Identification and characterization of zebrafish Tlr4 co-receptor Md-2

Andrea N. Loes[1,3], Melissa N. Hinman[1,4], Dylan R. Farnsworth[2,4], Adam C. Miller[2,4], Karen Guillemin[1,4,5], Michael J. Harms[1,3]

[1]Institute of Molecular Biology, [2]Institute of Neuroscience, [3]Department of Chemistry and Biochemistry, [4]Department of Biology, University of Oregon, Eugene, OR 97403, [5]Humans and the Microbiome Program, CIFAR, Toronto, Ontario M5G 1Z8, Canada

Funding sources:

This research was funded by grants from the American Heart Association (AHA-15BGIA22830013, MJH) and the National Institutes of Health (NIH-T32GM007413, ANL, NIH-F32DK107318, MNH, P50GM09891, KG, NIH-R24OD026591, ACM). MJH is a Pew Scholar in the Biomedical Sciences, supported by The Pew Charitable Trusts. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

SUPPLEMENTAL MATERIAL

ABSTRACT

The zebrafish (*Danio rerio*) is a powerful model organism for studies of the innate immune system. One apparent difference between human and zebrafish innate immunity is the cellular machinery for LPS-sensing. In amniotes, the protein complex formed by Toll-like receptor 4 and myeloid differentiation factor 2 (Tlr4/Md-2) recognizes the bacterial molecule lipopolysaccharide (LPS) and triggers an inflammatory response. It is believed that zebrafish have neither Md-2 nor Tlr4: Md-2 has not been identified outside of amniotes, while the zebrafish *tlr4* genes appear to be paralogs, not orthologs, of amniote *TLR4s*. We revisited these conclusions. We identified a zebrafish gene encoding Md-2, *ly96*. Using single-cell RNA-Seq, we found that *ly96* is transcribed in cells that also transcribe genes diagnostic for innate immune cells, including the zebrafish *tlr4*-like genes. Unlike amniote *LY96*, zebrafish *ly96* expression is restricted to a small number of macrophage-like cells. In a functional assay, zebrafish Md-2 and Tlr4a form a complex that activates NF-κB signaling in response to LPS, but *ly96* loss-of-function mutations gave little protection against LPS-toxicity in larval zebrafish. Finally, by analyzing the genomic context of *tlr4* genes in eleven jawed vertebrates, we found that *tlr4* arose prior to the divergence of teleosts and tetrapods. Thus, an LPS-sensitive Tlr4/Md-2 complex is likely an ancestral feature shared by mammals and zebrafish, rather than a *de novo* invention on the tetrapod lineage. We hypothesize that zebrafish retain an ancestral, low-sensitivity Tlr4/Md-2 complex that confers LPS-responsiveness to a specific subset of innate immune cells.

INTRODUCTION

Amniote innate immune systems are exquisitely sensitive to lipopolysaccharide (LPS), a component of the cell wall in Gram-negative bacteria (1–3). LPS is sensed by a protein complex composed of Toll-like receptor 4 (Tlr4) and Md-2 (also known as LY96 and ESOP-1) (1, 4). LPS binds in a pocket of Md-2, triggering dimerization of the Tlr4/Md-2 complex (Fig 1) (5). This, in turn, activates a Myd88-dependent NF-κB response (6). When properly regulated, the LPS activation of Tlr4/Md-2 regulates microbiome populations (7), recruits neutrophils to sites of infection (8), and induces angiogenesis (9). When dysregulated, Tlr4/Md-2 activity induces septic shock (10, 11), plays roles in inflammatory disorders (11, 12), and is a key player in the tissue remodeling that accompanies tumorigenesis (13, 14).

The role of Tlr4/Md-2 in LPS-sensing outside of amniotes remains poorly understood. Understanding this response in zebrafish (*Danio rerio*) is of particular interest, as the zebrafish is a powerful model organism for studies of vertebrate innate immunity (15). Zebrafish have mature genetic resources, rapid generation time, clear embryos, and facile germ-free derivation (16, 17). The zebrafish is increasingly popular as a model for understanding host-microbe interactions (18), as well as a tool to understand the development of the innate immune system (16).

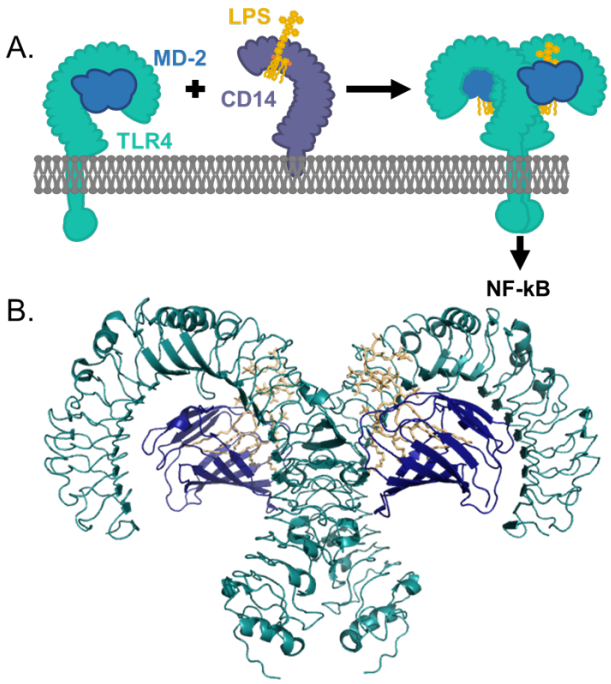

Fig 1. LPS activation of amniote Tlr4 requires cofactors Md-2 and Cd14. A) Schematic representation of LPS transfer from Cd14 to the Tlr4/Md-2 complex. LPS (yellow) is brought by Cd14 (purple) and loaded into Md-2 (navy blue). Md-2 is bound by Tlr4 (cyan). Binding of LPS to the Md-2 co-receptor causes dimerization of the Tlr4/Md-2 complex, activating a downstream inflammatory response. B) The interface between human Tlr4 (cyan) and Md-2 (navy blue) is extensive. Both are required to form a productive interaction with LPS (yellow). Structure shown was made from PDB 3FXI (19).

The zebrafish response to LPS is puzzling (20, 21). In some ways it is similar to amniotes. As in amniotes, LPS triggers Myd88-dependent NF-κB inflammatory response (22, 23). Further, the expression patterns of genes induced by LPS stimulation are highly similar between mouse and zebrafish (24). There are, however, several lines of evidence that suggest Tlr4/Md-2 is not involved. Most critically, the gene encoding the essential co-receptor Md-2 has not been identified in zebrafish and other ray-finned fishes (20, 21, 25). Further, zebrafish Tlr4 proteins do not activate NF-κB in response to LPS in *ex vivo* assays, even when complemented with a mouse or human Md-2 (20, 21). Finally, zebrafish do not have a direct ortholog to amniote *tlr4*. Rather, they possess three *tlr4*-like genes—*tlr4ba*, *tlr4bb*, and *tlr4a1*—that are

106 thought to have arisen from an ancestral Toll-like receptor lost in tetrapods but retained in ray-
107 finned fishes (21). These observations have led to the hypothesis that zebrafish respond to LPS
108 by a non-Tlr4/Md-2-dependent pathway.

109 We set out to carefully revisit these conclusions using resources unavailable when the
110 initial investigations of zebrafish Tlr4 were performed. Using careful bioinformatics, we found
111 an ortholog of the gene encoding Md-2 (*ly96*) in zebrafish and other ray-finned fishes. When co-
112 transfected into mammalian cells, the zebrafish *ly96* and *tlr4ba* genes activate NF-κB signaling
113 in response to LPS. Single-cell RNA-seq experiments on larval zebrafish revealed that the gene
114 is expressed in a small subset of cells that express the zebrafish *tlr4*-like genes and the
115 macrophage-specific gene *mpeg1.1* (26). This contrasts with amniotes, in which *TLR4* and *LY96*
116 are both broadly expressed (27). Further, unlike mammalian *TLR4* and *LY96* mutants that
117 exhibit increased resistance to systemic LPS challenge (4, 28), zebrafish larvae with loss of
118 function *ly96* mutations are not protected from LPS toxicity. Finally, we revisited the history of
119 the *tlr4* gene in zebrafish, finding that formation of an LPS-sensitive Tlr4/Md-2 complex is
120 likely an ancestral feature shared by mammals and zebrafish, rather than a *de novo* invention on
121 the tetrapod lineage. We hypothesize that zebrafish preserve an ancestral, low-sensitivity
122 Tlr4/Md-2 complex that plays an LPS-sensing role in a small population of innate immune cells.
123

124 **MATERIALS & METHODS**

125 ***Phylogenetic reconstruction analysis***

126 We constructed curated databases of Md-1, Md-2, Tlr4, and Cd180 protein sequences
127 from across the vertebrates. Cd180 and Md-1 are paralogs of Tlr4 and Md-2, respectively (29).
128 We obtained amino acid sequences of these proteins from NCBI, Ensembl, Fish1TK, amphibian
129 transcriptomes (30–33), UniProt, and ZFIN. We constructed a multiple sequence alignment for

Tlr4 and Cd180 and for Md-2 and Md-1 using MSAPROBS (34), followed by manual editing in MEGA (35). We trimmed the alignment to remove highly variable (and therefore unalignable) regions. We used PHYML (36, 37) with subtree pruning and re-grafting to construct the ML phylogeny. Pilot analyses revealed that the JTT substitution model with 8 rate categories and a floating gamma distribution parameter yielded the highest likelihood trees (38–40). An Akaike information criterion (AIC) test was used to control for overfitting (41). We rooted our trees at the duplication of these proteins in early vertebrates. Alignment figures in supplement were made with JalView (42).

Synteny analysis

For the *ly96* synteny analysis, we used the Ensembl synteny module (43) to map homologs onto the chromosomes of species of interest. For the tlr4 synteny analysis, we took the 22 genes flanking human *TLR4 (*11 on each side) and the 22 genes flanking zebrafish *tlr4*. We used tblastn with default parameters to BLAST these sequences against 11 vertebrate genomes. We discarded all hits with e-value < 0.001 and then calculated a running average of the log (e-value) along each chromosome with a sliding window of 10,000 bases. Finally, we divided this running average by the maximum observed log (e-value)/bp value across all genomes. This value occurs for the window centered on the zebrafish *tlr4* gene. On the final relative scale, 0.0 indicates no hits observed in a given window and 1.0 is the maximum e-value per base pair. The complete analysis pipeline is implemented in a collection of shell scripts and jupyter notebooks (https://github.com/harmslab/vertebrate-tlr4-synteny/).

