSGs in the seahorses, the same orthologous genes and tree topology produced from the PSGs analyses were used. Only the ancestral state of the seahorse species was different from that of all other background species genes were recognized as the true LSGs. To ensure result reliability, only ancestral amino acids with posterior probabilities ≥ 0.95 were used to determine lineage-specific mutation. Moreover, we discarded lineage-specific mutated sites within poorly aligned fragments: we calculated the pairwise sequence similarities of 10 amino acids centering each lineage-specific mutated site. If the mean similarity was lower than 0.7 or the lowest was lower than 0.35, this fragment was defined as a poorly aligned fragment[69]. GO and KEGG enrichment analyses of LSGs were then performed.

Local gene synteny and phylogenetic analysis of the *tlx* gene families of the lined seahorse

Syntenic genes in other species (fruit fly, amphioxus, human, and zebrafish) were acquired using the Ensembl genome browser from Genomicus (https://www.genomicus.biologie. ens.fr/genomicus-100.01). The output was sorted for genomic contig/scaffold/linkage and then for start position within each of these genes. Next, phylogenetic analysis based on Tlx1, Tlx2, and Tlx3 amino acids sequences was performed in vertebrates including lined seahorse. A total of 41 species of cyclostomes, fishes (26 species from 16 orders), amphibians, reptiles and mammals were included. Protein sequences were aligned using MAFFT (version 7.475) with minor manual adjustments. A maximum-likelihood tree was generated with IQtree v.2. Boot-strap support was established through 1000 iterations for Ultra-Fast Bootstrapping and SH-like approximate likelihood ratio test. The consensus tree was visualized by the FigTree (version 1.4.4) software.

Alignment of Tlx1 amino acids in different lineages

Next, Tlx1 amino acid sequences from representative vertebrates were compared using Clustal W 2.1. The GenBank accession numbers of these genes are listed in Supplementary Table 16. The results showed that the amino acid sequences of the homeobox were completely identical among all vertebrates, except for the seahorse-specific mutation in which the hydrophilic threonine (T) replaced the hydrophobic alanine (A) exclusively in the seahorse lineage (Supplementary Fig. 10a). Next, we manually Extended Data the scope of comparison to a broader taxa selection, including mammals, birds, reptiles, amphibians, and fishes. A total of 38 vertebrates containing 23 fish species from 15 orders were included, and we confirmed that among these species, only seahorses harbored the lineage-specific mutation (Supplementary Fig. 11a).

Validation of Tlx1 amino acids in 18 seahorse species and other Syngnathidae fishes

Further PCR analysis was conducted using a total of 18 seahorse species to validate this lineage-specific mutation. Genomic DNA was extracted from 18 seahorse species, namely *H. abdominalis*, *H. algiricus*, *H. barbouri*, *H. camelopardalis*, *H. capensis*, *H. comes*, *H. erectus*, *H. fuscus*, *H. hippocampus*, *H. histrix*, *H. ingens*, *H. jayakari*, *H. kelloggi*, *H. kuda*, *H. mohnikei*, *H. reidi*, *H. spinosissimus*, and *H. zosterae*. The primers (F: 5′-TCTCACCGCTCACTGTAAC−3′; R: 5′-GGTCATTT TGAGGGCTTT−3′) were designed to amplify the second exon of *tlx1* in these seahorses. The PCR procedure was conducted using standard PCR conditions. The PCR products were loaded on a 1.5% agarose gel and electrophoresed at 130 V for 25 min. The results were photographed under a UV light. We also investigated the mutation locus by comparison with the genomes of all other available members of the Syngnathinae subfamily, including *S. acus*, *S. typhle*, *S. scovelli*, *S. biaculeatus*, *P. taeniolatus*, *O. manadensis* and *Nerophis ophidion*.

