

# Defining the Origins of the NOD-Like Receptor System at the Base of Animal Evolution

Christina Lange,<sup>1</sup> Georg Hemmrich,<sup>1</sup> Ulrich C. Klostermeier,<sup>3</sup> Javier A. López-Quintero,<sup>1</sup> David J. Miller,<sup>2</sup> Tasja Rahn,<sup>1</sup> Yvonne Weiss,<sup>2</sup> Thomas C.G. Bosch,<sup>†1</sup> and Philip Rosenstiel<sup>†\*3</sup>

<sup>1</sup>Zoological Institute, Christian-Albrechts-University Kiel, Kiel, Germany

<sup>2</sup>ARC Centre of Excellence in Coral Reef Studies and Comparative Genomics Centre, James Cook University, Townsville, Queensland, Australia

<sup>3</sup>Institute of Clinical Molecular Biology, University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany

†Denotes shared senior authorship.

\*Corresponding author: E-mail: p.rosenstiel@mucosa.de.

Associate editor: Claudia Kappen

## Abstract

Distinguishing self from nonself and the onset of defense effector mechanisms upon recognition of pathogens are essential for the survival of all life forms in the animal kingdom. The family of nucleotide-binding and oligomerization domain-like receptors (NLRs) was first identified in vertebrates and comprises a group of pivotal sensor proteins of the innate immune system for microbial cell wall components or danger signals. Here, we provide first evidence that early diverging metazoans have large and complex NLR repertoires. The cnidarian NACHT/NB-ARC genes include novel combinations of domains, and the number of one specific type (NB-ARC and tetratricopeptide repeat containing) in *Hydra* is particularly large. We characterize the transcript structure and expression patterns of a selected HyNLR, HyNLR type 1 and describe putative interaction partners. In a heterologous expression system, we show induced proximity recruitment of an effector caspase (HyDD-Caspase) to the HyNLR type 1 protein upon oligomerization indicating a potential role of caspase activation downstream of NLR activation in *Hydra*. These results add substantially to our understanding of the ancestral innate immune repertoire as well as providing the first insights into putative cytoplasmic defense mechanisms at the base of animal evolution.

**Key words:** *Hydra*, innate immunity, epithelial defense, NOD-like receptors, Crohn's disease, inflammation.

## Introduction

All animals face constant challenges from infective agents, and this has led to the evolution of diverse mechanisms of immune response. Whereas the adaptive immune system is only used by the jawed vertebrates, the innate immune system is phylogenetically ancient and parts of it date back to early metazoans (Wiens et al. 2005, 2007; Hibino et al. 2006; Rast et al. 2006; Rosenstiel et al. 2009). Most of the current knowledge on innate immunity has been gathered in vertebrates and *Drosophila* (Lemaitre and Hoffmann 2007). Recent comparative analyses imply that the innate immune repertoire of early diverging animals such as cnidarians is remarkably complex and vertebrate like (Hemmrich et al. 2007; Miller et al. 2007; Augustin et al. 2010). Canonical toll-like receptors (TLRs) have been identified in cnidarians, and *Hydra* has been shown to synthesize antimicrobial peptides in response to immune challenge via a TLR-like system (Bosch et al. 2009).

In vertebrates, the TLRs respond to extracellular immune stimuli, whereas the primary intracellular pattern molecules are the nucleotide-binding and oligomerization domain-like receptors (NLRs) or NODs. Vertebrate NLR genes encode regulatory cytosolic proteins with a tripartite domain structure that are central players in innate immune

responses to viruses and other pathogen-associated molecular patterns (PAMPs). Two subgroups of NLRs are recognized: 1) the AP (apoptotic)-ATPases with, for example, apoptotic protease activating factor-1 (APAF1; Zou et al. 1997), which harbor the NB-ARC domain (van der Biezen and Jones 1998) and 2) the NACHT-(domain present in neuronal apoptosis inhibitor protein [NAIP], the major histocompatibility complex [MHC] transactivator [CIITA], HET-E, and TP1) NTPases that are mostly represented by the bona fide NLRs that contain a NACHT domain. Historically, this domain has been synonymously called NOD. Two features of NLR proteins are common—the central NOD or NACHT (Koonin and Aravind 2000) and C-terminal leucine-rich repeats (LRRs). NLRs also contain one of several different N-terminal effector binding domains that contain DEATH-like folds: either by a pyrin domain (PYD), a bona fide DEATH domain (Brakebusch et al. 1992; Itoh and Nagata 1993; Tartaglia et al. 1993), a caspase recruitment domain (CARD), or a baculovirus inhibitor of apoptosis protein repeat (BIR). Upon recognition of a variety of different cytosolic danger signals including PAMP from bacteria and viruses, NLRs form macromolecular protein complexes by self-oligomerization and recruit effector molecules via homotypic domain interactions (e.g., CARD/

CARD). In vertebrates, two major types of NLR complexes can be defined: 1) in the NODosome (NOD1/NOD2), the serine–threonine kinase receptor-interacting serine/threonine-protein kinase (RIPK2) is recruited (McCarthy et al. 1998; Inohara et al. 2000), which leads to the activation of proinflammatory pathways via canonical RIPK2/IKK/NF- $\kappa$ B signaling and 2) inflammasomes can be constituted by several different NLRP (NLR family, PYD containing) proteins, which mediate direct activation of proinflammatory caspases via formation of a large protein complexes, that enable the processing of interleukin (IL) 1 $\beta$ -like substrates (Martinon et al. 2002; Agostini et al. 2004; Rosenstiel, Till, et al. 2007).

Animals contain few NB-ARC proteins, but cell death (CED)-4/APAF1 is widely distributed throughout the Bilateria. Until recently, it was assumed that NLRs sensu stricto evolved at the level of teleost fish by exon shuffling events, as classical NLRs are not present in the genomes of either *Drosophila* or *Caenorhabditis* (Ting and Davis 2005). However, the discovery of over 200 NLRs in the genome of the sea urchin *Strongylocentrotus* (Hibino et al. 2006; Rast et al. 2006) indicates an earlier origin for this gene family. Most of the sea urchin NLRs contain one or two DEATH-folds as N-terminal effector–binding domains instead of a single CARD, PYD, or BIR domain. This principle is also realized in putative orthologous interacting proteins, for example, a RIPK2-like protein that also carries a C-terminal DEATH domain instead of a CARD.

Although NLRs sensu stricto are known only from members of the animal kingdom, they have functional analogs in plants and fungi that contain some of the same domains. Higher plants typically contain large families of pathogen-resistance (R-) genes (Nürnberg et al. 2004; Chisholm et al. 2006) that encode proteins with a C-terminal LRR domain and a central NB-ARC domain—a NACHT-related domain also found in the CED-4/APAF1 caspase regulators of metazoans (van der Biezen and Jones 1998). Activation of the encoded R-proteins by pathogen infection induces a hypersensitive response (Ausubel et al. 1995; Hulbert et al. 2001). The STAND family of fungal incompatibility proteins are also tripartite, the central NACHT domain being flanked by a C-terminal WD-40 repeat and N-terminal HET CED effector domains (Paoletti and Saupé 2009).

In this report, we have set out to systemically survey the repertoires of NACHT and NB-ARC domain genes in existing expressed sequence tag (EST) and genome data sets of early branching metazoans. We provide the first evidence that early diverging metazoans have large and complex NLR repertoires. The cnidarian NACHT/NB-ARC complements include novel combinations of domains, and the number of one specific type (NB-ARC and tetratricopeptide repeat [TPR] containing) in *Hydra* is particularly large. Surveying the *Hydra* genome allowed the identification of a number of potential NLR-interacting proteins. One of these, a caspase containing a DEATH domain, was shown to interact with a *Hydra* NLR-like protein in vitro. These results add substantially to our understanding of the ancestral innate immune repertoire as well as providing the first insights

into cytoplasmic defense mechanisms at the base of animal evolution.

## Materials and Methods

### Animals

We used *Hydra magnipapillata* strain 105 for all the work presented here. The animals were cultured according to standard procedures at 18 °C.

### In Silico Analysis

SMART (<http://smart.embl-heidelberg.de/>; Letunic et al. 2006) and a local install of HMMER (<http://hmmer.janelia.org/>) program were used for domain search. DNAMAN was used for general sequence analysis.

We screened the genomes of *Gallus gallus*, *Xenopus tropicalis*, *Oryzias latipes*, *Petromyzon marinus* (all available at <http://genome.ucsc.edu/>), and *Daphnia pulex* (<http://genome.jgi-psf.org/>) via Blast searches (Altschul et al. 1990) with domains of different NLRs and via text searches. For *Nematostella vectensis*, we screened the gene models available at StellaBase (<http://128.197.80.144/stellabase/>; Sullivan et al. 2006) for NACHT domains and performed Blast searches with a consensus sequence from all the NACHT domains from the gene models.

Due to the fact that the genome of *Amphimedon queenslandica* is not available online, we screened the whole-genome shotgun sequences via Blast searches with the NACHT consensus sequence from *N. vectensis*. The hits with a sequence identity above 94% were assembled using the Minimus assembler from the AMOS package (<http://amos.sourceforge.net>; Sommer et al. 2007). For the *Acropora millepora* data, we screened the EST database available at [www.compagen.org](http://www.compagen.org) (Hemrich and Bosch 2008). For *H. magnipapillata*, we used the NACHT domain consensus sequence from *N. vectensis* as a query sequence for Blast searches in the whole-genome shotgun sequences. All hits were assembled, the NACHT domain containing assembly sequences and the TPR-type NB-ARC-domain containing sequences were mapped to the TIGR *H. magnipapillata* genome scaffolds and screened for haplotypes. The remaining sequences were screened for DEATH-fold domains, LRRs, and TPRs in the right reading frame orientation to the nucleotide-binding domain (NBD) domain. For the calculation of the NACHT domain tree, SMART-predictions of rather complete NACHT domains were used. For organisms with a large amount of NLR proteins (*H. magnipapillata*, *A. queenslandica*, and *N. vectensis*), previous trees were calculated on amino acid alignments using the ClustalW Multiple Alignment Tool (Thompson et al. 1994) of BioEdit (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>; Hall 1999) and the neighbor joining algorithm (Saitou and Nei 1987) of MEGA 4.0 (Tamura et al. 2007) (Bootstrap:1,000). Representatives of all groups supported with a bootstrap value >50% were selected. For already published NLRs of *Strongylocentrotus purpuratus* and *Branchiostoma floridae* representatives were selected on the

base of the published trees (Hibino et al. 2006; Huang et al. 2008). These prototype amino acid sequences were aligned by hand on the basis of the NACHT domain alignment published by Koonin and Aravind (2000). As outgroup, the NB-ARC domain of an R-protein (RPS4) of *Arabidopsis thaliana* was selected. For the calculation of the tree, a local install of MrBayes 3.1.2 (Huelsenbeck and Ronquist 2001) was used. A total of 2.6 million generations were calculated for the tree shown in figure 3 and 4 million generations for the tree shown in figure S3, Supplementary Material online, using the general time reversible model and four chains with a burn-in of 25% and the invgamma rate variation. Only posterior probabilities >50% are shown. The tree was visualized using FigTree1.3.1 (<http://tree.bio.ed.ac.uk/software/figtree/>; Rambaut 2006–2009).

