Analysis of fish-specific NLRs in zebrafish, *Danio rerio*

Inaugural-Dissertation

zur
Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von

Julia Zielinski

Aus Dorsten

Berichterstatter: Prof. Dr. rer. nat. M. Leptin

und

PD. Dr. rer. nat. N. Gehring

Tag der mündlichen Prüfung: 30.01.2012

Abreviations

aa	amino acids
AGM	aorta-gonad-mesonephros
ALM	anterior lateral mesoderm
APAF1	apoptotic protease activating factor 1
ASC	apoptosis-associated speck-like protein containing a CARD
BCR	B cell receptor
BIR	baculoviral inhibitor of apoptosis repeat
bp	base pairs
CARD	caspase activation and recruitment domain
CIITA	major histocompatibility complex, class II, transactivator
cDNA	copy DNA
cfu	colony forming units
dpf	days post fertilization
dpi	days post infection
dsDNA	double stranded DNA
ER	endoplasmic reticulum
EST	expressed sequence tags
Fisna	fish-specific NACHT associated domain
HSC	hematopoietic stem cell
hpf	hours post fertilization
hpi	hours post infection
ICM	intermediate cell mass
IFN	interferon
Ig	immunoglobulin
IL	interleukin
IRG	immunity-related GTPase
LRR	leucine-rich repeat
LPS	lipopolysaccharide
MHC	major histocompatibility complex
mRNA	messenger RNA
n	number of animals
NAIP	neuronal apoptosis inhibitory protein
NALP	NACHT-LRR-PYD-containing protein
NACHT	NTPase domain found in NAIP, CIITA, HET-E and TP1

NB-ARC	nucleotide-binding domain shared by APAF1, certain plant R gene products and nematode CED-4
NFκB	nuclear factor kappa B
NK cell	natural killer cell
NLR	nucleotide-binding, leucine-rich repeat containing family of proteins
NOD	nucleotide-binding oligomerization domain containing protein
ORF	open reading frame
PAMPs	pathogen associated molecular patterns
PGN	peptidoglycan
PLM	posterior lateral mesoderm
PRR	pattern recognition receptor
poly I:C	polyinosinic-polycytidylic acid
RIG 1	retinoic acid-inducible gene 1
RLR	RIG-like receptor
RT	reverse transcriptase
SPRY	splA/ryanodine receptor domain
TCR	T cell receptor
TIR	Toll/IL-1 receptor
TICAM1	Toll-like receptor adaptor molecule 1
TLR	Toll-like receptor
Tn	*Tetraodon negrovidis*
Tr	*Takifugu rubripes*
TNFα	tumor necrosis factor α
VLR	variable lymphocyte receptor
WD40	repeat of ~ 40 aa, often terminating in a Trp-Asp (W-D) dipeptide
wpf	weeks post fertilization

1. Introduction

1.1 Innate and adaptive immune system

The vertebrate innate and adaptive immune systems defend the host against microbe predation and pathogen infection, which would otherwise destroy it. The innate immune system is the first and generic response to invasion by infectious agents (Parish and O'Neill 1997; Kvarnhammar, Petterson et al. 2011). The main components of the innate immune system are its cellular constituents, such as natural killer (NK) cells and phagocytes, and the complement system. On the molecular level the innate immune system is equipped with a limited set of pathogen-recognition receptors (PRRs) that activate mechanisms to minimize the detriment to the infected host (Leber, Crimmins et al. 2008). In contrast to the receptors of the adaptive immune system, PRRs build a genome encoded molecule family of a predictive number.

Activation of the adaptive immune system is a consequence of triggered innate immune mechanisms, such as activated PRRs (Kobayashi, Chamaillard et al. 2005; Fritz, Le Bourhis et al. 2007), an active complement system (Fang, Xu et al. 1998) and the release of cytokines and chemokines. In contrast to the more non-specific innate immunity, the adaptive immune system generates a highly specific response to the invading pathogens, owing to a widespread array of antigen receptor specificities that are obtained by gene rearrangements and receptor diversifications and provides an immunological memory. The cells of the adaptive immune system are the lymphocytes, which mainly comprise T-cells and B-cells (Thomas, Cobb et al. 2009).

1.2 Evolution of the immune systems in vertebrates

The innate immunity is thought to constitute a defense system and is found in all forms of plant and animal life (Janeway 2001). By contrast, defining hallmarks of the adaptive immune system—lymphocytes, antigen receptors, *major histocompatibility complex* (*MHC*) *I* and *II* genes, gene conversion and specialized primary (thymus and/or bone marrow) and secondary lymphoid tissues (lymph nodes and/or lymphoid follicles) (Flajnik and Du Pasquier 2004)—initially ocure in the gnathostome lineage (jawed vertebrates), as they are absent in the agnathans (jawless vertebrates) (Azumi, De Santis et al. 2003). This finding has led to the widespread acceptance that evolution of the adaptive immunity coincided with the emergence of jawed vertebrates. The birth of this system and, therefore, the immunoglobulin (Ig)-based antigen receptor is believed to have occurred when a transposable element containing RAG genes inserted an Ig superfamily gene of the variable type.

Whether this crucial event happened during two rounds of genome-wide duplications, before and after the split of gnathosomes and agnathans (Abi Rached 1999), or by gradual and slower changes over an extended period that started long before the separation of the two lineages (Klein and Nikolaidis 2005) is still an ongoing discussion.

Support for the latter hypothesis has come from studies on the phylogenetic roots of adaptive immunity in lamprey and hagfish, which uncovered a type of variable lymphocyte receptor (VLR). VLRs represent a non Ig-based receptor family, which consists of leucine-rich repeat (LRR) modules that are arranged by somatic recombination in lymphocyte-like cells, providing immunological memory. These LRR modules specifically recognize pathogen associated molecular patterns (PAMPs) (Shintani, Terzic et al. 2000; Pancer, Amemiya et al. 2004; Alder, Rogozin et al. 2005). Although VLRs and B/T-cell receptors are structurally and evolutionary unrelated, they have comparable specificity and affinity. Their independent development strongly supports the beneficial effect of a specific immune system on survival.

Danio rerio, or zebrafish, belongs to the class of ray-finned fish (Teleostei), which are among the earliest vertebrate groups to develop lymphocytes, lymphoid tissue and to express B-cell and T-cell receptors (Danilova, Bussmann et al. 2005; Castro, Bernard et al. 2011). As a model organism the zebrafish can be used to study the immune systems of jawed vertebrates.

1.3 Cellular and molecular components of the innate immune system

In addition to physical barriers, such as skin and mucous membranes, or chemical mechasims, for instance, urine excretion, the innate immune system comprises different cell types and secreted signaling molecules, here referred to as cellular and molecular components of the innate immunity.

Hematopoetic stem cells not only differentiate into lymphoid cells but also give rise to myeloid progenitors that differentiate into two important innate immune cells, macrophages and granulocytes (Akashi, Traver et al. 2000), amongst others. Macrophages are professional phagocytotic cells that are able to ingest and destroy infectious agents (Babior, Curnutte et al. 1976). They recognize PAMPs through specific PRRs (Meylan, Tschopp et al. 2006). Activated macrophages are able to engulf the microbially invaded cell and thereby clear the infection.

Granulocytes are characterized by the presence of granules in their cytoplasm. Upon infection, the granulocytes are activated, degranulate and release histamines, proteoglycans and proteolytic enzymes, which help to clear the infection (Giacomin, Gordon et al. 2008). They are divided into three groups according to the

immunohistochemical staining properties of the granules. Mature granulocytes are termed neutrophils, eosinophils and basophils. Adult zebrafish have at least two types of granulocyte lineages: neutrophils and eosinophils. The presence of basophilic granulocytes or tissue mast cells could not be identified as yet (Lieschke, Oates et al. 2001). Another aspect in which zebrafish immunity components differ from those of mammals is that erythrocytes contain a nucleus (Willett, Cortes et al. 1999).

The complement system is a further component of the innate immune system. It comprises a set of 40 proteins that are found in humans in all body tissues but is primarily expressed by hepathocytes. Once activated, the complement proteins coat invading pathogens, which targets them for phagocytosis by macrophages (Alexander, Anderson et al. 2008).

1.3.1 Cell autonomous innate immunity

Pathogens normally enter the host cell via phagocytosis and autophagy, where they are ultimately digested by degradative autolysosomes (Deretic 2011). However, in many cases, pathogens such as *Shigella flexneri* adapt to the host immune system and manage to escape or inhibit autophagy.

If pathogens managed to enter the cell, PRRs can recognize different motifs of invaded pathogens (Huang, Canadien et al. 2009; Deretic 2011). Similar to specialized phagocytes, some types of non-immune cells, for example, epithelial cells, are able to cope with invading pathogens, although the PRR expression pattern differs from that of immune cells (Neal, Leaphart et al. 2006; Kumar, Kawai et al. 2009; Richardson, Sodhi et al. 2010).

The recognition of specific molecular patterns released from pathogens by PRRs is crucial for the activation of the immune system.

PRRs are divided into membrane-bound receptors, for example, the Toll-like receptors (TLRs), and cytosolic, soluble receptor types, such as RIG-like receptors (RLRs) or NACHT domain and leucine-rich repeat containing receptors (NLRs). TLRs survey the extracellular environment or membrane-enclosed intracellular compartments for signs of microbial presence. NLRs and RLRs recognize the presence of intracellular invading pathogens in the cytosol. PAMP recognition through distinct PRRs results in activating downstream signaling cascades that lead to the expression of cytokines, such as interferons, and chemokines (Dunne 2011). NLRs are the major subject of this study.

1.4 NLR protein structure, oligomerization and signaling

As with most immunity factors, pioneer studies have been done with mammalian organisms, and most of the information results from these studies.

NLRs represent a major family of intracellular immune receptors. As long as no pathogen is sensed, NLRs are kept in the cytosol in an inactive, autoinhibitory state (Yu, Acehan et al. 2005). Most of the NLRs contain C-terminal LRRs that recognize cytosolic pathogen patterns. Although it is known that LRRs are necessary for PAMP recognition (Inohara, Ogura et al. 2001), in vertebrates, no evidence for a direct interaction exists to date. However, the only motif for which a direct interaction between the C-terminal effector domain and its ligand was shown for repeat of ~ 40 aa, often terminating in a Trp-Asp (W-D) dipeptide (WD40) domain of APAF1 and its activator cytochrome c (Hu, Benedict et al. 1999). APAF1 was the pioneer-subject in studying the structure of NLR proteins due to the solution of its three dimensional structure. Independent studies demonstrated that after binding to cytochrome c APAF1 gets activated, a series of steps is initiated that alter the protein conformation and lead to oligomerization in an ATP-dependent manner. After sensing pathogens, NLRs undergo a conformational change and arrange in oligomers, which is obligatory for downstream signaling and clearance of infection. Formation of such a signaling platform is unique to this PRR subtype. These oligomeres are namend dependent on the NLR type: The christalization of an APAF1 oligomer revealed a wheel-like structure with seven fold symmetry (Zou, Henzel et al. 1997; Qin, Srinivasula et al. 1999; Acehan, Jiang et al. 2002; Riedl, Li et al. 2005; Yu, Wang et al. 2006). Active APAF1 oligomers were called apoptosome because they recruit and activate Procaspase-9 and Procaspase-3 and thereby initiate the apoptotic pathway (Reubold, Wohlgemuth et al. 2011). In the case of NALP1, NALP3 and IPAF the oligomers were called inflammasome and in all cases Caspase-1 was recruited and activated as a result (see below) (Mariathasan, Newton et al. 2004). The NODs build another subfamily (NOD1, 2, 3 and 9). The oligomerization platform for NODs is referred to as nodosmome and in the case of NOD1 and 2 it activates the NF□B pathway (Tattoli, Travassos et al. 2007).

Upstream of their C-terminal LRRs, NLRs feature a central nucleotide-binding domain called NACHT domain, the helical domain 1 (HD1), the winged helix domain (WHD) and the super helical (SH) domain. The NACHT domain is evolutionarily highly conserved, as it is closely related to the NB-ARC domain in plant NB-LRR proteins (Bonardi, Tang et al. 2011).

Most of the NLRs contain an N-terminal effector-binding domain, which binds to downstream signaling molecules like receptor interacting protein-2 or caspases that mediate and activate innate immune responses. For further information on NLR architecture and domain function see the sections below.

The insight of the APAF1 structure led to speculations about intermolecular interactions between two proteins within an oligomer. Sequence and structure comparisons

between APAF1 and further members of the conserved NLR family identified putative interaction sites of conserved residues (Proell, Riedl et al. 2008; Danot, Marquenet et al. 2009). Conserved features between APAF1 and other NLRs led to the assumption of similar biochemical behaviours in catalytic activities and were often used as starting hypotheses in studies on NLRs (Proell, Riedl et al. 2008).

1.4.1 LRR-mediated pathogen recognition of NLRs

The LRRs are 20–30 amino acids long and contain a conserved LxxLxLxxN motif. In NLRs they are sequentially positioned in their C-terminal end in various numbers and assemble into a horseshoe-like shape, with the LxxLxLxxN motifs located on the inner concave surfaces (Kajava 1998).

LRRs occur in many proteins where they act as protein interaction domains (Kobe and Kajava 2001). For TLR1 and TLR2, a binding of the TLR heterodimer with its ligand at the concave surface of the LRR domain has been proposed (Jin, Kim et al. 2007). Yet, this assumption cannot simply be transferred to the cytosolic NLRs, since a different cellular localization of the receptors also requires different conditions regarding the proximity of its ligands. Although LRRs are characterized by the conserved motif, NLRs are known to have a broad range of different elicitors. For some members of the NLR family, the ligands have been identified: NOD2 LRRs recognize the peptidoglycan (PGN) subunit muramyl-dipeptide (MDP).

NOD1 is activated upon sensing of L-Ala-γ-D-Glu-meso-diaminopimelic acid (TriDAP) (Laroui, Yan et al. 2011). Members of the inflammasome-NLR subfamily, for example, NALP1 and NALP3, IPAF (Ice-protease activating factor) and NAIP5 (Neuronal apoptosis inhibitory protein 5), are more sensitive to changes in intracellular ion concentration or danger signals than to PAMPs (Kufer 2008).

Furthermore, a set of danger-associated molecular patterns (DAMPs) have been shown to induce NLR-mediated responses, such as ATP, uric acid, UV irradiation or changes of temperature (Mariathasan, Weiss et al. 2006; Feldmeyer, Keller et al. 2007; Levin, Wickliffe et al. 2008; Cerqueira, Boas et al. 2011).

Small truncations of the LRRs in NOD1 and NOD2, which affects the ligand-recognition sites, lead to inactive protein variants (Rosenstiel, Till et al. 2007).

1.4.2 The NACHT domain

NLR stands for “NACHT domain and LRR-containing” proteins, and the presence of the NACHT domain led to the classification of this new emerging protein family. The NACHT domain is central to NLR proteins, not only regarding the position but also their function. In the group of conserved NLRs and the Nalps the domain ranges from 160 to

195 amino acids (aa) and the NACHT domain in fish specific NLRs is 170 aa. Proteins that share this motif belong to the class of nucleoside triphosphatases (NTPases), referred to as the ATP-binding cassette (ABC) proteins. The ABC region in all NLR proteins includes catalytic residues with canonical binding site motifs for phosphate residues (Walker A motif) and magnesium ions (Walker B motif), as well as several motifs predicted to be involved in hydrolysis of nucleotides, ATP, deoxyATP (dATP) and/or GTP (Inohara, Chamaillard et al. 2005).

The role of ATP hydrolysis *per se* in the regulation of NLR protein activity is presently unknown, but it is assumed to be important for the induction of conformational changes required for the return of the NLR protein to the original inactive state. Furthermore, disease-associated sequence variants were pinpointed to residues near NTPase motifs, the Walker A and Walker B motifs (Ye, Lich et al. 2008). It was shown that the cycle of nucleotide binding and hydrolysis is pivotal for NLR-mediated signaling (Albrecht, Domingues et al. 2003).

1.4.3 Effector domains and downstream signalling pathways

The CARD and PYD classify the proteins as death-fold domain superfamily proteins. Some NLRs possess a baculoviral inhibitor of apoptosis repeat (BIR) domain or a transactivator domain.

The current model for the nodosome proposes that the caspase activation and recruitment domain (CARD)-containing NLRs (NOD1 and NOD2) interact with the CARD-containing kinase RIPK2, which leads to activation of the NFκB and MAPK pathways.

