Functional characterization of the cnidarian antiviral immune response reveals ancestral complexity

Magda Lewandowska[1,*], Ton Sharoni[1], Yael Admoni[1], Reuven Aharoni[1], Yehu Moran[1,*]

[1]Department of Ecology, Evolution and Behavior, Alexander Silberman Institute of Life Sciences, Faculty of Science, Hebrew University of Jerusalem, Jerusalem 9190401, Israel

[*]Correspondence: magda.lewandowska@mail.huji.ac.il (M.L.); yehu.moran@mail.huji.ac.il (Y.M.)

ABSTRACT

Animals developed a broad repertoire of innate immune sensors and downstream effector cascades for defense against RNA viruses. Yet, this system highly varies between different bilaterian animals, masking its ancestral state. In this study we aimed to characterize the antiviral immune response of the cnidarian *Nematostella vectensis* and decipher the function of the retinoic acid-inducible gene I-like receptors (RLRs) known to detect viral double-stranded RNA (dsRNA) in bilaterians, but activate different antiviral pathways in vertebrates and nematodes. We show that a mimic of long viral dsRNA triggers a complex antiviral immune response bearing features distinctive for both vertebrate and invertebrate systems. Furthermore, the results of affinity assays and knockdown experiments provide functional evidence for the conserved role of RLRs in initiating immune response to dsRNA that originated before the cnidarian-bilaterian split and lay a strong foundation for future research on the evolution of the immune responses to RNA viruses.

Introduction

The immune system has long been known for its remarkable patterns of rapid evolution owing to strong selective drivers such as fast-evolving pathogens[1,2]. Thus, revealing conservation among phylogenetically distant lineages can provide unprecedented insights into the evolution of these defense mechanisms. For instance, it has been recently reported that both eukaryotic antiviral DNA-sensing mechanism driven by cGAS-STING axis and the downstream inhibitors of virus replication called viperins have originated in procaryotes as anti-bacteriophage mechanisms[3-5]. Viruses are very often sensed by their nucleic acids which bear features not shared by their host cells[6,7]. Specifically, eukaryotes had to adapt to emerging RNA viruses by developing strategies to recognize such non-self genetic material. The best characterized foreign features are i) double-stranded RNA (dsRNA) structures and ii) triphosphate on 5' ends, both of which are mostly absent during host cell homeostasis but are accumulated in viral infection, either directly derived from the viral genomes or formed as the replication or transcription intermediates[8-10]. In plants, nematodes and arthropods, the presence of the cytoplasmatic dsRNA triggers RNA interference (RNAi) which involves slicing dsRNA into short interfering RNAs (siRNA) by the ribonuclease III Dicer, often followed by signal amplification by RNA-dependent RNA polymerases (RdRPs) and final silencing of viral RNA by Argonaute proteins[11-14]. In vertebrates, dsRNA is detected by several families of pattern-recognition receptors (PRRs) which trigger downstream expression of type I interferons (IFNs) and other proinflammatory cytokines[7]. Retinoic acid-inducible gene I (RIG-I) -like receptors (RLRs) is a family of metazoan-specific ATP-dependent DExD/H box RNA helicases that function as the major cytoplasmic PRRs binding dsRNA[15-17] (**Fig. 1a**). In vertebrates, ligands of RIG-I and its paralog melanoma differentiation-associated protein 5 (MDA5) include short, blunt-end dsRNA with 5' di- and triphosphate[9,18-22] and long irregular dsRNA[23-26], respectively. Caspase activation and recruitment domains (CARDs) of RIG-I and MDA5 (**Fig. 1b**) are necessary for regulation, oligomerization and subsequent interaction with adaptor molecules to trigger downstream effector cascades[27]. Absence of the CARD domain in the third vertebrate RLR – Laboratory of Genetics and Physiology 2 (LGP2) – prevents signal transduction and is correlated with its dual regulatory functions[28].

Although RLRs have been found in many animal phyla (**Fig. 1a**) and display structural conservation **(Fig. 1b)**, their function in invertebrate immune response remains understudied in animals other than vertebrates and nematodes, leaving a major gap in the understanding of RLRs evolution. In this study we aimed to characterize the immune response to viral dsRNA mimics in *Nematostella vectensis*, a model organism of phylum Cnidaria (sea anemones, corals, jellyfish and hydroids) separated from its sister group Bilateria (the majority of extant animals, including vertebrates and nematodes) by > 600 million years of evolution[29,30]. We observe in this cnidarian a strong immune response triggered by long, but not short 5' triphosphate-bearing dsRNA which supports our phylogenetic analyses of RLRs. We show that both *N. vectensis* RLRs (NveRLRs) are likely to take part in the antiviral immune response and that one of them is showing affinity to long dsRNA. Finally, knockdown of this RLR results in impaired downstream effector processes suggesting its key role in initiating immune response to dsRNA.

Results

Ancestral RLRs duplication likely predates the Bilateria-Cnidaria split

In order to gain a better understanding of the evolutionary fate of RLRs and the position of *N. vectensis* homologs within the family of these viral nucleic acid sensors, we reconstructed previous phylogenetic trees with an addition of numerous recently available sequences. Instead of including other distantly related DExD/H helicases, such as RNA-specific endoribonuclease Dicer or elongation initiation factor 4A (eIF4A), we performed phylogenetic analysis exclusively of RLRs with the sequences of sponges, one of the first two metazoan phyla to diverge[31,32], set as an outgroup (**Fig. 1a**). Similar to previous studies[33,34], we have not identified any RLRs homologs in Placozoa and Ctenophora. Within Cnidaria, we identified RLRs sequences in Hexacorallia (sea anemones and stony corals) while they are absent in the Meduzosoa clade, clearly indicating a loss. Interestingly, unlike in previous studies[33,34], we observed a well-supported clustering of all hexacorallian RLRs within bilaterian MDA5/LGP2 clade, which forms a sister group to bilaterian RIG-I sequences. This unexpected finding suggests that, in contrast to the previous hypothesis[35], RLRs paralogs duplicated

before the split of Bilateria and Cnidaria and all cnidarian RLRs paralogs originated from a MDA5/LGP2 ancestral protein. Furthermore, both *Nematostella* CARD-containing protein sequences – NveRLRa and NveRLRb – are positioned in separated clades with orthologs from other sea anemones suggesting their ancient duplication predating sea anemone divergence and most likely the functional non-redundancy. Clustering of RLRb sequences of sea anemones within one of the clades of stony corals which have split 320 million years ago[36] further supports the hypothesis of an ancient sub- or neofunctionalization of the sea anemone RLRs.

Lack of response in *Nematostella* to RIG-I specific ligand

To functionally support our observation that *Nematostella* RLRs are closer related to Bilateria MDA5 receptor, we decided to first employ known ligand affinity and test *Nematostella* response to MDA5 and RIG-I-specific ligands. To this end, we microinjected *N. vectensis* embryos with polyinosinic:polycytidylic acid (poly(I:C)), a mimic of long viral dsRNA and a potent agonist of MDA5[25,26], and a short dsRNA 19-mer with 5' triphosphate group (5'ppp-dsRNA) which is known to be detected by RIG-I[9,19]. Analysis of differentially expressed genes (DEG) upon the treatments with viral mimics revealed a strong response to poly(I:C) (**Fig. 2a, Supplementary File S1, Supplementary Fig. S1d,e**) with a peak of the differential expression at 24 hours post-injection (hpi) accounting for 67.26% of variance revealed by Principal Component 1 (**Fig. 3c**). Among 3 different time points, we have observed an almost complete lack of transcriptomic response in 6 hpi (*n* of DEG = 14) which agrees with a low transcript abundance at the onset of zygotic transcription in *Nematostella*[37]. Both at 24 and 48 hpi (*n* of DEG = 1475 and 524, respectively) the majority of DEG were upregulated (**Fig. 3a,b**) which is a common pattern of the innate immune response to viral ligands[38]. In contrast, the transcriptomic response to vertebrate RIG-I specific dsRNA ligand revealed a striking lack of signature of the antiviral immune processes (**Fig. 2b, Supplementary File 1, Supplementary Fig. S1a,b,c**) despite being applied at 90 – 180-fold higher concentration compared to concentrations used for vertebrates, suggesting that unlike in vertebrates[39-41], a triphosphate group on 5' blunt-end of short dsRNA is not triggering an immune reaction in *N. vectensis*.

Results of the gene-set enrichment analysis (GSEA) revealed the abundance of gene ontology (GO) terms related to the innate immunity and strengthened our inference on strong antiviral response triggered by poly(I:C) at 24 hpi (**Fig. 3d, Supplementary File S2**) and to a lesser extent at 48 hpi (**Fig. 3e, Supplementary File S2**). Importantly, the vast majority of responding genes at the later stage overlaps with the upregulated genes of the former one (**Fig. 3a**), suggesting a continuous attenuating immune response. In all tested groups enriched GO terms contained many vertebrate-specific terms, therefore we had to treat it as an approximation to a true gene function. Although the GSEA for the short 5'ppp-dsRNA had not revealed enriched GO terms which would pass the statistical threshold likely due to the low DEG abundance (**Supplementary File S2**), we decided to examine the only DEG group responding to this treatment i.e. genes downregulated at 6 hpi (**Fig. 2b**). Identified GO terms groups were predominantly related to the early-stage development (**Supplementary Fig. S1f**) which led us to the hypothesis that the presence of very high molarity of charged compounds might either directly or indirectly interfere with the onset of zygotic transcription, possibly by altering the cellular pH or disrupting physiological processes through the divalent cations chelating activity[42].

Response to poly(I:C) reveals patterns of both invertebrate and vertebrate antiviral innate immunity

Among poly(I:C)-upregulated genes at 24 hpi we identified both of *Nematostella* RLRs, with more significant increase for *NveRLRb* (edgeR-based $\log_2$ Fold Change (FC) = 3.275, False Discovery Rate (FDR) = 1.54e-20) than for *NveRLRa* (score below the fold change threshold, i.e. edgeR-based $\log_2$FC = 1.864, FDR = 6.55e-09). This increase suggests a possibility of a feed-forward loop similar to that observed in vertebrate antiviral immune response[43]. Moreover, many genes linked to RNAi (e.g. *NveDicer1*, *NveAGO2*, *NveRdRP1-3*) and numerous homologs of genes involved in antiviral innate immune response in both vertebrates and invertebrates animals[7,44] (e.g. Interferon regulatory factors (IRFs), RNAse L, guanylate-binding proteins (GBPs), 2'-5'-oligoadenylate synthetase 1 (OAS1), nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB), radical SAM domain-containing 2 (Viperin), to mention few, **Supplementary File S1**) were also detected.

Interestingly, we observed a significant upregulation from a previously undescribed factor (gene symbol: NVE23912) which is a cysteine-rich sequence (11 cysteine residues) with a predicted signal peptide and no significant homology to any known genes. Our search for homologs in Transcriptome Shotgun Assembly (TSA) and NCBI nr databases revealed that it is likely a secreted hexacorallian-specific protein (**Supplementary Fig. S2**) which resembles pattern of proteins under strong selective pressure displayed by the high conservation limited to the cysteine positions. Altogether, the described features make it a good candidate for further functional studies which could validate whether this novel factor is playing an important role in the innate immunity of *N. vectensis* and possibly other members of Hexacorallia.