The splenic phenotype of the Syngnathidae

To clarify the splenic phenotype of the Syngnathidaes, we detected the most available samples (including *H. abdominalis*, *H. erectus*, *S. typhle*, *S. biaculeatus* and *Nerophis ophidion*). Briefly, after euthanasia with 300 ng/ml MS-222 (Sigma-Aldrich, USA), the individuals were dissected and imaged by a MVX10 camera (Olympus, Japan) or Olympus SZX2 Stereo Microscopes (Olympus, Japan). In addition, the Micro CT scan of the *S. biaculeatus* was conducted on the Nemo Micro-CT (NMC-200) of PINGSENG Healthcare Inc, a high-resolution imaging technology based on cone beam the CT principle. Firstly, samples were soaked in the contrast agent for 12 days and then were placed into the sample chamber vertically, whose scanning tube voltage and current were set to 90 kV and 90uA, respectively. During the scanning process, the detector and bulb rotated 360° around the central axis of the sample chamber and performed 10,000 projections in the scanning area, which lasted 1000 s. After capturing images by the detector, it was reversely reconstructed using the FDK method on Avatar software (version 1.7.2, PINGSENG Healthcare Inc.), with a pixel size of 10 um × 10 um × 16 um. Histological analyses of the *S. biaculeatus* spleens were conducted, then the transcriptomic profiles of the *S. biaculeatus* spleens and the *H. erectus* small white organ were also sampled and sequenced (Supplementary Data 9).

Generation of *tlx1*$^{▲}$ zebrafish lines

Strain *AB zebrafish (*D. rerio*) were maintained under a 14-h light (8:00-22:00)/10-h dark (22:00-8:00) cycle at 26–28 °C to induce spawning. CRISPR/Cas9-mediated mutagenesis was used for generating indels in zebrafish *tlx1* (ENSDARG00000003965; http://ensembl.org) using sites identified by ZiFiT Targeter (http://zifit.partners.org/ZiFiT/). Mutagenesis targeting two regions in *tlx1* exon 1, GGACCTGGACTACGGTTT (site 1) and GGTCCTACAACATGAACTT (site 2), was induced. Purified gRNAs (~80 pg) were co-injected with Cas9 mRNA (~400 pg) into zebrafish embryos (F0 fish) at the one-cell stage. Two days after injection, 8–10 embryos were collected for genomic DNA extraction to check whether the targeted genomic fragment was mutated. The target genomic regions were amplified by PCR and subcloned into the pTZ57R/T vector. The adult founders were outcrossed with wild-type fish to obtain F1 fish, which were subsequently genotyped and outcrossed with wild-type fish to yield F2 fish. Next, heterozygous F2 individuals were intercrossed to produce homozygous F3 fish. Finally, we generated two *tlx1* knockout zebrafish lines. Two *tlx1* nonsense alleles with 122-bp deletion *tlx1*$^{▲}$ (−122) and 5-bp deletion *tlx1*$^{▲}$ (−5) in the first exon were generated (the red arrowheads and the yellow-backed sequence), which caused frame-shift mutations at the position 43 and 47 AA, respectively (the red-backed sequence), as well as premature transcription termination event at the position 47 and 58 AA (the green-backed sequence) (Supplementary Data 10).

Generation of *tlx1*^{A208T} and *tlx1*^{A207T} zebrafish lines

CRISPR/Cas9-mediated homologous recombination (HR) was used for generating point mutations in zebrafish *tlx1*. We also used sites identified by ZiFiT Targeter (http://zifit.partners.org/ZiFiT/) to design CRISPR/Cas9 targets near the point mutation sites. The designed sites targeted two regions in *tlx1* exon 2, GGCGCGTGAACGACGTCCG (site 3) and GGAGGTGTTCGGTTCTGGTA (site 4). HR donor plasmids were constructed using the Hieff Clone Plus Multi One Step Cloning Kit (YEASEN, China). The G622A HR donor was constructed by ligating four fragments (a left arm, a middle arm, a selective marker, and a right arm) with the pHRHG vector (XinJia, China). The fragments of the left arm (1081 bps), middle arm (923 bp), and right arm (950 bp) were amplified from the genomic DNA of AB/WT zebrafish by using the PrimeSTAR HS DNA polymerase (Takara, Japan). The point mutation (G622A) in the middle arm was introduced using the Fast Mutagenesis System (Transgen, China). To prevent the HR donor from being excised by the CRISPR/Cas9 system, the silent mutations of the two CRISPR/Cas9 target sites in the middle arm was also introduced using the Fast Mutagenesis System (Transgen). The selective marker was amplified from a heart-specific transgene plasmid including *myl7* promoter, DsRed and SV40 polyA signal. Furthermore, *frt* sites were added to both sides of the selective marker, which could be removed through excision by Flp recombinase. The G619A HR donor was constructed similarly.