153 ***Gene expression analysis***

154 Whole 6 days post fertilization (dpf) zebrafish were euthanized by tricaine methane
155 sulfonate overdose, flash frozen in 1 mL of Trizol (Ambion), thawed, and homogenized.
156 Chloroform (200 µL) was added to each tube followed by mixing, centrifugation at 12,000 g for
157 10 minutes at 4°C, transfer of the aqueous phase to a separate tube, addition of 200 µL ethanol,
158 and binding of sample to an RNeasy mini kit column (Qiagen). RNA was washed and eluted
159 according to the manufacturer's instructions and treated with RQ1 DNase (Promega). RNA was
160 reverse transcribed into cDNA using Superscript II Reverse Transcriptase (Invitrogen) and an
161 oligo dT (20) primer, then amplified by PCR using gene-specific primers for zebrafish *ly96* (5' -
162 TGTATGGCATCTGAGAAAGCAGA - 3' and 5' - AAGAGCAGGGGGAAACAGTC - 3') and
163 the housekeeping gene *b2m* (5' - ACGCTGCAGGTATATTCATC - 3' and 5' -
164 TCTCCATTGAACTGCTGAAG - 3'). PCR products were separated by electrophoresis on a 6%
165 *bis*-Acrylamide (19:1) gel that was stained with 1X SYBR green 1 nucleic acid gel stain
166 (Invitrogen) and imaged using an AlphaImagerHP (Alpha Innotech). The identity of the *ly96* RT-
167 PCR product was verified by Sanger sequencing.

168

169 ***Single-cell RNA-Seq***

170 Single-cell analysis of transcription patterns of *ly96*, *tlr4ba*, *tlr4bb*, and *tlr4a1* was performed
171 using the recently released Zebrafish Single-Cell Transcriptome Atlas (44)**.** Briefly, dissociated
172 cells were run on a 10X Chromium platform using v2 chemistry. Dissociated samples for each
173 time point (1, 2 and 5 dpf) were submitted in duplicate to determine technical reproducibility.
174 The resulting cDNA libraries were sequenced on either an Illumina Hi-seq or an Illumina Next-
175 seq. The resulting sequencing data were analyzed using the 10X Cellranger pipeline, version

176 2.2.0 (45) and the Seurat software package for R, v3.4.4 (46, 47) using standard quality control,
177 normalization, and analysis steps. We aligned reads to the zebrafish genome, GRCz11_93, and
178 counted expression of protein coding reads. The resulting matrices were read into Seurat where
179 we performed PCA and UMAP analysis on the resulting dataset with 178 dimensions and a
180 resolution of 13.0, which produced 220 clusters and one singleton. Differential gene expression
181 analysis was performed using the FindAllMarkers function in Seurat and Wilcoxon rank sum
182 test.
183
184 ***Cell Culture and Transfection Conditions***
185 Mammalian expression vectors containing human *TLR4* and mouse *Tlr4* were obtained
186 from Addgene (#13085 and #13086), originally deposited by Ruslan Medzhitov. Human *CD14*
187 and *ELAM-Luc* were also obtained from Addgene (#13645 and #13029) originally deposited by
188 Doug Golenbock. Human *MD-2* was obtained from the DNASU Repository (HsCD00439889)
189 and contains a C-terminal V5-tag. Mouse *Md-2* (UniProt #Q9JHF9) and *Cd14* (UniProt
190 #P10810), as well as opossum *Md-2* (UniProt #F6QBE6), *Cd14* (NCBI Accession
191 #XP_007473804.1) and chicken *Md-2* (UniProt #A0A1D5NZX9), and *Cd14* (UniProt
192 #B0BL87) were designed to be free of restriction sites, codon-optimized for human expression,
193 and purchased as mammalian expression vector constructs in pcDNA3.1 (+) from Genscript
194 (New Jersey, USA). Zebrafish *tlr4ba* and *ly96* were also obtained from Genscript in pcDNA3.1
195 (+). Zebrafish *tlr4bb* was a gift from Carol Kim. We re-cloned this protein from its original
196 vector into pcDNA3.1 (+) to limit variability in expression due to differences in vector size and
197 promoter.

198 Human embryonic kidney cells (HEK293T/17, ATCC CRL-11268) were maintained up
199 to 30 passages in DMEM supplemented with 10% FBS at 37° C with 5% CO_2. For each
200 transfection, a confluent 100 mm plate of HEK293T/17 cells was treated at room temperature
201 with 0.25% Trypsin-EDTA in HBSS and resuspended with an addition of DMEM + 10% FBS.
202 This was diluted 4-fold into fresh medium and 135 µL aliquots of resuspended cells were
203 transferred to a 96-well cell culture treated plate. Transfection mixes were made with 10 ng of
204 *tlr4*, 1 ng of *cd14*, 10 ng of *ly96*, 1 ng of *Renilla*, 20 ng of *ELAM-Luc*, and 58 ng of pcDNA3.1
205 (+) per well for a total of 100 ng of DNA, diluted in OptiMEM to a volume of 10 µL/well. To the
206 DNA mix, 0.5 µL per well of PLUS reagent was added followed by a brief vortex and room
207 temperature incubation for 10 min. Lipofectamine was diluted 0.5 µL into 9.5 µL OptiMEM per
208 well. This was added to the DNA + PLUS mix, vortexed briefly and incubated at RT for 15 min.
209 The transfection mix was diluted to 65 µL/well in OptiMEM and aliquoted onto a plate. Cells
210 were incubated with transfection mix overnight (20-22 hrs) and then treated with LPS. *E. coli* K-
211 12 lipopolysaccharide (LPS) (tlrl-eklps, Invivogen) was dissolved at 5 mg/mL in endotoxin-free
212 water, and aliquots were stored at -20° C. To avoid freeze-thaw cycles, working stocks of LPS
213 were prepared at 10 µg/mL and stored at 4° C. To prepare treatments, LPS was diluted in 25%
214 phosphate buffered saline and 75% DMEM. Cells were incubated with treatments for 4 hr. The
215 Dual-Glo Luciferase Assay System (Promega) was used to assay Firefly and Renilla luciferase
216 activity of individual wells. Each NF-κB induction value shown represents the Firefly luciferase
217 activity/Renilla luciferase activity, normalized to the buffer treated transfection control to
218 compare fold-change in NF-κB activation for treatments.
219
220

221 ***Generation of mutant zebrafish***

222 Zebrafish experiments were approved by the University of Oregon Institutional Animal
223 Care and Use Committee. Chop Chop (http://chopchop.cbu.uib.no) was used to design a guide
224 RNA (gRNA) targeting the first exon of zebrafish *ly96* (si:dkey-82k12.13, GRCz11). A gRNA
225 template was generated by a template-free Phusion polymerase (New England Biolabs) PCR
226 reaction using a scaffold primer (5'-
227 GATCCGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTA
228 ACTTGCTATTTCTAGCTCTAAAAC-3') and an *ly96*-specific primer (5'-
229 AATTAATACGACTCACTATAGGGTATCAGATATGGCGCTTGTTTTAGAGCTAGAAAT
230 AGC-3'), then cleaned using the QIAquick PCR Purification Kit (Qiagen), transcribed using a
231 MEGAscript kit (Ambion), and purified by phenol-chloroform extraction and isopropanol
232 precipitation. Cas9 RNA was made by linearizing the pT3TS-nls-zCas9-nls plasmid (41) with
233 XbaI, purifying it using the QIAquick Gel Extraction Kit (Qiagen), performing an in vitro
234 transcription reaction using the T3 mMESSAGE kit (Invitrogen), and purifying the RNA using
235 the RNeasy Mini kit (Qiagen). AB strain zebrafish embryos were microinjected at the one cell
236 stage with 1-2 nL of a mixture containing 100 ng/μL Cas9 mRNA, 50 ng/μL gRNA, and phenol
237 red, and raised to adulthood. Fin DNA was amplified by PCR using primers specific to the
238 targeted region (5'- CAAATTGGATTCACAACAGAGC -3' and 5' -
239 CCATGGAAAATCAATGAAAAGC - 3'). Mosaic mutants were identified based on loss of an
240 HaeII restriction site and were outcrossed to wildtype AB zebrafish to generate heterozygotes.
241 Fish with loss-of-function mutations were identified by Sanger sequencing and further crossed to
242 generate three independent homozygous *ly96* mutant lines.
243

Fish LPS survival assay

WT and homozygous *ly96* mutant zebrafish embryos were grown under standard conditions in separate 10 cm petri dishes at a density of one fish per mL of embryo medium (EM), with fifty fish total per dish. At 5 dpf, lipopolysaccharides (LPS) purified from *Escherichia coli 0111:B4* (Sigma L2630) was dissolved in EM and added to dishes at a final concentration of 0.6 mg/mL, and control fish were mock treated with EM alone. Dead larvae, as determined by lack of heartbeat, were counted and removed at regular intervals from 16 to 48 hours or from 16 to 72 hours after addition of LPS, at which time the experiment was terminated and surviving fish were humanely euthanized.

RESULTS

Zebrafish have a gene encoding MD-2

The strongest evidence against Tlr4/Md-2 performing LPS-sensing in zebrafish is the lack of Md-2. Md-2 is essential for LPS recognition by amniote Tlr4, as it contains the LPS binding pocket (Fig 1). We therefore asked whether we could find a gene encoding Md-2 in bony fishes.

We started by using the human MD-2 protein sequence to BLAST against the zebrafish genome and transcriptomes. This returned no significant hits, so we took a more phylogenetically informed strategy. Relative to humans, the earliest branching, functionally characterized Tlr4/Md-2 complex is from chicken (*Gallus gallus*). We therefore "walked out" from amniotes towards fishes, starting with amphibians. We BLASTed the human MD-2 protein sequence against the *Xenopus laevis* genome. This revealed a hit to a hypothetical protein with 30% identity (OCT74818.1). When reverse-BLASTed against the human proteome, this hit

267 returned Md-2. To validate the amino acid sequence, we compared it to the sequences of
268 functionally characterized Md-2 proteins from amniotes. We found that the *X. laevis* gene
269 appeared to be N-terminally truncated. Using XenBase, we identified the full-length transcript in
270 the transcriptome for *X. laevis*. By BLASTing against available amphibian transcriptomes (30–
271 33), we further identified putative Md-2 proteins in *Rhinella marina, Hynobius retardatus,*
272 *Odorrana margaretae*, and *Ichthyophis bannanicus (*Fig S1).

273 With these putative amphibian Md-2 sequences in hand, we returned to our search for a
274 zebrafish Md-2. A BLAST against a zebrafish transcriptome using the *X. laevis* sequence
275 revealed a likely transcript (si:dkey-82k12.13, 23% identity). We then searched additional fish
276 transcriptomes available from the Fish-T1K project (49) and identified a set of transcripts from
277 three species that matched Md-2 (Fig S1). The genes we identified in bony fishes that encode
278 putative Md-2 proteins were highly diverged. On average, they exhibited only 26% identity
279 against human Md-2, and only ~40% identity relative to one another.

280 We next set out to assign the orthology of these putative Md-2 sequences. Our primary
281 concern was that these newly identified sequences were paralogs of Md-2. We therefore built a
282 phylogenetic tree to elucidate whether these newly identified sequences grouped with Md-2 or
283 with its direct paralog, Md-1. We constructed an alignment of 294 Md-1 and Md-2 protein
284 sequences sampled from amniotes, amphibians, and bony fishes and then used this to infer a
285 maximum likelihood phylogeny (Fig 2A). The alignment is available in File S1.