To get wider view of the nature of poly(I:C)-induced DEG we examined promoter sequences of the induced genes by two different approaches. First, we screened the coding strand for the presence of the TATA-box in both the close proximity to the transcription starting site (TSS) (38 bp) and in a more permissive screening window (100 bp upstream and 100 bp downstream of TSS). It has been previously suggested that mammalian immune-related genes which are rapidly diverging and exhibit greater levels of expression variability across individual cells, such as cytokines and chemokines, share a common promoter architecture enriched in TATA-boxes[45]. Interestingly, *N. vectensis* displays a significant increase in abundance of TATA-box elements in poly(I:C)-upregulated genes when searching both window sizes which seems to correlate with the level of genes inducibility (**Supplementary Fig. S3a,b, Supplementary File S3**). Within protein sequences of TATA-box containing genes, we predicted a similar enrichment of signal peptides suggesting that many of these proteins might be involved in secretory pathways (**Supplementary Fig. S3c,d, Supplementary file S3**). Furthermore, the search of known transcription factor binding sites (TFBS) revealed numerous motifs known to be involved in regulating transcription of antiviral immune-related genes in vertebrates such as those recognized by STATs, IRFs, NF-κB or members of ETS family[46-49] (**Supplementary File S3**). In order to circumvent the limitation of using the vertebrate motif matrix, we scanned the *N. vectensis* genome for the presence of the homologs of vertebrate immune-related transcription factors. Importantly, we have identified numerous candidate homologs of these factors in

N. vectensis genome among which a large group showed upregulation in response to poly(I:C) treatment supporting the notion that they might play role in orchestrating the observed immune response (**Supplementary File S3**).

Role of NveRLRs in detecting long dsRNA

To confirm the results of our RNA-seq DEG analysis we assayed gene expression in independent biological replicates. RT-qPCR analysis at 24 hpi validated the upregulation of *NveRLR*s (relative expression$_{NveRLRa}$ = 1.98, 95% CI, 1.042 – 3.494, p-value = 0.0425, relative expression$_{NveRLRb}$ = 5.795, 95% CI, 3.992 – 8.411, p-value = 0.000643) (**Fig. 4a**), as well as several other putative immune-related genes (**Supplementary Fig. S4d, Supplementary File S4**) in response to poly(I:C) treatment, and an unaffected expression level of *NveRLR*s transcripts when treated with short 5'ppp-dsRNA (relative expression$_{NveRLRa}$ = 0.815, 95% CI, 0.468 –1.420, p-value = 0.325, relative expression$_{NveRLRb}$ = 1.071, 95% CI, 0.528 – 2.173, p-value = 0.778) (**Fig. 4b**). Importantly, the examination of the *NveRLR*s mRNA levels in response to the control treatments did not reveal a significant background upregulation which could distort the results of ligand specificity (**Supplementary Fig. 5a,b**). To confirm these results at the protein level, we generated custom polyclonal antibodies against *N. vectensis* RLRs which specificity has been tested beforehand. NveRLRs levels were tested at 48 hpi in order to diminish the effect of maternally deposited proteins. The result of Western blot confirmed strong upregulation of both NveRLRs after poly(I:C) stimulation (**Fig. 4c,d**) which correlates with the increased transcript abundance.

Next, we aimed to examine the ability of NveRLRs to bind poly(I:C). To this end, we generated two *N. vectensis* transgenic lines, each expressing FLAG-tagged *NveRLR* and a fluorescent mCherry gene under a ubiquitous promoter of the TATA-Box Binding Protein (TBP) gene (**Fig. 5a**). Progeny of F_1 female heterozygotes and wild-type animals was collected directly after fertilization (0 h) and the presence of maternally deposited FLAG-tagged RLRs was confirmed (**Fig. 5b**). In vitro binding assays of poly(I:C) covalently linked to biotin on wild-type protein extracts confirmed specificity of mouse FLAG antibody (**Fig. 5c**). The results of the in vitro poly(I:C) binding on the transgenic lines

revealed a significant enrichment of NveRLRb in poly(I:C)-biotin pulldown samples indicating specific binding of long dsRNA by NveRLRb (**Fig. 5c**). Unexpectedly, no poly(I:C) affinity was detected when assaying NveRLRa (**Fig. 5d**). In order to monitor how accurately the conditions of transgenic expression mimic the native proteome composition, we examined levels of NveRLRs levels in recently published mass spectrometry data spanning different developmental stages of *N. vectensis*[50]. Interestingly, we noticed that while NveRLRb displays relatively stable expression throughout the lifecycle, levels of NveRLRa in the unfertilized egg are below the detection threshold (in agreement with previous proteomic studies of *Nematostella* eggs[51,52]) and show significantly lower expression than NveRLRb across all developmental stages (**Supplementary Fig. S6**). It is therefore plausible that NveRLRa carries regulatory function or binds yet uncharacterized ligands or, alternatively, that its ligand specificity matures along with the development due to co-expression of other crucial factors.

Knockdown of *NveRLRb* interferes with the *in vivo* response to long dsRNA

Poly(I:C)-induced upregulation of *NveRLR*s both at the gene and the protein levels and FLAG:NveRLRb affinity to poly(I:C)-biotin led us to the assumption that both proteins might carry an important function in detecting viral dsRNA and hence, orchestrating downstream antiviral immune processes in *Nematostella.* To further corroborate this theory, we generated knockdown (KD) animals by microinjection of short hairpin RNA (shRNA) targeting 3 different regions of each of *NveRLR*s. The initial validation assays of KD efficiency and shRNAs immunogenicity revealed a strong (~85-90%) and moderate (~60%) effect of all *NveRLRb* and *NveRLRa* shRNAs, respectively (**Supplementary Fig. S5c,d**), and very low impact on the expression levels of putative immune-related genes of all shRNAs (**Supplementary Fig. S5e-j**). Due to the lack of strong knockdown effect by all candidate *NveRLRa* shRNAs, we decided to include all 3 tested variants for this gene and 2 shRNAs for *NveRLRb*. Following the assumption that NveRLRs might act as sensors in antiviral immune response, we co-injected each shRNAs with poly(I:C) and tested at 48 hpi the mRNA levels of candidate genes previously proved to respond to the poly(I:C) treatment. We classified genes as those affected by *NveRLR*s KD when their expression was significantly different when compared to

control shRNA and obtained by at least two different shRNAs. The first unexpected observation was that while *NveRLRb* KD efficiency remained comparable to the initial screening assays (~90%), *NveRLRa* KD level decreased to approximately 45% (**Fig. 6a,b, Supplementary File S4**). Of note, none of the *NveRLR*s KD experiments exerted a strong and ubiquitous reciprocal effect on the other sensor. Importantly, knockdown of *NveRLRb* resulted in noticeable downregulation of both tested components of RNAi i.e. *NveDicer1* (relative $expression_{shRNA1}$ = 0.370, 95% CI, 0.242 – 0.565, p-value = 0.005; relative $expression_{shRNA2}$ = 0.165, 95% CI, 0.106 – 0.257, p-value = 0.001) and *NveAGO2* (relative $expression_{shRNA1}$ = 0.663, 95% CI, 0.53 – 0.83, p-value = 0.01; relative $expression_{shRNA2}$ = 0.451, 95% CI, 0.272 – 0.547, p-value = 0.00095), as well as *NveIRF1* (relative $expression_{shRNA1}$ = 0.401, 95% CI, 0.236 – 0.683, p-value = 0.012; relative $expression_{shRNA2}$ = 0.262, 95% CI, 0.171 – 0.402, p-value = 0.00215) (**Fig. 6c,d,e, Supplementary File S4**) and an apparent but not significant decrease in expression of hexacorallian-specific factor NVE23912 (**Supplementary Fig. S4c, Supplementary File S4**). Interestingly, neither *NveOAS1* nor *NveGBP1* mRNA levels were significantly affected by the *NveRLRb* shRNA-poly(I:C) co-injection (**Supplementary Fig. S4a,b, Supplementary File S4**). In contrast to *NveRLRb* KD, response to *NveRLRa* shRNAs did not reveal any clear signature of the impaired downstream process in 5 out of 6 tested genes and displayed a general pattern of high expression variation (**Fig. 6c,d, Supplementary Fig. S4a,b,c, Supplementary File S4**). Only *NveDicer1* mRNA levels displayed a mild decrease in response to the treatment (relative $expression_{shRNA1}$ = 0.718, 95% CI, 0.470 – 1.098, p-value = 0.0966; relative $expression_{shRNA2}$ = 0.695, 95% CI, 0.581 – 0.832, p-value = 0.0049; relative $expression_{shRNA3}$ = 0.693, 95% CI, 0.493 – 0.972, p-value = 0.041) (**Fig. 6e, Supplementary File S4**). Altogether, our results indicate a strong link between the presence of NveRLRb and the ability to initiate downstream processes involving at least two key RNAi components i.e. NveDicer1 and NveAGO2 and a homolog of a known vertebrate IRF. Lack of effect of *NveRLRa* KD despite testing three shRNAs targeting different transcript regions together with the negative result of poly(I:C)-biotin binding assay suggest that NveRLRa might carry different functions.

Discussion

In this study, we examined transcriptomic response to two different viral dsRNA mimics in *N. vectensis* and aimed to elucidate the role of NveRLRs in the antiviral immune pathways. We observed a lack of any signature of the antiviral immune response to the canonical RIG-I agonist which supports our hypothesis about the evolution of cnidarian RLRs from ancestral MDA5/LGP2 precursor protein **(Fig. 1a)**. In vertebrates, RIG-I binds to 5' ends of dsRNA and recognizes the presence of di- and triphosphate on 2'-O-unmethylated nucleotide, with a strong preference to the base-paired blunt ends[9,18-22,53]. In contrast, MDA5 is known to require a stable oligomerization along the dsRNA molecule for effective downstream signaling and hence, it displays a strong affinity to long molecules with at least partial stretches of dsRNA[23-26]. Of note, poly(I:C) is known to carry 5'-diphosphate in at least a fraction of the molecules due to the synthesis process, however, uneven length of annealed strands results in single-stranded ends and long, irregular dsRNA structures[54]. In light of our results, it is likely that the activation of NveRLRs depends on the molecule length rather than the 5' end recognition. To the best of our knowledge, the distinctive features of an effective RIG-I agonist have so far been only functionally characterized in vertebrates despite RLRs homologs being found in many invertebrate genomes. Therefore, further research on such non-vertebrate homologs can provide key insights into the evolution of dsRNA 5' end recognition.