The G622A donor plasmid was purified before microinjection by the Gel Extraction Kit (Qiagen). The zCas9 protein, sgRNAs, and donor plasmid (15 pg) were co-injected into zebrafish embryos at the one-cell stage. Adult F0 zebrafish were outcrossed with wild-type fish to screen the founders with selective marker expression and correct genotyping. The founders were then outcrossed with wild-type fish to yield F2 fish, in which the selective marker was excised through injection of *flp* mRNA (50 pg) at the one-cell stage. Subsequently, heterozygous F2 individuals without selective marker were intercrossed to produce homozygous F3 fish. And the G619A line was also obtained similar to the G622A line. The primers were listed in Supplementary Data 10.

Genotyping and splenic phenotype detection

Zebrafish were euthanized with 200 ng/ml MS-222 (Sigma-Aldrich, USA) and genotyped via tail fin clipping. Genomic DNA was extracted from the fin tissues. Genotypes of the three mutation lines (*tlx1*$^{▲}$, *tlx1*^{A208T} and *tlx1*^{A207T}) were determined using standard PCR conditions and then sequenced by Sanger sequencing (Sangon Biotech Co., Ltd, China). The primers were as follows: *tlx1*$^{▲}$, F: 5′-ATCGTCTGTAGTTCCGTCTTTC-3′, R: 5′-ACCGCTTTAACCGCTGAGA-3′; *tlx1*^{A208T} and *tlx1*^{A207T}, F: 5′-ATTTGCCTTCCACTGCTTGG-3′, R: 5′-CCAGCTGACCTCACGGTTTAT-3′.

Whole-mounts of abdominal organs were carefully removed in accordance with a previous study by Xie et al.[24]. Briefly, after euthanasia with 200 ng/ml MS-222 (Sigma-Aldrich, USA), the whole-mounts of abdominal organs were carefully removed from wild-type (23 fish) zebrafish as well as *tlx1*$^{▲}$ (11 fish), *tlx1*^{A208T} (27 fish), and *tlx1*^{A207T} (17 fish) mutants (Fig. 3). The organs were then fixed in 4% PFA for 5 min. All zebrafish checked were sampled randomly and done blind with respect to mutant status. Representative images were obtained using a MVX10 camera (Olympus, Japan) (Fig. 3 and Supplementary Figs. 16–18).

Whole mount in situ hybridization

Lined seahorse embryos at different developmental stages (early stage (S1), ~6 days post fertilization, dpf; mid stage (S2), ~12 dpf; late stage (S3), ~18 dpf) were sampled and fixed in 4% paraformaldehyde (PFA)/phosphate-buffered saline (PBS) overnight at 4 °C, in accordance with the methods of Leerberg et al.[70]. Primers were designed for hybridization of the *tlx1* gene in lined seahorse (F: 5′-GATCACATGGGACTAGCGGCAC −3′; R: 5′- GGCGTAATACGACTCACTATAGGGGTGTTACAGTGAGCGGTGAGAGAG−3′). The RNA antisense probes were then yielded through in vitro transcription using the DIG RNA Labeling Kit (SP6/T7) (Roche, Mannheim, Germany). During hybridization, the seahorse embryos were digested with 10 μg/ml proteinase K solution for an appropriate time and subsequently fixed again in 4% PFA for 20 min. Next, the embryos were preincubated in antibody blocking solution for 1 h at 20 °C with PBSTw and then incubated for 2 h at 20 °C in 1:3000 dilution of anti-DIG-AP Fab. The samples were washed twice for 5 min each in PBSTw and developed in coloration solution (30 μl NBT/BCIP stock in 10 ml Coloration buffer) at 4 °C. The reaction was terminated by washing with 50% and 100% methanol for 5 min each. The embryos were then fixed with 4% PFA overnight in darkness, mounted in 100% glycerol, and imaged under an Olympus SZX2 Stereo Microscopes (Olympus, Japan).