By contrast, the PYD-containing NALPs drive caspase activation by binding to the adaptor protein ASC (Apoptosis-associated speck-like protein containing a CARD), which leads to the processing of proinflammatory cytokines or activating inflammasomes (Wagner, Proell et al. 2009). Mammalian NALP1, NALP3 and IPAF recruit and activate caspase-1, either via direct interaction through the CARD domains of IPAF and caspase-1 or with a connecting adaptor protein ASC for the NALPs (Motani, Kushiyama et al. 2011). It is assumed that oligomerized NLRs arrange a scaffold of tightly packed caspases that lead to activation and further processing of downstream signaling pathways. Proinflammatory caspase-1 is a cysteine protease that cleaves inactive cytokines, such as prointerleukin-1β (proIL-1β), to produce active IL-1β, an essential cytokine in mammalian inflammation processes (Thornberry, Bull et al. 1992). Active caspase-1 also cleaves cytokines like IL18, IL17b and IL33, which after they are secreted can bind to their receptors.

The activated TLR/IL receptors initiate NFκB or MAPK pathways, as well as induce additional cytokines like IL6, IFNγ and IL4. These promoted inflammatory respones lead to an activated innate immune system (Cerretti, Kozlosky et al. 1992).
NLR signaling can also cause programmed cell death, namely apoptosis and pyroptosis. For information abot the apoptosome see above.
One way of distinguishing between the two categories is their initiating pathways. Apoptosis is a non-lytic cell death process, which is mediated via the apoptosome. Lytic pyroptosis is characterized and initiated by a Caspase-1 inflammasome, which also involves ASC and absent in melanoma 2 (AIM2) protein (For further information read (Miao, Rajan et al. 2011)).

1.5 The zebrafish immune systems

A zebrafish embryo requires 6 days to develop a functional gastrointestinal system, use up the yolk sac and open its mouth (Kimmel, Ballard et al. 1995). At around 7 days post fertilization, zebrafish larvae start to feed and immediately provide bacteria and viruses with an augmented target, as the pathogens gain multiple routes of access into the fish body and can enter the digestive tract. The cells of the innate immune system are fully developed within the first 6 days and provide the embryo with the first line of defense that its environment and circumstances demand (Bertrand, Kim et al. 2007; Jing and Zon 2011).
In contrast to the early onset of innate immune cell development, cells of the adaptive immune system are not detected in zebrafish until 4 weeks post fertilization (Lam, Chua et al. 2004). As a model organism, the zebrafish, therefore, provides on the one hand an adaptive-immunity-free system during embryonic stages, which enables the exploration of isolated innate immunity; on the other hand, during later developmental stages, it can be used to investigate both the innate and adaptive immune systems of jawed vertebrates.

1.5.1 Development of zebrafish immune cells

Hematopoiesis of the zebrafish produces cell lineages nearly identical to those of mammals or other higher vertebrates (Berman, Kanki et al. 2005); as in all adult vertebrates, erythroid, myeloid and lymphoid cells share a common progenitor (Willett, Cortes et al. 1999) (See figure 1.1). However, a few differences in the hematopoietic systems do exist, the obvious one being that zebrafish lack bone marrow and lymph nodes. In teleosts, the kidney is the equivalent of bone marrow in other vertebrates and is the major site of hematopoiesis (Davidson and Zon 2004). The hematopoietic, pluripotent stem cells originate from the whole kidney marrow and differentiate into

blood cells and immune cells (Song, Sun et al. 2004). As in mammals, hematopoiesis in zebrafish occurs in three consecutive waves (Figure 1.1).

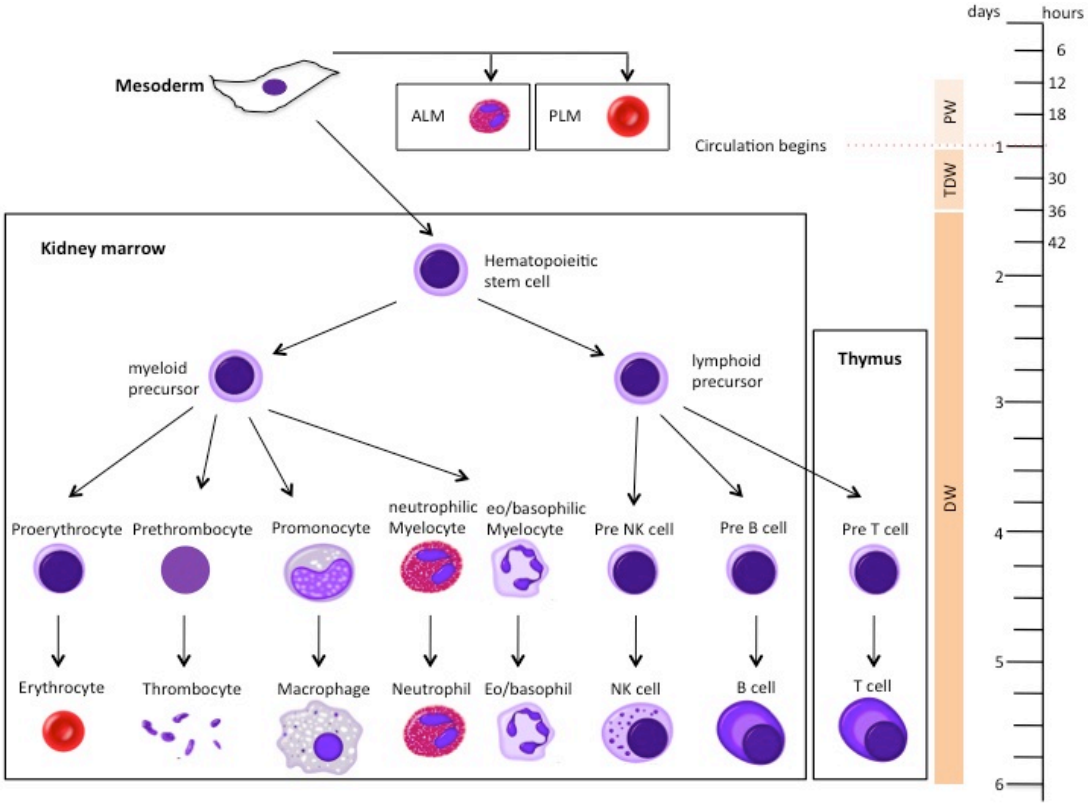

Figure 1.1: Hematopoiesis in zebrafish. Blood and immune cells in zebrafish develop in three consecutive waves. The primitive wave (PW) starts at 11 hours post fertilization (hpf) and gives rise to granulocytes and erythrocytes from the mesoderm tissue in the anterior lateral mesoderm (ALM) and the posterior lateral mesoderm (PLM), respectively. During a transient definitive wave (TDW) multi-lineage erythromyeloid progenitors appear in the posterior blood island. 36hpf, the definitive wave (DW) begins, when the first pluripotent hematopoietic stem cells start to migrate to the kidney, which is the site of myeloid and lymphoid cell development. Lymphoid precursors migrate to the thymus and develop into mature T-cells. Figure summarizes information reviewed by (Jing and Zon 2011).

Embryonic primitive hematopoiesis starts at around 11hpf when hemangioblasts—which have the potential to become either endothelial vascular cells or hematopoietic cells—appear in the ALM and PLM, which collectively are analogous to the blood islands in the mammalian yolk sac. Hemangioblasts in the PLM predominantly give rise to primitive erythrocytes.

The ALM, which later becomes the rostral blood island, is the major site for development of primitive myeloid cells. At around 24hpf, erythrocytes originating from the PLM enter the blood circulation (Jing and Zon 2011).

A transient definitive wave initiates shortly after multi-lineage erythromyeloid progenitors appear in the posterior blood island. Starting from 26hpf, definitive hematopoietic stem cells (HSCs) emerge from hemogenic endothelial cells of the

dorsal aorta in the aorta–gonad–mesonephros (AGM) region. Shortly thereafter, at 48hpf, HSCs migrate to and seed the caudal hematopoietic tissue, which is an expansion of the posterior blood island and acts as a transient hematopoietic site that gives rise to erythroid, myeloid and thromboid cells. The caudal hematopoietic tissue is equivalent to mouse fetal liver or placenta. HSCs from the AGM region colonize the kidney around 48hpf. Kidney marrow, which is functionally similar to the mammalian bone marrow, gives rise to all blood lineages, including erythroid, myeloid, thromboid and lymphoid cells for the larval and adult zebrafish (Davidson and Zon 2004). At around 54hpf, lymphoid progenitor cells from the AGM region seed the thymus, which is the site for maturation of lymphoid T-cells (Paik and Zon 2010).

In the kidney marrow, the myeloid cell lineage gives rise to macrophages and granulocytes, which are both components of the innate immune system, as well as thrombocytes and erythrocytes. Lymphoid precursor cells differentiate into the adaptive T-lymphocytes and B-lymphocytes and also into NK cells. Early T-cell progenitors migrate into the thymus where their differentiation into mature immune-competent T-cells takes place (Davidson and Zon 2004).

1.5.2 Innate immune components in the zebrafish embryo & larva

Zebrafish embryos feature a highly evolved complement system that is maternally derived and, therefore, already present from the beginning of embryonic development (Wang, Zhang et al. 2008). All three complement-activating pathways, the classical pathway, the alternative pathway and the lectin pathway, have been identified in fish. Compared to mammals, additional copies of components have been identified in zebrafish, such as for the C3 convertase, the factor B and the mannose-binding lectin (Gongora, Figueroa et al. 1998; Zarkadis, Mastellos et al. 2001).

Three different receptor types, TLRs, RLRs and NLRs, have mainly been studied in the context of mammalian immunity. These receptors also play a major role in the detection of microbic patterns in zebrafish embryos and are involved in immune-related pathways (Hall, Flores et al. 2009; Oehlers, Flores et al. 2011; Zhang, Sun et al. 2011).

1.5.2.1 TLRs

The TLR family is the best characterized of the three PRR types and is conserved from *Drosophila melanogaster* to mammals. The membrane-spanning receptors feature N-terminal LRRs, a transmembrane domain and a C-terminal Toll/IL-1 receptor (TIR) motif.

Phylogenetic analysis revealed 15 TLRs in *Danio rerio*. Sequence comparisons between zebrafish and human TLRs have uncovered set of seven genes (TLR1, 2, 3,

4b.a, 4b.b, 7, 8a, 8b, 9) that were homologous, but it also revealed a non-mammalian class of TLRs unique to teleostei (TLR5, 14, 19, 20a, 21 and 22) (Palti 2011). In zebrafish a *Mycobacterium* infection led to an upregulation of TLR 1, 2, 9 and 14 mRNA level (Meijer, Krens et al. 2004).

The mammalian TLR4 is a central protein in the receptors' complex, which responds to LPS, resulting in NF□B activation via myeloid-differentiation primary response gene 88 (MyD88) and TIR domain-containing adapter protein (TIRAP) (Hoshino, Takeuchi et al. 1999; Fitzgerald, Rowe et al. 2003; Palti 2011). For TLR4 two homologs were identified in zebrafish, which is one of a few fish-species that express TLR4 at all (Sepulcre, Alcaraz-Perez et al. 2009). TLR4 signaling in fish is fundamentally different: Recent studies showed that zebrafish TLR4 does not recognize LPS and that the TIR domain acts as a negative regulator of MyD88-dependent signaling. Due to this different function another nomenclature was proposed; the TLR4 orthologs in zebrafish were named TLR4b.a and TLR4b.b (Chluba, Jault et al. 2004; Sullivan, Charette et al. 2009; Palti 2011).

For TLR3 mRNA an upregulation upon infection with several gram-negative bacteria was shown (Phelan, Mellon et al. 2005; Bilodeau-Bourgeois, Bosworth et al. 2008). Only one study has provided insight on ligand specificity and cellular localization. (Matsuo, Oshiumi et al. 2008) showed that the fugu TLR22 recognizes dsRNA on the cell surface and upon activation with polyinosine-polycytidylic acid (poly IC) it recruits TICAM-1 and induces IFN expression in fish cells (Matsuo, Oshiumi et al. 2008).

Furthermore, in carp cells, lipoteichoic acid (LTA), peptidoglycan (PGN) from *Stapholococcus aureus* and Pam_3CSK_4 have been identified as ligands for the teleost TLR2 homolog (Ribeiro, Hermsen et al. 2010). Although for the majority of zebrafish TLRs the function and even the activating ligands remain to be elucidated, it is known that TLR expression is abundant already during the first 5 days of embryonic development (Chluba, Jault et al. 2004). Compared to mammals, the TLR diversity in teleosts can probably cover a larger variety of pathogen recognition. In contrast to human TLRs, it is not known which members are expressed at the cell membrane and which are primarily located to the endoplasmic reticulum (ER) or lysosomal-like vesicles.

1.5.2.2 RLRs

RLRs are cytosolic receptors that are important for the RNA-triggered interferon response (Perrot, Deauvieau et al. 2010). Three members belong to this PRR subtype, the retinoic acid-inducible gene 1 (RIG1), melanoma differentiation-associated gene 5

(MDA5) and laboratory of genetics and physiology 2 (LGP2) and all of them are expressed in zebrafish (Zou, Chang et al. 2009).

A C-terminal HELICc domain recognizes RNA of viral origin, the ATP-dependent DExD/H motif unwinds RNA, and in addition, RIG1 and MDA5 share two tandem-arranged CARDs involved in protein–protein interactions.

RIG1- and MDA5-mediated interferon signaling is directed by the mitochondria antiviral signaling protein (MAVS), a CARD-containing protein associated with mitochondria, and negatively regulated by LGP2. LGP2 has been shown to interfere with the binding of RIG1/MDA5 to viral RNAs (Zou, Chang et al. 2009). The laboratory of M. Brémont showed that RIG1 overexpression in a cyprinid cell line led to a strong cellular antiviral response (Biacchesi, LeBerre et al. 2009). Other groups demonstrated an elavated level of MDA5 and LGP2 protein after viral infection or IFN stimulation in rainbow trout (Chang, Collet et al. 2011) and a virus-induced expression of LGP2 mRNA in japanese flounder (Ohtani, Hikima et al. 2010). Although not many studies of RLRs in zebrafish embryos exist, in other teleosts, RLR function has been linked to the same pathways as in mammals.

1.5.2.3 NLRs

Seven NLR proteins were found in the zebrafish genome that phylogenetically cluster with the mammalian NLRs: NOD1, 2, 3, 9, CIITA (class II MHC transactivator) and APAF1 (apoptotic protease activator 1). Furthermore, a new NLR gene has been identified in the zebrafish genome database by our group and was named *NACHT-P1*. On a phylogenetic tree NACHT-P1 clustered with APAF1 and it also contained a WD40 domain instead of LRRs. NACHT-P1 has not been identified in any mammalian genome before, but a search in the human and mouse genome revealed NACHT-P1 orthologues to be present in both genomes (Stein, Caccamo et al. 2007) (all shown in figure 1.2).

In addition to the seven conserved NLRs in zebrafish three genes were found in zebrafish that were annotated as *nalps* but did not cluster with mammalian *nalps* on a phylogenetic tree (Stein, Caccamo et al. 2007).

Recently, it has been demonstrated by Oehlers et al. that NOD1 and NOD2, two members of the conserved NLR family, play a role in reducing intracellular bacterial burden after *Salmonella* infection in zebrafish larvae. Morpholino-mediated depletion of NOD1 and NOD2 expression decreased expression of dual oxidase in the intestinal epithelium and impaired the ability of larvae to reduce intracellular bacterial burden (Oehlers, Flores et al. 2011).

In addition to the conserved NLRs and the zebrafish Nalps, a new gene family was found that shared a similar structure with the conserved NLRs but presented additional features that were found only in fish species (for further information see below). Therefore, they were called the novel fish or the fish-specific NLRs.