286 The putative amphibian and bony fish Md-2 sequences grouped with the tetrapod MD-2
287 sequences with strong support (SH = 0.99). The Md-1/Md-2 protein tree largely reproduced the
288 species tree, with the exception of amphibians. On the Md-1 lineage, amphibians form a
289 polytomy with fishes at the base of the tree; on the Md-2 lineage, they are placed inside the

290 amniote clade with a relatively short internal branch. This is likely an artifact of the small

291 number of amphibian sequences, as well as the rapid evolution of the genes along these lineages.

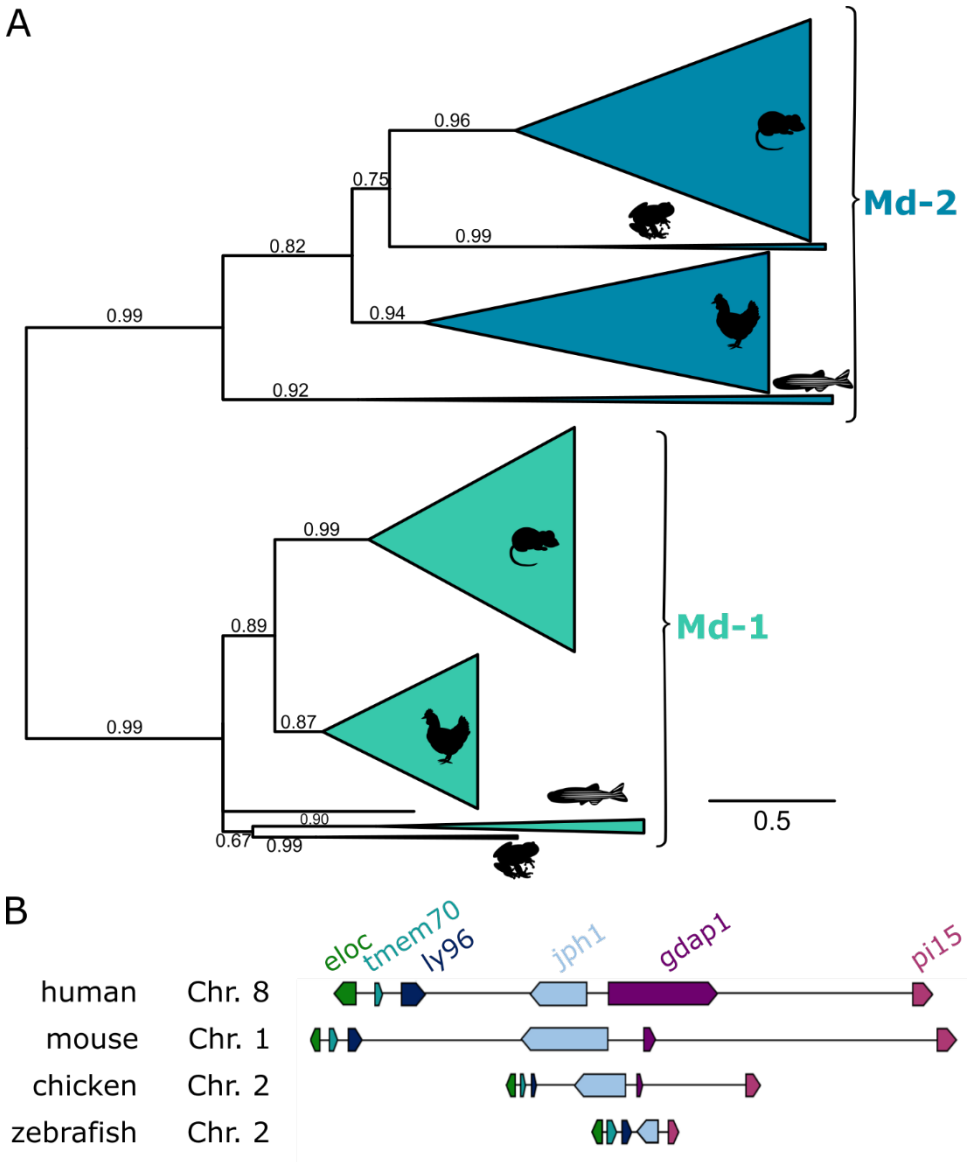

Fig 2. Phylogeny and synteny of the identified zebrafish protein support classifying it as an Md-2 (the *ly96* gene). A) Maximum likelihood phylogeny of Md-2 and Md-1 proteins. Wedges are collapsed clades of orthologs, with wedge height corresponding to the number of included taxa and wedge length indicating the longest branch length with the clade. Support values are SH-supports calculated using an approximate likelihood ratio test. Clades are colored to highlight Md-2 (blue) vs. Md-1 (green) classification. The taxa included in each clade are noted on the tree by silhouettes of mammals (mouse), sauropsids (chicken), amphibians (frog), and fish (zebrafish). B) Genomic organization of genes surrounding Md-2 in vertebrates. Arrows for genes represent the coding strand. Approximate distances between genes are represented by the length of line for the selected chromosome.

Overall, the tree is consistent with a single gene duplication event sometime before the evolution of bony vertebrates. Then, Md-1 and Md-2 were preserved along most descendant lineages. This said, the protein sequences of Md-1 and, particularly, Md-2 are evolving rapidly. The total branch lengths between the last common ancestor of Md-2 to its human and zebrafish descendants are 2.00 and 2.44, respectively. Put another way, the average site in the Md-2 sequence has changed its amino acid ~2 times over the last 430 million years. Only 7 of 160 positions in MD-2 are universally conserved across the clade.

To cross-validate the orthology of this newly identified gene, we next examined its locations in the *X. laevis* and *D. rerio* genomes. We found that the synteny is consistent with other bony vertebrates (Fig 2B). In five genomes sampled from across bony vertebrates—including *X. laevis* and *D. rerio*—the gene encoding Md-2 is located between *tmem70* and *jph1b* (Table S1). This provides strong evidence that these amphibian and fish genes are, in fact, orthologous to the amniote gene encoding Md-2. By convention, the gene encoding Md-2 is known as *ly96*. We therefore refer to this gene as zebrafish *ly96* hereafter.

Due to the genome duplication event that occurred along the zebrafish lineage (50), we also looked for a second copy of *ly96*. We examined the genomic location of the *jph1a* paralog, but we were unable to identify an additional gene with any similarity to *ly96*. It appears that an inversion may have occurred in this region, complicating identification by synteny alone. This said, no additional transcripts were identified within the zebrafish transcriptome with similarity to the identified zebrafish *ly96* sequence. This is consistent with a loss of the duplicate copy of this gene.

Finally, we attempted to identify an Md-2 sequence from earlier-diverging lineages including Chondrichthyes (cartilaginous fishes) and Agnatha (jawless fishes). Despite extensive

315 BLASTing, we were unable to identify an Md-2 protein sequence or *ly96* gene in either lineage.
316 This is consistent with *ly96* arising after the divergence of cartilaginous and bony fishes (~470
317 million years ago), but before the divergence of bony- and ray-finned fishes (~435 million years
318 ago). The sequence resources for cartilaginous and jawless fishes remain relatively sparse,
319 however, so we cannot exclude an earlier origin for *ly96*.
320
321 **Zebrafish transcribe *ly96* in innate immune cells**
322 We next asked whether zebrafish express *ly96*. To do so, we used the recently released
323 Zebrafish Single-Cell Transcriptome Atlas (44). This dataset consists of single-cell RNASeq
324 transcriptomes for 44,102 individual cells extracted from 1, 2 and 5 dpf zebrafish. The gray
325 points in Fig 3A and 3B shows the entire Atlas: each point is a cell, plotted such that cells with
326 similar transcription profiles appear near one another. Cluster identity can be established by
327 examining differentially expressed transcripts and using these marker genes to assess cell type
328 expression *in vivo* (44); this provides a means to assess which cell types express *ly96* simply by
329 asking which clusters possess *ly96* transcripts.