Transcriptome analyses

Transcriptome analysis of the brain and three immune-related organs (the liver, kidney, and intestine) of zebrafish. RNA sequencing libraries

of these tissues (each tissue including four biological replicates) of the wild-type, *tlx1*▲, and *tlx1*^{A208T} zebrafishes were constructed, and then sequenced on a flow cell using an Illumina HiSeq™ 2500 platform (Supplementary Table 17). Raw data (raw reads) of FASTQ format of these four tissues were first processed using in-house Perl scripts. Clean reads were mapped to the zebrafish genome assembly (GRCz11) using the TopHat2 program. Gene expression levels were estimated with fragments per kilo-base of exon per million fragments (FPKM values) using the Cufflinks program. Differential expression genes (DEGs) were checked using DEGSeq2 and identified based on corrected *p*-values (*Q*-value) and false discovery rates. Pair-wise comparisons across 12 combinations (brain: *tlx1*▲ vs. WT, *tlx1*^{A208T} vs. WT, *tlx1*▲ vs. *tlx1*^{A208T}; kidney: *tlx1*▲ vs. WT, *tlx1*^{A208T} vs. WT, *tlx1*▲ vs. *tlx1*^{A208T}; liver: *tlx1*▲ vs. WT, *tlx1*^{A208T} vs. WT, *tlx1*▲ vs. *tlx1*^{A208T}; intestine: *tlx1*▲ vs. WT, *tlx1*^{A208T} vs. WT, *tlx1*▲ vs. *tlx1*^{A208T}) were conducted. Only genes with an absolute value of log2 (fold change) ≥1 and false discovery rate significance score <0.05 were used for subsequent analysis. GO enrichment in four tissue combinations of *tlx1*▲ vs. *tlx1*^{A208T} were conducted using the GOseq R package, and corrected $P < 0.05$ indicated significant enrichment.

Comparative whole genome search for seahorse immune/pregnancy genes

We scanned the genomes for a subset of immune-related genes that might be involved in pregnancy in four seahorse species, as well as six other Syngnathidae species, including greater pipefish, broadnosed pipefish, gulf pipefish, alligator pipefish, weedy seadragon, and Manado pipefish. The query protein sequences of these genes were obtained from Ensembl Genome Browser 104 or NCBI with representatives from *Homo sapiens*, *T. rubripes*, *O. latipes*, and *G. aculeatus*. Firstly, the genomes of the 10 Syngnathidae species were compared with those of other vertebrates using the best-hit search in TBLASTN v2.9.0 to find matching reads in our genomic datasets. BLAST parameters were set to an expectation cutoff of 1E − 5, allowing a maximum number of 1,000 returned sequences. In addition, all above immune-related genes were predicted using the homology-based gene prediction tool GeMoMa v1.7.1[59,71], with four species as reference organisms. The selected species were zebrafish, medaka, fugu, and stickleback. Introns from the mapped RNA-seq reads were extracted and filtered by the GeMoMa modules ERE and DenoiseIntrons. Next, we independently ran the module GeMoMa pipeline for each reference species using MMseqs2[72] as alignment tools, on the mapped RNA-seq data. Finally, the eleven individual annotation results were combined into a final annotation by using the GeMoMa modules GAF and AnnotationFinalizer[73]. Moreover, for immunoglobulin heavy chain genes (*ighvs*), the genome of the tiger tail seahorse has been fully predicted in Ensembl. Taking the sequence of the tiger tail seahorse as a reference, we used Exonerate (version 2.2.0) to conduct homology-based gene prediction for other genomic datasets with the default parameters. Phylogenetic analyses of genes identified in seahorses and other representative vertebrates were conducted using the amino acid sequences. The GenBank accession numbers of these gene families used for phylogenetic analysis are listed in Supplementary Data 13.

Reconstruction of character state evolution

Seven non-Syngnathiformes *L. oculatus*, *D. rerio*, *T. rubripes*, *G. aculeatus*, *O. niloticus*, *O. latipes*, *X. maculatus*, and thirteen Syngnathiformes including *H. comes*, *H. erectus*, *H. zosterae*, *H. mohnikei*, *S. acus*, *S. typhle*, *S. scovelli*, *P. taeniolatus*, *S. biaculeatus*, *O, manadensis*, *N. ophidion*, *F. commersonii* and *A. strigatus* were used to conduct the characters' reconstruction analysis based on previous study[7]. Briefly, the phylogenetic reconstruction of these species was conducted with the methods in the section phylogenetic analysis with minor modifications. Available empirical data on four characters including amino acid replacement in *tlx1*, brood pouch development, presence of spleen and immune genes simplification for each species included in the phylogenetic analyses were mapped onto our molecular phylogeny in Mesquite Ver.3.70 (http://www.mesquiteproject.org/).