### Incubation in LPS or Flagellin

Animals were starved for 48 h. The medium was exchanged prior to the experiment and animals were kept in fresh medium overnight. Twenty polyps were incubated with lipopolysaccharide (LPS) (200 ng/ml) or with flagellin (2.5 µg/ml) for 4 h. Animals that were incubated in fresh medium served as negative control. To ensure that the microbe-associated molecular patterns (MAMPs) reach the gastric cavity, the two solutions and the pure medium were injected into the polyps using a 1 ml syringe and a micro cannulation needle (Hamilton Bonaduz AG, Bonaduz, Switzerland). As a control for wounding, we used untreated animals that were incubated in fresh medium. Afterward, RNA was isolated as described below.

### Real-Time Polymerase Chain Reaction

Real time polymerase chain reaction (RT PCR; TaqMan) was performed according to the manufacturer's guidelines using a 7900HT Fast Real Time PCR-System (Life Technologies Corporation). Expression levels of *HyNLR type 1* were compared with *actin*. Expression levels were calculated relative to *actin* using the standard-curve method (Livak and Schmittgen 2001) and performed in duplicate.

Oligos used for PCR: *HyNLR type 1*: 5'-GGG CTC CTT GAA AAG GTG TTT-3', 5'-TAG TAT GCA TGA ACA GAG ACA CAT AGC T-3', Probe: 6-FAM TTT GGT TAT CAT TAT CGG TTT TGT CAG ACA GCA-TAMRA

*actin*: 5'-TCC GTG TTG CTC CAG AAG AA-3', 5'-TTT GGG TCA TTT TTT CAC GAT TAG-3', Probe: 6-FAM CCC TGT CCT TCT TAC TGA AGC TCC CCT GA-TAMRA.

### Procaine Treatment

The procaine method originally described by Epp et al. (1979) and modified by Bode et al. (1987) was used for separating the ectodermal from the endodermal layer. The head and the foot of the animals were removed by sectioning. Afterward, the body columns were treated in solution A (1/3 1% procaine-HCl solution in distilled water, 1/3 DM solution (Gierer et al. 1972), 1/3 Hydra medium, pH 4.5) for 5 min. This was followed by a one in treatment in solution B

(as solution A, but pH 2.5). The body columns were then transferred to DM solution and kept at 18 °C. The ring-shaped ectoderm and the rod-shaped endoderm were separated with thin needles. Afterward, RNA was isolated as described below.

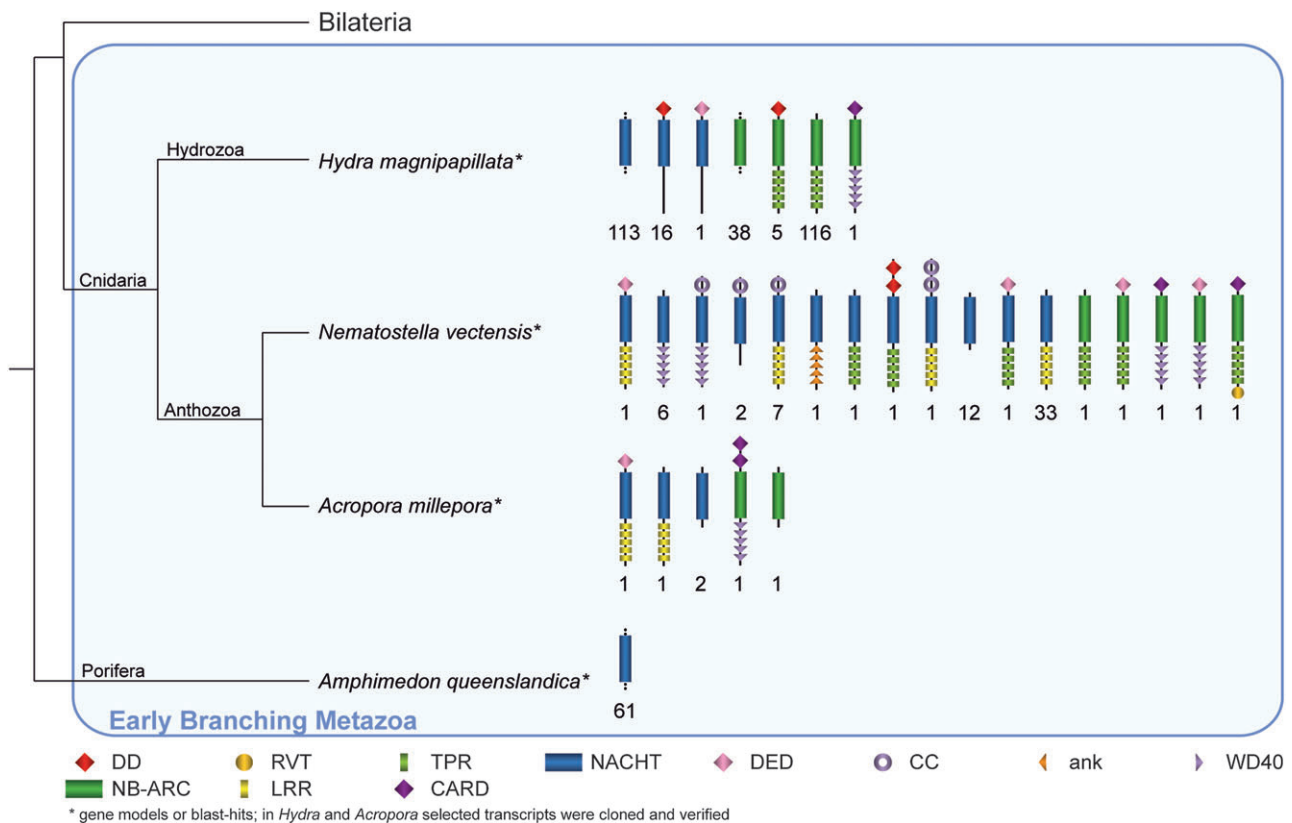
### RNA Isolation and PCR

Total RNA was isolated using TRIzol (Life Technologies Corporation, Carlsbad, CA). For ss cDNA synthesis, the first strand cDNA synthesis Kit (Fermentas GmbH, St.Leon-Rot, Germany) was used. For amplification of gene-specific fragments, the following primer sets were used:

*HyNLR type 1*: 5'-AGA TAT CTT TGA ATC ATG TGA TTG G-3', 5'-TTG TAC ACG TAC TAT ATT AAT ACC-3'  
 Various *HyNLRs*: 5'-TGG TAT TGG GAA AAC ATG GTT GC-3', 5'-GAT CAA TTG CAT AAA CTC TAC CAG C-3'  
*HyDD-Caspase*: 5'-CTA ATG AGA AGA ATG TAA ATA CGG-3', 5'-TTA AAG TAG ATT GAA GTT CTG GC-3'  
*HyDED-Caspase*: 5'-TTT CAT TCA AGC TTG TCA AAC C-3', 5'-GCC ATA TTA GAG TTA AGT TCC-3'  
*HyCARD-Caspase*: 5'-GAG TCT GAT GTC TCA AAT GG-3', 5'-TAA GAT TCA ATA ACA TAT CCT AAT TCC-3'  
*HySGT-1*: 5'-ATG AAA TCT TTT GTT GAA CAA GTT TCT AGT TAC G-3', 5'-CAT CAC TTC GAC CAT TTC TGA ATC C-3'  
*HyDODE*: 5'-GGA AAT ATA GTT CCA AGT AAT TGG-3', 5'-AAT AAC TAT TGC GCT ATG CTG C-3'  
*HyDED-DD-Kinase*: 5'-CTG ATC TTA TTT CAA TAA TCA ATG G-3', 5'-CAA CCT TCT TTT GAC TTC TGC-3'  
*HyNF-κB*: 5'-AAA CTT ATT AAT AGA CTA ACT GG-3', 5'-TTC ACT TTT ACA GGT TGA TCC-3'  
*HyJUN*: 5'-GCT TTT TAC GAT AGA TTG GTA ATG AG-3', 5'-GTG CTT CAA TAA AAC CTC GTG C-3'  
*nb042*: 5'-TTT TGG GTG TGG TTG GTT TT-3', 5'-TTC ACC TGC AAC TTC GTT TG-3'  
*HyDkk1/2/4-C*: 5'-TCC AAA ATG AAA TCA ACG GTT ATT-3', 5'-ACC ACA CGC AAA TGT TTG G-3'  
*actin*: 5'-CCA AAA TAG ATC CTC CGA TCC-3', 5'-AAG CTC TTC CCT CGA GAA ATC-3'

The fragments were cloned into the pGEM-T vector (Promega GmbH) and transformed into electrocompetent DH10B *Escherichia coli* cells (Invitrogen GmbH) prior to sequencing. Full-length cDNA sequences were obtained by 3'-rapid amplification of cDNA ends (RACE) PCR as described before (Fedders et al. 2004) and the 5' ends were amplified using the transspliced leader B sequence as primer-binding site (5'-TTT AGT CCC TGT GTA ATA AG-3'). All sequences were deposited at GenBank (*HyNLR type 1*: HQ660063; *HyNLR type 1 var 2*: HQ660068; *HyNLR type 1 var 3*: HQ660064; *HyNLR type 1 var 4*: HQ660065; *HyDODE*: HQ660066; *HySGT1*: HQ660067; *HyNF-kappaB*: HQ660069; *HyCARD-Caspase*: Q983384; *HyDD-Caspase*: HQ660070; *HyDED-Caspase*: HQ660071; *HyJUN*: HQ660072; *HyDED-DD-Kinase*: HQ660073). Gene models for *HyHSP90*: Hma2.226194, Hma2.216489, Hma2.217134. Gene model for *HyChp-1*: Hma1.134266.

RT PCR was performed on ss cDNA. The cDNA samples were equilibrated with primers against *actin* (5'-CCA AAA TAG ATC CTC CGA TCC-3' and 5'-AAG CTC TTC CCT



**Fig. 1.** Overview of structures of NBD-containing proteins in basal metazoans. For two NB-ARC domain containing models in *Nematostella vectensis*, a previous publication (Zmasek et al. 2007) was taken into consideration. Below the predicted protein structure, the numbers of proteins are mentioned. Abbreviations: TPR, tetratricopeptide repeat; NB-ARC, nucleotide-binding adaptor shared by APAF1, *R* gene products, and CED-4; DD, DEATH domain; NACHT, domain present in neuronal apoptosis inhibitor protein (NAIP), the major histocompatibility complex (MHC) transactivator (CIITA), HET-E, and TP1; LRR, Leucine-rich repeat; DED, DEATH effector domain; CARD, Caspase recruitment domain; CC, coiled coil; ank, ankyrin repeat; RVT, reverse transcriptase; FIIND, domain with function to find; and ucd, uncharacterized domain. The lengths of proteins and domains are not to scale. The number of repetitions of a repetitive domain (WD40, LRR, TPR, and ank) do not reflect the exact number of repeats.