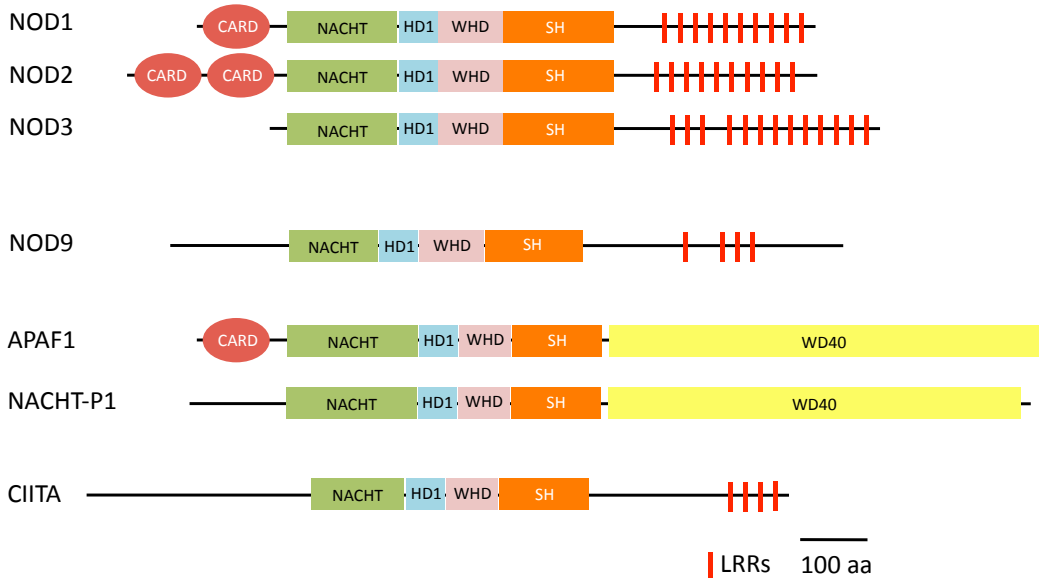

Figure 1.2: Protein structure of seven conserved zebrafish NLRs. NOD1, NOD2, NOD3, NOD9, APAF1, NACHT-P1 and CIITA proteins and the respective domains are depicted to scale. WD40 is short for ~ 40 aa, often terminating in a Trp-Asp (W-D) dipeptide. Information for protein sizes and structures were taken from ENSEMBL, (Tschopp, Martinon et al. 2003; Stein, Caccamo et al. 2007).

1.6 Novel fish NLRs in zebrafish

Initially, a genome-wide database screen was performed with the ENSEMBL release Zv6 to search for genes that contained a NACHT domain. TBLASTN was used to screen unfinished clones from the genome sequencing project and trace sequences from the whole genome shotgun project. About 200 genes containing a NACHT domain were found that were distributed throughout the genome of *Danio rerio*, mostly occurring in large gene tandems arrays, especially on chromosome 4 (Stein, Caccamo et al. 2007). Phylogenetic analysis of 277 NLR genes, including conserved mammalian and fish NLRs, Nalps and the newly identified NLR family in zebrafish, revealed that the new 200 genes clustered with each other and were highly conserved. Each of them was originally named after its genomic location (for example, 4.44 indicates that the gene resides on chromosome 4, and it is the 44th NLR-encoding gene on this chromosome) but since annotation of these genes is an ongoing process the names and numbers were considered as preliminary.

Like the conserved NLRs and the Nalps, the new zebrafish NLRs contained C-terminal LRRs, a NACHT domain, a HD1, the WHD and the SH domain (Figure 2). However, most of them had no characteristic N-terminal effector domain, such as pyrin domain (PYD) or CARD. Instead immediately upstream of the NACHT domain, they all share a sequence that is highly conserved among all fish-specific NLRs. This region was called fish-specific NACHT-associated (Fisna) domain because it was only found in association with the NACHT domain (Stein, Caccamo et al. 2007). A short peptide motif (DIY/FT) within this domain was also found in mammalian NOD2, but so far no biological function has been assigned to it.

Further genome analyses that also included the two pufferfish genomes of *Takifugu rubripes* and *Tetraodon negrovirides* led to the assumption that the Fisna domain has been recruited specifically by a common ancestor of the novel NLR proteins in the fish lineage (Stein, Caccamo et al. 2007).

Recently, we performed an additional screen for NACHT-domain containing genes using BioMart. BioMart is a web-based data mining tool provided by ENSEMBL which is customizable in contrast to their normalized databases. The dataset, filters and attributes can be individually chosen which makes it easier to evaluate the results. As the dataset we chose *Danio rerio* genes (Zv9) and selected an individual protein domain filter for the NACHT domain using its PFAM ID (PF05729). Furthermore, we chose attributes that included the ENSEMBL gene ID (ENSDARG entry) in the search output. This second search for fish-specific NLRs led to a new dataset of 366 NACHT-domain encoding genes with sizes ranging from 0,8 to 3,7 kilobases (kb). Their exon numbers range from 3 to 12. The positions of the first fish-NLR dataset of about 200 genes were mostly withdrawn within the last five years and adjusted to the newest genome assembly. The increase in gene number can be explained by the progress in zebrafish genome assembly and growing data on transcript evidence and expressed sequence tags (ESTs) since Zv6.

Present work on annotation of this new set of genes is done in colaboration with Dr. Kerstin Howe (The Wellcome Trust Sanger Institute, Hinxton), Prof. Dr. Thomas Wiehe and Giuliano Crispatzu (both Institute for Genetics, Cologne). The common goals are (1) to find all NACHT-domain containing genes in the zebrafish genome, (2) to combine exsisting predictions from National Center for Biotechnology Information (NCBI) and ENSEMBL databases on our own NLR databse, (3) to identify distinct coordinates for each gene and (4) to incorporate additional NLR sequences that are based on our own experimental data (ePCR and sequencing projects, both not subject to this thesis) into the new database. Creating such a new database should help to get an overview on

the whole fish-NLR gene family and to indicate which genes have to be re-annotated in existing and open databases.

In a further step a Hidden Markov Model (HMM) projections for the NACHT domain, Fisna domain and B30.2 domain on the identified fish-specific NLR genes were performed to look for the presence ot these domains within the genes. This work is still in progress and is mainly performed by Giuliano Crispatzu and Dr. Kerstin Howe.

For the genes or constructs that were the subject of this thesis the presence of a NACHT domain and a Fisna domain had supporting evidence by sequencing.

In 2009 a study on a different large gene family in teleost fish, the tripartite motif (TRIM or finTRIM) genes, has been published (van der Aa, Levraud et al. 2009). The authors report that some of these genes share a domain called B30.2 that has been assigned to RNA recognition of viral origin in rainbow trout. They showed the B30.2 domain was also among NLRs in zebrafish. Indeed, in many of our fish-specific NLR genes a B30.2 domain was present downstream of the LRRs, with the exception for NLR group IV (see below). Since Zv8 the B30.2 domain is included in the ENSEMBL protein maps, which facilitates the identification of this domain in contrast to the manual approaches. Yet, there are no experimental results available on the function of this domain in zebrafish NLRs.

1.6.1 Group-specificities of novel fish NLRs

Based on their sequence differences within the NACHT domain, Fisna domain and the N-terminal region, the zebrafish NLRs were divided into four subgroups (see figure 2). Each of these groups has further shared motifs upstream of the Fisna domain.

The N-terminal part of group I contains a domain which is akin to a pyrin domain. The N-terminus of group II shows a lower degree of similarity to the pyrin-like domain than group I. It is composed of a 101 amino acid stretch that is unique for the second group and is shown by all of its members.

Within the N-termini of group II, III and group IV similar sequence repeats occured. This sequence of about 35 to 41 amino acids appears in two, three or four copies per protein, or in one case, in ten copies (Figure 1.3, indicated in the left light blue boxes) (Stein, Caccamo et al. 2007). Not all N-termini of the novel zebrafish NLRs could be identified due to the absence of sufficient ESTs within these regions.

Almost all of the C-termini of group I, II and III contain a B30.2 (or PRY–SPRY) motif downstream of the LRRs that has been assigned to virus recognition (Gack, Shin et al. 2007; Wolf and Goff 2007; Uchil, Quinlan et al. 2008).

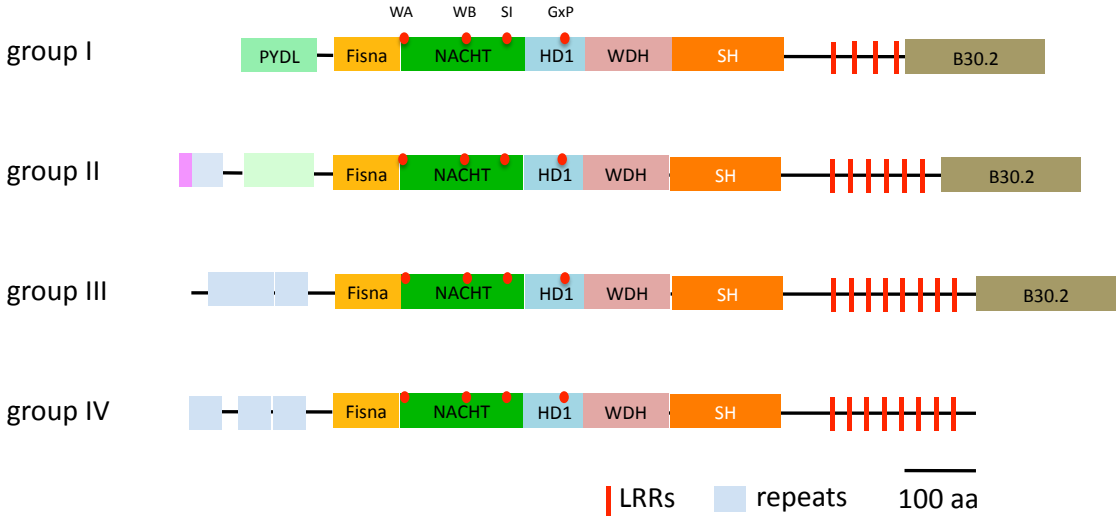

Figure 1.3: Domain structure of the four groups of novel fish NLRs in zebrafish. Sequence similarities in the NACHT domain (green box) and the Fisna domain (yellow box) led to a subdivision into four groups. The indicated domain structure is considered as representative for most of the respective group members, deviations are possible. Copy numbers of N-terminal repeats (light blue box) for group III and IV vary upon the single proteins. For group III occurence of two to ten repeats is possible. HD1 stands for helical domain 1, WHD is winged helical domain and SH super helical domain. Motifs within the NACHT- and HD1 domains are indicated as red filled circles. WA is Walker A and WB Walker B motif, SI is sensor I and GxP indicates a cysteine rich region.

1.7 Difficulties in analyzing novel fish NLRs

The extrem sequence similarities within the fish-specific NLR family pose a challenge in many aspects. On the one hand a reliable annotation of *nlrs* in genome databases is not available. In many cases genes are incomplete and/or overlapping with each other. A search for fish-specific NLR genes by the automated gene prediction algorithm of ENSEMBL was unsuccessful for most of the genes. In some cases a NACHT domain was found but sufficient EST evidence was missing therefore complete gene annotations were impossible. Furthermore, a complete prediction for the LRRs at the 3' ends was not completed due to repetitions of similar sequence streches and the absence of ESTs at this part of the genes.

Furthermore, mapping the genes on the chromososmes is rather difficult. Especially the analysis of gene organization on chromosome 4 led to complications since the entire chromosome is composed of heterochromatin. This led to a subgroup of fish-specific NLR genes that have transcriptional evidence but no assigned chromosome, hence are associated to "scaffold".

On the other hand the sequence similarities and number of genes made it rather impossible for reverse genetical approaches as e.g. the use of morpholinos or creating

transgenic knock-out animals. For expression analyses only a group-wise or single transcript characterization was feasible since it is not fully understood how many genes exsist and are expressed.

1.8 Aims

The aim of this project was to understand the function of novel fish-NLRs. We wanted to find out if this newly identified protein family was involved in the innate immune system of zebrafish embryos and if they play a role in defending the fish against bacterial and viral infections.

For that purpose, semiquantitative reverse transcriptase (RT) PCR, SYBR Green Real Time PCR, Southern Blot and Northern Blot analysis were used to compare basal expression with the protein expression after pathogen treatment.

In order to identify putative pathways fish NLRs are involved in we searched for binding partners. We applied a yeast two hybrid screen and comprehensive yeast two hybrid experiments for an unbiased evaluation of physical interaction of fish-specific NLRs. Furthermore, the putative interaction partners were tested in co-immunoprecipitation experiments.

2. Material and Methods

2.1 Materials

Chemicals used for this study were purchased from the following companies if not stated otherwise: Amersham-Pharmacia, Analytik Jena, Applied Biosystems, Biomol, Biozym, Bioline, Clontech, Fluka, New England Biolabs, Riedel-de-Häen, Serva, Invitrogen, Merck, Roche, Roth and Sigma-Aldrich.

2.1.1 Kits

InnuPREP Plasmid Mini Kit	Analytik Jena
InnuPREP DOUBLEpure Kit	Analytik Jena
Matchmaker Library Construction & Screening Kit	Clontech
Advantage® 2 PCR Kit	Clontech
TOPO TA Cloning Kit pCR□2.1-TOPO plasmid	Invitrogen
SuperScript III First Strand cDNA Synthesis Kit	Invitrogen
Illustra™ MicroSpin™ S-200 HR Columns	GE Healthcare
ECL™ Western Blotting Detection Reagents	GE Healthcare
µMACS mRNA Purification Kit	Miltenyi
Random primed DNA labeling Kit	Roche
RNeasy Mini Kit	Quiagen

2.1.2 Buffers and Solutions

Blocking solution for Western blotting	5 % low-fat milk powder (SUCOFIN)
Church Buffer	0.5 M Na_2HPO_4 1 mM EDTA 7 % SDS pH 7.2
Colloidal coomassie staining solution	0.08% Coomassie Brilliant Blue G250 1.6% ortho-phosphoric acid 8% ammonium sulfate add 20% of methanol before use
DNA-extraction Buffer	10 mM tris base pH 8.2 10 mM EDTA 200 mM NaCl 0.5 % SDS

	200 µg/ml proteinase K freshly added
Embryo media	40 mM NaCl 1 mM KCl 1.5 mM HEPES 2.5 mM $CaCl_2$ pH 7.21
Dropout media (10, 1 l)	200 mg L-adenine hemisulfate salt 200 mg L-arginine HCl 200 mg L-histidine HCl monohydrate 300 mg L-Isoleucine 1000 mg L-leucine 300 mg L-lysine HCl 200 mg L-methionine 500 mg L-phenylalanine 2000 mg L-threonine 200 mg L-tryptophan 300 mg L-tyrosine 200 mg L-uracil 1500 mg L-valine
Laemmli buffer (4 x)	8 % SDS 400 mM DTT 240 mM Tris-HCl pH 6.8 0.004 % bromophenol blue 40 % glycerol
LB-Medium (1 l)	10 g bactotrypton 5 g bacto yeast extract 10 g NaCl 15 g agar (for plates only) pH 7.2 with NaOH
PBS	130 mM NaCl 2.7 mM KCl 7 mM Na_2HPO_4 3 mM KH_2PO_4 pH 7.4
MOPS Buffer (10 x)	0.2 M MOPS 80 mM NaAc 10 mM EDTA pH 7.0
PBST	0.2 % Tween-20 in PBS

Ponceau S (10 x, 100 ml)	2 g Ponceau S 30 g trichloracetic acid 30 g sulfosalicyclic acid			
Prehybridisation Buffer	1 M NaCl 50 mM tris base pH 7.5 10% dextransulfate 1% SDS 250 µg salmon sperm, sonicated 30 ml aliquots at -20 °C			
RIPA Buffer (100 ml)	0.88 g NaCl 2.5 ml 1 M tris base pH 7.5 1 ml EDTA 1 ml NP-40 1 g Doc 0,01 % SDS			
RNA Loading Buffer	50% glycerol 1mM EDTA 0.25 % bromphenol blue			
RNA Sample Buffer (5 ml)	500 µl formamide 200 µl formaldehyde 100 µl MOPS (10 x) 2 µl ethidium bromide			
Running Buffer (10 x)	0.25 M tris base 1.9 M glycine 10 % SDS			
SD medium (1 l)	6.7 g yeast nitrogen base without aa 100 ml 10 x dropout solution 20 g agar (for plates only) 2 % glucose			
Separating gel buffer (4 x, 1 l)	1.5 M tris base pH 8.8 4 ml SDS (10%)			
Separating gel mix for SDS-PAGE (100 ml)		10 %	12 %	15 %
	H_2O (ml)	41.6	35	25.5
	30 % acrylamide (ml)	33.4	40	49.5

	4 x seperating gel buffer (ml)	25	25	25
	5 ml per mini gel plus 100 µl APS and 6 µl TEMED			
SSC (20 x)	3 M NaCl 0.3 M Na-Citrate pH 7.4 50 mM NaOH			
Stripping Buffer (Western blot analysis, 100 ml)	800 µl β-mercaptoethanol 2 % SDS 12.5 ml 0.5 M Tris pH 6.8			
Stripping Buffer (Southern blot analysis)	0.1% SDS			
Stacking gel buffer (4 x, 500 ml)	30.3 g tris base 20 ml 10% SDS pH 6.8			
Stacking gel mix for SDS-PAGE (100 ml)	17.2 ml 30 % acrylamide 25.5 ml 4 x stacking gel buffer 57.3 ml H_2O 2.5 ml per mini gel plus 50 µl APS and 3 µl TEMED			
TAE Buffer (50 x, 1 l)	1 M tris base 57.1 ml glacial acetic acid 100 ml 0.5 M EDTA pH 8			
Tris buffered saline (TBS) (1 l)	3 g tris base 8.8 g NaCl 0.2 g KCl pH 7.4			
TBST	TBS with 0.2 % Tween-20			
Transfer buffer for western blotting (1 l)	5.82 g tris base 2.93 g glycine 3.75 ml 10 % SDS 200 ml methanol			
Washbuffer 1 (Northern + southern blot analysis)	2 x SSC 0.1 % SDS			

Washbuffer 2 (Northern + southern blot analysis)	1 x SSC 0.1 % SDS
YPD (1 l)	20 g difco peptone 10 g yeast extract 20 g agar (for plates only) 2 % glucose

2.1.3 Antibodies

All antibodies listed below are stored in 50 % glycerol at -80 °C. The indicated numbers correspond to the numbers in the FileMaker Pro antibody list of the Leptin lab.