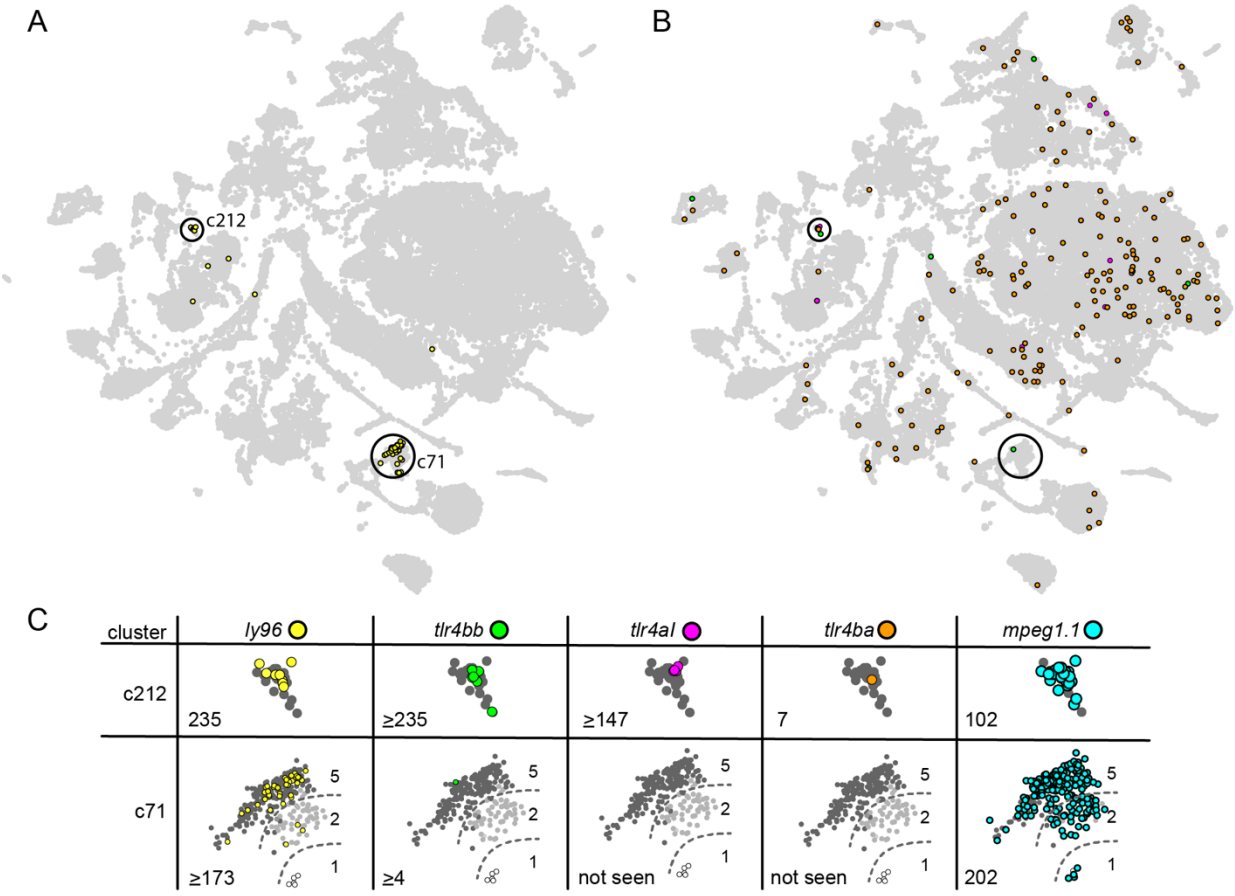

Fig 3. *ly96* and *tlr4* genes are expressed in macrophage cells. Each point in these plots is an individual cell characterized by single-cell RNA-Seq. The distance between the cells corresponds to the relative difference in their transcriptional profiles (44). A) Yellow points indicate cells expressing *ly96*, gray points show all 44,102 cells in the data set. The two clusters in which *ly96* is expressed (c71 and c212) are highlighted with black circles. B) Colored points indicate cells expressing *tlr4bb* (green), *tlr4al* (magenta), or *tlr4ba* (orange); gray points and circles are identical to panel A. C) Enlarged views of clusters c212 and c71, separated by gene of interest. This includes the genes shown in panels A and B, as well the macrophage marker, *mpeg1.1* (light blue) (26, 51). The number in the bottom left of each table entry is the expression level of the gene within the cluster divided by its expression level in all other cells in the dataset. If there was no expression in cells outside the cluster, expression within the cluster was divided by the detection threshold (0.001) giving a minimum estimate for the enrichment. The background cells are now colored by the developmental stage from which the cell was isolated: 1 dpf (white), 2 dpf (light gray), or 5 dpf (dark gray). The dashed lines in shown on c71 are approximate divisions between the age-dependent sub-clusters of c71.

We found that *ly96* is expressed in two clusters, denoted “c71” and “c212” (Fig 3A, yellow points). Both of these clusters are annotated in the Atlas as putative macrophage cells

based on their transcription profiles (44). *ly96* is highly enriched in these clusters relative to other clusters. This can be measured by taking the ratio of the average expression level of *ly96* for the cells in the cluster relative to the average expression level of *ly96* in all other cells. This ratio is 235 for cluster c212 and ≥173 for cluster c71, indicating that *ly96* is highly enriched in these clusters. For comparison, the well-established macrophage marker *mpeg1.1* (26, 51) has ratios of 102 and 202, respectively, for these same clusters (Fig 3C).

We next investigated the expression of the *tlr4bb*, *tlr4al*, and *tlr4ba* genes. We found that *tlr4bb* and *tlr4al* had quite limited expression patterns (Fig 3B, green and pink), while *tlr4ba* was expressed broadly (Fig 3B, orange). All three *tlr4* genes were found in cluster c212, but only *tlr4bb* was found in cluster c71 (Fig 3C).

The Atlas also has the potential to reveal time-course information for the expression of these genes, as it contains cells isolated from fish at 1, 2 and 5 dpf. We therefore shaded the cells within clusters c71 and c212 by their developmental time point (Fig 3C). Cluster c212, where we observed overlapping expression for *ly96* and all three *tlr4* genes, consists entirely of cells isolated from 5 dpf zebrafish (Fig 3C). Cluster c71 has three discrete sub-clusters corresponding to the age of the fish from which the cell was extracted. We see no *ly96* in the 1 dpf sub-cluster, a small amount in the 2 dpf sub-cluster, and the highest level in the 5 dpf sub-cluster (Fig 3C). Likewise, *tlr4bb* is expressed in the 5 dpf sub-cluster but no others. For comparison, the macrophage marker *mpeg1.1* is found in all cells within c71 and c212, regardless of the age of the fish from which the cell was extracted.

These observations suggest that *ly96* and all three *tlr4* genes are expressed together in a subset of macrophage cells by 5 dpf (Fig 3C, c212). Samples of later time points would be

372 necessary to establish if these genes are at their full expression level by 5 dpf, or if their
373 expression level and cell-type specificity continues to change as the fish develop.
374
375 **Zebrafish Tlr4a/Md-2 can activate NF-κB in response to lipopolysaccharide**
376 Given the low sequence similarity between the zebrafish Md-2 protein and its amniote
377 orthologs, it was not clear that the zebrafish Md-2 would be capable of mediating the Tlr4
378 response to Md-2. We therefore turned to an *ex vivo* cell culture assay to assess the ability of the
379 zebrafish Md-2 to partner with zebrafish Tlr4a and Tlr4b for LPS activation. In this assay, we co-
380 transfected genes encoding complex components into HEK293T cells and then used luciferase to
381 quantify NF-κB output in response to exogenously applied LPS (6).
382 We started by co-transfecting cells with zebrafish *ly96* and zebrafish *tlr4ba* or *tlr4bb* and
383 then measuring NF-κB activation in response to exogenously applied LPS. We saw no activation
384 (Fig 4A). This result was unsurprising, as this experiment attempted to activate a Tlr4/Md-2
385 complex without Cd14—an important peripheral protein that brings LPS to Tlr4/Md-2
386 complexes in amniotes, dramatically increasing the NF-κB response (Fig 1) (52–56). We thus
387 co-transfected *tlr4ba* or *tlr4bb* with zebrafish *ly96* and human *cd14*. In this context, we observed
388 potent activation of NF-κB in response to LPS for *tlr4ba*, but not *tlr4bb (*Fig 4A). To verify that
389 the activation of Tlr4a required Md-2, rather than merely Cd14, we tested the activation of Tlr4
390 and Cd14 without transfecting *ly96*—this complex did not respond to LPS (Fig 4A). We then
391 verified that the zebrafish Tlr4a/Md-2 complex, complemented with human Cd14, exhibited a
392 dose-dependent response to LPS (Fig 4B). The concentration of LPS needed for activation of the
393 zebrafish Tlr4a/Md-2 complex was much higher than that needed for activation of the human
394 proteins in these cells, but consistent with what has been observed for other species (57).

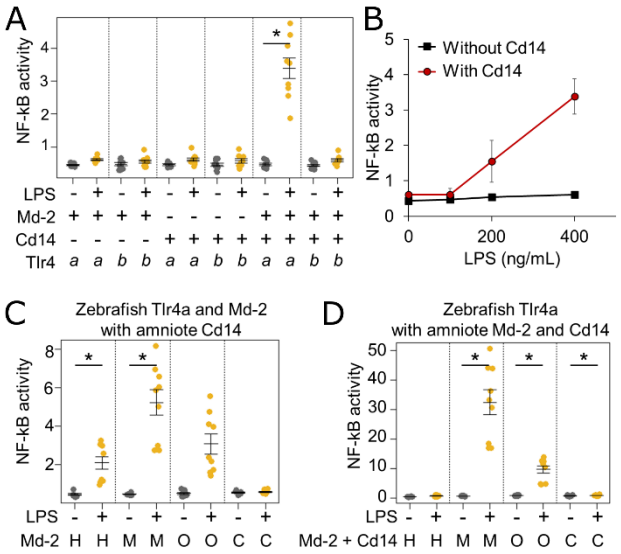

Fig 4. LPS activates the zebrafish Tlr4a/Md-2 in a functional assay. A) Activation of zebrafish Tlr4a and Tlr4b in the presence and absence of zebrafish Md-2 and human CD14. Points are the technical replicates from three biological replicates. Bold lines are the mean of the biological replicates. Error bars are a standard error on the mean of the biological replicates. B) Dose-dependence of LPS response by zebrafish Tlr4a/Md-2 in the presence (red) and absence (black) of human CD14. C) Zebrafish Tlr4a/Md-2 complemented with Cd14 proteins from amniotes. D) Zebrafish Tlr4a complemented with species-matched Md-2/Cd14 pairs taken from amniotes. Statistically significant differences (single-tailed Student's t-test) are noted on each panel (* $p < 0.05$)

Our results support the hypothesis that zebrafish Tlr4a/Md-2 can activate in response to LPS; however, this could only be done with the presence of a supporting mammalian protein (human Cd14). To determine if this was an artifact of the human protein, we tested the LPS activation of Tlr4a/Md-2 in the presence of human, mouse, opossum, and chicken Cd14. We found that all but the chicken Cd14 were able to support the activation of the complex (Fig 4C). Thus, the activity of the zebrafish Tlr4a/Md-2 complex does not depend exclusively on human Cd14 but can instead be supported by diverse Cd14 molecules. Given the importance of Cd14 in this assay, we looked for evidence of a zebrafish *cd14* gene; however, we were unable to locate such a gene. The inability to detect a *cd14* in fish may be due to rapid evolution of this gene since the most recent common ancestor, or, alternatively, Cd14 may have arisen as a supporting

417 molecule for LPS-recognition after the divergence of tetrapods. The requirement for Cd14 in
418 these experiments could be a problem with the heterologous cell line (these experiments were
419 done in human cells) or a missing alternate secondary cofactor (such as a fish LPS binding
420 protein).

421 Finally, to see if zebrafish Tlr4a behaved similarly to amniote Tlr4, we investigated
422 whether Md-2 from other species could act in concert with zebrafish Tlr4a. We co-transfected
423 *tlr4a* with human, mouse, or opossum *ly96* genes. We saw complementation by both mouse and
424 opossum Md-2 for LPS activation of zebrafish Tlr4a (Fig 4D). This suggests that the
425 requirements for activation by the Tlr4/Md-2 complex have been conserved for over 400 million
426 years and are shared across bony vertebrates.

427

428 **Md-2 is not required for LPS-induced death in 5 dpf larval zebrafish**

429 We next probed the physiological role of Md-2 in LPS-induced septic shock in larval 5
430 dpf zebrafish. We first treated 5 dpf larval WT zebrafish with LPS and followed their survival
431 over time. No treated WT fish survived more than 48 hours; the median survival time was 30 hrs
432 (Fig 5A). As a control, we also tested the LPS response for *myd88*$^{-/-}$ fish. As has been observed
433 previously (22), these showed a modest but significant increase in survival (Fig 5A). This was
434 consistent with LPS inducing a response that depends in part on a *myd88* dependent pathway.

435 To test the role of Md-2 in this response, we used CRISPR-Cas9-based mutagenesis to
436 establish three independent zebrafish lines with mutations in the first exon of the *ly96* gene. The
437 mutations were expected to induce a loss of function through removal of the start codon (*ly96*$^{A/A}$)
438 or through a frame shift and premature stop codon (*ly96*$^{B/B}$ and *ly96*$^{C/C}$) (Table S2). Using RT-
439 PCR primers downstream of the targeted region, we demonstrated that *ly96* mRNA is expressed
440 in mutant larval zebrafish (Fig 5E).