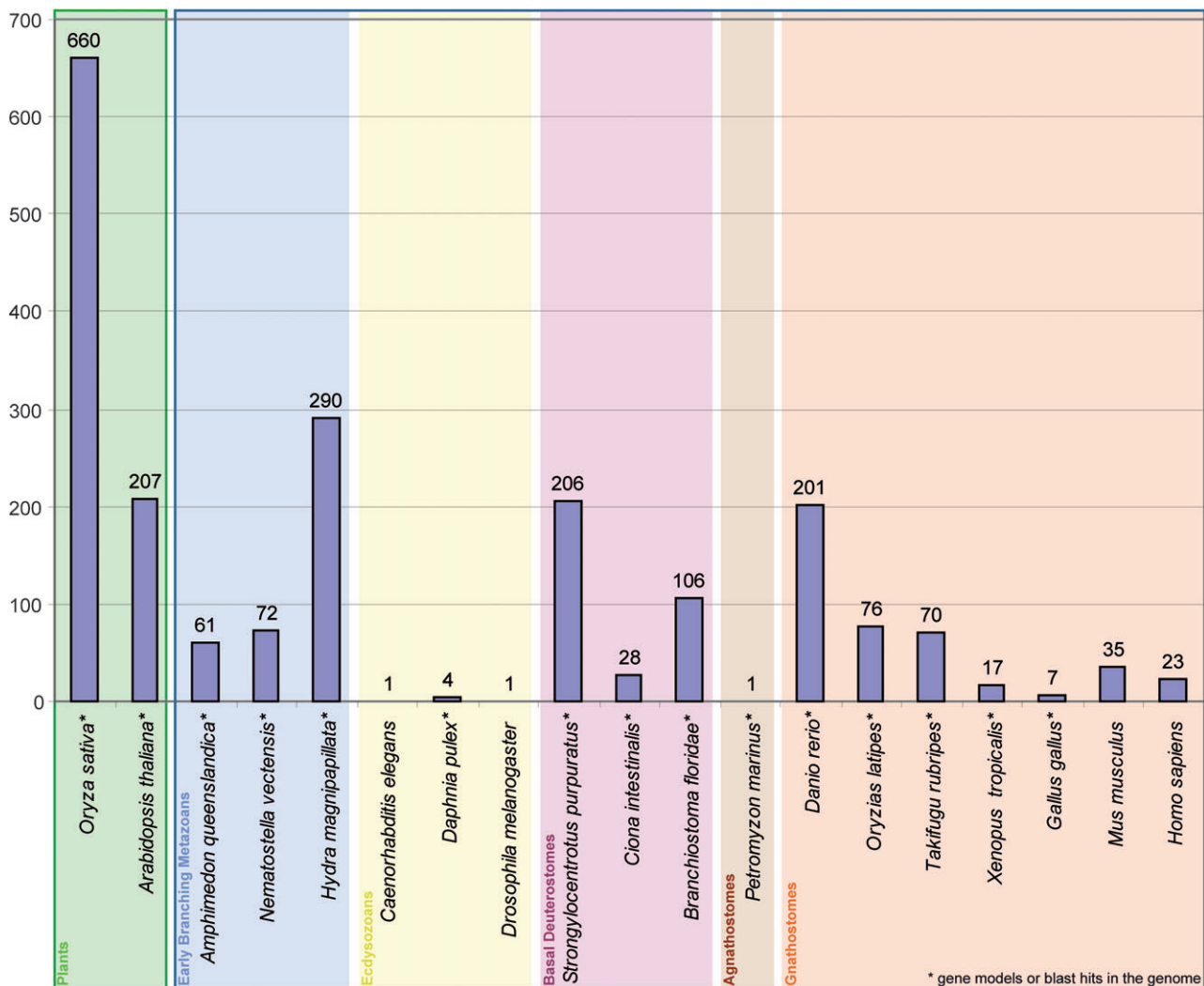
CGA GAA ATC-3'). For the procaine-treated animals, the nematocyte lineage-specific gene *nb042* (Milde et al. 2009) was used as a control ectodermal gene. The gland cell-specific gene *HyDkk1/2/4 C* (Augustin et al. 2006) was used as an endodermal control gene. Water control and equilibration controls were included in each reaction.

### In Situ Hybridization

In situ hybridizations were adapted from previous works (Grens et al. 1996). Riboprobes were prepared with the Dig-RNA labeling kit according to the manufacturer's instructions (Roche Applied Science, Germany) using sequence-specific primer combinations as follows: *HyNLR type 1*: 5'-TAG GTA TCT ATT TAG TCG AAG TG-3', 5'-GCT CAA GAA ATC TGT CTT GAC-3'; *HyDD-Caspase*: 5'-CTA ATG AGA AGA ATG TAA ATA CGG-3', 5'-TTA AAG TAG ATT GAA GTT CTG GC-3'. They were combined with an anti-Dig-AP antibody (Roche Applied Science). Alkaline phosphatase staining was carried out with Nitro Blue Tetrazolium/5-bromo-4-chloro-3-indolyl-phosphate p-toluidine salt (NBT/BCIP) in NTMT buffer (1:100). Animals were dehydrated prior to euparal embedding.

### DNA Isolation and Southern Blot

Genomic DNA from about 500 polyps was isolated by tissue homogenization in lysis buffer containing 100 mM Tris (pH 8), 100 mM ethylenediaminetetraacetic acid and 1% sodium dodecyl sulfate (SDS), followed by proteinase K treatment (1 mg/ml) for 2 h at 50 °C, chloroform-phenol extraction, and ethanol precipitation. After a washing step in 75% ethanol, DNA was redissolved in water and treated with RNase A (Fermentas GmbH). For Southern blot analysis, 20 µg of DNA were digested with *HindIII*, *XbaI*, and *EcoRV*, separated on a 0.6% agarose gel and transferred onto Hybond N+ nylon membrane (GE Healthcare, Buckinghamshire, UK) following standard procedures. The gene fragment of *HyNLR3* coding for the NBD domain was PCR amplified and used as probe template using the following primers: 5'-TGG TAT TGG GAA AAC ATG GTT GC-3', 5'-GAT CAA TTG CAT AAA CTC TAC CAG C-3'. The probes were radioactively labeled using Klenow fragment and  $\alpha^{32}$ -dCTP following standard procedure. Southern hybridization was performed in the hybridization solution containing 50% formamide, 4.8× SSC, 10 mM Tris-HCl (pH 7.5), 1% SDS, 1× Denhardt solution, and 10% Dextran sulfate at 42 °C overnight. Washing of Southern blots was



**Fig. 2.** Number of NBDs in selected multicellular organisms. The NACHT and NB-ARC domains were counted in *Gallus gallus*, *Xenopus tropicalis*, *Oryzias latipes*, *Petromyzon marinus*, *Daphnia pulex*, *Hydra magnipapillata*, *Nematostella vectensis*, and *Amphimedon queenslandica* at the basis of homology searches via Blast or HMMER or gene models. The total numbers of NBDs in *A. queenslandica* cannot be clearly estimated due to the fact that single genome reads were screened. For *Homo sapiens* (Zou et al. 1997; Schreiber et al. 2005), *Mus musculus* (Ceconi et al. 1998; Kanneganti et al. 2007), *Takifugu rubripes* (Stein et al. 2007), *Danio rerio* (Stein et al. 2007), *Branchiostoma floridae* (Zmasek et al. 2007; Huang et al. 2008), *Ciona intestinalis* (Sodergren et al. 2006), *Strongylocentrotus purpuratus* (Hibino et al. 2006; Rast et al. 2006; Zmasek et al. 2007), *Drosophila melanogaster* (Rodriguez et al. 1999), *Caenorhabditis elegans* (Yuan and Horvitz 1990), *Arabidopsis thaliana* (Meyers et al. 2003), and *Oryza sativa* (Bai et al. 2002) previous publications were taken into consideration.

performed in  $2\times$  SSC/0.1% SDS at room temperature and  $0.2\times$  SSC/0.1% SDS at  $65^\circ\text{C}$ .

### Expression Constructs

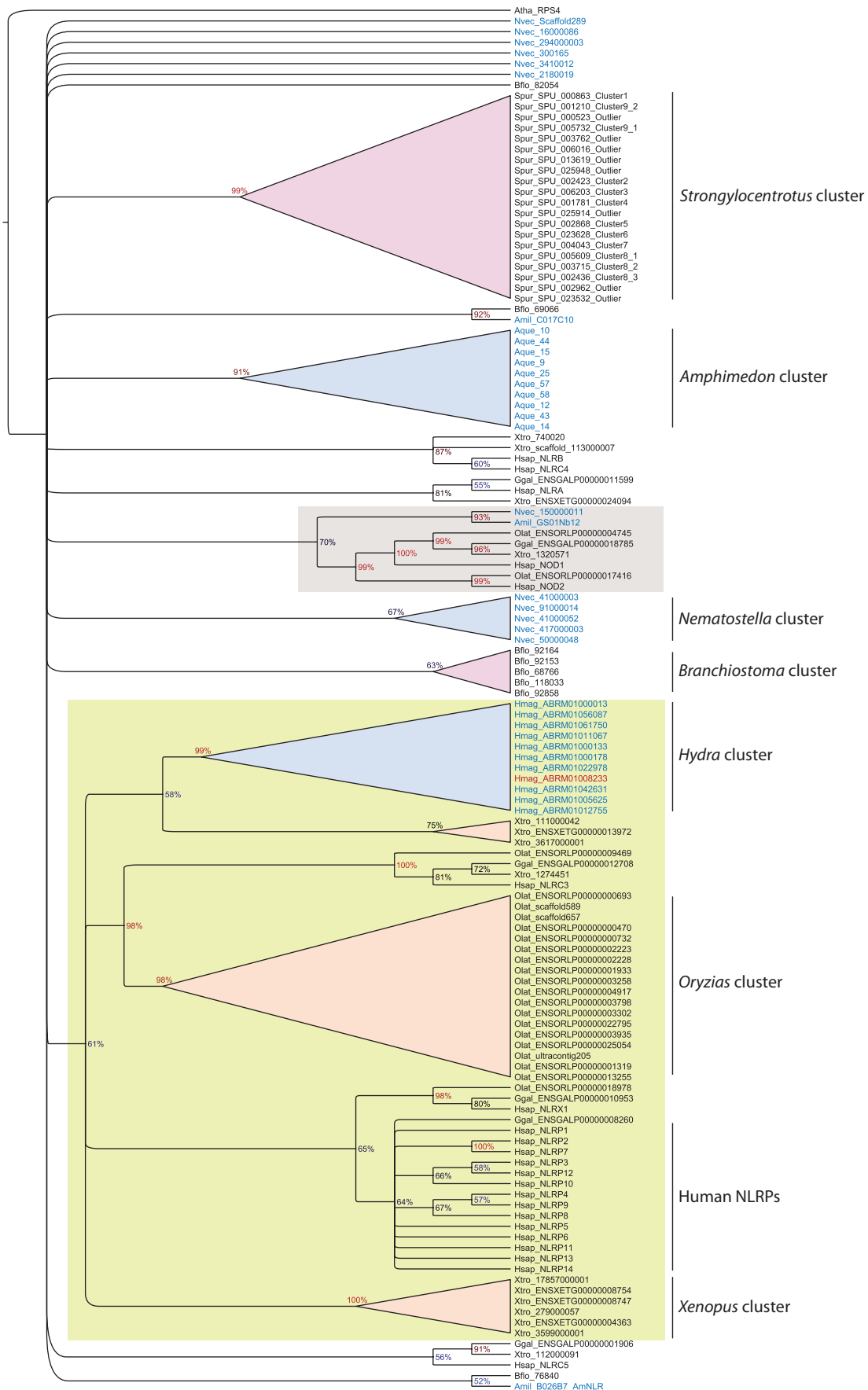
*HyDD-Caspase* was amplified from *H. magnipapillata* cDNA and cloned into the *EcoRI/XhoI* sites of pCMV-myc mammalian expression vector (Clontech Laboratories, Mountain View, CA). Using gateway cloning (Life Technologies Corporation), the HyNLR type 1-FKBP<sub>2</sub> chimeric sequence was transferred into a mammalian expression vector construct in our lab with an N-terminal Flag-tag.