Real-time quantitative PCR of genes involved in pregnancy

Adult male non-pregnant ($n = 4$) and pregnant ($n = 4$) lined seahorses collected from a fish farm (Zhangzhou, Fujian, China) were anesthetized with MS222 before brood pouches sampling. After the separation of embryos, the brood pouch was washed with PBS to remove embryonic contamination and total RNA was extracted using TRIzol reagent according to the manufacturer's instructions. 1 μg total RNA from the brood pouch sample was used to synthesize first-strand cDNA using the ReverAce qPCR RT Master Mix with gDNA Remover (Toyobo, Osaka, Japan). 'No RT' reactions were used as negative controls. The mRNA levels of genes involved in pregnancy, including *C3.2*, *cd8a*, *cd79a*, *gata3*, *il12a*, *il12b*, *T-bet*, and *tgfb1*, were determined by qRT-PCR (Supplementary Fig. 27). The primers used in this study are listed in Supplementary Table 18. qRT-PCR was performed on a Roche Light-Cycler 480 real time PCR system (Roche, Mannheim, Germany), using the SYBR Green I kit (Toyobo, Japan) according to the manufacturer's instructions. *β-actin* was used as an internal control.

Statistical analysis

Statistical analysis was performed using GraphPad Prism 7.0 (GraphPad Software, San Diego, CA, USA). All data are presented as the mean ± standard error of the mean (SEM). Statistical differences were estimated via unpaired Student's *t*-test and the significance level was set at 0.05.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The whole-genome raw reads and assemblies of *H. zosterae* and *H. mohnikei* have been deposited in the NCBI database under accession code PRJNA797939. The raw reads of the RNA-seq (including *Danio rerio*, *S. biaculeatus* and *H. erectus*) have been deposited in the NCBI database under accession code PRJNA799842. Accessions for previously published genomes used in this study were given in Supplementary Data 15. In addition, the data of the spleens phenotype and histology are available at Figshare (https://figshare.com/projects/Immunogenetic_losses_co-occurred_with_seahorse_male_pregnancy_and_mutation_in_tlx1_accompanied_functional_asplenia/153495). Source data are provided with this paper.

Code availability

Custom scripts employed for the analysis of the sequencing data are available at Figshare (https://figshare.com/projects/Immunogenetic_losses_co-occurred_with_seahorse_male_pregnancy_and_mutation_in_tlx1_accompanied_functional_asplenia/153495).

References

1. Kaufman, J. Evolution and immunity. *Immunology* **130**, 459–462 (2010).
2. Boehm, T. Evolution of vertebrate immunity. *Curr. Biol.* **22**, 722–732 (2012).
3. Schluter, S. F., Bernstein, R. M., Bernstein, H. & Marchalonis, J. J. 'Big Bang' emergence of the combinatorial immune system. *Dev. Comp. Immunol.* **23**, 107–111 (1999).
4. Flajnik, M. F. & Kasahara, M. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nat. Rev. Genet.* **11**, 47–59 (2010).