Table 1: Antibodies

No.	against	raised in	dilution	marked with	company
39	mouse	goat	1:3000	HRP	Jackson Immuno Research Lab
197	GFP	rabbit	1:2500	-	Torry pines Biolabs Inc.
199	rabbit	goat	1:1000	HRP	Dianova
287	HA	mouse	1:2000	-	Roche
254	c-myc	rabbit	1:1000	-	Santa Cruz
256	HA	rabbit	1:500	-	MBL
265	guinea pig	rabbit	1:2000	HRP	Abcam

2.1.4 *Escherichia coli* strains and genotypes used for cloning

DH5α — *sup*E44 Δ*lac*U169 (ϕ80*lacZ*ΔM15) *hsd*R17 *rec*A1 *end*A1 *gyr*A96 *thi*-1 *rel*A1

BL21(DE3) — *hsdS, gal [cI, ts*857*, cnd*1*, hsd*R17*, rec*A1*, end*A1*, gyr*A96*, thi*-1*, rel*A1*]*

2.1.5 *Saccharomyces cerevisiae* strains and genotypes

AH109 — *MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4*Δ, *gal80*Δ, *LYS2 : : GAL1UAS-GAL1TATA-HIS3, GAL2UAS-GAL2TATA-ADE2, URA3 : : MEL1UAS-MEL1TATA-lacZ, MEL1*

Y187 — *MAT*α*, ura3-52*, *his3-200, ade2-101*, *trp1-901, leu2-3*, *112*, *gal4*Δ, *met*–, *gal80*Δ, *URA3 : : GAL1*UAS-*GAL1*TATA-*lacZ, MEL1*

2.1.6 Primers

Primers used in this study were synthesized by Metabion, Invitrogen or Eurofins. They were stored at -20 °C in a concentration of 100 pmol/μl. The working concentration was 10 pmol/μl. All primers are listed below in 5'→3' direction. The indicated numbers correspond to the numbers in the primers list available in the Leptin lab.

Table 2: Primers

067	FisnaI for	GCGAATTCAGGAGAAGTTAAAGGAG
068	FisnaI rev	GGATCCTTTAAATATGTCGTTACACTT
052	FisnaII for	GCGAATTCAAATGTAGCCTGAAGA
053	FisnaII rev	GCGGATCCTTCAAACAAATGTTTGC
054	FisnaIII for	GCGAATTCAAAACCAGCATGAAGA
055	FisnaIII rev	GCGGATCCTTTAAAGATGTCATTGC
056	FisnaIV for	CGGAATTCAGATCAAATCTGCTGA
057	FisnaIV rev	TAGGATCCTGTAAAGATGTCTCTGC
028	upNACHTIII for	GCGAATTCATGGAAAATGTCAAACAGCT
029	upNACHTIII rev	GCGGATCCAAACATGTTGTTACACTTCA
021	NACHTI for	ATGAATTCTCCTGACACTGGGGATCGCAGG
062	NACHTI rev	GAGGATCCGTTTTTGATGAAGTACTGCTCC
006	NACHTII for	ATCTCGAGTCTCTTTCTGAAGTACTCCTCC
063	NACHTII rev	ATGGATCCTCTCTTTCTGAAGTACTCCTCC
017	NACHTIII for	CAGAATTCTTCTCACTAAAGGCATCGCT
064	NACHTIII rev	GCGGATCCTCTCTTCCTGAAATATTCCT
065	NACHTIV for	ATGAATTCTGCTGACGAAGGGAGTCGCT
066	NACHTIV rev	GCGGATCCGAGAACACTTCATAAGG
225	HAFisnaI for	CCCGAATTCATGTACCCATACGATGTTCCAGATTACGCTAAGGAGAAGTTA
226	HAFisnaII for	CCGGAATTCATGTACCCATACGATGTTCCAGATTACGCTAAATGTAGCCTG
227	HAFisnaIII for	CCCGAATTCATGTACCCATACGATGTTCCAGATTACGCTAAAACCAGC
228	HAFisnaIV for	GCGGAATTCATGTACCCATACGATGTTCCAGATTACGCTAGATCAAATCTG
229	HANACHTI for	CCCGAATTCATGTACCCATACGATGTTCCAGATTACGCTAGAGTCCTGACA
230	HANACHTII for	CCCGAATTCATGTACCCATACGATGTTCCAGATTACGCTAGAACTGTCCTG
231	HANACHTIII for	CCCGAATTCATGTACCCATACGATGTTCCAGATTACGCTAAGACTGTTCTC
232	HANACHTIV for	GGGGAATTCATGTACCCATACGATGTTCCAGATTACGCTAGAACTGTGCTG
209	actin for	CTCTTCCAGCCTTCCTTCCT
210	actin rev	CTCCTTCTGCATACGGTCAGCAA
046	SMART for	TTCCACCCAAGCAGTGGTATCAACGCAGAGTGG
047	SMART rev	GTATCGATGCCCACCCTCTAGAGGCCGAGGCGGCCGACA
048	5' AD Amplimer	CTATTCGATGATGAAGATACCCCACCAAACCC
049	3' AD Amplimer	GTGAACTTGCGGGGTTTTTCAGTATCTACGAT
093	T7 Primer	TAATACGACTCACTATAGGG
094	M13 rev primer	CAGGAAACAGCTATGAC
256	M13 for primer	TGTAAACGACGGCCAG

146	SP6	ATTTAGGTGACACTATAG
069	NLRI RT for	AAATCCCCTCCTGTTGATTACA
070	NLRI RT rev	CTGTGTCAGGCTGAACTTTAAA
071	NLRII RT for	CAGAAGCAGCTCTTGATCTCA
072	NLRII RT rev	CATGTTCAGTGTTGACCTGTT
073	NLRIII RT for	GCAGGATTCACTCCAACCAGAA
074	NLRIII RT rev	GATGTCATTGCAGTAGACTGG
075	NLRIV RT for	TCTGCTGAAGAAGTTTGAGTG
076	NLRIV RT rev	CCCTTCGTCAGCACAGTTCTG
275	HyNACHTI+II r	CATTCATCCAATCCATCAAGGATGAACATGACCTTACCTT
276	HyNACHTII+I f	GATTCAAAGTCCTGTTCATCTTTGATGGACTGGATGAATA
277	HyNACHTII+III r	TTTCATCCAGACCATCAAAGATGAACAGGACTTTGAATC
278	HyNACHTIII+II f	AGTGTAGAGTTGTGTTCATCCTTGATGGATTGGATGAAT
279	HyNACHTIII+IV r	GACGACACTCGTCCAGACCATCAAAGATGAACACAACTCTA
280	HyNACHTIV+III f	TGTTCATCTTTGATGGTCTGGATGAAAGCAGAATCACACTCA
281	HyNACHTI+IV r	GACGACACTCGTCCAGACCATCAAAGATGAACATGACCTTACCT
282	HyNACHTIV+I f	ATATAAAGTGCTGTTCATCTTTGATGGACTGGATGAATATCGC
283	HyNACHTII+IV r	AGACGACACTCGTCCAGACCATCAAAGATGAACAGGACTTTG
284	HyNACHTIV+II f	TAAAGTGCTGTTCATCCTTGATGGATTGGATGAATGTCGTCTTCCT
285	HyNACHTI+III r	CTTTCATCCAGACCATCAAAGATGAACATGACCTTACCTT
286	HyNACHTIII+I f	TATTCATCCAGTCCATCAAAGATGAACACAACTCTACACT

2.1.7 Plasmids

All plasmid sequences and variants are available at the Vector NTI Database in the Leptin lab.

pGADT7, pGBKT7	provided with the Matchmaker Library Construction & Screening Kit by Clontech
pGAD 424	kindly provided by AG Dohmen
pCR®2.1-TOPO	provided with the TOPO TA cloning Kit by Invitrogen

2.1.8 Software

Adobe Photoshop CS3

ClustalW

Jalview

Microsoft Office 2008

Vector NTI Advance 10

2.2 Methods

2.2.1 Methods in molecular biology and protein biochemnistry

2.2.1.1 Polymerase chain reaction (PCR)

PCRs were performed with the Biometra UNO-Thermoblock (Biotron) or the TProfessional Thermocycler (Biotron). For one reaction 1–1000 ng template DNA, 10 pmol forward primer and 10 pmol reverse primer, 40 mM deoxynucleotide triphosphate (dNTP)-mixture (Sigma), 1 x reactionbuffer and 1µl of either Taq DNA polymerase (Invitrogen) or REDTaq DNA polymerase (Sigma) was used. In the case of the Taq DNA polymerase 3 mM $MgCl_2$ was added. As DNA templates either genomic DNA, copied DNA (cDNA) synthesized via first strand synthesis (see 2.2.1.2) or plasmid DNA was taken. PCRs were standardly performed in 50µl reactions in 200µl reaction tubes (Starlab). If not noted otherwise the following PCR program was applied.

1. 5 min 95 °C denaturation
2. 15 sec 95 °C denaturation
3. 30 sec at 55 °C primer annealing
4. 1 min/kb at 72 °C elongation
5. 5 min at 72 °C elongation
6. 4 °C

x29

2.2.1.2 Quantitative real-time PCR

cDNA was amplified using SYBR Green I nucleic acid gel stain (Invitrogen). Quantitative real-time PCR was performed on a MyIQ single color real time PCR detection system (BioRad) supported by IQ5 optical system software (BioRad). 25 ng cDNA reverse transcribed from mRNA was used in a total reaction volume of 20 µl and the PCR was run on 96-well plates. The actual amplification reaction was followed by a melting curve analysis to specify cDNA amplification and exclude primer dimers. This was done by running a gradient from 55 °C to 95 °C, increasing the temperature by 0.5 °C every cycle. The following program has been used:

1. 3 min 95 °C denaturation
2. 10 sec 95 °C denaturation
3. 30 sec at 62 °C primer annealing
4. 40 sec at 72 °C elongation

x40

5. 1 min at 95 °C dissociation curve
6. 1 min 55 °C
7. 10 sec 55 °C + 0.5 °C/cycle melt curve ⌊ x80

β-actin was used as endogenous control. Each reaction was performed in triplicates to ensure statistical significance. With each gene a standard curve had to be performed in order to relatively quantify the threshold cycle value (Ct). Expression level of candidate genes taken from non-treated control embryos were normalized ($Ct_{normalized\ control}$) and set at 1 in a diagram. Relative fold (Rf) differences in the expression of candidate genes (gene of interest) to an adequate reference gene (endogenous control) and the normalized control was determined with the following model:

$$Ct_{normalized\ control} = Ct_{gene\ of\ interest\ non\text{-}inf} / Ct_{endogenous\ control}$$

$$Rf = (CT_{gene\ of\ interest\ inf} / Ct_{endogenous\ control}) / Ct_{normalized\ control}$$

Finally, Rf values for each gene were shown in a gene expression diagram in relation to the expression in a non-treated control ($Rf_{normalized\ control} = 1$).
For the standard curve 4 different dilutions of any zebrafish cDNA was taken (1:5, 1:10, 1:50, 1:100) and run in dublicats.

2.2.1.3 mRNA isolation and first strand cDNA synthesis

Dechorionised zebrafish embryos or adult zebrafish without head and tail fin were lysed in 1 ml RNA-lysis buffer (MiltnenYi) and pure mRNA was labeled with magnetic microbeads and isolated on magnetic columns. mRNA preparation using the μMACS mRNA Purification Kit (Miltenyi) was performed according to the manufacturer's instructions.
After elution from the magnetic column, the mRNA was precipitated by adding 1/10 vol. lithiume chloride solution (4 M), 2.5 vol ethanol (100%) and 1 μl glycogene followed by overnight incubation at -20 °C. Afterwards the precipitate was centrifuged for 20 min at 13.000 rpm at 4 °C and the pellet was washed for 5 min with 70% RNase-free ethanol. The air-dried RNA pellet was resuspended in 20μl RNase-free water and the RNA concentration was measured with a photometer (BIO Photometer, Eppendorf).
For reverse transcription of mRNA into single-stranded cDNA the SuperSkript III First-Strand cDNA Synthesis Kit (Invitrogen) was used according to the protocol. In a 20 μl reaction, about 500 ng cDNA was synthesized. cDNA quality was tested in a PCR reaction using the actin primer pair (see talbe 2).

2.2.1.4 Phenol - chlorophorm extraction

The method usind phenol and chlorophorm is commonly used to isolate total RNA from tissues. This way nucleic acids from cytoplasmic fractions were isolated without breaking the nuclei and proteins were removed. After compelty grinding the tissue in liquid nitrogen using mortar and pestle 1.5 ml Trizol® (Invitrogen) was added and homogenized. After transferring the mixture into an reaction tube it was incubated for 5 min at room temperature. 0.3 ml of chlorophorm was added (0.2 ml per ml Trizol®) and vortexed for 15 sec. It was incubated another 3 min at room temperature.

By centrifuging 12.000 g for 15 min at 4 °C an aquos phase on top was seperated from the organic phase on the bottom half of the tube. The aquos phase containing the nucleic acids was carefully transferred into a new tube and thereby seperated from proteins and cell compartments. If any parts of the organic phase was taken along by mistake the extraction was repeated. 0.5 ml isopropanol was added to the nucleic acid solution to precipitate RNA at -20 °C over night. Afterwards it was centrifuged at 12.000 g for 10 min at 4 °C. Pellet was washed with 70 % ethanol and centrifuged as in the previous step. After drying the pellet was dissolved in 50-100 µl RNAse-free water.

RNA samples were purified using the RNeasy Mini Kit from Quiagen and then traeted with DNAse to eliminate traces of DNA.

2.2.1.5 Formaldehyde-/Agarose gel electrophoresis

Agarose gel electrophoresis was used to separate nucleic acids depending on their size. Agarose gels for DNA fragment separation contained 0.8-1.2% agarose, 100 ml 1 x TAE buffer and 10 µl ethidium bromide solution (10 mg/ml). Resolution of DNA fragments occurred at an average voltage of 10 V/cm. The 1kb hyperladder 1 (Bioline) was used as a DNA standard.After electrophoresis, gels were illuminated with an UV lamp (BioRad) to visualise the DNA bands.

Formaldehyde agarose gel electrophoresis for RNA samples was performed with 1.05 % agarose in 140 ml H_2O, 20 ml MOPS buffer (10 x) and 40 ml formaldehyde. Resolution of RNA fragments ocurred at 44 V over night at 4 °C in 1.5 l MOPS buffer (1 x).

2.2.1.6 Extraktion of DNA fragments from agarose gels (gelextraktion)

The DNA fragment of interest and correct size was cut from the agarose gel using a sterile scalpel and transferred to an 1.5 ml reaction tube. The extraction was performed using the innuPREP DOUBLEpure Kit (Analytik Jena) according to their protocol.