### Immunoprecipitation and Western Blot

HEK293 epithelial cells (ACC305) were purchased from the German Collection of Microorganisms and Cell Cultures

(DSMZ, Braunschweig, Germany) and cultured as described (Rosenstiel, Sina, et al. 2007).

The HyDD-Caspase and the HyNLR type 1::FKBP<sub>2</sub> chimeric sequence were cotransfected into HEK293 cells; 24 h later FKBP homodimerization ligand AP20187 (Ariad Pharmaceuticals, Cambridge, MA) was added. Transfection cells were washed after 48 h with phosphate buffered saline and resuspended in lysis buffer containing 150 mM NaCl, 50 mM Tris-HCl, 1 v/v % NP40, 0.25 w/v % sodium deoxycholate, phosphatase inhibitor cocktail 2 (Sigma), and proteinase inhibitor complete (Roche). Afterward, samples were vortexed three times at maximum speed for 45 s. Cell debris was removed by centrifugation ( $16,000\times g$  for 15 min at  $4^\circ\text{C}$ ). Supernatants were incubated with rabbit monoclonal antibodies against Myc-tag (Cell Signaling Technology, Danvers, MA) and Protein A/G linked sepharose beads (Santa Cruz



**FIG. 3.** Phylogenetic analysis of representative NACHT domains across the animal kingdom. The tree was built by Bayesian inference of phylogeny. Percentages indicate posterior probabilities. Posterior probabilities below 50% are not shown; bright yellow box indicates that *Hydra*

Biotechnology, Santa Cruz, CA) and washed 3× with lysis buffer for 20 min. Supernatants and immunoprecipitation products were separated by SDS-polyacrylamide gel electrophoresis and blotted to a polyvinylidene fluoride membrane. For detection, mouse anti-flag M2 antibodies (Sigma) and rabbit antimyc antibodies (Cell Signaling Technology) were used.

## Results

### The Complexity of the NLR Repertoire in Nonbilateral Animals

In the following, we will subsume the NB-ARC domain and the NACHT domain under the term NBD. To elucidate the occurrence and diversity of NBD-containing proteins in early branching metazoans, we screened the available databases of various animal groups for NOD-like receptors and APAF1 orthologues via Blast analysis or HMMER searches and compared our genome-based results with the numbers of NBDs present in the genomes of other multicellular organisms published by other groups (Yuan and Horvitz 1990; Zou et al. 1997; Cecconi et al. 1998; Rodriguez et al. 1999; Bai et al. 2002; Meyers et al. 2003; Schreiber et al. 2005; Hibino et al. 2006; Rast et al. 2006; Kanneganti et al. 2007; Stein et al. 2007; Zmasek et al. 2007; Huang et al. 2008) (figs. 1 and 2; supplementary fig. S1, Supplementary Material online). Twenty-eight NACHT-only gene models have been proposed in the genome of the urochordate *Ciona intestinalis*, which we decided to leave out of the following analysis as no evidence for N- or C-terminal adaptor domains exists so far, which would support the presence of NLRs in this animal (Sodergren et al. 2006).

A first assembly of the sponge genome of *A. queenslandica* has been made available very recently (Srivastava et al. 2010). Using the trace archive of the genome project, we have identified 349 sequences carrying putative NACHT domains that were assembled to 61 genome contigs. This implied presence of a large number of NACHT proteins suggests a relatively complex NBD repertoire already in early diverging Metazoa.

In the case of cnidarians, assembled genomes are available for two representatives—the sea anemone *N. vectensis* (Putnam et al. 2007) and *H. magnipapillata* (Chapman et al. 2010). Scanning these resources and the EST data set for the coral *A. millepora* indicate the presence of NBD repertoires that are surprisingly complex in terms of both combinations of domains and absolute numbers, particularly in the case of NB-ARC proteins (figs. 1 and 2; supplementary fig. S1, Supplementary Material online). Clear counterparts of APAF1/CED-4 were identified in each of the cnidarians. We note that the anthozoan *Acropora* protein contains

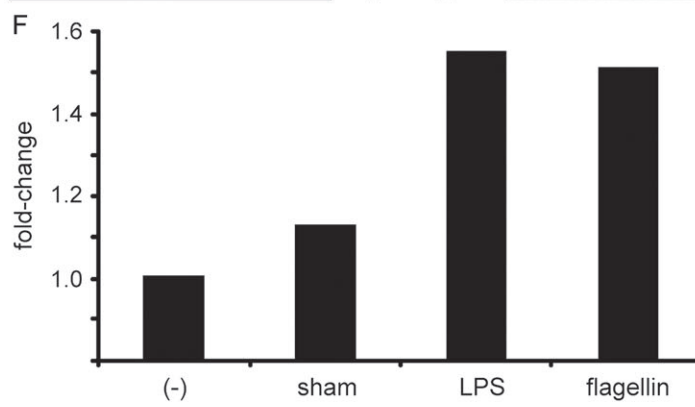
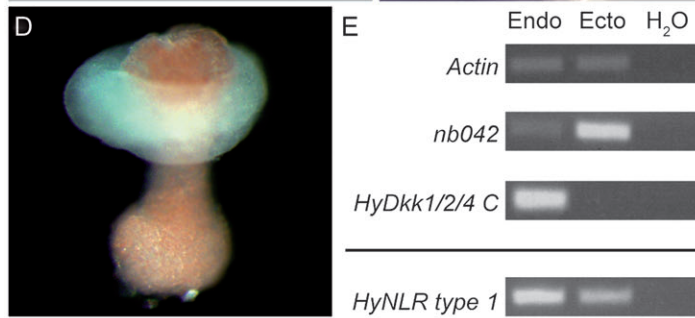
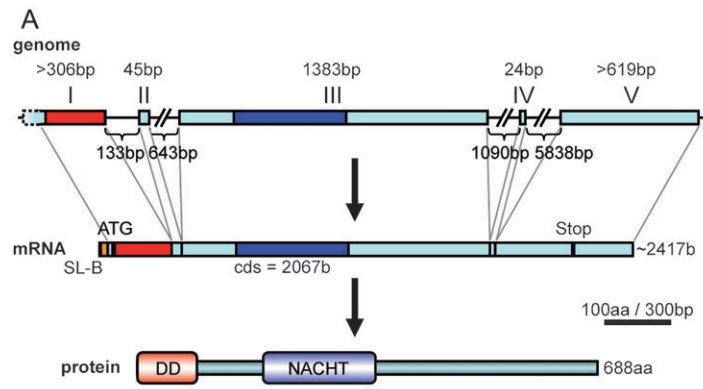
two N-terminal CARD domains. The domain structure of *Acropora* APAF1 protein was verified by cloning a full-length cDNA. In addition, an APAF1-like molecule with a C-terminal DEATH effector domain (DED) was identified in *Nematostella*. Several NB-ARC proteins with C-terminal TPRs were identified *Nematostella*, and the number of these encoded by the *Hydra* genome is surprisingly large (121). Among metazoans, the only known precedent for this novel domain architecture is from *B. floridae* (Zmasek et al. 2007). A subset of the NB-ARC/TPR proteins—two in *Nematostella* and five in *Hydra*—also have an N-terminal effector domain of a DEATH-like fold. This unusual architecture was experimentally verified for several of the *Hydra* gene predictions; we cloned and sequenced 12 transcript variants out of cDNA with primers for conserved regions showing that these genes are indeed transcribed.

Single classical NLRs were identified in both *Acropora* and *Nematostella*, but *Hydra* appears to lack such a gene; searching based on both HMMs or Blast analyses using canonical NLRs from a wide variety of species failed to identify a bona fide *Hydra* NLR. The typical tripartite domain architecture (N-terminal DED, Boldin et al. 1995; Chinnaiyan et al. 1995/NACHT domain/C-terminal LRRs) of the *Acropora* NLR was experimentally verified by cloning of the complete open reading frame (ORF). In addition to the classical NLRs, at least 33 other predicted NACHT/LRR proteins were identified in the genome of *Nematostella*, as well as NACHT proteins containing a wide variety of other domain architectures. Other orphaned NACHT and NB-ARC domains were identified in each of the cnidarians, and in *Hydra*, the numbers of such predicted proteins are particularly large (figs. 1 and 2; supplementary fig. S1, Supplementary Material online).

The striking complexity of the *Hydra* NACHT/NB-ARC repertoire is summarized in figures 1 and 2; supplementary figure S1, Supplementary Material online. In the *Hydra* genome, 290 putative NBD loci were predicted falling into two large groups: 130 of the NACHT domain type and 160 of the NB-ARC domain type (figs. 1 and 2; supplementary fig. S1, Supplementary Material online). One hundred and seventy fragments of these NLRs were cloned from cDNA that were assembled to 19 contigs and 30 singletons indicating that at least 49 genes are transcribed. The in silico findings are supported by the Southern blot hybridization using a probe for the NACHT domain resulting in a complex band pattern (supplementary fig. S4, Supplementary Material online).

A phylogenetic analysis using Bayesian inference with the amino acid sequences of selected NACHT domain representatives (fig. 3 and supplementary fig. S2, Supplementary Material online) revealed that frequently throughout the

← NACHT domains are located within one group with vertebrate sequences; gray box indicates a cluster of vertebrate NOD1 and NOD2 orthologues with similar cnidarian NACHT domains; as outgroup, the NB-ARC domain of RPS4 (*A. rabidopsis thaliana*) was used. Abbreviations: Amil (*Acropora millepora*), Atha (*A. thaliana*), Nvec (*Nematostella vectensis*), Spur (*Strongylocentrotus purpuratus*), Bflo (*B. ranchiostoma floridae*), Aque (*A. mphimedon queenslandica*), Xtro (*X. enopus tropicalis*), Olat (*O. ryzias latipes*), Hmag (*Hydra magnipapillata*), Hsap (*Homo sapiens*), and Ggal (*Gallus gallus*). HyNLR type 1, which was used for further analysis in detail, is highlighted in red. The nonbilaterian sequences are highlighted in blue.



entire animal kingdom vast species-specific NLR gene expansions occurred. This can be observed for *A. queenslandica*, *H. magnipapillata*, *S. purpuratus* and partially for *N. vectensis*, *B. floridae*, *X. tropicalis* and for the human NLRP genes. A more comprehensive tree including other NACHT domain containing proteins from fungi and humans (HET-E and TEP1), and NACHT-LRR proteins from *D. pulex* and two mosquito species, *Aedes aegypti* and *Culex quinquefasciatus*, were built using Bayesian inference analysis (supplementary fig. S3, Supplementary Material online). Interestingly, the NACHT domain of the fungal HET-E protein does not form a clear outgroup in the phylogenetic analysis, which may result from high diversity of the investigated NACHT domains or even may suggest lateral gene transfer. Whereas almost all human NLR genes and their vertebrate orthologues appear to be vertebrate-specific, NOD1, and NOD2 form a cluster not only with vertebrates but also with two cnidarian representatives indicating a putative phylogenetically old origin for NOD1 and NOD2, at least for their NACHT domains. Remarkably, all *Hydra* NACHT domains are located within one group of exclusively vertebrate sequences, indicating a relatively high sequence similarity in between these groups.