5. Matsunaga, T. & Rahman, A. What brought the adaptive immune system to vertebrates? - The jaw hypothesis and the seahorse. *Immunol. Rev.* **166**, 177–186 (1998).
6. Roth, O. et al. Evolution of male pregnancy associated with remodeling of canonical vertebrate immunity in seahorses and pipefishes. *Proc. Natl Acad. Sci. USA* **117**, 9431–9439 (2020).
7. Wilson, A. B., Ahnesjo, I., Vincent, A. C. & Meyer, A. The dynamics of male brooding, mating patterns, and sex roles in pipefishes and seahorses (family Syngnathidae). *Evolution* **57**, 1374–1386 (2003).
8. Whittington, C. M. & Friesen, C. R. The evolution and physiology of male pregnancy in syngnathid fishes. *Biol. Rev. Camb. Philos. Soc.* **95**, 1252–1272 (2020).
9. Wilson, A. B., Vincent, A., Ahnesjo, I. & Meyer, A. Male pregnancy in seahorses and pipefishes (family Syngnathidae): rapid diversification of paternal brood pouch morphology inferred from a molecular phylogeny. *J. Hered.* **92**, 159–166 (2001).
10. Skalkos, Z. G., Van, U. & Whittington, M. Paternal nutrient provisioning during male pregnancy in the seahorse *Hippocampus abdominalis*. *J. Comp. Physiol. B* **190**, 547–556 (2020).
11. Dudley, S. et al. Structural changes to the brood pouch of male pregnant seahorses (*Hippocampus abdominalis*) facilitate exchange between father and embryos. *Placenta* **114**, 115–123 (2021).
12. Stölting, K. N. & Wilson, A. B. Male pregnancy in seahorses and pipefish: beyond the mammalian model. *BioEssays* **29**, 884–896 (2007).
13. Bainbridge, D. R. The evolution of pregnancy. *Early Hum. Dev.* **90**, 741–745 (2014).
14. Reznick, D. N., Mateos, M. & Springer, M. S. Independent origins and rapid evolution of the placenta in the fish genus Poeciliopsis. *Science* **298**, 1018–1020 (2002).
15. Fernandez, N. et al. A critical review of the role of the major histocompatibility complex in fertilization, preimplantation development and feto-maternal interactions. *Hum. Reprod. Update* **5**, 234–248 (1999).
16. Lin, Q. et al. The seahorse genome and the evolution of its specialized morphology. *Nature* **540**, 395–399 (2016).
17. Li, C. Y. et al. Genome sequences reveal global dispersal routes and suggest convergent genetic adaptations in seahorse evolution. *Nat. Commun.* **12**, 1094 (2021).
18. Koss, M. et al. Congenital Asplenia in Mice and Humans with Mutations in a Pbx/Nkx2-5/p15 Module. *Dev. Cell* **22**, 913–926 (2012).
19. Roberts, C. W., Shutter, J. R. & Korsmeyer, S. J. Hox11 controls the genesis of the spleen. *Nature* **368**, 747–749 (1994).
20. Tribioli, C. & Lufkin, T. The murine Bapx1 homeobox gene plays a critical role in embryonic development of the axial skeleton and spleen. *Development* **126**, 5699–5711 (1999).
21. Langenau, D. M. et al. Molecular cloning and developmental expression of Tlx (Hox11) genes in zebrafish (*Danio rerio*). *Mech. Dev.* **117**, 243–248 (2002).
22. Hamilton, H. et al. Molecular phylogeny and patterns of diversification in syngnathid fishes. *Mol. Phylogenet Evol.* **107**, 388–403 (2017).
23. Stiller, J. et al. Phylogenomic analysis of Syngnathidae reveals novel relationships, origins of endemic diversity and variable diversification rates. *BMC Biol.* **20**, 75 (2022).
24. Xie, L. et al. Congenital asplenia due to a tlx1 mutation reduces resistance to Aeromonas hydrophila infection in zebrafish. *Fish. Shellfish Immun.* **95**, 538–545 (2019).
25. Logan, C., Wingate, R. J., McKay, I. J. & Lumsden, A. Tlx-1 and Tlx-3 homeobox gene expression in cranial sensory ganglia and hindbrain of the chick embryo: markers of patterned connectivity. *J. Neurosci.* **18**, 5389–5402 (1998).
26. Berger, J., Dorninger, F., Forss-Petter, S. & Kunze, M. Peroxisomes in brain development and function. *Biochim Biophys. Acta* **1863**, 934–955 (2016).
27. Lee, K. H., Cha, M. & Lee, B. H. Neuroprotective effect of antioxidants in the brain. *Int. J. Mol. Sci.* **21**, 7152 (2020).
28. Wei, R. et al. Dendritic cells in pregnancy and pregnancy-associated diseases. *Biomed. Pharmacother.* **133**, 110921 (2021).
29. Tilburgs, T. & Strominger, J. L. CD8+ effector T cells at the fetal-maternal interface, balancing fetal tolerance and antiviral immunity. *Am. J. Reprod. Immunol.* **69**, 395–407 (2013).
30. Yang, X., Zhang, C., Chen, G., Sun, C. & Li, J. Antibodies: The major participants in maternal-fetal interaction. *J. Obstet. Gynaecol. Res.* **45**, 39–46 (2019).
31. Girardi, G., Lingo, J. J., Fleming, S. D. & Regal, J. F. Essential role of complement in pregnancy: from implantation to parturition and beyond. *Front Immunol.* **11**, 1681 (2020).
32. Lu, L., Barbi, J. & Pan, F. The regulation of immune tolerance by FOXP3. *Nat. Rev. Immunol.* **17**, 703–717 (2017).
33. Tussiwand, R. et al. Compensatory dendritic cell development mediated by BATF-IRF interactions. *Nature* **490**, 502–507 (2012).
34. Haas, K. M., Poe, J. C., Steeber, D. A. & Tedder, T. F. B-1a and b-1b cells exhibit distinct developmental requirements and have unique functional roles in innate and adaptive immunity to S-pneumoniae. *Immunity* **23**, 7–18 (2005).
35. O'Neill, S. K. et al. Monophosphorylation of CD79a and CD79b ITAM motifs initiates a SHIP-1 phosphatase-mediated inhibitory signaling cascade required for B cell anergy. *Immunity* **35**, 746–756 (2011).
36. Schroeder, H. W. & Cavacini, L. Structure and function of immunoglobulins. *J. Allergy Clin. Immun.* **125**, 41–52 (2010).
37. Colvin, R. B. & Smith, R. N. Antibody-mediated organ-allograft rejection. *Nat. Rev. Immunol.* **5**, 807–817 (2005).
38. Girardi, G., Redecha, P. & Salmon, J. E. Heparin prevents antiphospholipid antibody-induced fetal loss by inhibiting complement activation. *Nat. Med.* **10**, 1222–1226 (2004).
39. Kahn, D. A. & Baltimore, D. Pregnancy induces a fetal antigen-specific maternal T regulatory cell response that contributes to tolerance. *P Natl Acad. Sci. USA* **107**, 9299–9304 (2010).
40. Romano, A. et al. FOXP3+ regulatory T cells in hepatic fibrosis and splenomegaly caused by schistosoma japonicum: the spleen may be a major source of tregs in subjects with splenomegaly. *PLoS Negl. Trop. Dis.* **10**, e0004306 (2016).
41. La Rocca, C., Carbone, F., Longobardi, S. & Matarese, G. The immunology of pregnancy: regulatory T cells control maternal immune tolerance toward the fetus. *Immunol. Lett.* **162**, 41–48 (2014).
42. Parker, J. et al. Immunological tolerance in the evolution of male pregnancy. *Mol. Ecol.* (2021).
43. Star, B. et al. The genome sequence of Atlantic cod reveals a unique immune system. *Nature* **477**, 207–210 (2011).
44. Malmstrøm, M. et al. Evolution of the immune system influences speciation rates in teleost fishes. *Nat. Genet* **48**, 1204–1210 (2016).
45. Dubin, A. et al. Complete loss of the MHC II pathway in an anglerfish, *Lophius piscatorius*. *Biol. lett.* **15**, 20190594 (2019).
46. Guslund, C. et al. Lymphocyte subsets in Atlantic cod (*Gadus morhua*) interrogated by single-cell sequencing. *Commun. Biol.* **5**, 1–9 (2022).
47. Xie, L. et al. Congenital Asplenia Interrupts Immune Homeostasis and Leads to Excessive Systemic Inflammation in Zebrafish. *Front Cell Infect. Microbiol* **11**, 668859 (2021).
48. Swann, J. B., Holland, S. J., Petersen, M., Pietsch, T. W. & Boehm, T. The immunogenetics of sexual parasitism. *Science* **369**, 1608–1615 (2020).
49. Parker, J., Guslund, N., Jentoft, S. & Roth, O. Characterization of pipefish immune cell populations through single-cell transcriptomics. *Front Immunol.* **13**, 820152 (2022).