2.2.1.7 DNA digestion with restriction enzymes

Digestion of plasmid DNA or PCR products was performed according instructions provided by the enzyme suppliers (NEB, Boehringer-Roche) using appropriate buffers, BSA and temperature. In general, the reactions were performed in a volume of 50 µl for 1.5-2 h at 37 °C. The digestion products of interest were isolated using gelextraction (see 2.2.1.4).

2.2.1.8 Ligation of DNA fragments

For ligations plasmid and insert were combined in a molar ratio of 1:3 in final reaction volume of 20 µl. Additionally, 1 µl of Ligase enzyme (NEB; 400,000 units/ml) and 2 µl Ligasebuffer (NEB, 20 x) were added and the reaction was incubated for 1 h at room temperature (RT).

2.2.1.9 Transformation of chemical competent *E. coli* cells

Chemical competent DH5α E. coli cells were purchased from Invitrogen and stored at -80 °C. 50µl aliquots of the bacteria were thawed on ice. 30-100 ng of plasmid DNA was added, gently mixed and chilled on ice for about 30 min. Transformation by heat shock occurred at 42 °C for 90 sec. Cells were transferred immediately to ice for 1 min and resuspended in 0.6 ml LB medium. Cells were incubated for 1 h at 37 °C at 125 rpm and then spread on LB agar plates containing the respective antibiotics for selective bacteria growth. Plates were incubated upside-down overnight at 37 °C to get single bacterial colonies.

2.2.1.10 Isolation of plasmid DNA from bacteria (Mini-prep)

One single colony was cultured in 4 ml LB media with the appropriate antibiotics and grown overnight at 37 °C and 125 rpm. Further isolation and purification of plasmid DNA was performed with the innuPREP Plasmid Mini Kit (Analytik Jena) according to their protocol.

2.2.1.11 Sequencing

Sequencing of DNA was performed according to the Sanger method using the Big Dye Terminator Sequencing Kit. For one sequencing reaction 25-250 ng DNA was required. Sequencing was done either by the CCG sequencing facility at the University of Cologne or by GATC Biotech in Konstanz.

2.2.1.12 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE is commonly used to seperate proteins or petides according to their length and molecular weight under denaturating circumstances. Gels were cast in a way that ¾ were composed of a 10-15 % seperating gel and ¼ of a 5 % stacking gel. Gel preparation and running was done using the Mini Protean 3 System (BioRad). Gels were run with constant voltage of 100 V in the stacking gel and 140 V in the seperating gel.

2.2.1.13 Western blot analysis

For western blot analysis, proteins were transferred onto Hybond nitrocellulose membranes (Amersham) using a dry blotting system (BioRad) for 45 min with constant current and an average voltage of 10 V. After blotting the amount of protein and uniformity of the bands on the membrane could be visualized by staining for 5 min with Ponceau S. After destaining with tap water membranes were blocked for 1 h in blocking solution (5% milk powder in TBST) at room temperature and incubated with primary antibodies diluted in blocking solution overnight at 4 °C. After washing three times with TBST for 15 min, membranes were incubated with secondary antibodies (coupled with horseradish peroxidase) diluted in blocking solution 2 h at room temperature. Following three wash steps in TBST, the membrane was incubated in ECL western blotting substrate (Amersham) for 1 min. The major component of ECL luminol is the substrate of horseradish peroxidase and is oxidated. This reaction induces local chemical luminescence at the binding site of the secondary antibody. This luminescence reaction can be captured by exposure to an x-ray film.

2.2.1.14 Northern blot analysis and isotope-labeled DNA-probe synthesis

2 µg of purified total RNA per lane was size-separated by formaldehyde agarose gel electrophoresis (see 2.2.1.5). Afterwards, the gel was prepared for blotting by incubating in H_2O 2 times for 15 min, 5 min in 50 mM NaOH and again in H_2O for 15 min and 40 min in SSC (20 x). The separated RNA was transferred onto a positively charged nylon membrane (Hybond-XL, GE Healthcare). 20 x SSC served as blotting solution, blotting procedure took place over night at room temperature. A clean glas plate was put on top of a tray filled with 1 l of 20 x SSC. Whatman paper was put on the plate with its ends touching the SSC. The nylon membran pre-incubated with water was put on top of the agarose gel which was placed on the whatman paper. The blotting-setup was ballasted with paper towels and filled bottles. Blotting occured over night. After blotting the RNA was immobilized in a UV crosslinker.

Pre-hybridization was done at 42°C for 1h in 25ml Church buffer. In between DNA-probes were synthesized according to the "procedure for labeling with [α-^{32}P] dCTP"-protocol of the random primed DNA labeling kit (Roche) using 50ng template DNA. After stopping the reaction 100µl TE was added und the probe was purified using the IllustraTM MicroSpinTM S-200 HR Columns (GE Healthcare). Purified probes were added to the pre-hybridization buffer and incubated on the membran over night at 65°C. Using washbuffer 1 and 2 (see 2.1.2) the isotope-labeled DNA probe was washed of the membrane until 50 Bq.

2.2.1.15 Southern blot analysis

10 µg genomic DNA was digested in 10µl of each restriction enzyme Bgl II and Pvu II (see 2.2.1.6). After performing agarose gel electrophoresis for 4h at 40 – 50V the gel was transferred into 0.25M HCL for depurination and incubated for 45min. The blotting setup was similar to that of the Northern blot, using 0.4M NaOH as blotting solution. DNA probe synthesis and incubation was done as described above (2.2.1.13).

2.2.2 Zebrafish methods

2.2.2.1 Zebrafish strain

Natural occurence of the fresh water fish *Danio rerio* is the Ganges. Animals that are kept in the facility originate from pet shops in Cologne and Göttingen and are referred to the Cologne strain.

2.2.2.2 Keeping and raising zebrafish

The water temperature in the fish facility was adjusted between 26°C and 28°C. 1/10 of the water volume in the fish facility was exchanged daily whereas one half of the water was normal tap water supplied by the RheinEnergie AG and one half was transmitted from a reverse osmosis plant. The water was circulated by a pump system, depris was sieved by integrated filter units and filtered water was sterilized by UV radiation to eliminate pathogens. Accumulation of toxic substances was prevented by bacterial filters. The day-night rythm was adjusted to 14h light and 10h darkness. The fish were fed twice a day. Independent of their age they were fed with Artemia (Rebbie) in the evening. In the morning they were fed with Tetramin with particle sizes increasing with the age of the fish.

The starting point of the light phase triggered the spawning. If embryos were required for experiments they were collected from special plastic boxes, the male and female fish were kept in over night, about half an hour after the the beginning of the day phase. Embryos were kept in embryo media in petri dishes at 28.5°C until day 6. After one week they started feeding and were transferred to the fish facility.

2.2.2.3 Anesthetization of embryos

0.02% Ethyl-4-aminobenzonate or Tricaine (Sigma) was used for anesthetization of embryos to adult fish. 0.4% tricaine in 50 mM Tris-HCL (pH 7.5) stock solution was kept in the dark.

2.2.2.4 Infection assays

For infection experiments *E. coli* (DH5□), *Y. ruckeri* (4015726Q) and *M. marinum* (ATCC927) were applied, all carrying the dsRED-expressing pGEMs3 plasmid *(Bitter, van der Sar et al. 2003) E. coli* were grown in standard LB medium with ampicillin (50 µg/ml) over night at 37°C. *Y. ruckeri* were also cultivated in LB medium containing ampicillin at 30.5 °C for 16 h. *M. marinum* were grown on Middlebrook-Crohn-7H10 agar (Becton Dickinson) with was enriched with ADC (2g glucose, 5g BSA, 0.85g NaCl in 100ml H_2O) on 30.5°C for 16h. Bacterial cultures were centrifuged for 3min at 4.000 g, supernatants were removed and pellets were washed in PBS or in case of *M. marinum* in PBST. Bacterial solutions of an optical density OD_{600}=1 was diluted 1:5 in PBS(T) containing 1/10 phenol red. 200 ng/µl synthetic double-stranded RNA poly IC in PBS and 1/10 phenol red was used to mimic viral infection.

Dechorinized embryos between 24 and 32hpf were arranged on an agar plate. To cause an infection the bacterial solutions or poly IC suspensions were injected into the caudal vein. Infected embryos were incubated at 28.5°C until RNA isolation.

2.2.2.5 Isolation of genomic DNA from *Danio rerio*

Genomic DNA was isolated from embryos or fin tissue of adult zebrafish. To isolate a sufficient amount of genomic DNA from zebrafish 100 72hpf embryos were collected in one reaction tube, 990µl DNA-extraction buffer and proteinase K were added (see 2.1.2 buffers and solutions). Finclips were taken from adult fish after anesthesizing them in 10 % tricaine. Tissues were and put on 55°C for 3h and vortexed every 20min. Subsequently, 750µl 100% EtOH was added and the reaction tube was inverted a few times. By adding EtOH the solubilised DNA was precipitated and visible as a white cocoon.

The DNA was caught with a glass capillary and transferred into fresh 70% ethanol. The ethanol was removed completely by pipetting and the pellet was dried at room temperature. Afterwards the DNA was dissolved in 5mM Tris (pH 8). Genomic DNA was stored at room temperature or 4°C.

2.2.2.6 Protein lysates from zebrafish

Protein lysate preparation was freshly prepared prior to SDS-GAGE and Western blot analysis as described in (Link 2006).

2.2.3 Methods with Saccharomyces cerevisiae

2.2.3.1 cDNA library construction and Yeast two Hybrid (Y2H) screen

cDNA library construction and Y2H screen was performed using the Matchmaker Library Construction & Screening Kit purchased from Clontech (Cat. No. 630445). If not stated otherwise all steps were performed according to the manufacturer's instructions.

2.2.3.1.1 Construction of the DNA-binding domain (BD) fusion vector

The 500 bp N-terminal part upstream of the NACHT domain of the group III 18.03 zebrafish NLR gene was cloned into pGKT7 as the bait. The required fragment was amplified from cDNA of embryos 8hpf using primers which added the restriction sites for EcoRI at the 5' prime end and XhoI at the 3' end to the bait insert. After DNA digestion the fragment inserted in frame downstream of the GAL4 DNA binding domain of pGKT7. Sequencing reaction of both strands was performed to verify the correct bait sequence and the reading frame. Subsequently the bait vector was transformed into Y187 yeast strain by small-scale yeast transformation.

2.2.3.1.2 Transformation of competent yeast cells

10ml of YPD medium was inoculated with a freshly grown yeast single colony picked from an agar plate. The over night grown yeast culture was diluted to OD_{600}=0.1 and incubated until an OD_{600} between 0.6 and 1 was obtained. Yeast transformation was continued as described in the manual.

2.2.3.1.3 Generating a zebrafish SMART double stranded cDNA library for the Y2H screen

First-strand cDNA synthesis was performed using 500 ng poly A^+ RNA of an adult male zebrafish and 1 µl random primers (CDSIII/6-primer, supplied by the kit).

Amplification of ds cDNA by Long Distance PCR: PCR reactions were performed using the Advantage® 2 PCR Kit by Clontech with the forward primer annealing within the SMART sequence and the reverse primer annealing within the CDS III anchor site. For 0.5µg of poly A+ RNA 22 thermal cycles were used for cDNA amplification. cDNA library synthesis was performed as depicted in the manual.

2.2.3.1.4 Constructing the GAL4 activation domain (AD) fusion library

The GAL4 AD fusion library is produced by cotransforming competent yeast cells with SMART ds cDNA and the linearized pGADT/-Rec vector.
SMART ds cDNA can recombine with the AD cloning vector in vivo since its 5' and 3' end have homologous sequences to the respective ends of the Sma I-linearized pGADT7 vector. Afterwards the transformants were pooled and the 1ml aliquots freezed at -80°C in 50% glycerol. (For further instructions see manual).

2.2.3.1.5 Mating library host strain with the bait & selection for yeast diploids expressing interaction proteins

For instructions see the manufacturer's protocols.

2.2.3.2 Analysis of positive interactions by colony PCR and sequencing

A few cells of each selected colony were picked with a sterile pipet tip and put into a PCR tube, respectively. The cell membranes were destroyed by placing the tubes containing the yeast cells into the microwave (700 W) for 1min to make the DNA accessible to the Taq polymerase. A master mix for 50µl PCR reactions was prepared containing a primer pair annealing 100bp up- and downstream of the library insert, respectively, using the standard PCR program.
5 µl of each PCR sample were analyzed via agarose gel electrophoresis on a 1% gel to check the size of the library insert and the amount of DNA per sample. In a further step the forward strand of the PCR fragments was sequenced and the results were compared to the genes in the databases by using the BLAST program on the ENSEMBL homepage.

2.2.3.3 Y2H to test protein interactions

To test interactions between two proteins both interaction candidates were cloned one time in the pGAD424 vector for expression of an DNA activation domain fusion and one time in the pGBKT7 vector to construct a DNA binding domain fusion protein to test a putative interaction in both directions. One possibility is to cotransform both vectors into the AH109 strain and select for interaction properties on appropriate

dropout media. Another way to control protein interactions with Y2H is to transform the pGBKT7 vector expressiong the bait fusion protein into Y187 yeast cells of the mating type and mating these cells with AH109 cells of mating type a expressing the prey protein fusion. For appropriate dropout media and incubation times see the protocol supplied with the kit. To verify positive interactions Co-immunoprecipitation experiments were performed.

2.2.3.4 Co-immunoprecipitation with yeast cells

Cotransformed AH109 yeast cells expressing both fusion proteins were used as starting material. Cells were incubated in 20ml SD –LEU/-TRP dropout medium from OD_{600}=0.1 until OD_{600}=1 at 30°C and 126rpm. Cells were peletted for 1min at 3.000g and resuspended in 500µl cold RIPA lysis buffer. Cellmembranes were disrupted using a sonicator. Cell fragments were spun down at 13.000rpm at 4°C for 15min and the supernatant was transferred into a fresh tube, 20µl were kept for the input.

For immunoprecipitation either protein A agarose (Roche) or EZview™ Red Anti-HA Affinity Gel (Sigma) which are protein A agarose beads precoupled to anti-HA monoclonal mouse antibodies. 50µl uncoupled protein A agarose beads were added to anti-cmyc antibody and myc-tagged protein complex and incubated for 2h at 4°C and shaken at 1.000rpm every 5min for 30sec. For preparation of the antigen-antibody complex 470µl input was incubated with 3µg of anti-cmyc antibody over night at 4°C.

3. Results

3.1 Expression analysis of novel NLRs in zebrafish

To understand the expression of fish-specific *nlrs* in zebrafish embryos and larvae semiquantitative RT-PCR, quantitative RT-PCR (qRT-PCR) and Northern blot analyses were performed. Since it was unkown which genes were expressed our expression analyses were generated in a way that they targeted a group of *nlr* genes in contrast to target single genes. Therefore, group-specific primers and probes were used which means that the sequences that were chosen for amplification or transcript targeting were unique for *nlr* groups *I* to *IV*.

Figure 3.1 shows four multiple alignments of 61 *nlr* nucleotide sequences. It depicts sequences for forward and reverse primers (in black frames) that were used to amplify each of the four *nlr* groups in all RT-PCR experiments. The alignments comprise a selection of the fish-specific *nlr* genes, sequences from groups II and III that showed 99 to 100 % redundancies to the depicted ones were removed. Thus 23 sequences of group II and 27 of group III are left out in figure 3.1. They also reflect the sequence similarities within the groups and the differences between the four nlr groups in general.

The figure 3.1 presents the specificities of each primer pair for one group and in contrast, the distinction of these sequences to the other three groups which are not targeted by the same primers. However, it also demonstrates that not every single gene of one group can be amplified by the chosen pair of primers but a major subset of each group. In figure 3.1 (B) the positions of the amplified sequences with regard to the Fisna and NACHT domain sequences in the genes are shown. The cDNA amplification of group I had a length of 310 nt, of group II it was 250 nt and the PCR products for grops III and IV was 302 and 320 nt . Forward and reverse primers were located on two neighbouring exons to recognize contamination with genomic DNA. The complete structure and domain positions of each NLR groups are depicted in Figure 1.3.

On gene 14.39 which belongs to the *nlr* group I the forward primer starts at nucleotide (nt) 302 and the reverse primer ends at nt 614. The position in a representative *nlr II* gene (14.27) was nt 406 to 643, for *nlr III* (up12f) nt 457 to 759 and for a *nlr IV* gene (15.08b) the position was nt 367 to 687. The primer sequences are also given in table 2.