An alignment with representative NACHT domains that were used to build the phylogenetic tree (fig. 3) shows that the various NACHT domain motifs, like the P-loop and the Mg<sup>2+</sup>-binding site are conserved within all investigated animal groups, although the different NACHT sequences are very heterogeneous (supplementary fig. S2, Supplementary Material online). A phylogenetic analysis using only the conserved motifs was calculated according to Koonin and Aravind (2000) but did not result in an informative tree (data not shown).

#### Expression Profile of *HyNLR Type 1*, a *Hydra*-NLR Representative Gene

In contrast to anthozoans, *Hydra* lacks a classical NLR but contains at least 16 proteins with an N-terminal DEATH domain followed by a NACHT domain (fig. 1 and supplementary fig. S1, Supplementary Material online). To provide insights into possible function of the *Hydra* NLR-like proteins, the transcript architecture and expression pattern of one of these *Hydra* NLR-type genes, *HyNLR type 1*, was selected and analyzed in detail (fig. 4 and in red in fig. 3). The *HyNLR type 1* gene consists of five exons that are transcribed to mRNA of about 2,417 bases containing the

2,067 bp ORF (fig. 4A). After transcription the original 5' untranslated region of the mRNA is replaced by the *Hydra*-specific transspliced leader B sequence (Stover and Steele 2001; fig. 4A). Blast searches with the complete 688 aa predicted *HyNLR type 1* protein against the mammalian protein databases result in NLRPs as best hits, resulting from similarity in DEATH and NACHT domains. However, the C-terminal part of the *Hydra* protein gave no significant matches with the National Center for Biotechnology Information NR database, implying that this part of the protein is unique.

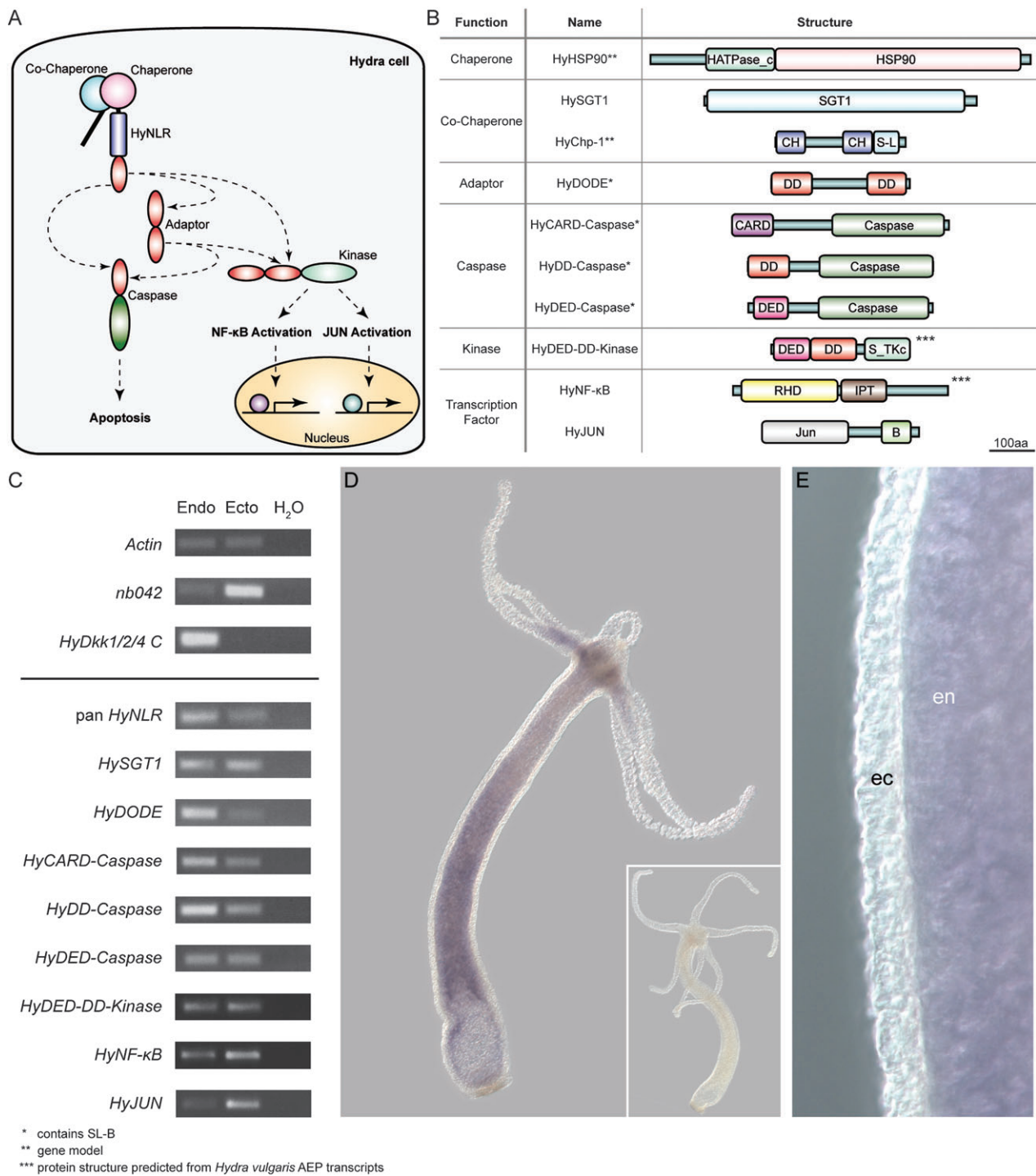
As assessed by procaine separation experiments and in situ hybridization, the gene is predominantly expressed in the endodermal layer (fig. 4B–E). We were able to clone several isoforms of the transcript indicating abundant alternative splicing (supplementary fig. S5, Supplementary Material online). To address the question whether transcript levels are altered upon stimulation with MAMPs, we incubated polyps in LPS or flagellin and furthermore injected these MAMPs into the gastric cavity of these animals. After 4 h, we isolated RNA for RT PCR analysis. As negative control, we used animals with injected culture-medium and untreated animals. As shown in figure 4F, we detected a slight upregulation of the transcript level of *HyNLR type 1* indicating a potential regulation in response to bacterial stimuli in vivo.

#### The Putative Interactome of *HyNLR Type 1* and Other *HyNLRs*

Upon activation NLRs recruit several supporting proteins that are necessary to form large molecular complexes, for example, the inflammasome comprised of NLRP1 or NLRP3 monomers, apoptosis-associated speck-like protein containing a CARD (ASC; Masumoto et al. 1999) and caspase 1 (Cerretti et al. 1992; Thornberry et al. 1992) or the NODosome formed by NOD1 or NOD2 in association with the serine–threonine kinase RIPK2. Additionally, chaperones and cochaperones are required to foster and modulate the functional state of the complexes. To elucidate the conservation of the NLR-associated molecular pathways, we screened the *H. magnipapillata* genome for genes whose protein products may interact with the *HyNLR type 1* protein in an ancient NLR signaling cascade (fig. 5A and B). Three orthologues of the NLR-chaperone HSP90 (Mayor et al. 2007), and orthologues of the NLR cochaperones

←

**Fig. 4.** *HyNLR type 1*. (A) The gene structure of *HyNLR type 1*. The gene consists of five exons of various sizes and encodes a transcript of about 2,417 bases. At the 5' end the transspliced leader B is added. The 688 aa long protein contains a predicted DD and a NACHT domain but lacks LRRs. (B) The in situ hybridization shows an expression in the whole endoderm of *Hydra* along the whole-body axis. (C) Enlargement of the in situ hybridization showing the endodermal expression. (D) A procaine-treated animal showing the clear separation of ectodermal (white) and endodermal (brown) tissue. The colors depict the natural pigmentation of the animal. (E) Semiquantitative reverse transcriptase PCR performed on RNA extracted from the ectodermal and the endodermal cell layers of *Hydra* after procaine treatment showing a predominant expression of *HyNLR type 1* in the endodermal layer. As control genes the housekeeping gene *actin*, the ectodermally expressed nematoblast gene *nb042* and the endodermally expressed gland cell gene *HyDkk1/2/4 C* are used. (F) Quantitative RT PCR showing a slightly upregulation of *HyNLR type 1* after stimulation with LPS and Flagellin. (-) denotes untreated animals, in the sham-treated animals an injection wound was performed, but the animals were not exposed to the bacterial stimuli. For details, please refer to the Materials and Methods section. The fold change of the expression of *HyNLR type 1* is compared with the expression levels of the *actin* transcript. Abbreviations: en/endo (endoderm), ec/ecto (ectoderm).



**FIG. 5.** The putative interactome of HyNLRs. (A) A putative model how the signal transduction of an NLR in *Hydra magnipapillata* could be performed. The NLR could either activate a caspase or a kinase directly or indirectly via an adaptor protein. The caspase could perform apoptosis and the kinase could activate transcription factors like NF-κB or JUN that may regulate the expression of antimicrobial peptides. (B) Table of the putative interactome of HyNLRs. We cloned the complete orfs for *HySGT1*, *HyDODE*, *HyCARD-Caspase*, *HyDD-Caspase*, *HyDED-Caspase*, and *HyJUN*. Abbreviations: DD, DEATH domain; DED, DEATH effector domain; CARD, Caspase recruitment domain; CH, CHORD; S-L, SGT1-like domain; S\_TKc, serine/threonine protein kinase; RHD, rel homology domain; IPT, Ig like, plexins, transcription factors; and B, basic region leucine zipper; The *HyDD-Caspase*, the *HyDED-Caspase*, and *HyDODE* transcripts contain the transspliced leader B, the protein structure of HyNF-κB and HyDED-DD-Kinase are derived from transcripts from *H. vulgaris* AEP. (C) Semiquantitative reverse transcriptase PCR performed on RNA extracted from the ectodermal and the endodermal cell layers of *Hydra* after procaine treatment. As control genes the housekeeping gene *actin*, the ectodermally expressed nematoblast gene *nb042* and the endodermally expressed gland cell gene *HyDkk1/2/4 C* are used. The *HyDD-Caspase*, *HyDODE*, and *HyCARD-Caspase* are predominantly expressed in the endodermal layer, *HySGT1*, the *HyDED-Caspase* and the *HyDED-DD-Kinase* are equally expressed, *HyNF-κB* and *HyJUN* are predominantly expressed in the ectodermal layer. (D) In situ hybridization performed for the *HyDD-Caspase* showing an endodermal expression in the whole-body column. (E) Enlargement of the in situ hybridization of the *HyDD-Caspase* showing the endodermal expression. Abbreviations: ec, ectoderm; en, endoderm.