Transcriptomic response to poly(I:C) revealed that many canonical vertebrate antiviral factors triggered by IFN, known as interferon-stimulated genes (ISGs)[43], are also taking part in *Nematostella* immune response. Interestingly, we observed several intriguing features of promoter region architecture such as enrichment in the TATA box sequence in poly(I:C)-upregulated genes. These elements were previously shown to display analogies in orchestrating expression of rapidly diverging and transcriptionally variable genes in phylogenetically distant groups, such as mammals[45] and yeast[55,56]. On the other hand, response to poly(I:C) and KD experiments revealed similarities to antiviral invertebrate systems and suggested a link between NveRLRb and the RNAi pathway. Of note, a similar level of complexity and involvement of diverse antiviral mechanisms was previously suggested for the Pacific oyster *Crassostrea gigas*[57-60], although the response to the canonical RLRs

ligands presented here has not yet been comprehensively characterized in this molluscan species. The interdependence of RLRs and RNAi has been functionally demonstrated in the model nematode *Caenorhabditis elegans*, where RLRs were shown to interact with Dicer and provide crucial assistance for RNAi machinery to produce primary and secondary antiviral siRNAs[61-63]. However, unlike most other bilaterian and cnidarian RLRs, the nematode receptors lack any CARD domains (**Fig. 1b**) that typify action via oligomerization rather than association with Dicer. Importantly, there is growing evidence that virus-host interactions involve other classes of small RNAs including Dicer- and AGO-dependent microRNAs (miRNAs)[64] and several studies in chordates suggested differential expression of host miRNAs in response to poly(I:C)[65-68]. We have recently demonstrated that the cargo of NveAGO1 is restricted to miRNAs, whereas NveAGO2 can carry both miRNAs and siRNAs[69]. This hints that poly(I:C)-upregulated NveAGO2 could function as the antiviral RNAi effector protein. Further studies will help to decipher the role of RNAi components in the antiviral immune response of *N. vectensis*.

The results of *NveRLRb* KD indicate that there are likely alternative immune cascades triggered by poly(I:C) administration which might be initiated by other dsRNA sensors. Among these, Toll-like receptors (TLRs) are obvious candidates due to their well-known role as PRRs[70]. However, the only TLR of *N. vectensis* has been recently shown to mediate immune response in NFκB-dependent way in response to *Vibrio coralliilyticus* and flagellin[71] which indicates its involvement in recognizing bacterial rather than viral pathogens. An intriguing question for future studies is whether NveRLRa is acting as a nucleic acid sensor. On one hand the stable co-existence of two separately-clustering RLRs paralogs in sea anemones (**Fig. 1a**) and the clear increase in *NveRLRa* expression upon poly(I:C) challenge (**Fig. 4a,c**) suggest that it is likely a functional component of antiviral immune response which might display affinity to yet uncharacterized ligands. Nonetheless, the short truncation of helicase domain and aberrant KD patterns suggest an alternative but not mutually exclusive hypothesis that NveRLRa might carry some regulatory functions involved in complex feedback mechanisms.

289 To the best of our knowledge, our study provides the first functional insights into the role of
290 RLRs in a non-bilaterian animal. The initial results suggest that RLRs capacity to sense 5' end of
291 dsRNA evolved in Bilateria, although further studies involving invertebrate RLRs will provide key
292 answers on this matter. We show that *N. vectensis* response to a viral dsRNA mimic is characterized
293 by high complexity and includes both vertebrate-like features, as well as invertebrate-like
294 involvement of RNAi machinery in an RLR-dependent manner. This hints that key elements of both
295 extant antiviral systems were already present in a cnidarian-bilaterian common ancestor. Our results
296 lay the foundation for further functional studies on downstream effector mechanisms in *N. vectensis*
297 which might provide key insights into the evolution of the antiviral immune response in Metazoa.

298

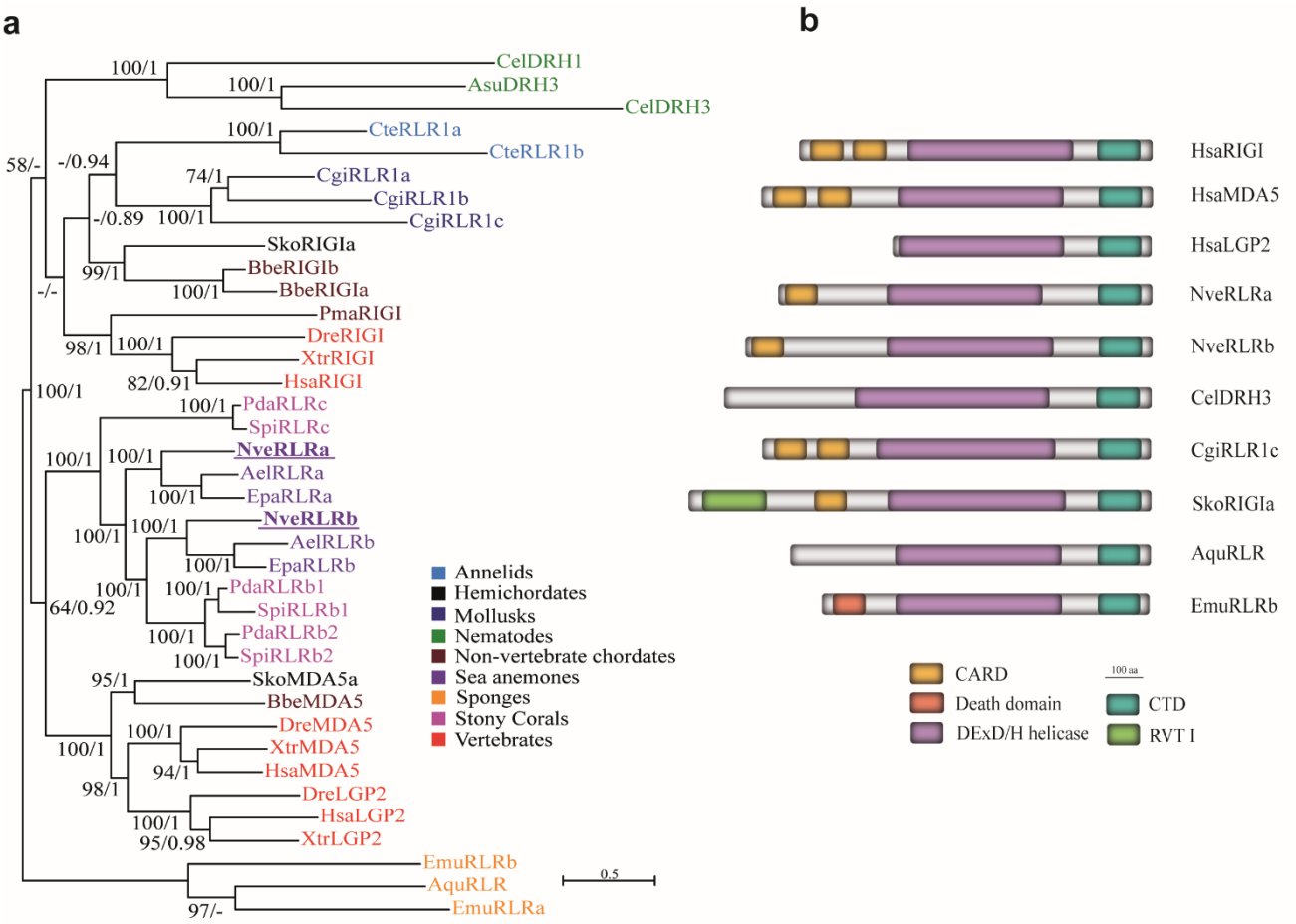


Figure 1. Phylogenetic relationship of metazoan RLRs. (a) Maximum likelihood and Bayesian inference consensus phylogenetic tree of representative RLR sequences, bootstrap values above 50% are presented for each node. Posterior probability values of a Bayesian tree of the same topology between 0.85-1.0 are indicated for each node. Ael, *Anthopleura elegantissima*, Aqu, *Amphimedon queenslandica*, Asu, *Ascaris suum*, Bbe, *Branchiostoma belcheri*, Cte, *Capitella teleta*, Cgi, *Crassostrea gigas*, Cel, *Caenorhabditis elegans*, Dre, *Danio rerio*, Emu, *Ephydatia muelleri*, Epa, *Exaiptasia pallida*, Has, *Homo sapiens*, Nve, *Nematostella vectensis*, Pda, *Pocillopora damicronis*, Pma, *Petromyzon marinus*, Sti, *Stylophora pistillata*, Xtr, *Xenopus tropicalis*. **(b)** Schematic representation of selected RLR representatives of major phylogenetic groups. CARD - caspase recruitment domain; CTD - C-terminal domain; RVT I – reverse transcriptase.