We then tested the three *ly96* mutant zebrafish lines for their susceptibility to LPS-induced septic shock. The results were mixed. Compared to matched WT controls, *ly96*$^{A/A}$ zebrafish survived for slightly longer (Fig 5B); *ly96*$^{B/B}$ zebrafish survived similarly (Fig 5C), and *ly96*$^{C/C}$ zebrafish survived shorter (Fig 5D). This is consistent with some other pathway rather than Tlr4/Md-2 being the primary route for LPS-induced death in larval zebrafish.

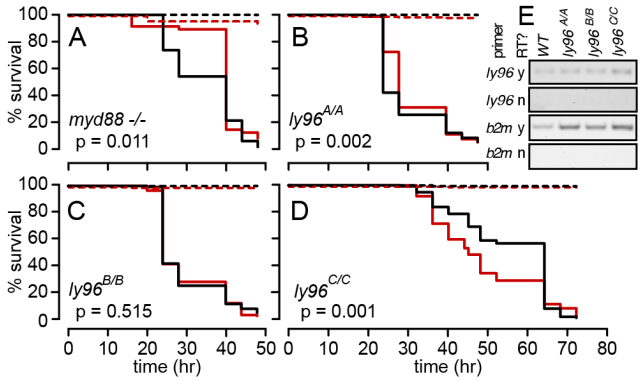

Fig 5. *ly96* mutations only moderately affect LPS survival in larval zebrafish. A-D) Curves show survival of wildtype (black) and mutant (red) zebrafish in the presence of 0.6 mg/mL LPS (solid line) or mock treatment (dashed line). The genotype is indicated on each panel. The p-value was determined by comparing the matched survival curves by a log-rank Mantel-Cox test. The experiments shown in panels A-C were performed in parallel, while the experiments in panel D were performed at a later date with an LPS lot that showed lower potency, necessitating a longer treatment time. Panels A-D represent averages of one, five, five, and three experimental repeats, respectively. Panel E shows mRNA transcript level for each mutant zebrafish. Rows use different primers (*ly96* or *b2m*) with and without reverse transcriptase (RT). Columns show fish genotype.

The zebrafish *tlr4* paralog arose after the evolution of *ly96*

Finally, we revisited the idea that the evolutionary history of zebrafish *tlr4* genes implies that they do not act as LPS-sensing molecules. Previous authors suggested that an ancestral *TLR* gene duplicated in the ancestor of bony vertebrates (~450 million years ago), and that the two paralogs were then differentially lost on the mammalian and bony fish lineages (21), respectively. This early divergence, before the evolution of *ly96*, may suggest very different functional roles for mammalian versus fish *tlr4*s.

We set out to better resolve when the zebrafish *tlr4* paralog arose relative to its mammalian counterparts, particularly with regard to the evolution of *ly96*. As with our analysis of Md-2, we started with a phylogenetic tree and then turned to synteny. For the phylogenetic tree, we constructed a multiple sequence alignment containing 263 Tlr4 sequences and 190 Cd180 protein sequences as an outgroup. (Cd180 is the most closely related paralog to Tlr4) (58). The alignment is available in File S2. In the resulting maximum likelihood tree, Tlr4 and Cd180 form distinct, well-supported clades (Fig 6A). Within the Tlr4 clade, zebrafish Tlr4a and Tlr4b are part of a monophyletic group with other Tlr4s.

We next investigated the genomic context for *tlr4* genes in eleven genomes, each with a complete chromosome assembly (Table S3). We selected a set of 22 genes flanking human *tlr4* and a set of 22 genes flanking the three zebrafish *tlr4* genes (Table S4). Notably, there were no shared homologs between the sets, demonstrating the radical difference between the genomic contexts of human and zebrafish *tlr4*. We then used these sets of genes to BLAST each of the eleven genomes and calculated a running average for the BLAST e-values along each chromosome. This allowed us to assess the overall similarity of genomic regions to either the human or zebrafish *tlr4* context. Fig 6B-G shows representative traces for six chromosomes taken from five species, with results for all genomes in Fig S2-S13. We were able to distinguish two distinct contexts for *tlr4* genes. In some organisms—human and frog, for example—*tlr4* is surrounded by hits from the human gene set (Fig 6B and D). In other organisms—zebrafish and pike, for example—*tlr4* is surrounded by hits from the zebrafish gene set (Fig 6C and E).

To place our results in their evolutionary context, we plotted our BLAST output against the phylogeny for our chosen species. For each species, we displayed the chromosome with the most hits from the human set (Fig 6H) and the chromosome with the most hits from the zebrafish

489 set (Fig 6I). We made an exception for the pike, displaying the chromosome with the *tlr4* gene
490 (linkage group 5), not the chromosome with the most zebrafish hits (linkage group 6). We
491 indicated whether a gene from the human or zebrafish set was seen somewhere on that
492 chromosome by coloring the square corresponding to that gene.

493 Four species had *tlr4* in a human-like context: human, chicken, frog and gar. None of
494 these species—including the gar—had a duplicate copy of *tlr4* in a zebrafish-like context. The
495 human-like context of the gar gene is shown Fig 6F, while the lack of *tlr4* in the most zebrafish-
496 like region of the gar genome is shown in Fig 6G. The remainder of the ray-finned fishes had *tlr4*
497 in either a zebrafish-like context (catfish, zebrafish, and pike) or had no *tlr4* gene at all
498 (bonytongue, cod, and puffer).

499 The most parsimonious history consistent with the observed distribution across genomes
500 is shown in Fig 6J. In this scenario, *tlr4* arose in a genomic context similar to the one preserved
501 in humans. This occurred after the divergence of bony and cartilaginous fishes (~475 million
502 years ago), but before the divergence of ray-finned and lobe-finned fishes (~430 million years
503 ago). The ancestral genomic context was preserved in tetrapods, including humans. It was also
504 maintained in the ray-finned fishes for ~130 million years, as indicated by the location of the *tlr4*
505 gene in the gar genome. Then, sometime between 300 and 250 million years ago, the *tlr4* gene
506 was both duplicated into the genomic context observed in zebrafish, as well as lost from the
507 ancestral context. Finally, between 150 and 100 million years ago, a tandem duplication occurred
508 within the *Cypriniformes* fishes, leading to the tandem copies of *tlr4* observed in zebrafish, carp,
509 and other *Cypriniformes* fishes.
510
511

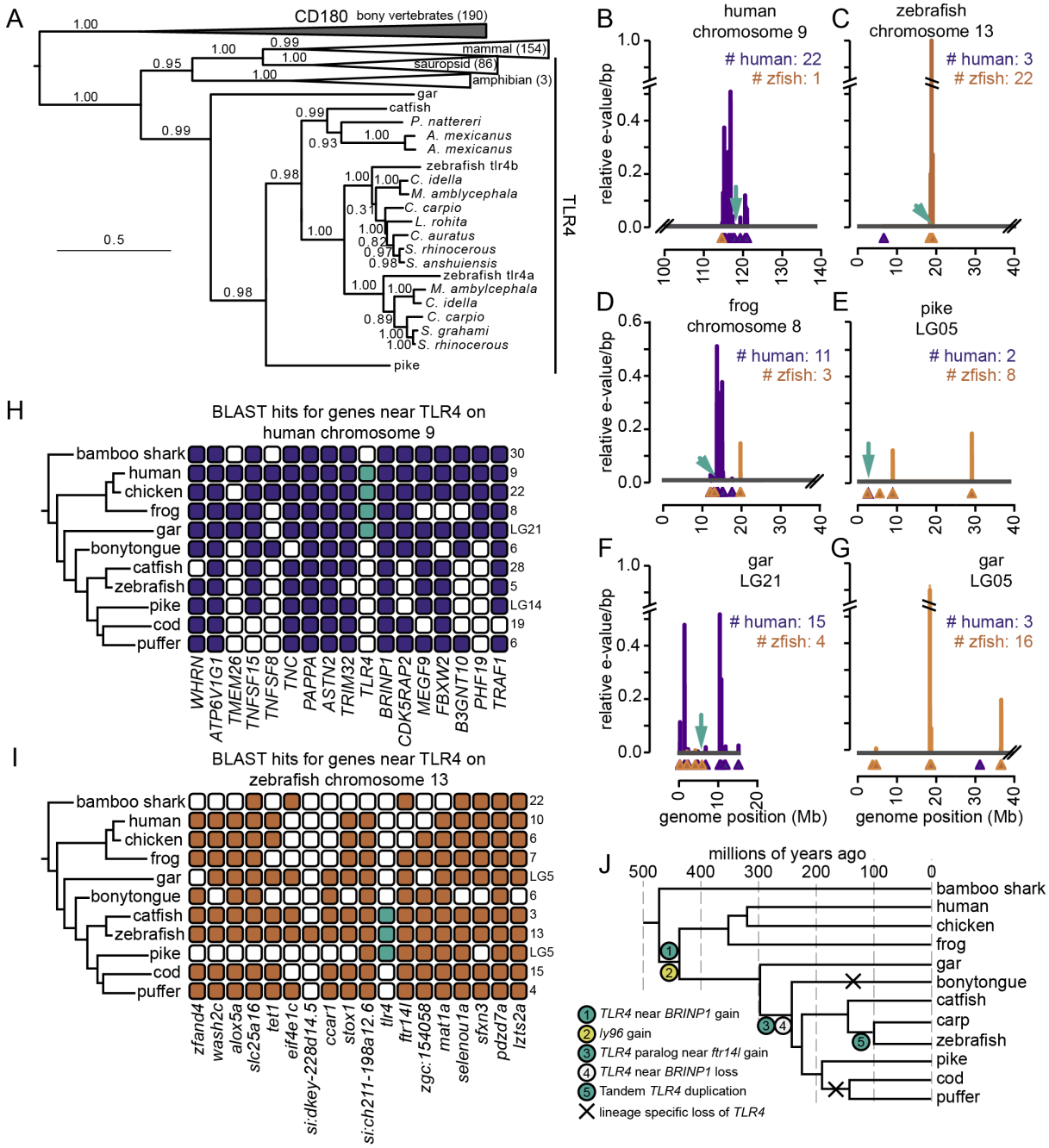

Fig 6. Zebrafish *tlr4* paralogs evolved within the ray-finned fishes. A) Maximum likelihood phylogeny for 453 Tlr4 and Cd180 protein sequences. SH supports are indicated on the tree. Wedges are clades, with the length indicating the maximum branch length from the ancestor of the clade. The taxonomic distribution and number of genes within each wedge are indicated on the plot. B-G) Hits for human (purple) and zebrafish (orange) gene sets on six representative chromosomes taken from five species. The species and chromosome are indicated at the top of each plot. The x-axis denotes position on the chromosome. Triangles indicate gene start positions. The green arrow indicates the location of a Tlr4 gene. The y-axis is a running average of the BLAST e-value for each gene set along the genome (see methods). The numbers on the plot indicate the number of human and zebrafish hits within the region shown. H,I) Each row

shows the chromosome with the most BLAST hits from the human (panel H) or zebrafish (panel I) gene set. Columns indicate specific genes from the set, with names denoted below. A colored square indicates a gene found somewhere on the chromosome. A green square is a *tlr4* gene. The species tree is shown on the left; the chromosome number is on the right. J) Schematic representation of a plausible scenario for the history of the *tlr4* gene. Times are taken from Hughes et al. (59) and timetree.org (60).