Chp-1 (Hahn 2005) and SGT1 (Mayor et al. 2007) were identified and the full ORF of one of the cochaperones, *HySGT1*, was cloned. To identify the initiators of downstream signals elicited by NLR activation, we screened the database for genes whose protein products contain an effector domain of the DEATH-fold domain family for homotypic interactions with HyNLRs. We identified and cloned a gene encoding a protein consisting of two DEATH domains (fig. 5B). This Double-DEATH protein HyDODE could function as an adaptor protein-like ASC in the human NLRP1 and NLRP3 inflammasomes and could serve as a bridge for the interaction of *Hydra* NLRs with further effector proteins like caspases or kinases.

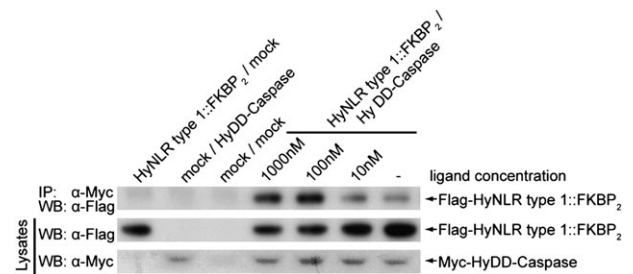
We further identified the complete coding sequence for a caspase containing a CARD at the N-terminus and experimentally verified the *H. magnipapillata* orthologues of two suggested transcripts encoding *Hydra* caspases that contain an N-terminal DED (HyDED-Caspase) or a DEATH domain (HyDD-Caspase) (Böttger and Alexandrova 2007; fig. 5B). HyNLRs could form larger complexes with these caspases directly or indirectly via homotypic complexation of DEATH-like domains.

In order to search putative kinases that could link HyNLRs to MAPK or NF- $\kappa$ B signaling pathways, we screened for kinase domains linked to DEATH, CARD, or DED domains. Although we could not identify genes that were orthologous to RIPK1 or RIPK2, several genes were identified that encode proteins with a DED at the N-terminus followed by a DEATH domain and a protein kinase domain (fig. 5B). The sequence of a single representative, *HyDED-DD-Kinase 1*, was experimentally verified by full-length cloning. This protein structure seems to be a taxonomically restricted gene in *Hydra* and is not orthologous to RIPK1 or RIPK2 but may link HyNLRs to a putative NF- $\kappa$ B or JNK pathway.

Using the sea urchin NF- $\kappa$ B as template, we could identify a derived NF- $\kappa$ B transcript in *Hydra*. The translated ORF contains a conserved Rel homology domain and an immunoglobulin-like fold (fig. 5B). It seems that the encoded protein is shorter than other NF- $\kappa$ B proteins like the previously published NF- $\kappa$ B orthologue of the sea anemone *N. vectensis* (Sullivan et al. 2007).

Furthermore, we detected an orthologue for the JUN transcription factor that may also interact with HyNLR signaling (Girardin et al. 2001).

In order to determine the sites of expression along the body column, the transcript levels of *HySGT1*, *HyDD-Caspase*, *HyDED-Caspase*, *HyDODE*, the *HyDED-DD-Kinase*, *HyNF- $\kappa$ B*, and *HyJUN* were assessed via RT-PCR after separating the two tissue layers of *Hydra*, ectoderm and endoderm, with procaine (fig. 4D). Primers derived from conserved regions of the NACHT domains of HyNLRs were employed in order to amplify a large proportion of NACHT domain coding parts of NLR-transcripts (pan HyNLR). Whereas *HySGT1* and the *HyDED-Caspase* show a uniform expression level in both ecto- and endoderm, HyNLRs, the *HyDD-Caspase* and *HyDODE*, are predominantly expressed in the endodermal layer (fig. 5C). This pattern is supported



**Fig. 6.** Coimmunoprecipitation of the HyNLR type 1 chimeric protein and the HyDD-Caspase. Line 1: Immunoprecipitation was performed using an  $\alpha$ -Myc antibody to catch the HyDD-Caspase; Western blot was performed using an  $\alpha$ -Flag antibody to detect the coprecipitated HyNLR type 1::FKBP<sub>2</sub> fusion protein. The amount of the precipitated HyNLR type 1::FKBP<sub>2</sub> chimera increases with the ligand concentration. Line 2: Expression control for the Flag-tagged HyNLR type 1::FKBP<sub>2</sub> chimeric protein in HEK293 cells; Line 3: Expression control for the Myc-tagged HyDD-Caspase.

for the *HyDD-Caspase* by in situ hybridization (fig. 5D and E). The *HyDED-DD-Kinase* appears to be similarly expressed in the endodermal and in the ectodermal tissue, whereas *HyNF- $\kappa$ B* and *HyJUN* are predominantly expressed in the ectodermal layer but also albeit at a lower level within the endoderm of *Hydra*.

### HyNLR Type 1 and the HyDD-Caspase Interact In Vitro

To investigate whether the HyNLR type 1 protein may directly interact with the HyDD-Caspase, a heterologous expression system was used. Human HEK293 cells were cotransfected with expression-constructs for Flag-tagged chimeric HyNLR type 1 in which the N-terminal DEATH and the NACHT domain were fused to two domains of FKBP, which can be oligomerized by the cell-permeable ligand AP20187 and Myc-tagged HyDD-Caspase. The method was already successfully used to investigate induced proximity signaling of vertebrate NLRs (Inohara et al. 2000; Ogura et al. 2001). Immunoblotting analysis showed that the chimeric HyNLR type 1::FKBP<sub>2</sub> and the HyDD-Caspase constructs were expressed when transfected in HEK293 cells (fig. 6). Remarkably, we were able to coprecipitate the HyNLR type 1 protein with the HyDD-Caspase upon forced oligomerization in a ligand concentration-dependent manner. Addition of 100 nM AP20187 led to a robust coprecipitation of both overexpressed proteins, whereas in the absence of the ligand only a weak interaction was detectable. Our results give first hints that the HyDD-Caspase could be an endogenous interaction partner for the HyNLR type 1 protein and may be involved in the formation of an inflammasome-like HyNLR protein complex.

### Discussion

The salient finding of the present study is that complex NLR repertoires evolved very early in animal evolution. NACHT/NB-ARC protein diversity is high in all three cnidarians examined, and detailed analysis of trace archive

data shows the presence of NLRs also in the sponge genome (Srivastava et al. 2010). Our data suggest that NLRs are ancient genes with a putative immune function in basal metazoans. It is tempting to speculate that the common metazoan ancestor might have used NACHT and putatively also NB-ARC domain proteins for cytoplasmic defense as well.

Finding classical NLRs in anthozoan cnidarians directly contradicts the assumption that NLRs evolved at the level of the teleost fishes (Ting and Davis 2005) and led us to reinvestigate their apparent absence from ecdysozoans. Two genes encoding proteins with a NACHT domain followed by LRRs were identified in the freshwater crustacean *D. pulex* and each of two mosquito species, *A. aegypti* and *C. quinquefasciatus* (XP\_001658101, XP\_00184881), suggesting that classical NLRs may also be present in ecdysozoans. This finding warrants a further in-depth analysis, especially with regard to the N-terminal adaptor domains once more complete sequences also from other ecdysozoan genomes become available. The absence of NLRs from *Drosophila* and *Caenorhabditis* is consistent with particularly high levels of gene loss in these model ecdysozoans (Miller et al. 2005).

In a previous publication (Zmasek et al. 2007), four gene models encoding NB-ARC domains were reported. Our analyses appear to support two of these models. As the complete data set of the earlier study is not publically available, we could not confirm the other suggested models, including a putative gene model containing a reverse transcriptase domain, which would be unique in the entire animal kingdom.

Whereas classical NLRs were detected in anthozoan cnidarians, no proteins with this tripartite architecture could be identified in *Hydra*. Note, however, that the presence of classical NLRs in *Hydra* cannot yet be rigorously excluded due to the limited size of the genomic contigs in the current *H. magnipapillata* genome assembly (Chapman et al. 2010). The Anthozoa are regarded as the most basal cnidarian class, whereas it is thought that hydrozoans diverged more recently within the cnidarians, indicating that the NACHT-LRR structure is ancient and that the LRR region might have been secondarily lost. A number of NLR-like proteins were detected in *Hydra*, each with an N-terminal DEATH and central NACHT domains. The relatively long C-terminal regions of these proteins are devoid of known domains and give no significant matches to known proteins. There are at least two possible scenarios under which these NLR-like proteins could function as a receptor: 1) via heterotypic interactions with discrete LRR proteins as has been demonstrated in the case of the TLR system of *Hydra* (Bosch et al. 2009) or 2) *Hydra*-specific PAMP recognition domains that are as yet unidentified may be present in the C-terminal regions of these NLR-like proteins.

Although a tripartite NLR sensu stricto is missing in *Hydra*, we detected and cloned another type of NBD-containing molecule that contains other repetitive domains, the TPRs, at the C-terminal end. The NB-ARC-TPR-structure was already bioinformatically predicted in *Branchiostoma* and *Nematostella* (Zmasek et al. 2007) and can also be found

in fungi and bacteria (Leipe et al. 2004), albeit the biological function of this subfamily is completely unclear. Although it may be discussed whether the NB-ARC-TPR proteins of bacteria were acquired by horizontal gene transfer (Budd et al. 2004; Guljamow et al. 2007), the latter finding could serve as another hint for an ancient origin of the NBD-containing receptors.

A major difference of the newly detected invertebrate NLRs compared with their vertebrate orthologues is the presence of a DEATH or DED domain instead of a CARD or PYD as an N-terminal effector-binding domain (supplementary fig. S1, Supplementary Material online). The phenomenon was first detected in the genome of the sea urchin *S. purpuratus* (Hibino et al. 2006; Rast et al. 2006). It is important to note that all the aforementioned domains belong to the DD superfamily that share a similar DEATH-like fold structure and seem to have evolved from a primordial DD superfamily member. Whereas the PYD domain of NLRPs seems to be vertebrate specific, we show that CARD-domains being usual effector domains in vertebrate NLRs and APAF1 are present in the basal orthologues of APAF1 but not in NLRs of basal metazoans. Nevertheless, our results indicate that both CARD and DEATH/DED domains were already present within NBD domain containing proteins at the basis of animal evolution.

The phylogenetic analysis shows that in *Hydra* as well as in many other animal species NLR gene expansions occurred. With regard to a possible function in immunity, this may indicate species-specific adaptations to a variety of ecological niches. Interestingly, in these analyses, we observed weakly supported clustering of sequences from distant species (e.g., *Hydra* and *Xenopus*). Whether this is due to the high diversity and numbers of NACHT sequences or due to convergent evolution remains an interesting question for future investigation.

Linking these findings with the prior observation of the different N- and C-terminal domains, it is important to note that the suggested common role of invertebrate and vertebrate NLRs in immune responses could be the result of homology of the entire gene, but it may also result from a homology of the NACHT domain and a secondary convergent evolution of the domain architecture (e.g., by domain shuffling).