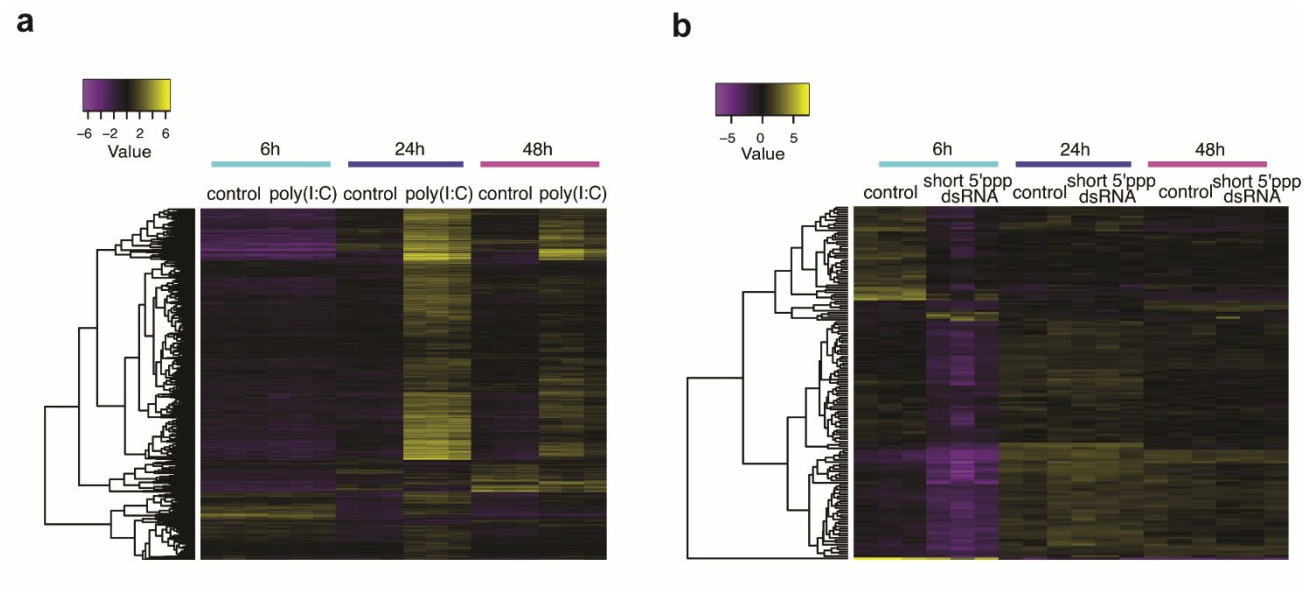
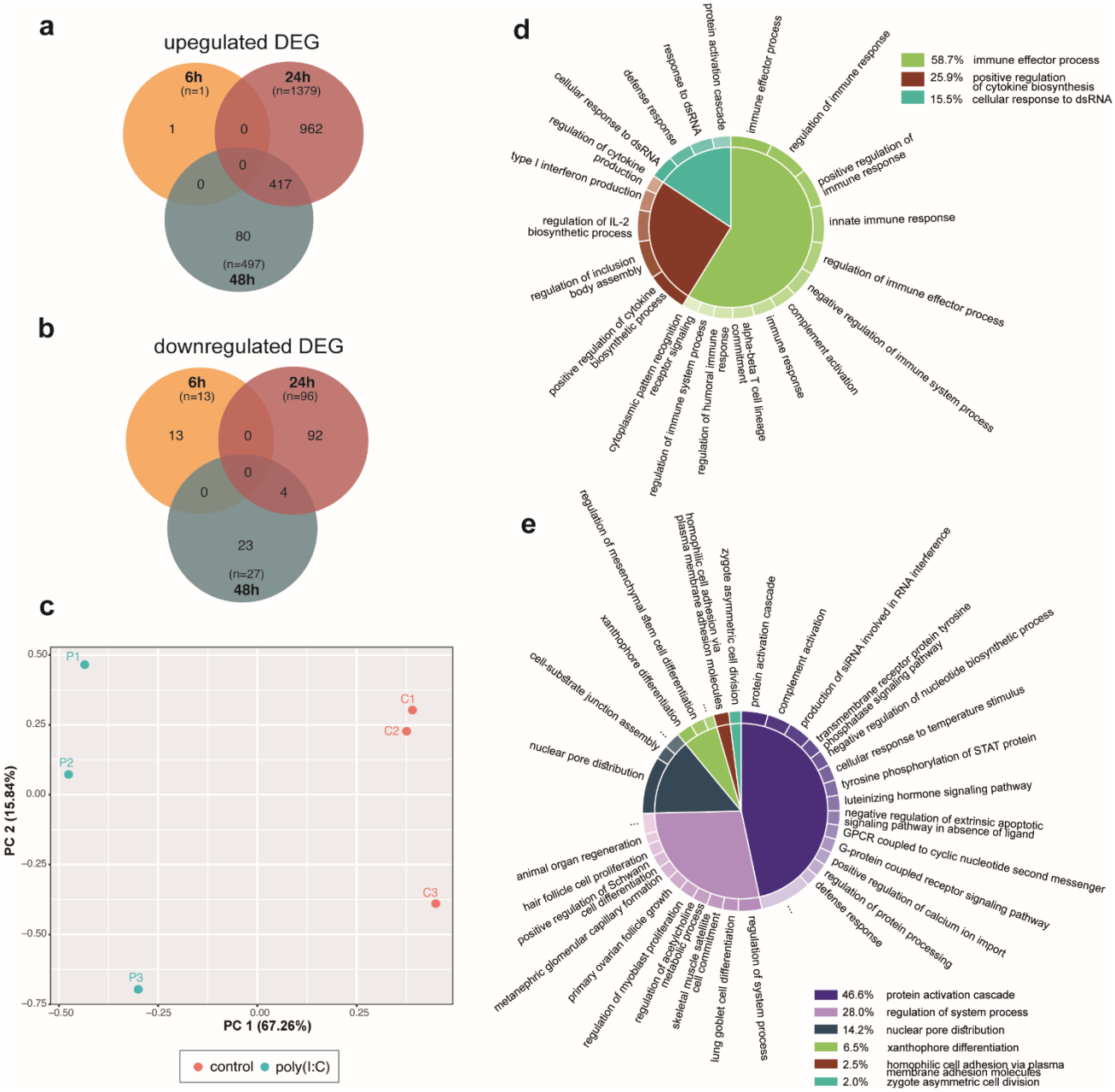


Figure 2. Differential gene expression after microinjections of viral mimics. Heatmap of differentially expressed genes upon administration of **(a)** poly(I:C) vs 0.9% NaCl serving as a control, and **(b)** 19-mer dsRNA with 5' triphosphate and 19-mer dsRNA with 5' hydroxyl group serving as a control.



316 **Figure 3. Signature of the innate immune response to poly(I:C).** Venn diagram of differentially
317 expressed genes which were **(a)** upregulated and **(b)** downregulated after poly(I:C) administration. **(c)**
318 PCA plot representing whole transcriptome of poly(I:C)-injected animals at 24 hpi. **(d)** GO terms
319 enrichment results of DEG upregulated at 24 hpi and **(e)** 48 hpi.

320

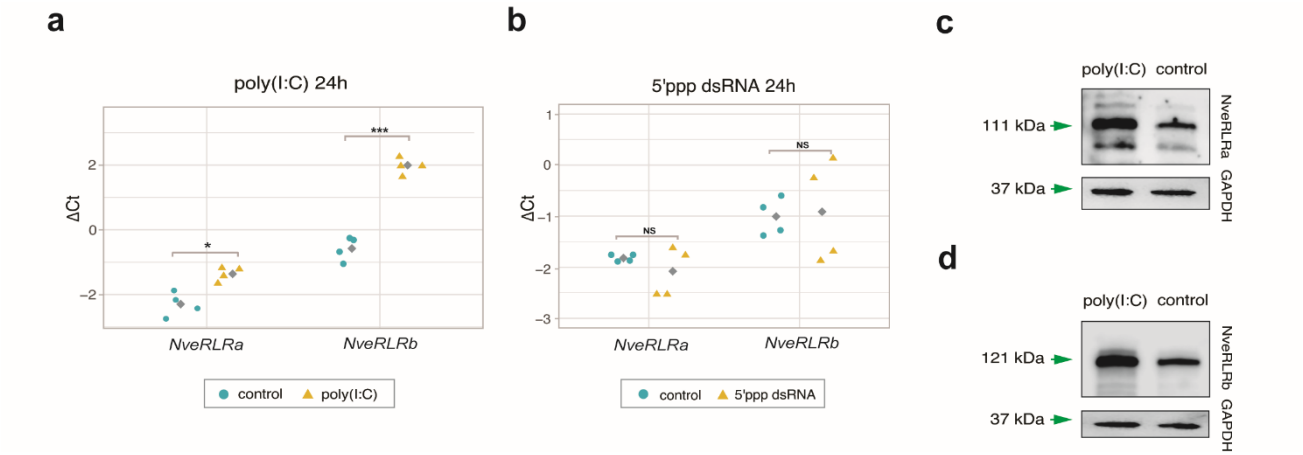


Figure 4. Response of *Nematostella* putative dsRNA helicases to viral mimics. *NveRLR*s mRNA expression level measured by RT-qPCR in response to **(a)** poly(I:C) and **(b)** short 5' ppp dsRNA. Grey squares represent mean values. Western blot validation of **(c)** NveRLRa and **(d)** NveRLRb protein level in response to poly(I:C) at 48 hpi. Significance level was assessed by paired two-tailed Student's t-test; * p value < 0.05, ** p value <0.01, *** p value <0.001, NS – not significant.

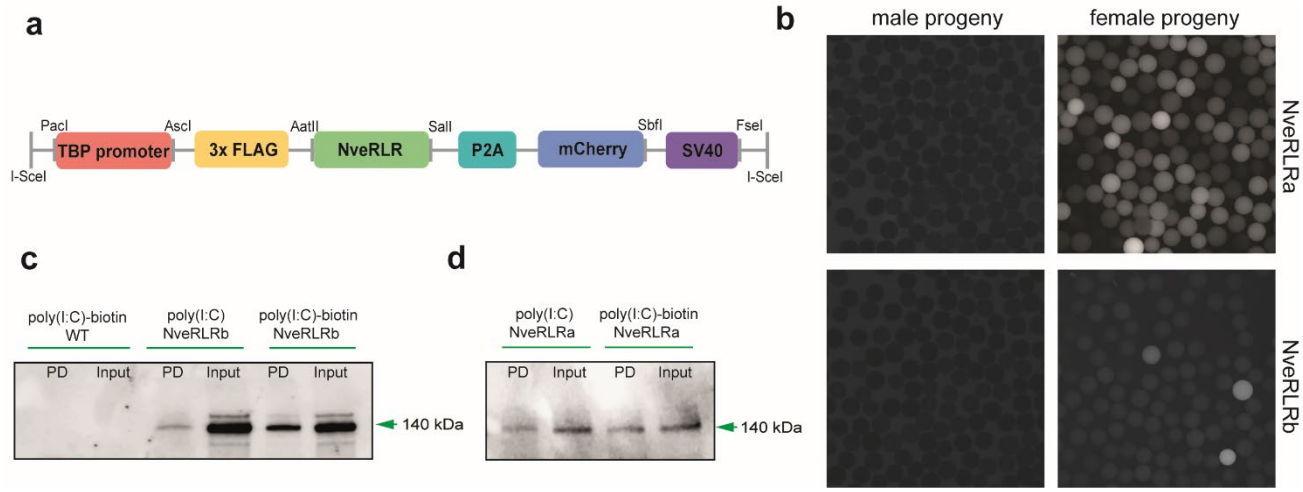


Figure 5. NveRLRs affinity to poly(I:C). (a) Schematic representation of the FLAG-*NveRLR* construct (7,071 bp and 7,643 bp for *NveRLRa* and *NveRLRb*, respectively) used for transgenesis. TBP promoter, self-cleaving P2A sequence, mCherry gene and polyadenylation signal SV40 are also shown. **(b)** Maternal deposition of the FLAG-NveRLR observed after crossing transgenic females (right panels) with WT males; fluorescent protein is missing in transgenic male progeny (left panels). **(c)** Results of poly(I:C)-biotin *in vitro* binding assay showing affinity of FLAG-NveRLRb but not **(d)** FLAG-NveRLRa to poly(I:C); PD – pulldown.