50. Klein, S. L. & Flanagan, K. L. Sex differences in immune responses. *Nat. Rev. Immunol.* **16**, 626–638 (2016).
51. Gnerre, S. et al. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc. Natl Acad. Sci. USA* **108**, 1513–1518 (2011).
52. Servant, N. et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **16**, 259 (2015).
53. Gross, J. B. et al. Synteny and candidate gene prediction using an anchored linkage map of *Astyanax mexicanus*. *Proc. Natl Acad. Sci. USA* **105**, 20106–20111 (2008).
54. Edgar, R. C. & Myers, E. W. PILER: identification and classification of genomic repeats. *Bioinformatics* **21**, 152–158 (2005).
55. Price, A. L., Jones, N. C. & Pevzner, P. A. De novo identification of repeat families in large genomes. *Bioinformatics* **21**, 351–358 (2005).
56. Stanke, M. & Waack, S. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* **19**, 215–225 (2003).
57. Majoros, W. H., Pertea, M. & Salzberg, S. L. TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics* **20**, 2878–2879 (2004).
58. Korf, I. Gene finding in novel genomes. *BMC Bioinforma.* **5**, 59 (2004).
59. Keilwagen, J. et al. Using intron position conservation for homology-based gene prediction. *Nucleic Acids Res.* **44**, e89 (2016).
60. Campbell, M. A., Haas, B. J., Hamilton, J. P., Mount, S. M. & Buell, C. R. Comprehensive analysis of alternative splicing in rice and comparative analyses with Arabidopsis. *BMC Genomics* **7**, 327 (2006).
61. Haas, B. J. et al. Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments. *Genome Biol.* **9**, R7 (2008).
62. Torresen, O. K. et al. An improved genome assembly uncovers prolific tandem repeats in Atlantic cod. *BMC Genomics* **18**, 95 (2017).
63. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
64. Minh, B. Q., Nguyen, M. A. & von Haeseler, A. Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* **30**, 1188–1195 (2013).
65. De Bie, T., Cristianini, N., Demuth, J. P. & Hahn, M. W. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* **22**, 1269–1271 (2006).
66. Huerta, J. et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, 309–314 (2019).
67. Fernandez, R. & Gabaldon, T. Gene gain and loss across the metazoan tree of life. *Nat. Ecol. Evol.* **4**, 524–533 (2020).
68. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
69. He, K. et al. Echolocation in soft-furred tree mice. *Science* **372**, 1305 (2021).
70. Leerberg, D. M., Hopton, R. E. & Draper, B. W. Fibroblast growth factor receptors function redundantly during zebrafish embryonic development. *Genetics* **212**, 1301–1319 (2019).
71. Keilwagen, J., Hartung, F., Paulini, M., Twardziok, S. O. & Grau, J. Combining RNA-seq data and homology-based gene prediction for plants, animals and fungi. *BMC Bioinforma.* **19**, 189 (2018).
72. Steinegger, M. & Soding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).
73. Chueca, L. J. et al. De novo Genome Assembly of the Raccoon Dog (Nyctereutes procyonoides). *Front. Genet.* **12**, 658256 (2021).