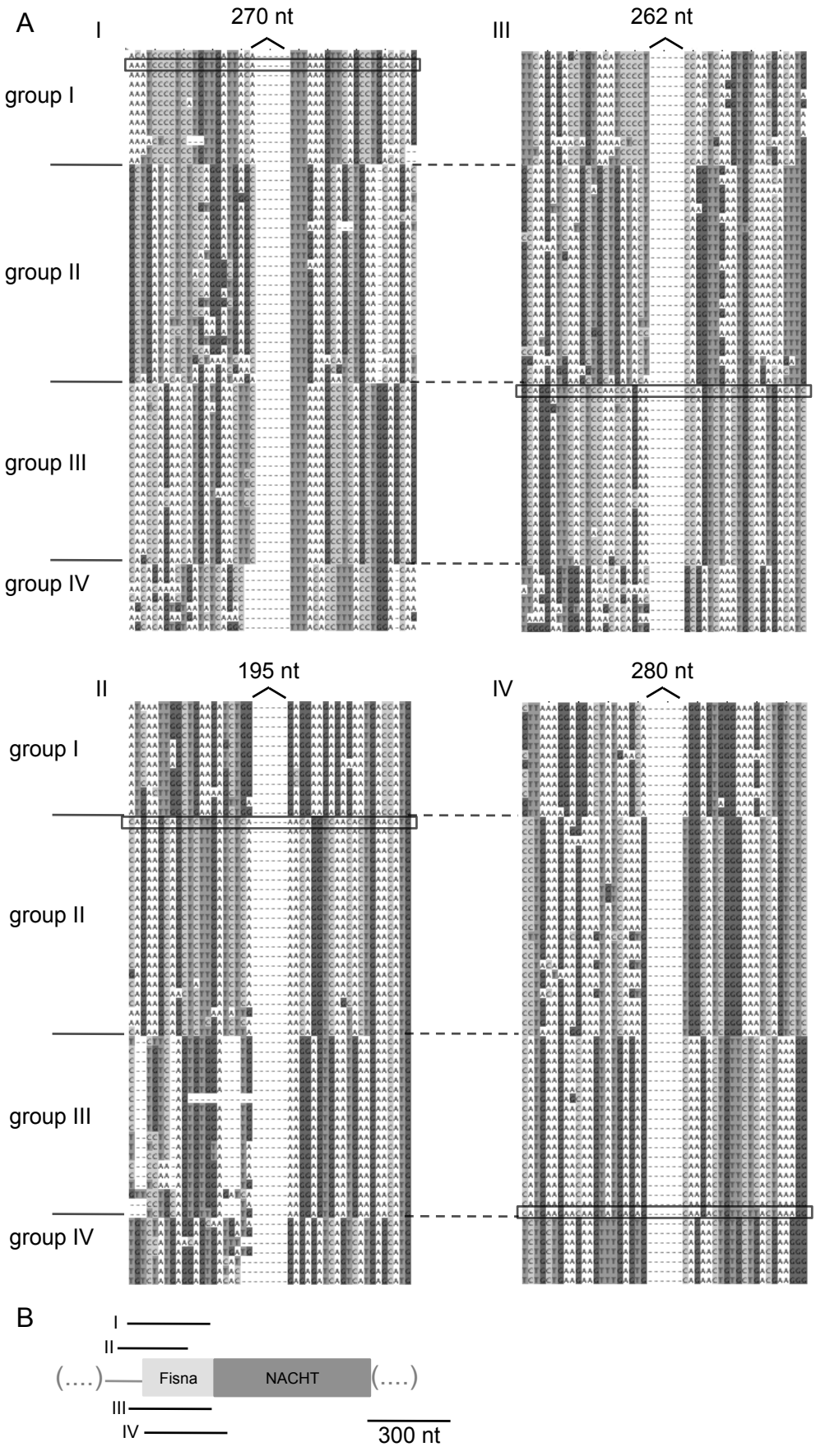

Figure 3.1 (see next page).

Figure 3.1: Multiple alignments of RT-PCR primer sequences used for *nlr* amplifications. (A) shows four primer pairs (I, II, III and IV) and *nlr* sequences aligning at the respective positions. The forward and reverse primers of the four *nlr* gene groups are framed in black. Forward primers (left sides) and the reverse primers (right sides) of the respective group enclose sequence stretches from 195 to 280 nucleotides (nt, indicated above) that are not shown in this figure. The nucleotides are illustrated in a colour code: adenine in white, cytosine in light gray, thymine in middle gray and guanine in dark gray. In (B) the positions of the amplified sequences by the group-specific primers are depicted in regard to the Fisna domain. The figure does not show C-terminal or N-terminal parts of the genes, left out parts are symbolized by brackets.

3.1.1 Reverse transcription analyses

To analyse *nlr* gene expression mRNA was isolated from embryos or larvae from different developmental stages. DNAse treatment of the mRNA samples was used to avoid amplification of genomic DNA during PCR. Since the primers were located on two different exons a contamination withn genomic DNA would have been visible by larger amplifications. This was not the case. Reverse transcriptase (RT) was used to synthesize cDNA, which was amplified using PCR. A semiquantitative comparison between amplified *nlr* gene sequences and the β-*actin* level of the sample shows the relative amount of *nlr* mRNA expression. For each developmental stage the same cDNA sample was used as a template for the PCR.

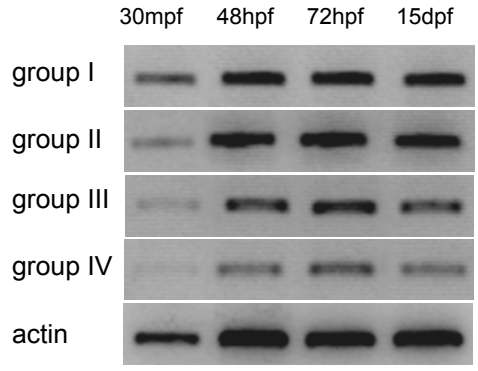

Figure 3.2: *nlr I-IV* expression during embryonic development in zebrafish analyzed by RT-PCR. mRNA was isolated at 30mpf, 48hpf, 72hpf and 15dpf to analyze nlr expression in embryos. 12ng of each cDNA sample was amplified in 29 PCR cycles. For PCR protocol see section 2.2.1.1; for primers see table 2. 10μl of each reaction was run on a 1% agarose gel. The picture was taken on a gel documentation system from BioRad.

Group-specific primers, as described above, were used during these RT-PCR experiments. Figure 3.2 shows the expression of *nlr I-IV* genes from embryos 30 mpf, 48hpf, 72hpf and larvae 15dpf. The results indicate a maternal expression for all four *nlr* groups which remain expressed during further embryonic and larval development. Compared to the level of *β-actin nlr I* and *II* appeared to be equally expressed during the first two weeks of development. For *nlr* groups *III* and *IV* the expression level peaks at 72hpf and weakens towards 15dpf.

Quantitative Real Time PCR (qRT-PCR) using the SYBR Green method was additionally applied to compare the relative expression of the four NLR groups. cDNA of embryos 30hpf was used for the qRT-PCR analyses. According to the qRT-PCR results *nlr III* shows the highest expression level of all four groups. *nlr IV* was five times less expressed and showed the lowest expression level. *nlr I* expression was only a little elevated compared to *nlr IV.* The *nlr II* level was about two times elevated compared to *I* and *IV.*

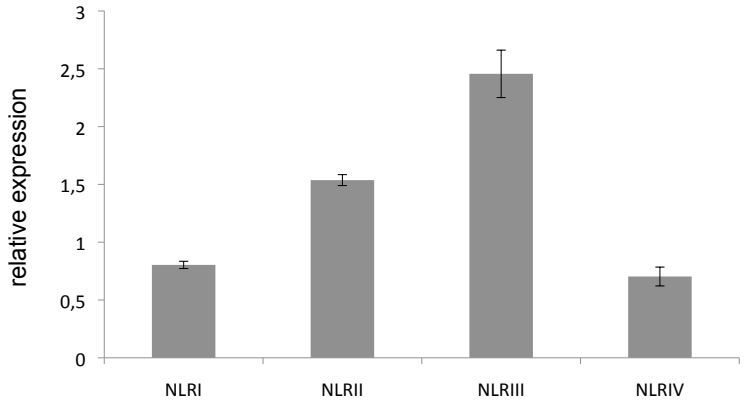

Figure 3.3: Comparison of *nlr I-IV* relative expression using qRT-PCR. The SYBR Green qRT-PCR method was used to to compare the relative expression differences between the four *nlr* groups. cDNA was taken from embryos 30hpf. For the endogenes control and reference for relative gene expression β-actin was used. The relative expression level is represented in the y-axis.

3.2.2 Northern blot analysis of *nlr I-IV* expression

A Northern blot analysis was additionally performed to investigate fish-NLR expression in zebrafish embryos, larvae and adult fish.

First, to test the specificities of the DNA probes for the *nlr* groups a Southern blot was performed. The hybridization of probes FN I, II, III and IV was tested on *nlr* clones for each of the four groups. pGBKT7 plasmids containing FisnaNACHT sequences from

the *nlr* groups I to IV were linearized and load on a agarose gel in different titers (see figure 3.4).

Figure 3.4 shows that the probes bind with higher affinity to their own group but they also bind with weaker affinity to the foreign three groups. Probe FN III shows a significant degree of cross-hybridization with FisnaNACHT I. Regarding the signals for the different plasmid-titers probe I, II and IV show at least a ten times stronger signal on the construct from their own group compared to the other groups.

In the rows for FN III and FN IV additional bands (black arrows) were visible below the expected size of 9kb for the linearized plasmid. They presumably result from uncut and supercoiled plasmids that migrate faster in an agarose gel than linearized DNA.

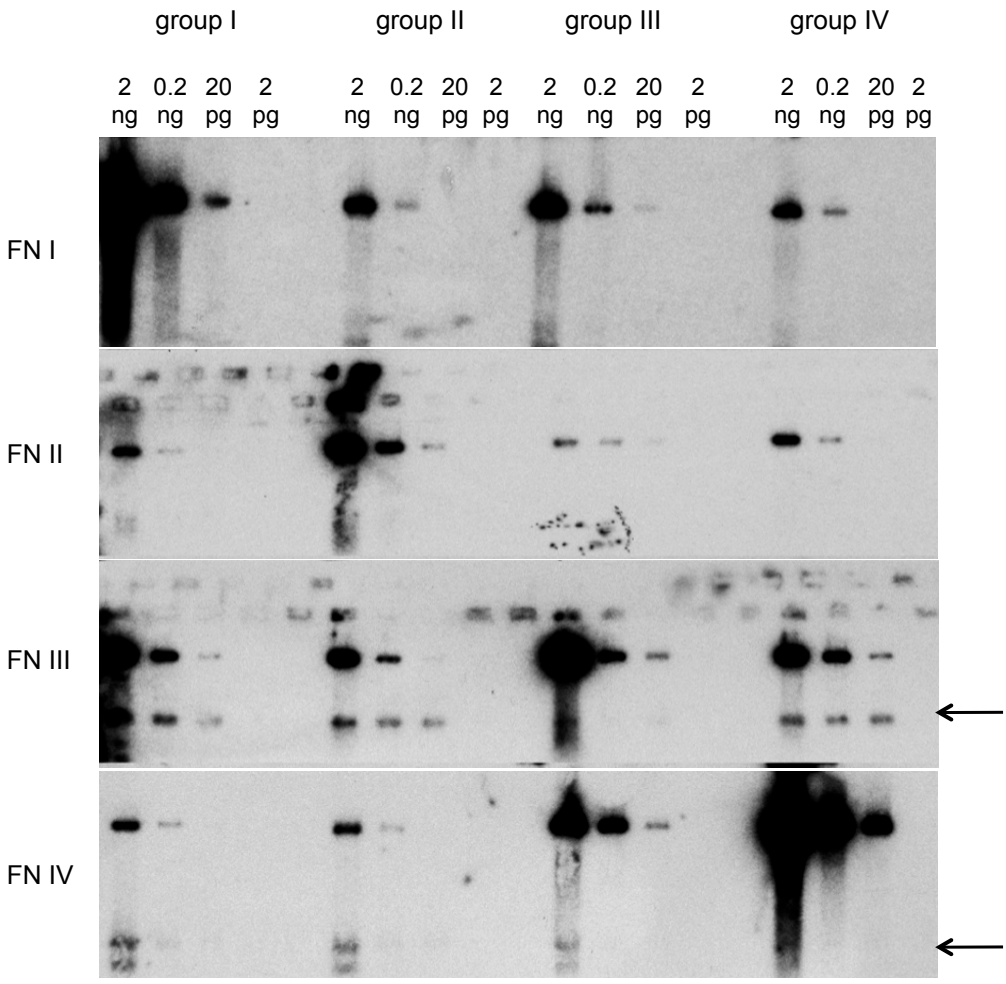

Figure 3.4: Group-specificity of probes FN I to IV. The pGBKT7 plasmids containing FisnaNACHT cDNA of the four nlr groups were liniearized using the restriction enzyme EcoRI. The samples were loaded on four gels in a serial dilution from 2pg to 2ng. The black arrows indicate bands that propably resulted from uncut, supercoiled plasmid DNA.

During the Northern blot analysis the tested DNA probes were used to detect *nlr I-IV* RNA molecules at different stages of development (see figure 3.5). total RNA was isolated from the fish and treated with DNAse to avoid DNA contamination. The same amounts of RNA from every stage was load on each gel. The blots were named group I to IV according to the probes FN I-IV that were used to detect *nlr* RNA transcripts.

According to the results summarized in figure 3.5 *nlr I* genes showed only weak expression at 48hpf and even weaker in adult zebrafish. Both signals that were produced by the probe FN I indicate a size of 8 to 9kb (sizes were estimated with an RNA marker and by sizes of ribosomal RNA which runs at about 1.87kb and approximately 5kb; not shown). For 2wpf embryos and 4wpf larvae no signal was visible.

Probe FN II showed a strong signal at 8kb for the adult zebrafish and a weak band at the same size for embryos 48hpf. The original X-ray films also shows a blurry signal for 2 and 4wpf between 3kb and 9kb, but no distinct band was observed for *nlr II* during these stages. In adult zebrafish two weaker bands were detected between 4.7kb and 5kb in addition to the band observed at 8kb.

nlr III genes are more abundant at all stages analyzed compared to the other *nlr* genes. The highest expression was observed in samples from 48hpf and from adult fish. The strongest signals were obtained at about 8kb, as observed for group II. Faint bands in both samples were detected below 3kb (black arrows) that are not easily visible on the picture but on the original X-rays. At a similar size a signal detected by the FN III probe for larvae 2wpf. For this stage a strong signal at a size of 8kb was not detected in contrast to embryos 48hpf and adult zebrafish. For 4wpf only a low signal was visible but no distinct band.

nlr IV expression was present at 48hpf and in adult fish according to the Northern blot results. One faint band at 8kb was detected in both samples. No signals were picked up by the FN IV probe for the middle stages, 2wpf and 4wpf. Figure 4.3 (B) shows 1□g of total RNA samples that were run in parallel to the gel fort he Northern blot analysis. The two bands show the two subunits of ribosomal RNA, the upper band was the large 28S subunit, the lower band the small 18S subunit. The distinct bands indicate that no RNA degradation procesess were ongoing during that time.

A former performed Northern blot analysis with different total RNA samples of 48hpf, 4wpf (1mpf) and adult fish showed different results for the *nlr II* expression (Figure supp.1 in supplementary data) than shown in figure 3.4. According to these former results the expression of *nlr II* was much stronger at 48hpf than obtained in the Northern blot analysis shown here. The older results showed a similar expression for *nlr II* and *nlr III* at 48hpf and in adult fish. Furthermore, the differences between *nlr II*

and *nlr III* expression in 48hpf embryos compared to adult fish was not as strong as the signal shown in figure 3.5. The results for the groups I and IV were similar to the more recently performed analysis.