This revised evolutionary history places the evolution of the zebrafish *tlr4* paralogs much later than was previously hypothesized (21). Importantly, the duplication of TLR4 occurred *after* the evolution of Md-2, meaning that the formation of the Tlr4/Md-2 complex likely pre-dates the duplication event. Thus, the interaction with Md-2 and the ability to activate with LPS were an ancestral feature of zebrafish Tlr4 rather than something that could only be gained in parallel along the tetrapod and bony fish lineages.

DISCUSSION

Our observations led us to reevaluate the decade-old idea that Tlr4 does not participate in the LPS-induced inflammatory response in zebrafish. We have identified the zebrafish gene encoding the Tlr4 co-receptor Md-2 (*ly96*). The gene, like *tlr4ba* and *tlr4bb,* is transcribed in zebrafish cells that transcribe a collection of macrophage genes. In concert with zebrafish Tlr4a, zebrafish Md-2 is capable of activating NF-κB signaling in an *ex vivo* functional assay. Finally, a careful phylogenetic analysis suggests that the mammalian and zebrafish *tlr4* genes are not as evolutionarily distinct as previously thought. While not direct orthologs, the zebrafish paralogs evolved well after *ly96* and likely preserve an ancestral LPS recognition activity.

Our work demonstrates that, given the correct context, zebrafish Tlr4a and Md-2 form a functional complex that recognizes LPS and activates NF-κB signaling. Further, the molecular basis for the interaction between the partners appears to have been conserved for the last 450 million years - zebrafish Tlr4a is compatible with mouse and opossum Md-2 (Fig 4D). This is

despite the fact that the orthologous proteins from each species have only ~20% identity at the amino acid sequence level. The simplest explanation for this observation is that the ability of Tlr4/Md-2 to activate in response to LPS is an ancestral feature of the protein complex that has been conserved across the bony vertebrates—from mammals to bony fishes.

We have not shown, however, that LPS-induced Tlr4/Md-2 signaling actually occurs in zebrafish. Our two attempts to do so—our cell culture functional assay and analysis of zebrafish *ly96* loss of function mutants—both gave mixed results. We will discuss each in turn.

LPS activation of Tlr4a/Md-2 in human cells requires supporting molecules

In our functional assays, we had to add a mammalian Cd14 to activate NF-κB signaling through zebrafish Tlr4a/Md-2 (Fig 4C). In amniotes, Cd14 delivers LPS directly to Md-2 (Fig 1). We could find no ortholog to *cd14* in the zebrafish genome.

One possibility is that the human cell line used for the functional assays is missing some critical component for the delivery of LPS and assembly of the active dimer. Tlr4, Md-2, and Cd14 are the necessary and sufficient set of amniote proteins that confer an LPS-dependent NF-κB response in HEK293T cells. It could be that some other non-homologous protein plays the role of Cd14 in zebrafish.

Another possibility is that LPS is not a zebrafish Tlr4a/Md-2 agonist *in vivo*. We showed that we can activate the complex in a human cell line given an appropriate delivery molecule and a high enough LPS concentration. But, under physiological conditions, the Tlr4a/Md-2 complex could respond to some other chemically similar ligand. This would not be surprising: changes in ligand specificity have been observed across Md-2 in the amniotes (61). There is also some evidence that zebrafish Tlr4a may be antagonized by LPS *in vivo* (20). This would be compatible

573 with another ligand activating the complex and LPS competing and activating at a lower level
574 than can be achieved by the native ligand.

575 Finally, our observation that Tlr4a activates NF-κB with both mouse and opossum Md-2
576 directly contrasts previous work that showed the complex could not activate (Fig 4C) (20, 21).
577 The key difference between our experiments and those done previously is the sequence of *tlr4a*
578 used. Previous investigators used a construct that was ~75 amino acids shorter than tetrapod
579 Tlr4s. This construct is missing both the signal peptide required to target Tlr4a to the cell surface
580 and a region of the protein that is likely critical for Md-2 binding (Fig S13). In contrast, we used
581 a full-length ORF (ENSDART00000044697.6, GRCz10). The difference in our constructs arises
582 because the previous analysis relied on cDNA that, apparently, captured an alternate splice
583 variant of *tlr4a*.
584

585 **Multiple pathways contribute to LPS-induced death in larval zebrafish**

586 Larval zebrafish *ly96* loss of function mutants did not exhibit appreciably altered death
587 rates upon exposure to LPS compared to WT (Fig 5B-D). This is consistent with a previous
588 morpholino study that knocked down *tlr4a* and observed no change in sensitivity to LPS (20).
589 This contrasts with mice, however, where knockout of *LY96* is protective against endotoxic
590 shock (4) and disruption or knockout of *Tlr4* leads to hypo-responsiveness to LPS (1, 62).

591 We cannot rule out the possibility that this lack of response is due to an experimental
592 artifact. First, zebrafish may have retained a second copy of the *ly96* gene from the teleost
593 genome duplication that maintained function even after deletion of the targeted copy. We were
594 unable find any evidence of such a gene; however, the challenge of finding the original *ly96* gene
595 means that we cannot rule this out. A second possibility is that the mutants that we generated

596 may not represent a complete loss of function. For example, use of a potential alternative start
597 codon 17 amino acids downstream of the normal start codon could produce a truncated protein
598 (*ly96*$^{A/A}$). Although this would be missing N-terminal amino acids that are known to be critical
599 for Md-2 function in other systems (Table S2), these amino acids may not be necessary in
600 zebrafish. Finally, we tested a single developmental time point. It could be that Tlr4a/Md-2,
601 while expressed in larvae (Fig 3C), is not yet a large player in LPS sensing. Further experiments
602 on zebrafish at different time points may help clarify this point.

603 Another challenge is that LPS-mediated death is a relatively blunt instrument to test for
604 the activity of the Tlr4/Md-2 complex. We observed that addition of LPS dramatically increased
605 death rate (Fig 5A), even in a *myd88*$^{-/-}$ background. This indicates that at least one other non-
606 Toll-like pathway contributes to LPS-induced. One possibility is that this occurs by intracellular
607 sensing of LPS via caspases and inflammasomes (63). Various studies have shown
608 inflammasome signaling to be widespread in zebrafish larvae (64) and Il-1r to be required to
609 prevent cell death in response to infection in multiple cells (65). Intracellular sensing may be
610 much more important in fish than mammals: zebrafish have 385 of these putative intracellular
611 sensors whereas humans have 22 (66).

612 As a result of such alternate pathways, even if the Tlr4a/Md-2 complex contributes to the
613 LPS-induced inflammatory response, removing it might not lead to a measurable difference in
614 death rate. Compounding this difficulty, our expression analysis revealed that zebrafish *ly96* is
615 much more restricted in its expression than the corresponding mammalian genes (27). There
616 may, in fact, be specific subtypes of macrophages that express *ly96* and *tlr4s*—and are defective
617 in LPS sensing in the *ly96* mutants—but remain invisible at the level of LPS-induced death.
618 Higher-resolution studies of LPS-induced inflammation will be required to sort this out.

619

620 **Snapshot in the evolutionary history of this complex**

621 The presence of Md-2 in zebrafish indicates that both Tlr4 and Md-2 existed, together, in
622 the last common ancestor of bony vertebrates. Because descendants along both the tetrapod and
623 ray-finned fish lineages activate with LPS, the ability to respond to LPS is likely an ancestral
624 function that has been conserved for 435 million years.

625 That said, these proteins have evolved significantly since this shared ancestor. Along the
626 tetrapod lineage, a supporting collection of proteins evolved. Cd14 arose through a duplication
627 within the Toll-like receptor family and is now an essential component of the Tlr4/Md-2
628 complex, delivering LPS to Md-2 in a coordinated fashion (53, 54). Tetrapods also acquired
629 Lipid Binding Protein (LBP), improving LPS delivery (52, 67). Amniotes then further adjusted
630 the Tlr4/Md-2 pro-inflammatory response through addition of amniote-specific Damage-
631 Associated Molecular Pattern (DAMP) molecules such as S100A9 (68), which amplify LPS-
632 induced inflammation (69). All the while, mutations to Md-2 changed its specificity for LPS and
633 its chemical analogs (70). For example, humans acquired unique lipid IVa antagonism sometime
634 after the divergence of humans and mice (71, 72).

635 The changes that occurred along the ray-finned fish lineages are not yet clear. Did they
636 acquire supporting LPS delivery molecules analogous to Cd14? Has the specificity of Md-2
637 fluctuated in ray-finned fishes as it has along the tetrapod lineage? Further work is needed to
638 answer these questions.

639 We hypothesize, however, that ray-finned fishes maintain an ancestral, low-sensitivity
640 Tlr4/Md-2 LPS sensing complex. Fish have previously been shown to be relatively resistant to
641 septic shock (73, 74), with high concentrations of LPS needed to activate teleost leukocytes (25,

642 75–77). This parallels the observation that early diverging tetrapods, such as amphibians, also
643 require high doses of LPS to trigger an inflammatory response (78). This could be explained if
644 ray-finned fishes do not have specialized machinery to deliver LPS to the complex, but instead
645 use Tlr4/Md-2 as a simple LPS sensor. Other observations consistent with a relatively primitive
646 Tlr4/Md-2 LPS response in zebrafish are the fact that Tlr4 was lost, independently, along
647 multiple fish lineages (20, 21, 73) (Fig 6J), as well as the existence of parallel LPS sensing
648 pathways in zebrafish (64). If the Tlr4/Md-2 complex is peripheral to the LPS response in ray-
649 finned fishes, it could be lost with minimal fitness consequences. In contrast, Tlr4/Md-2 became
650 progressively more central to the LPS response along the mammalian lineage—and as a result
651 has been highly conserved.

652 Our work suggests that we should re-visit our understanding of LPS signaling through
653 Tlr4/Md-2 in zebrafish. We hypothesize that zebrafish preserve an ancestral, low-sensitivity
654 Tlr4/Md-2 complex. In contrast to mammals—in which the Tlr4/Md-2 complex is the primary
655 LPS receptor—the zebrafish Tlr4/Md-2 complex acts in parallel with several LPS-sensitive
656 pathways, likely playing roles in a small population of innate immune cells.

657

658 **ACKNOWLEDGMENTS**

659 We thank Kristi Hamilton and Lila Kaye for assistance with zebrafish LPS survival assays, and
660 Rose Sockol and the University of Oregon Aquatic Animal Care Services staff for fish
661 husbandry. We thank Prof. Carol Kim for sharing the *tlr4bb* plasmid.

662

663

664 **REFERENCES**

665 1. Poltorak, A. 1998. Defective LPS Signaling in C3H/HeJ and C57BL/10ScCr Mice: Mutations
666 in Tlr4 Gene. *Science (80-.).* 282: 2085–2088.