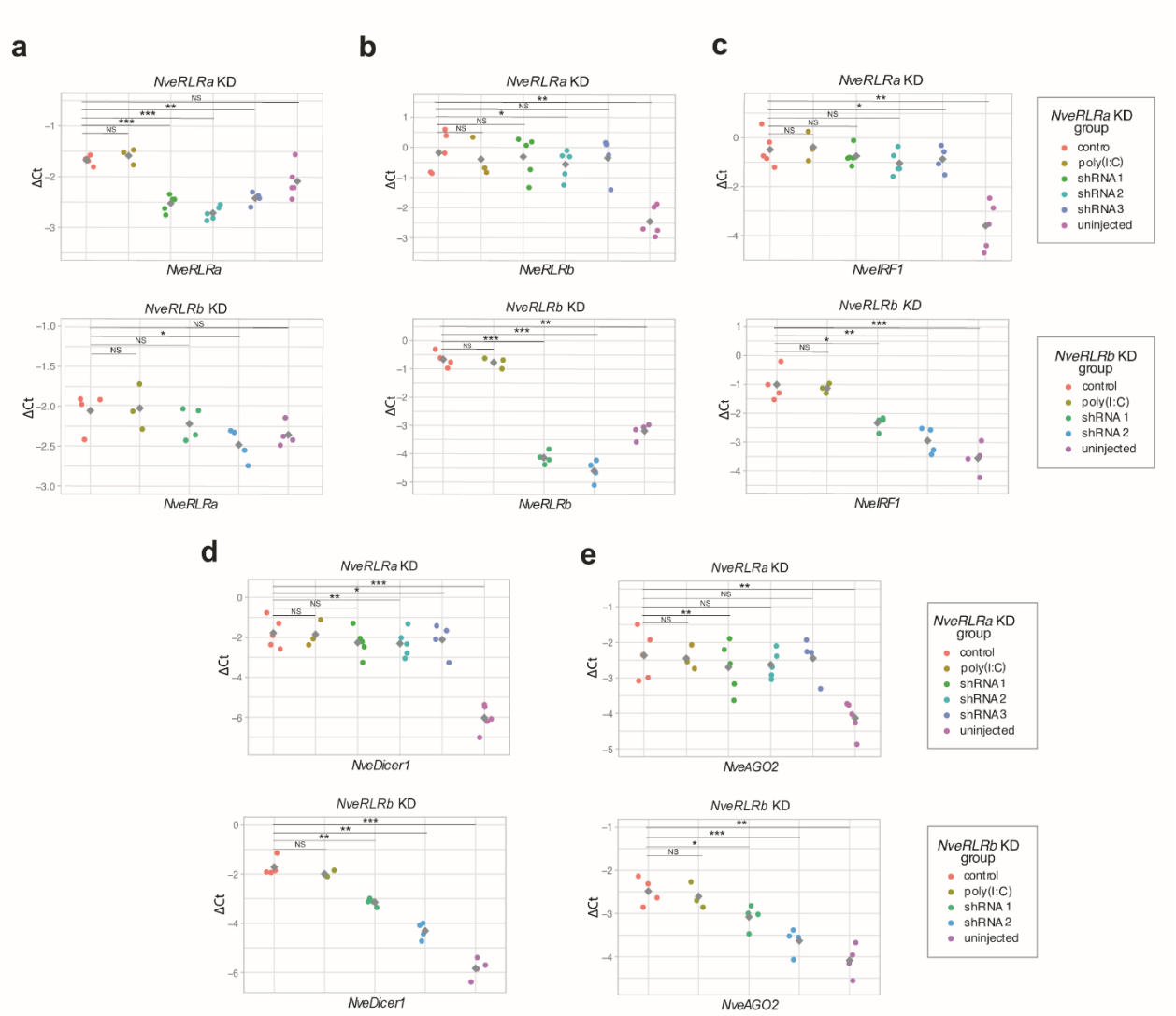


Figure 6. Downregulation of putative antiviral innate immunity-related genes in response to *NveRLR*s knockdown (KD) combined with poly(I:C) administration. RT-qPCR results of shRNA targeting *NveRLRa* and *NveRLRb* (upper and lower panels of each section, respectively) measuring the expression of **(a)** *NveRLRa*, **(b)** *NveRLRb*, **(c)** *NveIRF1*, **(d)** *NveDicer1*, **(e)** *NveAGO2*. Grey squares represent mean values. All comparisons were done by paired two-tailed Student's t-test against the control shRNA. Significance level: * p value < 0.05, ** p value <0.01, *** p value <0.001, NS – not significant.

346 ## Materials and Methods

347 ### Sea anemone culture

348 *Nematostella* embryos, larvae and juveniles were grown in the dark at 22 °C in 16‰ artificial
349 seawater, while polyps were grown at 18 °C and fed with *Artemia salina* nauplii three times a week.
350 The induction of gamete spawning was performed as previously described[72]. The gelatinous egg sack
351 was removed using 3% L-Cysteine (Merck Millipore, Burlington, MA, USA) and followed by
352 microinjection of viral mimics or shRNAs.

353 ### Injection of viral mimics

354 To stimulate the antiviral immune response in *Nematostella*, we used two types of synthetic
355 dsRNA. To mimic the presence of long dsRNA, we used 3.125 ng/µl of high molecular weight
356 (HMW) poly(I:C) in 0.9% NaCl (Invivogen, San Diego, CA, USA) with an average size of 1.5 – 8 kb,
357 and 0.9% NaCl as a control. The second type of ligand was a synthetic dsRNA 19-mer with 5'
358 triphosphate (5'ppp-dsRNA) and a control dsRNA 19-mer with 5' hydroxyl group (5'ppp-dsRNA
359 control), both suspended in sterile RNase-free endotoxin-free water to a final concentration of 90
360 ng/µl (Invivogen). Each experiment was performed in triplicates and each biological replicate was
361 composed of 100-150 injected zygotes per time point. Within each biological replicate zygotes were
362 collected at 6, 24 and 48 hpi, flash frozen in liquid nitrogen and stored at -80 °C until processed.

363 ### Transcriptome library preparation and sequencing

364 Total RNA was extracted with Tri-Reagent (Sigma-Aldrich, St. Louis, MO, USA) according to
365 manufacturer's protocol, treated with 2 µl of Turbo DNAse (Thermo Fisher Scientific, Waltham, MA,
366 USA) and re-extracted with Tri-Reagent and 20 ug of RNA-grade glycogen (Thermo Fisher
367 Scientific). The quality of total RNA was assessed on Bioanalyzer Nanochip (Agilent, Santa Clara,
368 CA, USA) and only samples with RNA Integrity Number (RIN) > 7 were retained. Libraries were
369 constructed from 226 ng and 300 ng of total RNA from poly(I:C) and 5'ppp-dsRNA injected samples,
370 respectively. RNA-seq libraries were generated using SENSE Total RNA-seq Library Prep Kit v2

(Lexogen, Vienna, Austria) following the manufacturer's protocol and sequenced on NextSeq 500 (Illumina, San Diego, CA, USA) using single-end 75 bp chemistry.

Raw reads processing and differential gene expression analysis

Quality of raw reads was assessed and visualized with FastQC software[73]. Reads were trimmed and quality filtered by Trimmomatic with the following parameters (HEADCROP:9 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:20 MINLEN:36)[74] and the quality of the filtered reads was re-assessed in FastQC. Reads were mapped to *N. vectensis* genome (NCBI accession: GCA_000209225.1)[75] with STAR alignment program[76]. Gene counts were obtained with RSEM[77] (genes models, protein models and annotations are available at: https://figshare.com/articles/Nematostella_vectensis_transcriptome_and_gene_models_v2_0/807696). Differential gene expression analysis was carried out with edgeR[78] and DESeq2[79] implemented in the Trinity pipeline[80]. Treatment samples within each time point were compared to the corresponding control samples. Differentially expressed genes were defined by FDR < 0.05 and log2|fold change| ≥ 2. Only genes identified by both edgeR and DESeq2 methods were reported as differentially expressed. GO groups were identified by GSEA using GOseq Bioconductor package[81] implemented in the in-built Trinity pipeline[80]. An FDR cut-off of 0.05 was considered significant for the enriched or depleted GO terms. To reduce redundancy, GO terms were group based on semantic similarity using REVIGO[82] and visualized by CirGO v2.0[83].

shRNA generation and knockdown experiments

Three shRNA precursors from three different regions of each *NveRLR* gene as well as control shRNAs were designed and prepared as previously described[84] with minor modifications. In brief, 19 bp gene targeting motif size was chosen for each shRNA (minimum GC% content > 35%). We have introduced 2-3 mismatches to the star strand, which corresponds to the coding strand, to create the bulges in shRNA precursors following the structure of native miRNA in *Nematostella*[69,85]. Reverse complement sequence of shRNA precursors was synthesized as ultramer oligo by Integrated DNA Technologies (Coralville, IA, USA), mixed with T7 promoter primer in 1:1 ratio in a final

concentration of 25µM, denatured at 98 °C for 5 min and cooled to 24 °C. shRNAs were synthesized with AmpliScribe™ T7-Flash™ Transcription Kit (Epicentre, Charlotte, NC, USA) for 15 h followed by 15 min treatment with 1 µl of DNAse I. The in-vitro transcribed products were purified using the Quick-RNA Miniprep Kit (Zymo Research, Irvine, CA, USA). The quality and size of each precursor were checked on 1.5% agarose gel and its concentration was measured by spectrophotometer. The sequences of shRNAs precursors are provided in **Supplementary File S5**.

Initial screening of shRNA knockdown efficiency and toxicity revealed that microinjections of shRNAs of *NveRLRa* and *NveRLRb* proved effective and non-toxic at 48 hpi in 750-1000 ng/µl and 350-500 ng/µl concentration range, respectively. 3 shRNAs for *NveRLRa* (750 ng/µl, 750 ng/µl and 1000 ng/µl) and 2 shRNAs for *NveRLRb* (500 ng/µl each) were microinjected to *Nematostella* zygotes in a 10 µl mixture containing additionally 3.125 ng/µl of HMW poly(I:C), 1 µl of 9% NaCl and RNAse-free endotoxin-free water. Identically prepared 1000 ng/µl and 500 ng/µl of the control shRNA was included in each microinjection of *NveRLRa* and *NveRLRb* shRNAs, respectively. Moreover, in each microinjection experiment we included a subset of animals treated only with poly(I:C) 3.125 ng/µl to monitor the cytotoxic effect of shRNA control. Zygotes were collected at 48 hpi, flash frozen in liquid nitrogen and stored at -80 °C until further processed.

Reverse-transcription quantitative PCR

To validate the results of the RNA-seq and knockdown experiments, we assayed the expression levels of several candidate immune-related genes from the mammalian RLR pathway (*NveRLRa*, *NveRLRb*, *NveOAS1*, *NveIRF1*, *NveGBP1*), RNAi pathway (*NveDicer1* and *NveAGO2*) and a representative of hexacorallian-specific gene (NVE23912) by reverse-transcription quantitative PCR (RT-qPCR) at 24 hpi (RNA-seq) or 48 hpi (knockdown experiments). 3-5 biological replicates were used to validate the results of transcriptomics and poly(I:C)-shRNAs experiments, while one biological replicate was used to assess the efficiency and background immune response to shRNAs and poly(I:C) control. RNA was extracted from injected embryos following the same protocol used for RNA-seq libraries construction. cDNA was constructed using SuperScript III (Thermo Fisher

Scientific) for RNA-seq validation and iScript cDNA Synthesis Kit (Bio-Rad, Hercules, CA, USA) for knockdown experiments, according to the manufacturer's protocol. Real-Time PCR was prepared with Fast SYBR™ Green Master Mix (Thermo Fisher Scientific) on the StepOnePlus Real-Time PCR System (ABI, Thermo Fisher Scientific). The qPCR mixture contained cDNA template (1 µl), 2× Fast SYBR™ Green Master Mix (5 µl), primers (1 µl) and nuclease-free water to make up 10 µl total volume. qPCR thermocycling conditions were 95 °C for 20 s, 40 cycles of 95 °C for 3 s, 60 °C for 30 s. Melt curve analysis was initiated with 95 °C for 15 s and performed from 60 to 95 °C in 0.5 °C increments. The expression levels of candidate genes were normalized to the NVE5273 gene ($\Delta Ct = Ct_{reference\ gene} - Ct_{gene\ of\ interest}$) and the relative expression was calculated by the $2^{\Delta\Delta Ct}$ method. The significance level was calculated by applying paired two-tailed Student's t-test to ΔCt values for each of the pairwise comparisons. Sequences of all primers are shown in **Supplementary File S5**.