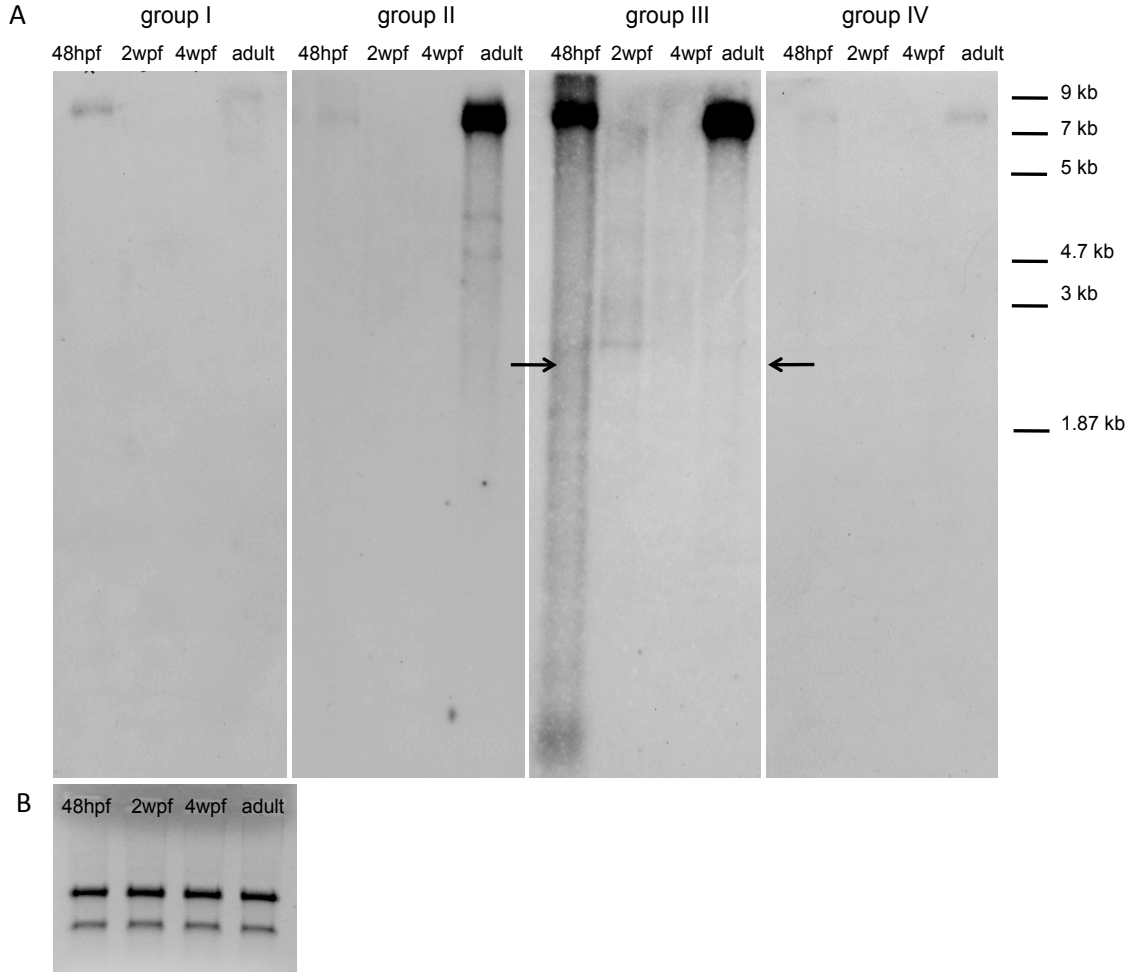

Figure 3.5: *nlr I-IV* expression during zebrafish development and in adulthood analyzed by Northern Blot. (A) and (B) 2μg totalRNA of zebrafish at 48hpf, 2wpf, 4wpf and of adult fish was run on a gel, respectively. (A) Membranes were hybridized with probes for Fisna and NACHT sequences of NLR group I to IV. The black arrows mark faint bands that are better visible in the x-ray films than on the picture**.** (B) shows agarose gel with size-separated total RNA with the large subunit (28S at ~5kb) of ribosomal RNA visible at the higher band and the small subunit (18S at 1.87kb) of ribosomal RNA in the lower band. It shows no degradation process of the RNA samples.

3.2.3 Transcription analyses of fish-specific NLRs after pathogen treatment

To figure out a potential involvement of fish-specific NLRs in pathogen defense mechanisms, zebrafish embryos were injected with either *Eschericha coli* (*E. coli*), *Yersinia ruckeri* (*Y. ruckeri*) or poly IC. *E. coli* is not a natural fish-pathogen and it has been reported that even an infection with 3000 colony forming units (cfu) did not cause

an elevated death rate (Sieger, Stein et al. 2009). This is propably due to an *E. coli*-induced *tumor necrosis factor alpha* (*tnfα*) expression.

Y. ruckeri is a fish pathogen and is a relative of *Y. pestis*, the causative agent for human pest. After injection of 2 to 3 cfu *Y. ruckeri* the expression of cytokines like *tnfα*, *Interferon−φ1* (*ifn-φ1*) and *immunity-related GTPase e3* (*irge3*) was elevated 24 hours post infection (hpi) (Sieger, Stein et al. 2009) and wild type embryos injected with 30 cfu died after 48 h (Aggad, Stein et al. 2010). Poly IC is a synthetic RNA which is often used to mimic virus infection. Furthermore, poly IC injection led to an upregulation of *tnfα* and *ifn-φ1* expression (Levraud, Boudinot et al. 2007; Dios, Romero et al. 2010).

For infections of zebrafish embryos the pathogen-containing solution was injected into the blood system. At around 24hpf the blood circulation starts which is the time point when the embryos were injected. This way injected pathogens were immediately distributed throughout the embryo.

3.2.3.1 RT-PCR analysis on fish-specific NLRs after pathogen treatment

For our experiments 3000cfu *E. coli*, 30cfu *Y. ruckeri* or 1mg/ml poly IC were injected into the caudal vein of zebrafish embryos 24hpf. The mRNA for cDNA synthesis was isolated 3hpi (28 to 30hpf) or 24hpi (49 to 51hpf). Expression of *nlr I-IV* was tested with semiquantitative RT-PCR as described above (see 3.2.1). The results are depicted in figure 3.6.

In 3.6 (A) the expression of *nlr I-IV* and *β–actin* is shown. For *nlr I, II, IV* and *β–actin* the expression level was not altered upon *E. coli* infection within 24 hours. In contrast, embryos 3hpi and 24hpi showed elevated *nlr III* expression compared to the non-treated controls.

Semiquantitative RT-PCR results of *nlr I-IV* and *β–actin* expression after *Y. ruckeri* infection is summarized in figure 3.6 (B). It shows that an infection with *Y. ruckeri* had no effect on any of the four *nlr* gene expressions. The signal for *nlr* II in embryos 3hpi was slightly elavated compared to the non-infected or 24hpi embryos, but for this sample also the *β*-actin signal was stronger than in the other samples.

Figure 3.6 (C) shows *nlr* expression in embryos after poly IC injections and for the non-treated control embryos at the same age. A difference was visible for *nlr I* expression after 3hpi and after 24hpi. The signal difference between 3hpi and control embryos was even stronger than between 24hpi and the control. The results show an elevated *nlr I* expression after poly IC injection. No differences in expression were registered for *nlr II, III* and *IV.*

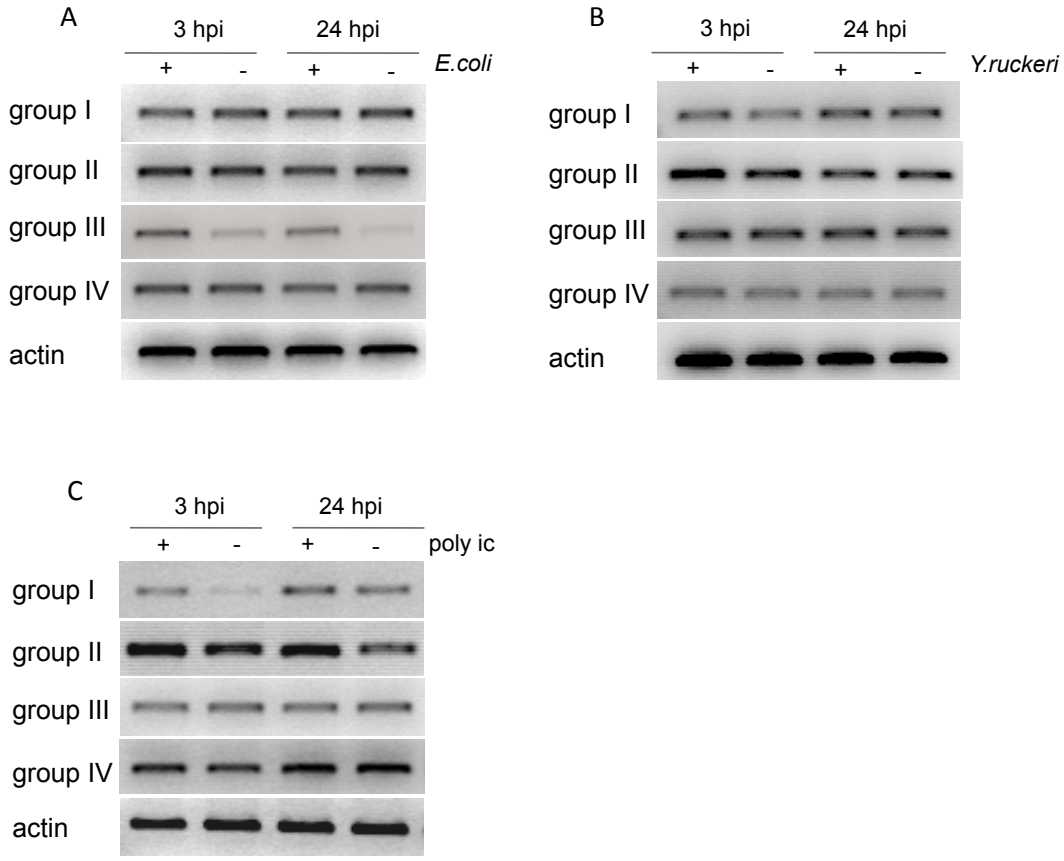

Figure 3.6: *nlr I-IV* expression after *E. coli* (A), *Y. ruckeri* (B) or poly IC (C) injection. mRNA was isolated 3hpi and 24hpi to analyze the effect of pathogen injection and poly IC on *nlr* expression. 12ng of each cDNA sample was amplified in 29 PCR cycles. For PCR protocol see section 2.2.1.1; for primers see table 2. 10μl of each reaction was run on a 1% agarose gel. The picture was taken on a gel documentation system from BioRad.

These results suggest an expressional upregulation of *nlr I* expression 3 hours after poly IC infection and a response on *E. coli* infection of *nlr III* that shows elevated gene expression 3hpi and 24hpi. The results depicted in figure 3.6 represent results obtained from three independently performed experiments with similar outcomes.

3.2.3.2 qRT-PCR analyses on fish-specific NLRs after pathogen treatment

Additional expression analyses of *nlr I-IV* was performed to quantify differences in gene expression observed from semiquantitative RT-PCR results (see figure 3.7). qRT-PCR was performed on embryos 3hpi since *nlr I* only showed a poly IC induced response at this timepoint according to the results obtained before. Relative transcript levels of the *nlr* genes were determined compared to the expression of the endogenous control *β-actin*. The relative gene expression levels in the non-treated embryo were set as reference or calibrator to monitor expression changes of genes in response to

pathogen introduction. The resulting relative fold (Rf) differences of *nlr I-V* were determined as described in 2.2.1.2 and applied on the y-axis of the diagrams. Expression of *nlr* genes in non-infected embryos were set to 1 to have a uniform comparison for all genes ($Rf_{normalized\ control}$ = 1). *nlr I* expression was about 2.2 fold increased upon poly IC injection compared to the non-infeceted control.

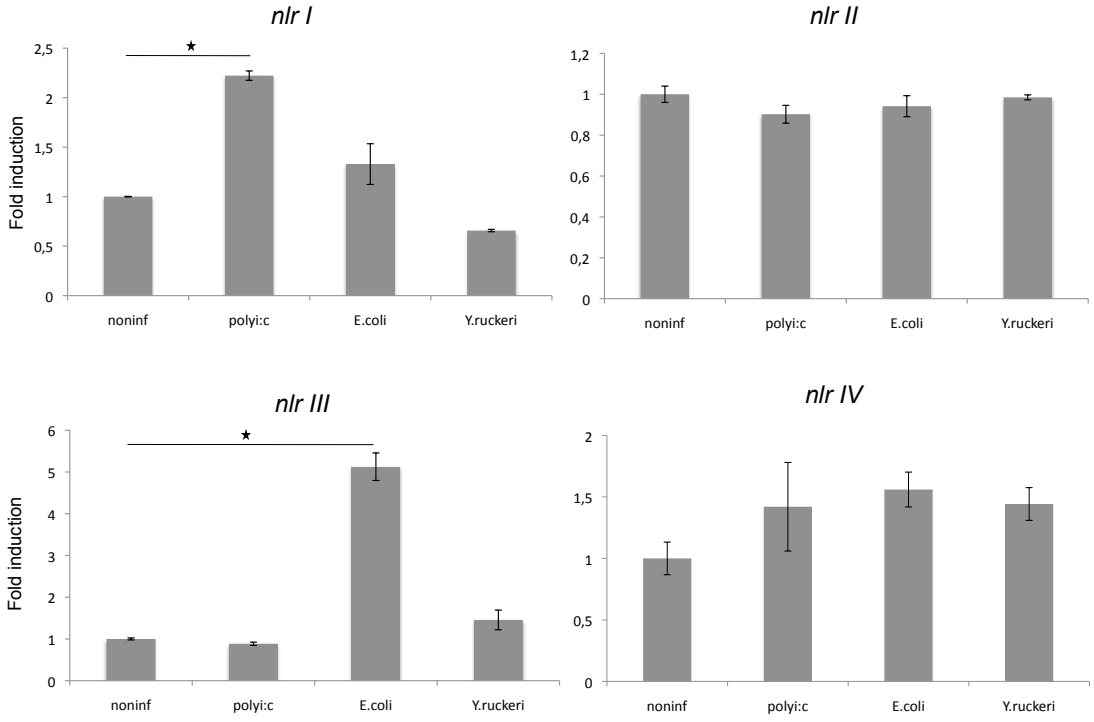

Figure 3.7: Gene expression analysis of fish-specifiv *nlr I-IV* genes after pathogen treatment by qRT-PCR. Gene expression was analyzed with cDNA 3hpi by the SYBR green qRT-PCR method. The expression of nlr *I-IV* genes after injection with *E. coli*, *Y. ruckeri* or poly IC was compared to expression in embryos without pathogen infections. β-actin was used as endogenous control and reference for relative gene expression. Rf differences were determined in comparison to the *nlr* expression level of non-treated embryos. Fold induction is represented in the y-axes. The line with the star indicates a significant difference between two levels of gene expression with a p-value ≥ 0.05 and n = 50.

nlr II gene expression was not changed after any pathogen treatment compared to the non-infected controls.

For *nlr III* the expression was strongly induced after *E. coli* infection. At 3 hpi the level was about 5.1 fold higher than in the non-infected control embryos. This difference in expression level was statistically significant with a p-value ≥ 0.05 (see fig. 3.7).

The results obtained from qRT-PCR confirmed the observations from the semiquantitative RT-PCR experiments. The data shown in figure 3.7 summarize results

obtained from three independently performed experiments. Each cDNA sample was taken from 50 embryos and was analyzed in triplets.
To conclude, the data on *nlr* expression revealed an early maternal expression of all *nlr* groups which is continued during embryonic and larval development. Nevertheless they display a variation in the amount of RNA expression of the different groups. Furthermore, an expression response of *nlr I* after poly IC injection and of *nlr III* after *E. coli* treatment was observed.

3.3 Protein interactions of novel NLR proteins

In order to find interaction partners for fish NLRs that would shed light on putative signalling pathways or protein function a yeast two hybrid screen was performed. For the screen the N-terminal part of a NLR III protein served as the bait (see figure 3.8). To screen for interaction partners a zebrafish cDNA library of one adult fish was used.

3.3.1 cDNA library for the yeast two hybrid screen

The cDNA library, which was used for the yeast two hybrid screen, was generated as described in sections 2.2.3.1.3 and 2.2.3.1.4.
To check the fragment sizes yeast clones from the cDNA library were randomly picked for colony PCR. Most of the fragments ranged between 0.1 to 0.2kb. Overall, colony PCRs of 200 transformed yeast clones (of a total of 806 clones) have been performed which revealed a distribution of cDNA insert sizes of 68% ≤ 0.2kb, 25% ≥ 0.2 to 0.6kb and 7% = 0.8 to 1kb.