Alternative mRNA splicing is known to be involved in regulation of NLR function. Extensive splicing was observed in plant *R* gene transcripts (Gassmann 2008), where ratios of certain splice isoforms change drastically upon pathogen invasion and the homotypic interaction of the encoded proteins seems to be obligatory for an effective immune responses (Dinesh-Kumar and Baker 2000; Dinesh-Kumar et al. 2000). In addition, alternative splicing was discovered in the human NOD2 transcript leading to a shorter transcript isoform that encodes a protein consisting of a single CARD domain that is able to negatively regulate NOD2 signaling (Rosenstiel et al. 2006). In addition to the full-length *HyNLR type 1* transcript corresponding to the 688 aa protein, we were able to detect three alternatively spliced isoforms (supplementary fig.

S5, Supplementary Material online) that could be translated to truncated proteins with a similar modulatory role for NLR function. Although the direct targets of splicing in the described cases are different, our observation provides evidence that splicing might be an important regulatory mechanism in NLR-mediated protein function in basal metazoans as well.

The original 5' end of the *HyNLR type 1* transcript is replaced by the transspliced leader B published by Stover and Steele (2001). Although the phenomenon of transsplicing is common in hydrozoans, the exact function of transspliced leader sequences is not completely understood (Stover and Steele 2001; Derelle et al. 2010), yet it has been proposed that the small RNA pieces added to the 5' end may be involved in translational regulation. The *HyNLR type 1* transcript contains two splice acceptor sites at the 5' end and both are used for the addition of the transspliced leader resulting in two different isoforms of the transcript differing in the length of the 5' UTR. Interestingly, in addition to *HyNLR type 1*, the transcripts encoding *HyDODE*, the *HyDD-Caspase* and the *HyDED-Caspase*, also contain the transspliced leader B that could be a hint for a translational coregulation.

All tested candidate genes are expressed in the endodermal epithelium of *Hydra*. We have shown previously that the endoderm is a remarkable potent chemical barrier against pathogenic invaders, for example, by orchestrated expression of antimicrobial peptides (AMPs) (Augustin, Anton-Erxleben, et al. 2009; Augustin, Siebert, et al. 2009; Bosch et al. 2009) and it is tempting to speculate that it is also a place for host–beneficial microbe interaction. Our data indicate that in addition to these AMPs, the endodermal epithelium expresses a large amount of NLRs (fig. 5C) that may serve as cytosolic MAMP sensors.

Because all tested candidates are coexpressed in the same cell layer, a first criterion for a putative interaction of their gene products in vivo is fulfilled. We would like to propose two different signaling cascades (fig. 5A). In a first scenario, the *HyNLR type 1* protein could interact with a caspase, for example, the *HyDD-Caspase*, either directly via homophilic binding of the two effector domains or via an indirect binding that is mediated by the *HyDODE* protein that could act as an adaptor analogous to the human NLRP1 or NLRP3 inflammasomes that lead to the activation of the IL 1 $\beta$  converting caspase 1. The situation schematically shown in figure 5A depicts a putative protein complex that resembles the human inflammasome in structure and putative function. We were able to support this scenario by a coimmunoprecipitation experiment in HEK293 cells leading to a linker-dependent coprecipitation of the *HyNLR type 1* chimera and the *HyDD-Caspase*. Thus, this protein interaction could lead to apoptosis as an ancient mechanism of immune defense in *Hydra*. Figure 5A also shows that an alternative pathway could employ a kinase. In the genomic scaffolds available for the *Hydra* genome, we have identified several kinases that harbor a DED/DD domain

within its predicted reading frame. Although these kinases have a reversed domain order and are not homologous to RIPK2 that interacts with the human NOD2 protein, one could speculate that the DD domain could mediate a direct or indirect interaction with the *HyNLR* proteins. In the human system, activation of the serine–threonine kinase RIPK2 leads to the activation of NF- $\kappa$ B or to the activation of MAP-Kinase signaling (Girardin et al. 2001). The *Hydra* orthologue of JNK has been discovered by Philipp et al. (2005). In the present paper, we show the presence of a JUN orthologue and a short orthologue for NF- $\kappa$ B (fig. 5). The remaining components for the NF- $\kappa$ B signaling cascade are also present in *Hydra* (Hemmrich et al. 2007; Miller et al. 2007). Thus, we hypothesize that both the JNK/AP-1 system and NF- $\kappa$ B transcription factors could be activated by the *HyDED-DD-Kinase*, which could result in the induction of distinct cellular responses downstream of *HyNLRs*.

Although the primary sequences and the domain compositions differ from *Hydra* to other animals like *Homo sapiens* or to plants, NLRs and putative interaction partners are already present in *Hydra* and therefore phylogenetically older than previously thought. Despite these findings and the fact that bona fide NACHT domains in vertebrates and deuterostomes so far have only been identified in NLR genes with a role in immunity, it is important to note that rigorous functional studies are required to assess the exact role of the basal metazoan NLRs. Building comparative single nucleotide polymorphism maps of NLRs in different basal species subjected to high selective pressure and short generation times will help to systematically elucidate the extent of sequence variability or conservation of genes, respectively. This will lead to a better understanding of the interplay between barrier and environment and contribute to unravel the evolutionary pressures that have shaped genetic diversity profiles of a gene family relevant for cytoplasmic immune defense across the entire animal kingdom.

## Supplementary Material

Supplementary table S1 and figures S1–S5 are available at *Molecular Biology and Evolution* online (<http://www.mbe.oxfordjournals.org/>).

## Acknowledgments

We thank Angelika Böttger for providing the DD-Caspase and the DED-Caspase sequences and Jonathan Rast for providing the gene models from *S. purpuratus* and helpful discussion. We would like to acknowledge Stefan Siebert for his work on the *HyJUN* transcript and René Augustin for beneficial discussions. We thank Robert Häsler for designing TaqMan probes and oligos. Tanja Kaacksteen, Melanie Schlapkohl, and Dorina Oelsner are gratefully acknowledged for expert technical assistance T.C.G.B. and P.R. are supported by the DFG Clusters of Excellence Future Ocean and Inflammation at Interfaces.

## References

- Agostini L, Martinon F, Burns K, McDermott MF, Hawkins PN, Tschopp J. 2004. NALP3 forms an IL-1 $\beta$ -processing inflammasome with increased activity in Muckle-Wells autoinflammatory disorder. *Immunity* 20:319–325.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol.* 215:403–410.
- Augustin R, Anton-Erxleben F, Jungnickel S, Hemmrich G, Spudy B, Podschun R, Bosch TCG. 2009. Activity of the novel peptide arminin against multiresistant human pathogens shows the considerable potential of phylogenetically ancient organisms as drug sources. *Antimicrob Agents Chemother.* 53:5245–5250.
- Augustin R, Franke A, Khalturin K, Kiko R, Siebert S, Hemmrich G, Bosch TCG. 2006. Dickkopf related genes are components of the positional value gradient in Hydra. *Dev Biol.* 296:62–70.
- Augustin R, Fraune S, Bosch TCG. 2010. How Hydra senses and destroys microbes. *Semin Immunol.* 22:54–58.
- Augustin R, Siebert S, Bosch TCG. 2009. Identification of a kazal-type serine protease inhibitor with potent anti-staphylococcal activity as part of Hydra's innate immune system. *Dev Comp Immunol.* 33:830–837.
- Ausubel FM, Katagiri F, Mindrinos M, Glazebrook J. 1995. Use of Arabidopsis thaliana defense-related mutants to dissect the plant response to pathogens. *Proc Natl Acad Sci U S A.* 92:4189–4196.
- Bai J, Pennill LA, Ning J, et al. (11 co-authors). 2002. Diversity in nucleotide binding site-leucine-rich repeat genes in cereals. *Genome Res.* 12:1871–1884.
- Bode HR, Heimfeld S, Chow MA, Huang LW. 1987. Gland cells arise by differentiation from interstitial cells in Hydra attenuata. *Dev Biol.* 122:577–585.
- Boldin MP, Varfolomeev EE, Pancer Z, Mett IL, Camonis JH, Wallach D. 1995. A novel protein that interacts with the death domain of Fas/APO1 contains a sequence motif related to the death domain. *J Biol Chem.* 270:7795–7798.
- Bosch TCG, Augustin R, Anton-Erxleben F, et al. (17 co-authors). 2009. Uncovering the evolutionary history of innate immunity: the simple metazoan Hydra uses epithelial cells for host defence. *Dev Comp Immunol.* 33:559–569.
- Böttger A, Alexandrova O. 2007. Programmed cell death in Hydra. *Semin Cancer Biol.* 17:134–146.
- Brakebusch C, Nophar Y, Kemper O, Engelmann H, Wallach D. 1992. Cytoplasmic truncation of the p55 tumour necrosis factor (TNF) receptor abolishes signalling, but not induced shedding of the receptor. *EMBO J.* 11:943–950.
- Budd A, Blandin S, Levashina EA, Gibson TJ. 2004. Bacterial alpha2-macroglobulins: colonization factors acquired by horizontal gene transfer from the metazoan genome? *Genome Biol.* 5:R38.
- Cecconi F, Alvarez-Bolado G, Meyer BI, Roth KA, Gruss P. 1998. Apaf1 (CED-4 homolog) regulates programmed cell death in mammalian development. *Cell* 94:727–737.
- Cerretti DP, Kozlosky CJ, Mosley B, et al. (12 co-authors). 1992. Molecular cloning of the interleukin-1 beta converting enzyme. *Science* 256:97–100.
- Chapman JA, Kirkness EF, Simakov O, et al. (74 co-authors). 2010. The dynamic genome of Hydra. *Nature* 464:592–596.
- Chinnaiyan AM, O'Rourke K, Tewari M, Dixit VM. 1995. FADD, a novel death domain-containing protein, interacts with the death domain of Fas and initiates apoptosis. *Cell* 81:505–512.
- Chisholm ST, Coaker G, Day B, Staskawicz BJ. 2006. Host-microbe interactions: shaping the evolution of the plant immune response. *Cell* 124:803–814.
- Derelle R, Momose T, Manuel M, Wincker CP, Houlston E. 2010. Convergent origins and rapid evolution of spliced leader translicing in metazoa: insights from the ctenophora and hydrozoa. *RNA* 16:696–707.
- Dinesh-Kumar SP, Baker BJ. 2000. Alternatively spliced N resistance gene transcripts: their possible role in tobacco mosaic virus resistance. *Proc Natl Acad Sci U S A.* 97:1908–1913.
- Dinesh-Kumar SP, Tham WH, Baker BJ. 2000. Structure-function analysis of the tobacco mosaic virus resistance gene N. *Proc Natl Acad Sci U S A.* 97:14789–14794.
- Epp LG, Tardent P, Bänninger R. 1979. Isolation and observation of tissue layers in Hydra attenuata Pall (Cnidaria, Hydrozoa). *Trans Am Microscop Soc.* 98:392–400.
- Fedders H, Augustin R, Bosch TCG. 2004. A Dickkopf-3-related gene is expressed in differentiating nematocytes in the basal metazoan Hydra. *Dev Genes Evol.* 214:72–80.
- Gassmann W. 2008. Alternative splicing in plant defense. *Curr Topics Microbiol Immunol.* 326:219–233.
- Gierer A, Berking S, Bode H, David CN, Flick K, Hansmann G, Schaller H, Trenkner E. 1972. Regeneration of hydra from reaggregated cells. *Nature* 239:98–101.
- Girardin SE, Tournebize R, Mavris M, et al. (11 co-authors). 2001. CARD4/Nod1 mediates NF- $\kappa$ B and JNK activation by invasive Shigella flexneri. *EMBO Rep.* 2:736–742.
- Grens A, Gee L, Fisher DA, Bode HR. 1996. CnNK-2, an NK-2 homeobox gene, has a role in patterning the basal end of the axis in hydra. *Dev Biol.* 180:473–488.
- Guljamow A, Jenke-Kodama H, Saumweber H, Quillardet P, Frangeul L, Castets AM, Bouchier C, Tandeau de Marsac N, Dittmann E. 2007. Horizontal gene transfer of two cytoskeletal elements from a eukaryote to a cyanobacterium. *Curr Biol.* 17:R757–R759.
- Hahn J-S. 2005. Regulation of Nod1 by Hsp90 chaperone complex. *FEBS Lett.* 579:4513–4519.
- Hall T. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser.* 41:95–98.
- Hemmrich G, Bosch TCG. 2008. Compagen, a comparative genomics platform for early branching metazoan animals, reveals early origins of genes regulating stem-cell differentiation. *BioEssays* 30:1010–1018.
- Hemmrich G, Miller DJ, Bosch TCG. 2007. The evolution of immunity: a low-life perspective. *Trends Immunol.* 28:449–454.
- Hibino T, Loza-Coll M, Messier C, et al. (16 co-authors). 2006. The immune gene repertoire encoded in the purple sea urchin genome. *Dev Biol.* 300:349–365.
- Huang S, Yuan S, Guo L, et al. (17 co-authors). 2008. Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. *Genome Res.* 18:1112–1126.
- Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–755.
- Hulbert SH, Webb CA, Smith SM, Sun Q. 2001. Resistance gene complexes: evolution and utilization. *Annu Rev Phytopathol.* 39:285–312.
- Inohara N, Koseki T, Lin J, del Peso L, Lucas PC, Chen FF, Ogura Y, Núñez G. 2000. An induced proximity model for NF- $\kappa$ B activation in the Nod1/RICK and RIP signaling pathways. *J Biol Chem.* 275:27823–27831.
- Itoh N, Nagata S. 1993. A novel protein domain required for apoptosis. Mutational analysis of human Fas antigen. *J Biol Chem.* 268:10932–10937.
- Kanneganti T-D, Lamkanfi M, Núñez G. 2007. Intracellular NOD-like receptors in host defense and disease. *Immunity* 27:549–559.
- Koonin EV, Aravind L. 2000. The NACHT family—a new group of predicted NTPases implicated in apoptosis and MHC transcription activation. *Trends Biochem Sci.* 25:223–224.