Antibody generation

For NveRLRa and NveRLRb Western blots following poly(I:C) stimulation, we used custom polyclonal antibodies raised against recombinant fragment antigens generated by rabbits' immunization (GenScript, Piscataway Township, NJ, USA). Each recombinant fragment was injected into three rabbits. After the third round of immunization, pre-immune and post-immune sera were sent to us for screening by Western blot against *Nematostella* lysate to identify sera specifically positive for NveRLRa and NveRLRb (bands of ~111 and ~121 kDa respectively). Finally, the antigens were used by the company for affinity purification from the relevant rabbits. Amino acid sequences of NveRLRa and NveRLRb fragments used for immunization are presented in **Supplementary File S5**.

Western blot

Equal amounts of protein were run on 4 – 15% Mini-PROTEAN® TGX™ Precast Protein Gel (Bio-Rad) followed by blotting to a Polyvinylidene fluoride (PVDF) membrane (Bio-Rad). Next, the membrane was washed with TBST buffer (20 mM Tris pH 7.6, 150 mM NaCl, 0.1% Tween 20) and blocked (5% skim milk in TBST) for 1 hour on the shaker at room temperature. Polyclonal antibody against NveRLRa or NveRLRb or monoclonal mouse anti-FLAG M2 antibody (Sigma-Aldrich) or

449 monoclonal mouse anti-GAPDH (Abcam, Cambridge, UK) serving as loading control was diluted
450 1:1000 in TBST containing 5% BSA (MP Biomedicals, Irvine, CA, USA) and incubated with the
451 membrane in a sealed sterile plastic bag at 4°C overnight. The membrane was washed three times
452 with TBST for 10 min and incubated for 1 hour with 1:10,000 diluted peroxidase-conjugated anti-
453 mouse or anti-rabbit antibody (Jackson ImmunoResearch, West Grove, PA, USA) in 5% skim milk in
454 TBST. Finally, the membrane was washed three times with TBST and detection was performed with
455 the Clarity™ Max ECL kit for pulldown experiments (Bio-Rad) and Clarity™ ECL kit for all other
456 experiments (Bio-Rad) according to the manufacturer's instructions and visualized with a CCD
457 camera of the Odyssey Fc imaging system (Li-COR Biosciences, USA). Size determination was
458 carried out by simultaneously running Precision Plus Protein™ Dual Color Protein Ladder (Bio-Rad)
459 and scanning at 700 nm wavelength.

460 **Cloning and transgenesis**

461 Synthetic genes (Gene Universal, Newark, DE, USA) including CDS of *NveRLRa* and *NveRLRb*
462 (scaffold_15:1090025-1101489 and scaffold_40:683898-697394, respectively), self-cleaving porcine
463 teschovirus-1 2A sequence (P2A)[86] and mCherry sequence[87] were amplified with Q5® Hot Start
464 High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, USA), visualized on 1%
465 agarose gel and purified by NucleoSpin Gel and PCR Clean-up (Macherey-Nagel, Düren, Germany).
466 Following digestion with restriction enzymes, PCR fragments were ligated to a pER242[88] vector
467 containing a TBP promoter previously proved to drive ubiquitous expression in *Nematostella*[89], three
468 N-terminal FLAG tags and SV40 polyadenylation signal. Plasmids were transformed into the *E. coli*
469 DH5α (New England Biolabs) strain and outsourced for Sanger sequencing (HyLabs, Rehovot,
470 Israel). Each *NveRLR* plasmid was subsequently injected into *N. vectensis* zygotes along with the
471 yeast meganuclease I-*Sce*I (New England Biolabs) to enable genomic integration[88,90]. Transgenic
472 animals were visualized under an SMZ18 stereomicroscope equipped with a DS-Qi2 camera (Nikon,
473 Tokyo, Japan) and positive animals were reared to the adult stage. At approximately 4 months old F_0
474 individuals were induced for gametes and crossed with wild-type animals to generate F_1 FLAG-
475 tagged TBP::NveRLR::mCherry heterozygotes. Positive F_1 individuals were selected and grown to

the adult stage. For the in vitro binding assay, only F_1 females descending from a single F_0 founder of each NveRLR line were used. Sequences of all used primers are provided in **Supplementary File S5.**

In vitro binding assay

Maternal deposition of FLAG-tagged TBP::NveRLR::mCherry transgene in F_2 animals was visualized under an SMZ18 stereomicroscope equipped with a DS-Qi2 camera (Nikon) and confirmed by Western blotting. Following fertilization with wild-type gametes, F_2 FLAG-tagged TBP::NveRLR::mCherry and wild-type zygotes were treated with 3% L-Cysteine (Merck Millipore), washed and snap frozen in liquid nitrogen. Next, animals were mechanically homogenized in the following lysis buffer: 50 mM Tris-HCl (pH 7.4), 150 mM KCl, 0.5% NP-40, 10% glycerol, Protease inhibitor cOmplete ULTRA tablets (Roche, Basel, Switzerland) and Protease Inhibitor Cocktail Set III, EDTA-Free (Merck Millipore). Protease inhibitors were added fresh just before use. After 1 h rotation in 4ºC the samples were centrifuged at 16000 × *g*, 15 min, 4 ºC and supernatant was collected. Protein concentration was measured using Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). Next, the lysate was pre-cleared as following: 100 µl of streptavidin magnetic beads (New England Biolabs) were washed in 1 ml of 1×PBS for 3 times and the FLAG-tagged TBP::NveRLR/wild-type lysate was added to the washed beads. Lysis buffer was added to make up 1.2 ml and samples were incubated at 4 ºC rotation for 1 hour. After the incubation, the pre-cleared lysates were collected and mixed with the HMW poly(I:C) (Invivogen) or HMW poly(I:C)-biotin (Invivogen) in the final concentration of 30 ng/ml and ATP (New England Biolabs) in the final concentration of 0.5 mM. Samples were incubated for 1 h in rotation at room temperature. Simultaneously, 100 µl of fresh streptavidin magnetic beads were blocked with wild-type lysates alike in the pre-clearing step. poly(I:C) samples were added to the blocked beads and incubated for 2 h in rotation at 4 ºC for poly(I:C)-biotin pulldown. After the incubation, the lysates were discarded and the beads were washed 3 times with 500 µl of the following wash buffer: 50 mM Tris-HCl (pH 7.4), Protease inhibitor cOmplete ULTRA tablets (Roche) and Protease Inhibitor Cocktail Set III, EDTA-Free (Merck Millipore). Subsequently, 40 µl of filtered double-distilled water and 20 µl of Blue Protein Loading Dye (New England Biolabs) were added to

the beads. The samples were heated at 100 °C for 8 min and placed on ice for 1 min, then centrifuged 1 min at 21,000 × *g* at 4 °C, and the supernatant was collected for Western blot.

Phylogenetic analysis

To construct an informative phylogenetic tree we selected representatives of major groups carrying RLRs: vertebrates (a fish, an amphibian, and a mammal), two non-vertebrate chordates (a lancelet and a lamprey), nematodes (*C. elegans* and *A. suum*), two lophotrochozoans (an oyster and an annelid) and hexacorallians (three sea anemones, each representing a different major sea anemone clade and two-reef building corals). Sponges RLRs sequences were chosen as an outgroup. The RLRs amino acid sequences were aligned using MUSCLE[91] and low certainty alignment regions were removed by TrimAl[92] using the –automatic1 for heuristic model selection. The maximum-likelihood phylogenetic trees were constructed using IQ-Tree[93] with the LG+F+R5 model which was the best fitting model both according to the Bayesian information criterion (BIC) and corrected Akaike information criterion. Support values of the ML tree were calculated using 1,000 ultrafast bootstrap replicates[94]. A Bayesian tree was constructed using MrBayes[95] with the WAG +I +G model and the run lasted 5,000,000 generations with every 100th generation being sampled. The Bayesian analysis was estimated to reach convergence when the potential scale reduction factor (PSRF) reached 1.0. Consensus domain composition was predicted by simultaneous search in Pfam[96] and NCBI Conserved Domains[97] databases run with default parameters.

Homologs of NVE23912 sequences were identified through a search in TSA and NCBI nr databases and *Nematostella* gene models. Amino acid sequences were aligned using MUSCLE[91] and visualized by CLC Genomics Workbench. Details of RLRs and NVE23912 homolog sequences used in the analysis are available in **Supplementary File S6**.

Promoter sequence analysis of DEG

Analysis of promoter sequences was performed as previously described[45] with minor modifications. In brief, coordinates of the TSS were retrieved from

nveGenes.vienna130208.nemVec1.bed file. We subset the upregulated DEG identified by poly(I:C) microinjection (n=1379) and the fraction of top 10% genes (n=138) and top 20% genes (n=276), setting the whole transcriptome as the background (n=18831). TATA box-containing genes were identified using FIMO[98] by having at least one statistically significant match (p-value cut-off of <0.05) to the TATA box consensus motif (MA0108.1) retrieved from JASPAR server[99]. Due to uncertainty in TSS calling, we have scanned the coding strand in two ways: a) narrow search included 38 bp upstream of TSS; b) wide search spanned both 100 bp upstream and 100 bp downstream of putative TSS whenever fitted in the scaffold boundaries. To estimate motifs enrichment in the same groups, we used the non-redundant JASPAR core motif matrix (pfm_vertebrates.txt) and run AME[100] in one-tailed Fisher's exact test mode. The searching region included 500 bp upstream of the putative TSS, the first exon and the first intron of the gene. For motif identification, the cut-off of adjusted by Bonferroni correction p-value < 0.05 was considered significant statistically significant. The presence of the signal peptide in each protein sequence was predicted by SignalP 4.1 Server with default settings[101].

Data availability

All sequencing data that support the findings of this study have been deposited in the National Center for Biotechnology Information Sequence Read Archive (SRA) and are accessible through the BioProject accession number PRJNA673983. Source data for Fig. 2,3c and Supplementary Fig. 1a,b,c,d,e have been provided in Supplementary File S1. Source data for Fig. 4a,b, 6 and Supplementary Fig. 6 have been provided in Supplementary File S4. All other relevant data are available from the corresponding authors on request.

References

1 Koonin, E. V. & Dolja, V. V. A virocentric perspective on the evolution of life. *Curr Opin Virol* **3**, 546-557, doi:10.1016/j.coviro.2013.06.008 (2013).

2 tenOever, B. R. The Evolution of Antiviral Defense Systems. *Cell Host Microbe* **19**, 142-149, doi:10.1016/j.chom.2016.01.006 (2016).