667 2. Schletter, J., H. Heine, A. J. Ulmer, and E. T. Rietschel. 1995. Molecular mechanisms of
668 endotoxin activity. *Arch. Microbiol.* 164: 383–389.

669 3. Ulevitch, R. J., and P. S. Tobias. 1995. Receptor-Dependent Mechanisms of Cell Stimulation
670 by Bacterial Endotoxin. *Annu. Rev. Immunol.* 13: 437–457.

671 4. Nagai, Y., S. Akashi, M. Nagafuku, M. Ogata, Y. Iwakura, S. Akira, T. Kitamura, A. Kosugi,
672 M. Kimoto, and K. Miyake. 2002. Essential role of MD-2 in LPS responsiveness and TLR4
673 distribution. *Nat. Immunol.* 3: 667.

674 5. Park, B. S., and J.-O. Lee. 2013. Recognition of lipopolysaccharide pattern by TLR4
675 complexes. *Exp. Mol. Med.* 45: e66.

676 6. Chow, J. C., D. W. Young, D. T. Golenbock, W. J. Christ, and F. Gusovsky. 1999. Toll-like
677 receptor-4 mediates lipopolysaccharide-induced signal transduction. *J. Biol. Chem.* 274: 10689–
678 92.

679 7. Yiu, J. H. C., B. Dorweiler, and C. W. Woo. 2017. Interaction between gut microbiota and
680 toll-like receptor: from immunity to metabolism. *J. Mol. Med. (Berl).* 95: 13–20.

681 8. Fan, J., and A. B. Malik. 2003. Toll-like receptor-4 (TLR4) signaling augments chemokine-
682 induced neutrophil migration by modulating cell surface expression of chemokine receptors. *Nat.*
683 *Med.* 9: 315–321.

684 9. Murad, S. 2014. Toll-Like Receptor 4 in Inflammation and Angiogenesis: A Double-Edged
685 Sword. *Front. Immunol.* 5: 313.

686 10. Leon, C. G., R. Tory, J. Jia, O. Sivak, and K. M. Wasan. 2008. Discovery and development
687 of toll-like receptor 4 (TLR4) antagonists: a new paradigm for treating sepsis and other diseases.
688 *Pharm. Res.* 25: 1751–61.

689 11. Kuzmich, N. N., K. V Sivak, V. N. Chubarev, Y. B. Porozov, T. N. Savateeva-Lyubimova,
690 and F. Peri. 2017. TLR4 Signaling Pathway Modulators as Potential Therapeutics in
691 Inflammation and Sepsis. *Vaccines* 5.

692 12. Zettel, K., S. Korff, R. Zamora, A. E. Morelli, S. Darwiche, P. A. Loughran, G. Elson, L.
693 Shang, S. Salgado-Pires, M. J. Scott, Y. Vodovotz, and T. R. Billiar. 2017. Toll-Like Receptor 4
694 on both Myeloid Cells and Dendritic Cells Is Required for Systemic Inflammation and Organ
695 Damage after Hemorrhagic Shock with Tissue Trauma in Mice. *Front. Immunol.* 8: 1672.

696 13. Gruffaz, M., K. Vasan, B. Tan, S. Ramos da Silva, and S.-J. Gao. 2017. TLR4-Mediated
697 Inflammation Promotes KSHV-Induced Cellular Transformation and Tumorigenesis by
698 Activating the STAT3 Pathway. *Cancer Res.* 77: 7094–7108.

699 14. Koliaraki, V., N. Chalkidi, A. Henriques, C. Tzaferis, A. Polykratis, A. Waisman, W. Muller,
700 D. J. Hackam, M. Pasparakis, and G. Kollias. 2019. Innate Sensing through Mesenchymal
701 TLR4/MyD88 Signals Promotes Spontaneous Intestinal Tumorigenesis. *Cell Rep.* 26: 536–
702 545.e4.

703 15. Yoder, J. A., M. E. Nielsen, C. T. Amemiya, and G. W. Litman. 2002. Zebrafish as an
704 immunological model system. *Microbes Infect.* 4: 1469–1478.

705 16. Renshaw, S. A., and N. S. Trede. 2012. A model 450 million years in the making: zebrafish
706 and vertebrate immunity. *Dis. Model. Mech.* 5: 38–47.

17. Hall, C., M. V. Flores, A. Chien, A. Davidson, K. Crosier, and P. Crosier. 2009. Transgenic zebrafish reporter lines reveal conserved Toll-like receptor signaling potential in embryonic myeloid leukocytes and adult immune cell lineages. *J. Leukoc. Biol.* 85: 751–765.

18. Burns, A. R., and K. Guillemin. 2017. The scales of the zebrafish: host-microbiota interactions from proteins to populations. *Curr. Opin. Microbiol.* 38: 137–141.

19. Park, B. S., D. H. Song, H. M. Kim, B.-S. Choi, H. Lee, and J.-O. Lee. 2009. The structural basis of lipopolysaccharide recognition by the TLR4–MD-2 complex. *Nature* 458: 1191–1195.

20. Sepulcre, M. P., F. Alcaraz-Pérez, A. López-Muñoz, F. J. Roca, J. Meseguer, M. L. Cayuela, and V. Mulero. 2009. Evolution of lipopolysaccharide (LPS) recognition and signaling: fish TLR4 does not recognize LPS and negatively regulates NF-kappaB activation. *J. Immunol.* 182: 1836–45.

21. Sullivan, C., J. Charette, J. Catchen, C. R. Lage, G. Giasson, J. H. Postlethwait, P. J. Millard, and C. H. Kim. 2009. The Gene History of Zebrafish tlr4a and tlr4b Is Predictive of Their Divergent Functions. *J. Immunol.* 183: 5896–5908.

22. Bates, J. M., J. Akerlund, E. Mittge, and K. Guillemin. 2007. Intestinal Alkaline Phosphatase Detoxifies Lipopolysaccharide and Prevents Inflammation in Zebrafish in Response to the Gut Microbiota. *Cell Host Microbe* 2: 371–382.

23. van der Sar, A. M., O. W. Stockhammer, C. van der Laan, H. P. Spaink, W. Bitter, and A. H. Meijer. 2006. MyD88 innate immune function in a zebrafish embryo infection model. *Infect. Immun.* 74: 2436–41.

24. Forn-Cuní, G., M. Varela, P. Pereiro, B. Novoa, and A. Figueras. 2017. Conserved gene

728 regulation during acute inflammation between zebrafish and mammals. *Sci. Rep.* 7: 41905.

729 25. Sepulcre, M. P., G. López-Castejón, J. Meseguer, and V. Mulero. 2007. The activation of
730 gilthead seabream professional phagocytes by different PAMPs underlines the behavioural
731 diversity of the main innate immune cells of bony fish. *Mol. Immunol.* 44: 2009–2016.

732 26. Ellett, F., L. Pase, J. W. Hayman, A. Andrianopoulos, and G. J. Lieschke. 2011. mpeg1
733 promoter transgenes direct macrophage-lineage expression in zebrafish. *Blood* 117: e49–e56.

734 27. Vaure, C., and Y. Liu. 2014. A Comparative Review of Toll-Like Receptor 4 Expression and
735 Functionality in Different Animal Species. *Front. Immunol.* 5: 316.

736 28. Lu, Y.-C., W.-C. Yeh, and P. S. Ohashi. 2008. LPS/TLR4 signal transduction pathway.
737 *Cytokine* 42: 145–151.

738 29. Kimoto, M., K. Nagasawa, and K. Miyake. 2003. Role of TLR4/MD-2 and RP105/MD-1 in
739 Innate Recognition of Lipopolysaccharide. *Scand. J. Infect. Dis.* 35: 568–572.

740 30. Qiao, L., W. Yang, J. Fu, and Z. Song. 2013. Transcriptome Profile of the Green Odorous
741 Frog (Odorrana margaretae). *PLoS One* 8: e75211.

742 31. Matsunami, M., J. Kitano, O. Kishida, H. Michimae, T. Miura, and K. Nishimura. 2015.
743 Transcriptome analysis of predator- and prey-induced phenotypic plasticity in the Hokkaido
744 salamander (*Hynobius retardatus*). *Mol. Ecol.* 24: 3064–3076.

745 32. Nourisson, C., M. Carneiro, M. Vallinoto, and F. Sequeira. 2014. De novo transcriptome
746 assembly and polymorphism detection in ecologically important widely distributed Neotropical
747 toads from the Rhinella marina species complex (Anura: Bufonidade. *Genomic Resour. Notes*
748 *Accept. 1 August 2014-30 Sept. 2014* .

749 33. Genomic Resources Development Consortium, W. Arthofer, B. L. Banbury, M. Carneiro, F.
750 Cicconardi, T. F. Duda, R. B. Harris, D. S. Kang, A. D. Leaché, V. Nolte, C. Nourisson, N.
751 Palmieri, B. C. Schlick-Steiner, C. Schlötterer, F. Sequeira, C. Sim, F. M. Steiner, M. Vallinoto,
752 and D. A. Weese. 2015. Genomic Resources Notes Accepted 1 August 2014-30 September 2014.
753 *Mol. Ecol. Resour.* 15: 228–229.

754 34. Liu, Y., B. Schmidt, and D. L. Maskell. 2010. MSAProbs: multiple sequence alignment
755 based on pair hidden Markov models and partition function posterior probabilities.
756 *Bioinformatics* 26: 1958–1964.

757 35. Tamura, K., G. Stecher, D. Peterson, A. Filipski, and S. Kumar. 2013. MEGA6: Molecular
758 Evolutionary Genetics Analysis Version 6.0. *Mol. Biol. Evol.* 30: 2725–2729.

759 36. Guindon, S., J. F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O. Gascuel. 2010.
760 New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the
761 Performance of PhyML 3.0. *Syst. Biol.* 59: 307–321.

762 37. Le, S. Q., and O. Gascuel. 2010. Accounting for Solvent Accessibility and Secondary
763 Structure in Protein Phylogenetics Is Clearly Beneficial. *Syst. Biol.* 59: 277–287.

764 38. Le, S. Q., and O. Gascuel. 2008. An Improved General Amino Acid Replacement Matrix.
765 *Mol. Biol. Evol.* 25: 1307–1320.

766 39. Hordijk, W., and O. Gascuel. 2005. Improving the efficiency of SPR moves in phylogenetic
767 tree search methods based on maximum likelihood. *Bioinformatics* 21: 4338–4347.

768 40. Anisimova, M., and O. Gascuel. 2006. Approximate Likelihood-Ratio Test for Branches: A
769 Fast, Accurate, and Powerful Alternative. *Syst. Biol.* 55: 539–552.

770 41. Akaike, H. 1998. Information Theory and an Extension of the Maximum Likelihood
771 Principle. In Springer New York. 199–213.

772 42. Waterhouse, A. M., J. B. Procter, D. M. A. Martin, M. Clamp, and G. J. Barton. 2009.
773 Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics*
774 25: 1189–1191.