- Leipe DD, Koonin EV, Aravind L. 2004. STAND, a class of P-loop NTPases including animal and plant regulators of programmed cell death: multiple, complex domain architectures, unusual phyletic patterns, and evolution by horizontal gene transfer. *J Mol Biol.* 343:1–28.
- Lemaitre B, Hoffmann J. 2007. The host defense of *Drosophila melanogaster*. *Annu Rev Immunol.* 25:697–743.
- Letunic I, Copley RR, Pils B, Pinkert S, Schultz J, Bork P. 2006. SMART 5: domains in the context of genomes and networks. *Nucleic Acids Res.* 34:D257–D260.
- Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>(Delta Delta C(T)) method. *Methods.* 25:402–408.
- Martinon F, Burns K, Tschopp J. 2002. The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-beta. *Mol Cell.* 10:417–426.
- Masumoto J, Taniguchi S, Ayukawa K, Sarvotham H, Kishino T, Niikawa N, Hidaka E, Katsuyama T, Higuchi T, Sagara J. 1999. ASC, a novel 22-kDa protein, aggregates during apoptosis of human promyelocytic leukemia HL-60 cells. *J Biol Chem.* 274:33835–33838.
- Mayor A, Martinon F, De Smedt T, Pétrilli V, Tschopp J. 2007. A crucial function of SGT1 and HSP90 in inflammasome activity links mammalian and plant innate immune responses. *Nat Immunol.* 8:497–503.
- McCarthy JV, Ni J, Dixit VM. 1998. RIP2 is a novel NF-kappaB-activating and cell death-inducing kinase. *J Biol Chem.* 273:16968–16975.
- Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW. 2003. Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *Plant Cell.* 15:809–834.
- Milde S, Hemmrich G, Anton-Erxleben F, Khalturin K, Wittlieb J, Bosch TCG. 2009. Characterization of taxonomically restricted genes in a phylum-restricted cell type. *Genome Biol.* 10:R8.
- Miller DJ, Ball EE, Technau U. 2005. Cnidarians and ancestral genetic complexity in the animal kingdom. *Trends Genet.* 21:536–539.
- Miller DJ, Hemmrich G, Ball EE, Hayward DC, Khalturin K, Funayama N, Agata K, Bosch TCG. 2007. The innate immune repertoire in cnidaria—ancestral complexity and stochastic gene loss. *Genome Biol.* 8:R59.
- Nürnberg T, Brunner F, Kemmerling B, Piater L. 2004. Innate immunity in plants and animals: striking similarities and obvious differences. *Immunol Rev.* 198:249–266.
- Ogura Y, Inohara N, Benito A, Chen FF, Yamaoka S, Nunez G. 2001. Nod2, a Nod1/Apaf-1 family member that is restricted to monocytes and activates NF-kappaB. *J Biol Chem.* 276:4812–4818.
- Paoletti M, Saupé SJ. 2009. Fungal incompatibility: evolutionary origin in pathogen defense? *BioEssays* 31:1201–1210.
- Philipp I, Holstein TW, Hobmayer B. 2005. HvJNK, a Hydra member of the c-Jun NH2-terminal kinase gene family, is expressed during nematocyte differentiation. *Gene Expr Patterns.* 5:397–402.
- Putnam NH, Srivastava M, Hellsten U, et al. (19 co-authors). 2007. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* 317:86–94.
- Rambaut A. 2006–2009. FigTree. Tree Figure Drawing Tool, Version 1.2.2. Available from: <http://tree.bio.ed.ac.uk/software/figtree/>.
- Rast JP, Smith LC, Loza-Coll M, Hibino T, Litman GW. 2006. Genomic insights into the immune system of the sea urchin. *Science* 314:952–956.
- Rodriguez A, Oliver H, Zou H, Chen P, Wang X, Abrams JM. 1999. Dark is a *Drosophila* homologue of Apaf-1/CED-4 and functions in an evolutionarily conserved death pathway. *Nature* 1:272–279.
- Rosenstiel P, Huse K, Till A, et al. (12 co-author). 2006. A short isoform of NOD2/CARD15, NOD2-S, is an endogenous inhibitor of NOD2/receptor-interacting protein kinase 2-induced signaling pathways. *Proc Natl Acad Sci U S A.* 103:3280–3285.
- Rosenstiel P, Philipp EE, Schreiber S, Bosch TC. 2009. Evolution and function of innate immune receptors—insights from marine invertebrates. *J Innate Immunity.* 1:291–300.
- Rosenstiel P, Sina C, End C, et al. (17 co-authors). 2007. Regulation of DMBT1 via NOD2 and TLR4 in intestinal epithelial cells modulates bacterial recognition and invasion. *J Immunol.* 178:8203–8211.
- Rosenstiel P, Till A, Schreiber S. 2007. NOD-like receptors and human diseases. *Microbes Infect/Institut Pasteur.* 9:648–657.
- Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol.* 4: 406–425.
- Schreiber S, Rosenstiel P, Albrecht M, Hampe J, Krawczak M. 2005. Genetics of Crohn disease, an archetypal inflammatory barrier disease. *Nat Rev Genet.* 6:376–388.
- Sodergren EG, Weinstock M, Davidson EH, et al. (228 co-authors). 2006. The genome of the sea urchin *Strongylocentrotus purpuratus*. *Science* 314:941–952.
- Sommer DD, Delcher AL, Salzberg SL, Pop M. 2007. Minimus: a fast, lightweight genome assembler. *BMC Bioinformatics.* 8:64.
- Srivastava M, Simakov O, Chapman J, et al. (33 co-authors). 2010. The Amphimedon queenslandica genome and the evolution of animal complexity. *Nature* 466:720–726.
- Stein C, Caccamo M, Laird G, Leptin M. 2007. Conservation and divergence of gene families encoding components of innate immune response systems in zebrafish. *Genome Biol.* 8:R251.
- Stover NA, Steele RE. 2001. Trans-spliced leader addition to mRNAs in a cnidarian. *Proc Natl Acad Sci U S A.* 98:5693–5698.
- Sullivan JC, Kalaitzidis D, Gilmore TD, Finnerty JR. 2007. Rel homology domain-containing transcription factors in the cnidarian *Nematostella vectensis*. *Dev Evol.* 217:63–72.
- Sullivan JC, Ryan JF, Watson JA, Webb J, Mullikin JC, Rokhsar D, Finnerty JR. 2006. StellaBase: the *Nematostella vectensis* Genomics Database. *Nucleic Acids Res.* 34:D495–D499.
- Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol.* 24:1596–1599.
- Tartaglia LA, Ayres TM, Wong GH, Goeddel DV. 1993. A novel domain within the 55 kd TNF receptor signals cell death. *Cell* 74:845–853.
- Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673–4680.
- Thornberry NA, Bull HG, Calaycay JR, et al. (29 co-authors). 1992. A novel heterodimeric cysteine protease is required for interleukin-1 beta processing in monocytes. *Nature* 356: 768–774.
- Ting JP-Y, Davis BK. 2005. CATERPILLER: a novel gene family important in immunity, cell death, and diseases. *Annu Rev Immunol.* 23:387–414.
- van der Biezen EA, Jones JD. 1998. The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Curr Biol.* 8: R226–R227.
- Wiens M, Korzhnev M, Krasko A, Thakur NL, Perovic-Ottstadt S, Breter HJ, Ushijima H, Diehl-Seifert B, Müller IM, Müller WEG. 2005. Innate immune defense of the sponge *Suberites domuncula* against bacteria involves a MyD88-dependent signaling pathway. Induction of a perforin-like molecule. *J Biol Chem.* 280:27949–27959.

- Wiens M, Korzhev M, Perovic-Ottstadt S, Luthringer B, Brandt D, Klein S, Müller WEG. 2007. Toll-like receptors are part of the innate immune defense system of sponges (demospongiae: Porifera). *Mol Biol Evol.* 24:792–804.
- Yuan JY, Horvitz HR. 1990. The *Caenorhabditis elegans* genes *ced-3* and *ced-4* act cell autonomously to cause programmed cell death. *Dev Biol.* 138:33–41.
- Zmasek CM, Zhang Q, Ye Y, Godzik A. 2007. Surprising complexity of the ancestral apoptosis network. *Genome Biol.* 8: R226.
- Zou H, Henzel WJ, Liu X, Lutschg A, Wang X. 1997. Apaf-1, a human protein homologous to *C. elegans* CED-4, participates in cytochrome *c*-dependent activation of caspase-3. *Cell* 90:405–413.