553 3 Bernheim, A. *et al.* Prokaryotic viperins produce diverse antiviral molecules. *Nature*,
554 doi:10.1038/s41586-020-2762-2 (2020).
555 4 Cohen, D. *et al.* Cyclic GMP-AMP signalling protects bacteria against viral infection. *Nature* **574**,
556 691-695, doi:10.1038/s41586-019-1605-5 (2019).
557 5 Morehouse, B. R. *et al.* STING cyclic dinucleotide sensing originated in bacteria. *Nature*,
558 doi:10.1038/s41586-020-2719-5 (2020).
559 6 Barrat, F. J., Elkon, K. B. & Fitzgerald, K. A. Importance of Nucleic Acid Recognition in
560 Inflammation and Autoimmunity. *Annu Rev Med* **67**, 323-336, doi:10.1146/annurev-med-052814-
561 023338 (2016).
562 7 Hartmann, G. Nucleic Acid Immunity. *Adv Immunol* **133**, 121-169, doi:10.1016/bs.ai.2016.11.001
563 (2017).
564 8 Weber, F., Wagner, V., Rasmussen, S. B., Hartmann, R. & Paludan, S. R. Double-stranded RNA is
565 produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by
566 negative-strand RNA viruses. *J Virol* **80**, 5059-5064, doi:10.1128/JVI.80.10.5059-5064.2006
567 (2006).
568 9 Schlee, M. *et al.* Recognition of 5' triphosphate by RIG-I helicase requires short blunt double-
569 stranded RNA as contained in panhandle of negative-strand virus. *Immunity* **31**, 25-34,
570 doi:10.1016/j.immuni.2009.05.008 (2009).
571 10 Liu, G., Park, H. S., Pyo, H. M., Liu, Q. & Zhou, Y. Influenza A Virus Panhandle Structure Is
572 Directly Involved in RIG-I Activation and Interferon Induction. *J Virol* **89**, 6067-6079,
573 doi:10.1128/JVI.00232-15 (2015).
574 11 Wang, X. H. *et al.* RNA interference directs innate immunity against viruses in adult Drosophila.
575 *Science* **312**, 452-454, doi:10.1126/science.1125694 (2006).
576 12 Szittya, G. & Burgyan, J. RNA interference-mediated intrinsic antiviral immunity in plants. *Curr*
577 *Top Microbiol Immunol* **371**, 153-181, doi:10.1007/978-3-642-37765-5_6 (2013).
578 13 Felix, M. A. *et al.* Natural and experimental infection of Caenorhabditis nematodes by novel
579 viruses related to nodaviruses. *PLoS Biol* **9**, e1000586, doi:10.1371/journal.pbio.1000586 (2011).

14 Lewis, S. H. *et al.* Pan-arthropod analysis reveals somatic piRNAs as an ancestral defence against transposable elements. *Nat Ecol Evol* **2**, 174-181, doi:10.1038/s41559-017-0403-4 (2018).

15 Yoneyama, M. *et al.* The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol* **5**, 730-737, doi:10.1038/ni1087 (2004).

16 Yoneyama, M. *et al.* Shared and Unique Functions of the DExD/H-Box Helicases RIG-I, MDA5, and LGP2 in Antiviral Innate Immunity. *The Journal of Immunology* **175**, 2851, doi:10.4049/jimmunol.175.5.2851 (2005).

17 Kang, D.-c. *et al.* mda-5: An interferon-inducible putative RNA helicase with double-stranded RNA-dependent ATPase activity and melanoma growth-suppressive properties. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 637-642, doi:10.1073/pnas.022637199 (2002).

18 Ren, X., Linehan, M. M., Iwasaki, A. & Pyle, A. M. RIG-I Recognition of RNA Targets: The Influence of Terminal Base Pair Sequence and Overhangs on Affinity and Signaling. *Cell Rep* **29**, 3807-3815 e3803, doi:10.1016/j.celrep.2019.11.052 (2019).

19 Schmidt, A. *et al.* 5'-triphosphate RNA requires base-paired structures to activate antiviral signaling via RIG-I. *Proc Natl Acad Sci U S A* **106**, 12067-12072, doi:10.1073/pnas.0900971106 (2009).

20 Hornung, V. *et al.* 5'-Triphosphate RNA is the ligand for RIG-I. *Science* **314**, 994-997, doi:10.1126/science.1132505 (2006).

21 Goubau, D. *et al.* Antiviral immunity via RIG-I-mediated recognition of RNA bearing 5'-diphosphates. *Nature* **514**, 372-375, doi:10.1038/nature13590 (2014).

22 Pichlmair, A. *et al.* RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates. *Science* **314**, 997-1001, doi:10.1126/science.1132998 (2006).

23 Pichlmair, A. *et al.* Activation of MDA5 requires higher-order RNA structures generated during virus infection. *J Virol* **83**, 10761-10769, doi:10.1128/JVI.00770-09 (2009).

24 Peisley, A. *et al.* Kinetic mechanism for viral dsRNA length discrimination by MDA5 filaments. *Proc Natl Acad Sci U S A* **109**, E3340-3349, doi:10.1073/pnas.1208618109 (2012).

607 25 Kato, H. *et al.* Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-
608 inducible gene-I and melanoma differentiation-associated gene 5. *J Exp Med* **205**, 1601-1610,
609 doi:10.1084/jem.20080091 (2008).
610 26 Kato, H. *et al.* Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses.
611 *Nature* **441**, 101-105, doi:10.1038/nature04734 (2006).
612 27 Rehwinkel, J. & Gack, M. U. RIG-I-like receptors: their regulation and roles in RNA sensing. *Nat*
613 *Rev Immunol* **20**, 537-551, doi:10.1038/s41577-020-0288-3 (2020).
614 28 Rodriguez, K. R., Bruns, A. M. & Horvath, C. M. MDA5 and LGP2: accomplices and antagonists
615 of antiviral signal transduction. *J Virol* **88**, 8194-8200, doi:10.1128/JVI.00640-14 (2014).
616 29 Technau, U. & Steele, R. E. Evolutionary crossroads in developmental biology: Cnidaria.
617 *Development* **138**, 1447-1458, doi:10.1242/dev.048959 (2011).
618 30 Layden, M. J., Rentzsch, F. & Rottinger, E. The rise of the starlet sea anemone Nematostella
619 vectensis as a model system to investigate development and regeneration. *Wiley Interdiscip Rev*
620 *Dev Biol* **5**, 408-428, doi:10.1002/wdev.222 (2016).
621 31 Kenny, N. J. *et al.* Tracing animal genomic evolution with the chromosomal-level assembly of the
622 freshwater sponge Ephydatia muelleri. *Nat Commun* **11**, 3676, doi:10.1038/s41467-020-17397-w
623 (2020).
624 32 Whelan, N. V. *et al.* Ctenophore relationships and their placement as the sister group to all other
625 animals. *Nat Ecol Evol* **1**, 1737-1746, doi:10.1038/s41559-017-0331-3 (2017).
626 33 Mukherjee, K., Korithoski, B. & Kolaczkowski, B. Ancient origins of vertebrate-specific innate
627 antiviral immunity. *Mol Biol Evol* **31**, 140-153, doi:10.1093/molbev/mst184 (2014).
628 34 Zou, J., Chang, M., Nie, P. & Secombes, C. J. Origin and evolution of the RIG-I like RNA
629 helicase gene family. *BMC Evol Biol* **9**, 85, doi:10.1186/1471-2148-9-85 (2009).
630 35 Sarkar, D., Desalle, R. & Fisher, P. B. Evolution of MDA-5/RIG-I-dependent innate immunity:
631 independent evolution by domain grafting. *Proc Natl Acad Sci U S A* **105**, 17040-17045,
632 doi:10.1073/pnas.0804956105 (2008).
633 36 Quattrini, A. M. *et al.* Palaeoclimate ocean conditions shaped the evolution of corals and their
634 skeletons through deep time. *Nat Ecol Evol*, doi:10.1038/s41559-020-01291-1 (2020).

635 37 Helm, R. R., Siebert, S., Tulin, S., Smith, J. & Dunn, C. W. Characterization of differential
636 transcript abundance through time during Nematostella vectensis development. *BMC Genomics*
637 **14**, 266, doi:10.1186/1471-2164-14-266 (2013).
638 38 Andresen, A. M. S., Boudinot, P. & Gjoen, T. Kinetics of transcriptional response against poly
639 (I:C) and infectious salmon anemia virus (ISAV) in Atlantic salmon kidney (ASK) cell line. *Dev*
640 *Comp Immunol* **110**, 103716, doi:10.1016/j.dci.2020.103716 (2020).
641 39 Chen, Y. *et al.* Gene expression profile after activation of RIG-I in 5'ppp-dsRNA challenged DF1.
642 *Dev Comp Immunol* **65**, 191-200, doi:10.1016/j.dci.2016.07.009 (2016).
643 40 Wang, Y., Wang, X., Li, J., Zhou, Y. & Ho, W. RIG-I activation inhibits HIV replication in
644 macrophages. *J Leukoc Biol* **94**, 337-341, doi:10.1189/jlb.0313158 (2013).
645 41 Kulkarni, R. R. *et al.* Activation of the RIG-I pathway during influenza vaccination enhances the
646 germinal center reaction, promotes T follicular helper cell induction, and provides a dose-sparing
647 effect and protective immunity. *Journal of virology* **88**, 13990-14001, doi:10.1128/JVI.02273-14
648 (2014).
649 42 Draper, D. E. A guide to ions and RNA structure. *RNA* **10**, 335-343, doi:10.1261/rna.5205404
650 (2004).
651 43 Schneider, W. M., Chevillotte, M. D. & Rice, C. M. Interferon-stimulated genes: a complex web of
652 host defenses. *Annu Rev Immunol* **32**, 513-545, doi:10.1146/annurev-immunol-032713-120231
653 (2014).
654 44 Wang, P. H. & He, J. G. Nucleic Acid Sensing in Invertebrate Antiviral Immunity. *Int Rev Cell*
655 *Mol Biol* **345**, 287-360, doi:10.1016/bs.ircmb.2018.11.002 (2019).
656 45 Hagai, T. *et al.* Gene expression variability across cells and species shapes innate immunity.
657 *Nature* **563**, 197-202, doi:10.1038/s41586-018-0657-2 (2018).
658 46 Zaslavsky, E. *et al.* Antiviral response dictated by choreographed cascade of transcription factors.
659 *J Immunol* **184**, 2908-2917, doi:10.4049/jimmunol.0903453 (2010).
660 47 Seifert, L. L. *et al.* The ETS transcription factor ELF1 regulates a broadly antiviral program
661 distinct from the type I interferon response. *PLoS Pathog* **15**, e1007634,
662 doi:10.1371/journal.ppat.1007634 (2019).

48 Chiang, H. S. & Liu, H. M. The Molecular Basis of Viral Inhibition of IRF- and STAT-Dependent Immune Responses. *Front Immunol* **9**, 3086, doi:10.3389/fimmu.2018.03086 (2018).

49 Gallant, S. & Gilkeson, G. ETS transcription factors and regulation of immunity. *Archivum immunologiae et therapiae experimentalis* **54**, 149-163, doi:10.1007/s00005-006-0017-z (2006).