775 43. Herrero, J., M. Muffato, K. Beal, S. Fitzgerald, L. Gordon, M. Pignatelli, A. J. Vilella, S. M.
776 J. Searle, R. Amode, S. Brent, W. Spooner, E. Kulesha, A. Yates, and P. Flicek. 2016. Ensembl
777 comparative genomics resources. *Database* 2016: bav096.

778 44. Farnsworth, D. R., L. Saunders, and A. C. Miller. 2019. A Single-Cell Transcriptome Atlas
779 for Zebrafish Development. *bioRxiv* 738344.

780 45. Zheng, G. X. Y., J. M. Terry, P. Belgrader, P. Ryvkin, Z. W. Bent, R. Wilson, S. B. Ziraldo,
781 T. D. Wheeler, G. P. McDermott, J. Zhu, M. T. Gregory, J. Shuga, L. Montesclaros, J. G.
782 Underwood, D. A. Masquelier, S. Y. Nishimura, M. Schnall-Levin, P. W. Wyatt, C. M. Hindson,
783 R. Bharadwaj, A. Wong, K. D. Ness, L. W. Beppu, H. J. Deeg, C. McFarland, K. R. Loeb, W. J.
784 Valente, N. G. Ericson, E. A. Stevens, J. P. Radich, T. S. Mikkelsen, B. J. Hindson, and J. H.
785 Bielas. 2017. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* 8:
786 14049.

787 46. Satija, R., J. A. Farrell, D. Gennert, A. F. Schier, and A. Regev. 2015. Spatial reconstruction
788 of single-cell gene expression data. *Nat. Biotechnol.* 33: 495–502.

789 47. R Core Team. 2017. R: A Language and Environment for Statistical Computing. .

790 48. Jao, L.-E., S. R. Wente, and W. Chen. 2013. Efficient multiplex biallelic zebrafish genome

791 editing using a CRISPR nuclease system. *Proc. Natl. Acad. Sci. U. S. A.* 110: 13904–9.

792 49. Sun, Y., Y. Huang, X. Li, C. C. Baldwin, Z. Zhou, Z. Yan, K. A. Crandall, Y. Zhang, X.
793 Zhao, M. Wang, A. Wong, C. Fang, X. Zhang, H. Huang, J. V. Lopez, K. Kilfoyle, Y. Zhang, G.
794 Ortí, B. Venkatesh, and Q. Shi. 2016. Fish-T1K (Transcriptomes of 1,000 Fishes) Project: Large-
795 scale transcriptome data for fish evolution studies. *Gigascience* 5.

796 50. Woods, I. G., C. Wilson, B. Friedlander, P. Chang, D. K. Reyes, R. Nix, P. D. Kelly, F. Chu,
797 J. H. Postlethwait, and W. S. Talbot. 2005. The zebrafish gene map defines ancestral vertebrate
798 chromosomes. *Genome Res.* 15: 1307–14.

799 51. Bernut, A., J.-L. Herrmann, K. Kissa, J.-F. Dubremetz, J.-L. Gaillard, G. Lutfalla, and L.
800 Kremer. 2014. *Mycobacterium abscessus* cording prevents phagocytosis and promotes abscess
801 formation. *Proc. Natl. Acad. Sci.* 111: E943–E952.

802 52. Heumann, D., R. Lauener, and B. Ryffel. 2003. The dual role of LBP and CD14 in response
803 to Gram-negative bacteria or Gram-negative compounds. *J. Endotoxin Res.* 9: 381–384.

804 53. Triantafilou, M., and K. Triantafilou. 2002. Lipopolysaccharide recognition: CD14, TLRs
805 and the LPS-activation cluster. *Trends Immunol.* 23: 301–304.

806 54. Kim, S. J., and H. M. Kim. 2017. Dynamic lipopolysaccharide transfer cascade to
807 TLR4/MD2 complex via LBP and CD14. *BMB Rep.* 50: 55–57.

808 55. Shapiro, R. A., M. D. Cunningham, K. Ratcliffe, C. Seachord, J. Blake, J. Bajorath, A.
809 Aruffo, and R. P. Darveau. 1997. Identification of CD14 residues involved in specific
810 lipopolysaccharide recognition. *Infect. Immun.* 65: 293–7.

811 56. Juan, T. S.-C., E. Hailman, M. J. Kelley, S. D. Wright, and H. S. Lichenstein. 1995.

812 Identification of a Domain in Soluble CD14 Essential for Lipopolysaccharide (LPS) Signaling
813 but Not LPS Binding. *J. Biol. Chem.* 270: 17237–17242.

814 57. Zhou, Y., Q. Liang, W. Li, Y. Gu, X. Liao, W. Fang, and X. Li. 2016. Characterization and
815 functional analysis of toll-like receptor 4 in Chinese soft-shelled turtle Pelodiscus sinensis. *Dev.*
816 *Comp. Immunol.* 63: 128–135.

817 58. Madan, R., D. Golenbock, S. Divanovic, A. Trompette, S. F. Atabani, D. T. Golenbock, A.
818 Visintin, R. W. Finberg, A. Tarakhovsky, S. N. Vogel, Y. Belkaid, E. A. Kurt-Jones, and C. L.
819 Karp. 2005. Negative Regulation of TLR4 Signaling by RP105. .

820 59. Hughes, L. C., G. Ortí, Y. Huang, Y. Sun, C. C. Baldwin, A. W. Thompson, D. Arcila, R.
821 Betancur-R., C. Li, L. Becker, N. Bellora, X. Zhao, X. Li, M. Wang, C. Fang, B. Xie, Z. Zhou,
822 H. Huang, S. Chen, B. Venkatesh, and Q. Shi. 2018. Comprehensive phylogeny of ray-finned
823 fishes (Actinopterygii) based on transcriptomic and genomic data. *Proc. Natl. Acad. Sci.* 115:
824 6249–6254.

825 60. Hedges, S. B., J. Dudley, and S. Kumar. 2006. TimeTree: a public knowledge-base of
826 divergence times among organisms. *Bioinformatics* 22: 2971–2972.

827 61. Muroi, M., T. Ohnishi, and K.-I. Tanamoto. 2002. MD-2, a Novel Accessory Molecule, Is
828 Involved in Species-Specific Actions of Salmonella Lipid A. *Infect. Immun.* 70: 3546–3550.

829 62. Hoshino, K., O. Takeuchi, T. Kawai, H. Sanjo, T. Ogawa, Y. Takeda, K. Takeda, and S.
830 Akira. 1999. Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to
831 lipopolysaccharide: evidence for TLR4 as the Lps gene product. *J. Immunol.* 162: 3749–52.

832 63. Yang, D., X. Zheng, S. Chen, Z. Wang, W. Xu, J. Tan, T. Hu, M. Hou, W. Wang, Z. Gu, Q.

833 Wang, R. Zhang, Y. Zhang, and Q. Liu. 2018. Sensing of cytosolic LPS through caspy2 pyrin
834 domain mediates noncanonical inflammasome activation in zebrafish. *Nat. Commun.* 9: 3052.

835 64. Forn-Cuní, G., A. H. Meijer, and M. Varela. 2019. Zebrafish in Inflammasome Research.
836 *Cells* 8: 901.

837 65. Mazon-Moya, M. J., A. R. Willis, V. Torraca, L. Boucontet, A. R. Shenoy, E. Colucci-
838 Guyon, and S. Mostowy. 2017. Septins restrict inflammation and protect zebrafish larvae from
839 Shigella infection. *PLOS Pathog.* 13: e1006467.

840 66. Jones, J. D. G., R. E. Vance, and J. L. Dangl. 2016. Intracellular innate immune surveillance
841 devices in plants and animals. *Science (80-.).* 354: aaf6395–aaf6395.

842 67. Bernheiden, M., J.-M. Heinrich, G. Minigo, C. Schutt, F. Stelter, M. Freeman, D. Golenbock,
843 and R. S. Jack. 2001. LBP, CD14, TLR4 and the murine innate immune response to a peritoneal
844 Salmonella infection. *J. Endotoxin Res.* 7: 447–450.

845 68. Loes, A. N., J. T. Bridgham, and M. J. Harms. 2018. Coevolution of the Toll-Like Receptor 4
846 Complex with Calgranulins and Lipopolysaccharide. *Front. Immunol.* 9: 304.

847 69. Ehrchen, J. M., C. Sunderkotter, D. Foell, T. Vogl, and J. Roth. 2009. The endogenous Toll-
848 like receptor 4 agonist S100A8/S100A9 (calprotectin) as innate amplifier of infection,
849 autoimmunity, and cancer. *J. Leukoc. Biol.* 86: 557–566.

850 70. Ohto, U., K. Fukase, K. Miyake, and T. Shimizu. 2012. Structural basis of species-specific
851 endotoxin sensing by innate immune receptor TLR4/MD-2. *Proc. Natl. Acad. Sci.* 109: 7421–
852 7426.

853 71. Walsh, C., M. Gangloff, T. Monie, T. Smyth, B. Wei, T. J. McKinley, D. Maskell, N. Gay,

854 and C. Bryant. 2008. Elucidation of the MD-2/TLR4 Interface Required for Signaling by Lipid
855 IVa. *J. Immunol.* 181: 1245–1254.

856 72. Anderson, J. A., A. N. Loes, G. L. Waddell, and M. J. Harms. 2019. Tracing the evolution of
857 novel features of human Toll-like receptor 4. *Protein Sci.* 28: pro.3644.

858 73. Iliev, D. B., J. C. Roach, S. Mackenzie, J. V Planas, and F. W. Goetz. 2005. Endotoxin
859 recognition: in fish or not in fish? *FEBS Lett.* 579: 6519–28.

860 74. Berczi, I., L. Bertók, and T. Bereznai. 1966. Comparative studies on the toxicity of
861 Escherichia coli lipopolysaccaride endotoxin in various animal species. *Can. J. Microbiol.* 12:
862 1070–1071.

863 75. Stafford, J. L., K. K. Ellestad, K. E. Magor, M. Belosevic, and B. G. Magor. 2003. A toll-like
864 receptor (TLR) gene that is up-regulated in activated goldfish macrophages. *Dev. Comp.*
865 *Immunol.* 27: 685–698.

866 76. MacKenzie, S., J. V. Planas, and F. W. Goetz. 2003. LPS-stimulated expression of a tumor
867 necrosis factor-alpha mRNA in primary trout monocytes and in vitro differentiated macrophages.
868 *Dev. Comp. Immunol.* 27: 393–400.

869 77. Pelegrín, P., J. García-Castillo, V. Mulero, and J. Meseguer. 2001. Interleukin-1β isolated
870 from a marine fish up-regulated expression in macrophages following activation with
871 lipopolysaccharide and lymphokines. *Cytokine* 16: 67–72.

872 78. Ishii, A., M. Kawasaki, M. Matsumoto, S. Tochinai, and T. Seya. 2007. Phylogenetic and
873 expression analysis of amphibian Xenopus Toll-like receptors. *Immunogenetics* 59: 281–293.

874