Acknowledgements

The authors acknowledge the Marine Biodiversity Collections of South China Sea, CAS for providing samples; the Laboratory Animal Center of Huazhong Agricultural University for providing experimental facilities and the Agency for Science, S. Fan and X. Lu for insightful comments and suggestions that improved the manuscript. This research was supported by the National Natural Science Foundation of China (41890853 to S.Z., 41825013 to Q.L., 42006108 to M.Q.), the Key Research Program of Frontier Sciences of CAS (ZDBS-LY-DQC004 to Q.L.), Guangdong Basic and Applied Basic Research Foundation (2021A1515011380 to Y.L.), the European Research Council (ERC) under the European Union's Horizon research and innovation program (EC/H2020/755659 to O.R.), the Strategic Priority Research Program of CAS (XDB42030204 to Q.L.), and the Key Deployment Project of COMS of CAS (COMS2020Q14 to Y.Z.).

Author contributions

Q.L., A.M., O.R. and B.V. conceived the project. Q.L. and A.M. supervised the study. X.W. and G.Q. collected the samples. M.Q., R.S., H.Y., X.W., Y.Z., H.Z., Z.Z. and J.Y. performed genome analyses. Y.L., H.Y. and G.Q. performed transcriptomic analysis. Y.L., W.L., B.Z. and H.J. performed CRISPR/Cas9 genomic editing analyses. Y.W. and H.J. performed the Quantitative Real-time PCR. H.J. and Yingyi Z. preformed in situ hybridization. S.Z. performed the Micro-CT analysis. Y.L., M.Q., R.S. and Q.L. wrote the manuscript with input from all other authors. All authors reviewed and contributed to the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-35338-7.

Correspondence and requests for materials should be addressed to Olivia Roth, Axel Meyer or Qiang Lin.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.