50 Columbus-Shenkar, Y. Y. *et al.* Dynamics of venom composition across a complex life cycle. *Elife* **7**, e35014, doi:10.7554/eLife.35014 (2018).

51 Lotan, T. *et al.* Evolutionary conservation of the mature oocyte proteome. *EuPA Open Proteomics* **3**, 27-36, doi:https://doi.org/10.1016/j.euprot.2014.01.003 (2014).

52 Levitan, S. *et al.* The making of an embryo in a basal metazoan: Proteomic analysis in the sea anemone Nematostella vectensis. *Proteomics* **15**, 4096-4104, doi:10.1002/pmic.201500255 (2015).

53 Schuberth-Wagner, C. *et al.* A Conserved Histidine in the RNA Sensor RIG-I Controls Immune Tolerance to N1-2'O-Methylated Self RNA. *Immunity* **43**, 41-51, doi:10.1016/j.immuni.2015.06.015 (2015).

54 Grunberg-Manago, M., Ortiz, P. J. & Ochoa, S. Enzymic synthesis of polynucleotides I. polynucleotide phosphorylase of Azotobacter vinelandii. *Biochimica et Biophysica Acta* **20**, 269-285, doi:https://doi.org/10.1016/0006-3002(56)90286-4 (1956).

55 Tirosh, I. & Barkai, N. Two strategies for gene regulation by promoter nucleosomes. *Genome Res* **18**, 1084-1091, doi:10.1101/gr.076059.108 (2008).

56 Newman, J. R. S. *et al.* Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* **441**, 840-846, doi:10.1038/nature04785 (2006).

57 Huang, B. *et al.* Characterization of the Mollusc RIG-I/MAVS Pathway Reveals an Archaic Antiviral Signalling Framework in Invertebrates. *Sci Rep* **7**, 8217, doi:10.1038/s41598-017-08566-x (2017).

58 He, Y. *et al.* Transcriptome analysis reveals strong and complex antiviral response in a mollusc. *Fish Shellfish Immunol* **46**, 131-144, doi:10.1016/j.fsi.2015.05.023 (2015).

59 Lv, X. *et al.* The Dicer from oyster Crassostrea gigas functions as an intracellular recognition molecule and effector in anti-viral immunity. *Fish Shellfish Immunol* **95**, 584-594, doi:10.1016/j.fsi.2019.10.067 (2019).

691 60 Lu, M. *et al.* A conserved interferon regulation factor 1 (IRF-1) from Pacific oyster Crassostrea
692 gigas functioned as an activator of IFN pathway. *Fish Shellfish Immunol* **76**, 68-77,
693 doi:10.1016/j.fsi.2018.02.024 (2018).
694 61 Guo, X., Zhang, R., Wang, J., Ding, S. W. & Lu, R. Homologous RIG-I-like helicase proteins
695 direct RNAi-mediated antiviral immunity in C. elegans by distinct mechanisms. *Proc Natl Acad*
696 *Sci U S A* **110**, 16085-16090, doi:10.1073/pnas.1307453110 (2013).
697 62 Lu, R., Yigit, E., Li, W.-X. & Ding, S.-W. An RIG-I-Like RNA helicase mediates antiviral RNAi
698 downstream of viral siRNA biogenesis in Caenorhabditis elegans. *PLoS pathogens* **5**, e1000286-
699 e1000286, doi:10.1371/journal.ppat.1000286 (2009).
700 63 Ashe, A. *et al.* A deletion polymorphism in the Caenorhabditis elegans RIG-I homolog disables
701 viral RNA dicing and antiviral immunity. *eLife* **2**, e00994-e00994, doi:10.7554/eLife.00994
702 (2013).
703 64 Bernier, A. & Sagan, S. M. The Diverse Roles of microRNAs at the Host–Virus Interface. *Viruses*
704 **10**, 440, doi:10.3390/v10080440 (2018).
705 65 Zhang, Q.-L. *et al.* Transcriptome-wide analysis of immune-responsive microRNAs against poly
706 (I:C) challenge in Branchiostoma belcheri by deep sequencing and bioinformatics. *Oncotarget* **8**,
707 73590-73602, doi:10.18632/oncotarget.20570 (2017).
708 66 Wang, J. *et al.* MicroRNA Transcriptome of Poly I:C-Stimulated Peripheral Blood Mononuclear
709 Cells Reveals Evidence for MicroRNAs in Regulating Host Response to RNA Viruses in Pigs. *Int*
710 *J Mol Sci* **17**, 1601, doi:10.3390/ijms17101601 (2016).
711 67 Wu, J. *et al.* MicroRNA transcriptome analysis of poly I:C-stimulated and PRRSV-infected
712 porcine alveolar macrophages. *Journal of applied genetics* **60**, 375-383, doi:10.1007/s13353-019-
713 00500-3 (2019).
714 68 Singaravelu, R. *et al.* A conserved miRNA-183 cluster regulates the innate antiviral response. *J*
715 *Biol Chem* **294**, 19785-19794, doi:10.1074/jbc.RA119.010858 (2019).
716 69 Fridrich, A., Modepalli, V., Lewandowska, M., Aharoni, R. & Moran, Y. Unravelling the
717 developmental and functional significance of an ancient Argonaute duplication. *bioRxiv*,
718 2020.2002.2004.933887, doi:10.1101/2020.02.04.933887 (2020).

719 70 Kawasaki, T. & Kawai, T. Toll-like receptor signaling pathways. *Front Immunol* **5**, 461,
720 doi:10.3389/fimmu.2014.00461 (2014).
721 71 Brennan, J. J. *et al.* Sea anemone model has a single Toll-like receptor that can function in
722 pathogen detection, NF-kappaB signal transduction, and development. *Proc Natl Acad Sci U S A*
723 **114**, E10122-E10131, doi:10.1073/pnas.1711530114 (2017).
724 72 Genikhovich, G. & Technau, U. Induction of spawning in the starlet sea anemone Nematostella
725 vectensis, in vitro fertilization of gametes, and dejellying of zygotes. *Cold Spring Harb Protoc*
726 **2009**, pdb prot5281, doi:10.1101/pdb.prot5281 (2009).
727 73 Andrew, S. *FastQC: A quality control tool for high throughput sequence data*,
728 <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (2010).
729 74 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence
730 data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).
731 75 Putnam, N. H. *et al.* Sea anemone genome reveals ancestral eumetazoan gene repertoire and
732 genomic organization. *Science* **317**, 86-94, doi:10.1126/science.1139158 (2007).
733 76 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21,
734 doi:10.1093/bioinformatics/bts635 (2013).
735 77 Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or
736 without a reference genome. *BMC Bioinformatics* **12**, 323, doi:10.1186/1471-2105-12-323 (2011).
737 78 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential
738 expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140,
739 doi:10.1093/bioinformatics/btp616 (2010).
740 79 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for
741 RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
742 80 Haas, B. J. *et al.* De novo transcript sequence reconstruction from RNA-seq using the Trinity
743 platform for reference generation and analysis. *Nat Protoc* **8**, 1494-1512,
744 doi:10.1038/nprot.2013.084 (2013).
745 81 Young, M. D., Wakefield, M. J., Smyth, G. K. & Oshlack, A. Gene ontology analysis for RNA-
746 seq: accounting for selection bias. *Genome Biol* **11**, R14, doi:10.1186/gb-2010-11-2-r14 (2010).

747 82 Supek, F., Bosnjak, M., Skunca, N. & Smuc, T. REVIGO summarizes and visualizes long lists of
748 gene ontology terms. *PLoS One* **6**, e21800, doi:10.1371/journal.pone.0021800 (2011).
749 83 Kuznetsova, I., Lugmayr, A., Siira, S. J., Rackham, O. & Filipovska, A. CirGO: an alternative
750 circular way of visualising gene ontology terms. *BMC Bioinformatics* **20**, 84, doi:10.1186/s12859-
751 019-2671-2 (2019).
752 84 Karabulut, A., He, S., Chen, C. Y., McKinney, S. A. & Gibson, M. C. Electroporation of short
753 hairpin RNAs for rapid and efficient gene knockdown in the starlet sea anemone, Nematostella
754 vectensis. *Dev Biol* **448**, 7-15, doi:10.1016/j.ydbio.2019.01.005 (2019).
755 85 Moran, Y., Agron, M., Praher, D. & Technau, U. The evolutionary origin of plant and animal
756 microRNAs. *Nat Ecol Evol* **1**, 27, doi:10.1038/s41559-016-0027 (2017).
757 86 Kim, J. H. *et al.* High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in
758 human cell lines, zebrafish and mice. *PLoS One* **6**, e18556, doi:10.1371/journal.pone.0018556
759 (2011).
760 87 Shaner, N. C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from
761 Discosoma sp. red fluorescent protein. *Nature biotechnology* **22**, 1567-1572, doi:10.1038/nbt1037
762 (2004).
763 88 Renfer, E. & Technau, U. Meganuclease-assisted generation of stable transgenics in the sea
764 anemone Nematostella vectensis. *Nat Protoc* **12**, 1844-1854, doi:10.1038/nprot.2017.075 (2017).
765 89 Admoni, Y., Kozlovski, I., Lewandowska, M. & Moran, Y. TATA Binding Protein (TBP)
766 Promoter Drives Ubiquitous Expression of Marker Transgene in the Adult Sea Anemone
767 Nematostella vectensis. *Genes (Basel)* **11**, doi:10.3390/genes11091081 (2020).
768 90 Renfer, E., Amon-Hassenzahl, A., Steinmetz, P. R. & Technau, U. A muscle-specific transgenic
769 reporter line of the sea anemone, Nematostella vectensis. *Proc Natl Acad Sci U S A* **107**, 104-108,
770 doi:10.1073/pnas.0909148107 (2010).
771 91 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput.
772 *Nucleic acids research* **32**, 1792-1797, doi:10.1093/nar/gkh340 (2004).

92 Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973, doi:10.1093/bioinformatics/btp348 (2009).

93 Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**, 268-274, doi:10.1093/molbev/msu300 (2015).

94 Minh, B. Q., Nguyen, M. A. & von Haeseler, A. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol* **30**, 1188-1195, doi:10.1093/molbev/mst024 (2013).

95 Ronquist, F. *et al.* MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* **61**, 539-542, doi:10.1093/sysbio/sys029 (2012).

96 El-Gebali, S. *et al.* The Pfam protein families database in 2019. *Nucleic acids research* **47**, D427-D432, doi:10.1093/nar/gky995 (2019).

97 Lu, S. *et al.* CDD/SPARCLE: the conserved domain database in 2020. *Nucleic acids research* **48**, D265-D268, doi:10.1093/nar/gkz991 (2020).

98 Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017-1018, doi:10.1093/bioinformatics/btr064 (2011).

99 Fornes, O. *et al.* JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic acids research* **48**, D87-D92, doi:10.1093/nar/gkz1001 (2020).

100 McLeay, R. C. & Bailey, T. L. Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data. *BMC Bioinformatics* **11**, 165, doi:10.1186/1471-2105-11-165 (2010).

101 Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods* **8**, 785-786, doi:10.1038/nmeth.1701 (2011).