3.3.2 Yeast two hybrid screen with N-terminus of NLR III protein

The N-termini of conserved NLRs function as effector binding domains and for most of them a downstream pathway has been identified. Except for group I, for which a pyrin-like domain has been identified, none of the fish-specific NLRs has such a typical domain, as a PYD or a CARD (compare to figures 1.2 and 1.3).
As a bait protein for the screen the N-terminus including the Fisna domain of an NLR III protein was used. Since the Fisna domain is a hallmark for this protein class without an assigned function it was interesting to also look for its interaction partners.
Based on our data NLR group III is the group with the most members which is the reason for choosing one of these as bait (Figure 3.8).

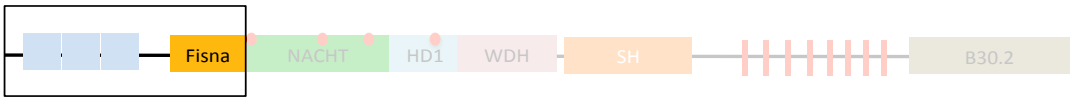

Figure 3.8: Bait construct in Yeast two hybrid screen. The bait was the N-terminus, including the Fisna domain, of an NLR group III protein and is marked by the black frame (aa 1-175). The light blue boxes (on the left) indicate the N-terminal repeats that occur in all group III NLRs in various numbers and in group IV NLRs (compare to figure 1.5.21).

The screen was performed by mating two yeast strains to coexpress bait construct with cDNA library. Yeast cells were spread on three different conditions of stringency to descriminate between weak, medium and strong protein interactions: Histidine represents low nutritional selection, additional selection for l*acZ* expressing cells elevates the stringency and adenine (*ade2* expression) provides strong nutritional selection. The incidence of false-positives was reduced with the increase of stringency. 3 x 10^8 yeast cells containing the bait sequence were mated with 4 x 10^7 yeast cells containing the cDNA library. The cell type with the lower number of cells was the limiting factor of forming diploid cells. Thus one can assume that around 4 x 10^7 diploid cells were screened for interaction partners. Table 3 shows the number of yeast colonies that grew after spreading the mated diploid cells on different selection media. Around 800 colonies appeared after 3 days on the media plates with the lowest nutritional selection (SD-LEU/-TRP/-HIS). After transferring the colonies on SD -LEU/-TRP/-HIS + x-Gal selection media to test the expression of further reporter genes it turned out that 25 out of the 806 clones also expressed *lacZ* but none of them expressed the reporter gene *ade2*.

Table 3: Distribution of yeast colonies on selection media

stringency	selection media	number of colonies
low	SD -LEU/-TRP/-HIS	806
medium	SD -LEU/-TRP/-HIS + x-Gal	25
strong	SD -LEU/-TRP/-ADE	0

The screen results obtained from the lowest stringency conditions are summarized in table supp 1 and in the supplementary data. All clones were sequenced and and blast searches were performed using the ENSEMBL database. 14 of the 25 clones gave not

hit in blast search analyses, neither for zebrafish nor for any different organism. A blast search analysis for one clone resulted in an intron sequence of the *myoblastosis oncogene-like 2* gene. Therefore it was not further analyzed.
Table 4 lists 11 of the 25 clones that showed *his3* and *lacZ* expression. Sequence lengths indicate the number of amino acids of the prey inserts.

Table 4: Results from BLAST searches of 1 sequences from HIS/lac-Z-positive yeast clones.

Group	**Clone**	**BLAST hit**	**Sequence length [bp]**	**Genomic location** (chromosome, nucleotide number)
A	5	Ribosomal protein S12/23	108	11, 44480808 to 44484374 (+)
A	6	Ribosomal protein S12/23	107	11, 44480811 to 44484374 (+)
A	10	Ribosomal protein S12/23	40	11, 44481584 to 44484368 (+)
B	9	Ribosomal protein S20	106	7, 57739579 to 57740887 (+)
B	11	Ribosomal protein S20	106	7, 57739579 to 57740887 (+)
C	2	NACHT NTPase	195	16, 51835993 to 51838105 (+)
C	7	NACHT NTPase	28	22, 26805065 to 26805148 (+)
C	12	NACHT NTPase	37	22, 26805059 to 26805169 (+)
D	17	Thiolase	46	24, 12809963 to 12810100 (+)
E	3	Endopeptidase inhibitor like protein (Sb:cb26)	40	Zv7_scaffold2597 114903 to 115022 (+)
F	4	Proteinase inhibitor 14 (serpina 1l)	118	20, 10342774 to 10342863 (+)

Matches for the same genes or genes which are related to each other were divided into subgroups. For group A, B, and C we found multiple clones. Group A and B each contain two identical clones, and group A and C also contain one clone, which is independent from the others, respectively.

Figure 3.9 shows the amino acid sequences of the identified binding partners. The colored boxes indicate the domains of the proteins annotated by the Pfam database.

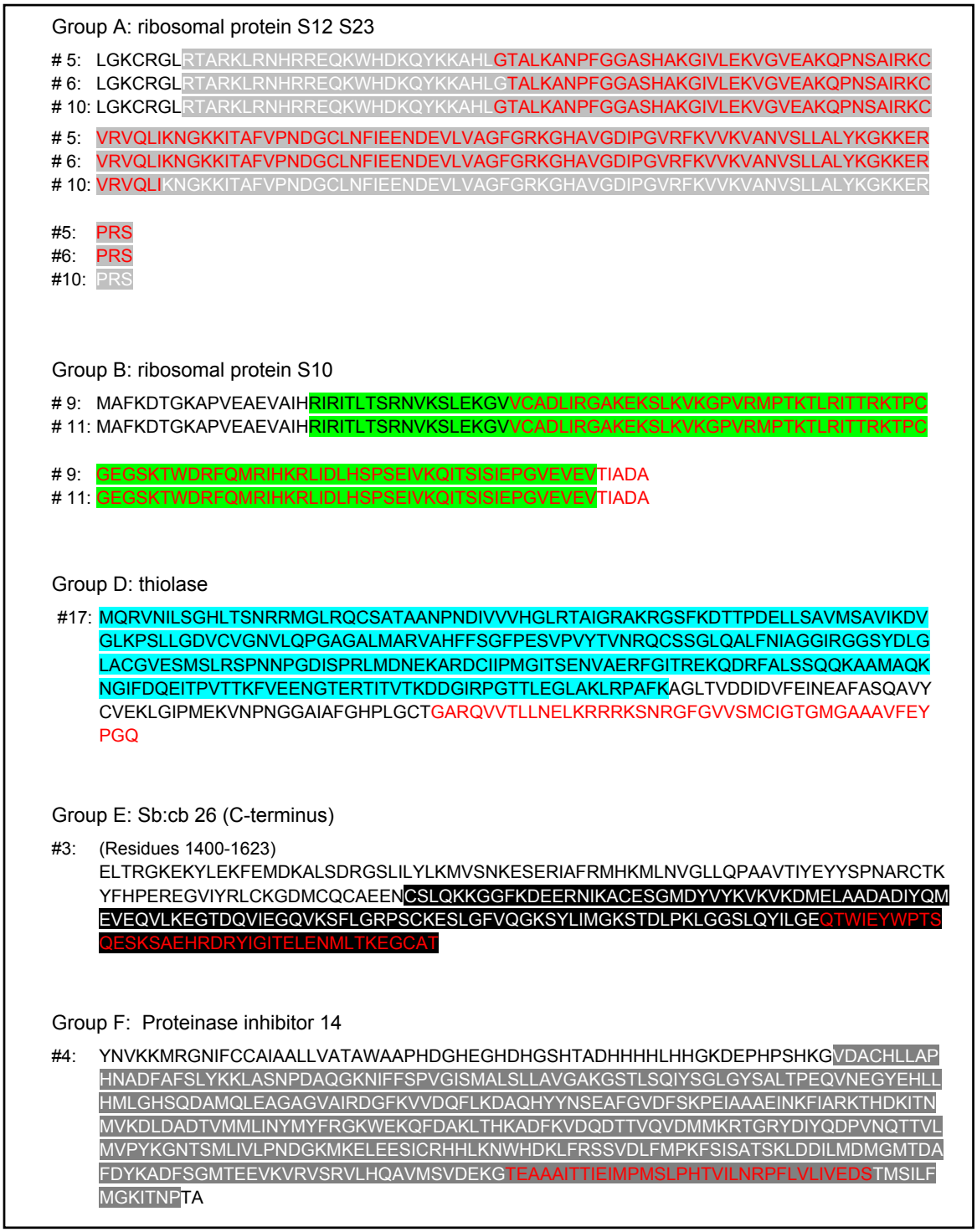

Figure 3.9: Prey sequences within the predicted proteins from the Ensembl and Pfam databases. The prey sequences are indicated by red letters, protein domains are marked in coloured boxes: light gray box: ribosomal_S12_23 Pfam domain; green box: ribosomal_S10 Pfam domain; blue box: thiolase Pfam domain; black box: Netrin C Pfam domain; dark grey box: inhibitor_14_serpin Pfam domain. Red letters: prey sequence.

As mentioned above, for group A and B and C multiple interacting proteins were detected. Group A and B summarize two different types of ribosomal proteins. Group A contains 3 different clones from one single gene, indicated by overlapping sequences (compare to genomic location in the right column) but different sequence length. Group B presents two copies of one identical clone (for both see table 4).

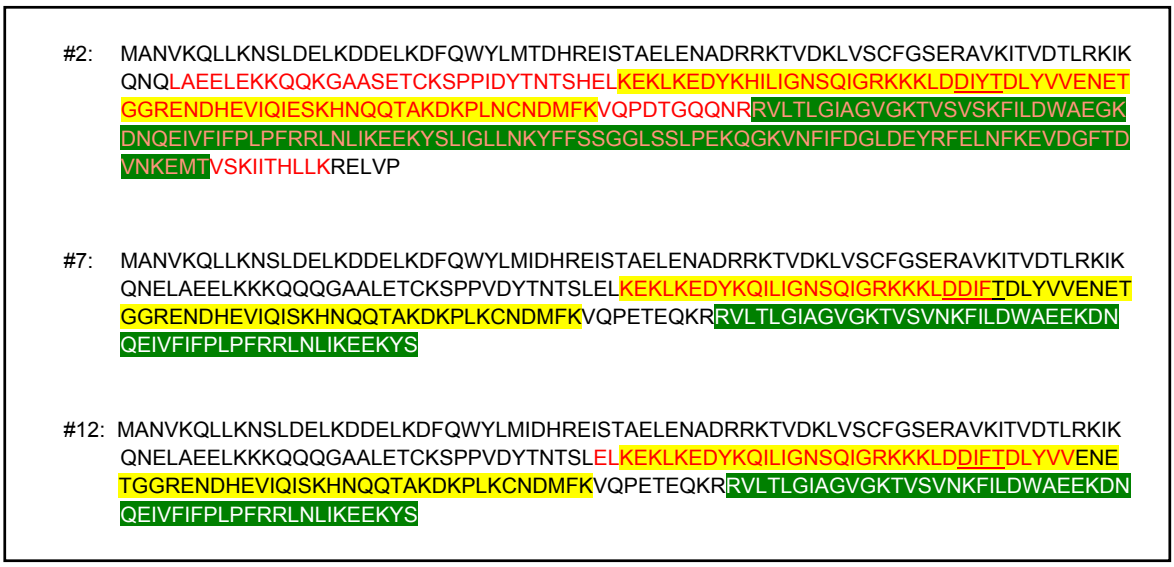

Figure 3.10: Group C – NACHT NTPAse sequences with indicated domains and prey sequences. Prey sequences are written in red letters, the Fisna domain is marked in yellow and the NACHT domain in green. In sequence 2 the entire NACHT domain is included in the prey clone in contrast to sequences 7 and 12 that only show the N-terminal half of the domain. The conserved motif DIY/FT within the Fisna domain was underlined.

As summarized in table 4 the bait peptide bound three times three different NACHT NTPase, which are no. 2, 7 and 12 (shown in figure 3.10). Clone no. 2 is the longest and its gene is located on chromosome 16. Clones no. 7 and 12 represent the same gene. Yet, proteins expressed from both clones bound to the bait protein which minimizes the chance of a false positive result. All NACHT NTPases found in the screen belong to NLR group I which is indicated by its initial 6 amino acids MANVKQ, a unique sequence for group I. For conserved NLRs it has been shown that recruitment and activation of their downstream interaction partners is dependent on self-oligomerization of the NLRs (Rosenstiel and Schreiber 2009). This perfectly fits with the results from the yeast two hybrid screen that NLRs bind to each other. Prey sequence 2 begins 23 amino acids upstream of the Fisna domain and ends 10 amino acids downstream of the NACHT domain. Sequences 7 and 12 overlap only partially with the Fisna domain whereas the length of prey sequence 7 suggests that binding

requires only the indicated 28 amino acids within the Fisna domain. Furthermore, the conserved motif DIY/FT was detected in all prey sequences which is present in all Fisna domains and in human NOD2 (Stein, Caccamo et al. 2007). This result makes the Fisna domain a candidate for a binding-site to the bait protein. For this reason, it was important to analyze further the binding properties between NLRs in order to narrow down a specific interaction site or region.

In addition to the three NLR interaction partners that bound under medium stringency condition 27 NLR group I proteins were identified as interaction partners under low stringency conditions (see supplementary data).

3.3.3 Interactions between fish-specific NLR protein domains

The Y2H screen revealed protein binding between NLR group III and NLR group I proteins. In order to verify this result and to test further binding of construct combinations, interaction specifically between NLR proteins was tested in further yeast two hybrid experiments. Constructs that were used for yeast two hybrid experiments were taken from cDNA isolated from zebrafish embryos to assure that the binding candidates were expressed in the embryo under normal conditions. Interaction between the original bait construct to Fisna domains of group I to IV was tested; the results are summerized in figure 3.11. The picture shows cell growth of yeast cells expressing the bait construct and either an AD fusion construct containing Fisna I, Fisna II, Fisna III or Fisna IV under strong selection conditions (compare to table 3). After transformation of the bait construct and the Fisna-AD construct, the cells were harvested in a medium selecting for the presence of the plasmids (SD-LEU/-TRP) but not for protein interaction. The amount of cells per sample was determined by measuring the optical density (OD) at 600 nm and the cells were plated in a serial dilution (1OD to 0,001OD; 1 OD corresponds to a cell number in 1 ml cell suspension with an optical density measured at 600 nm of 1; OD_{600} = 1). Results depicted in figure 3.12 originate from experiments that were performed in parralel to avoid biassing factors as ages of yeast cells or plates.

As it was obtained in the screen, Fisna I bound to the the bait protein, however, yeast growth was only detected for a cell density of 1OD. Compared to the yeast growth of clones expressing bait construct and either Fisna II or Fisna IV the interaction was about the same as for Fisna I.

Interaction between the bait and Fisna III was much stronger, compared to the Fisna domains belonging to NLR group I, II and IV, as the yeast cell growth was obtained at every step of the serial dilution.

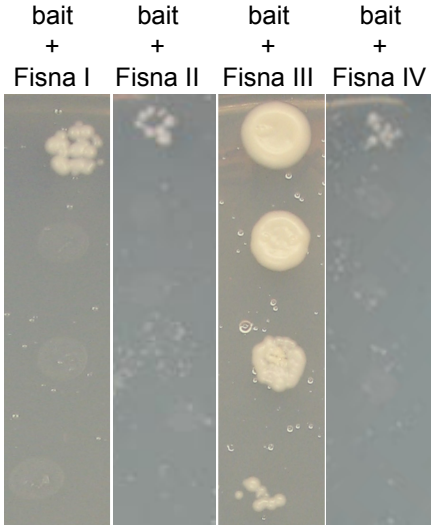

Figure 3.11: Interaction between the bait construct (N-terminus including Fisna III) and Fisna domains of NLR group I, II and IV. Interaction was tested by co-transformation of the bait-DNA-BD construct and the indivated Fisna-AD fusion constructs. From top to bottom a dilution from 1 OD, 0.1 OD, 0.01 OD and 0.001 OD was plateed to compare binding intensities. The selective medium was SD -LEU/-TRP/-ADE.

This result indicates a binding preference between proteins or domains that both belong to the group III of fish-specific NLRs. Since the bait as well as the prey fusion construct contain a Fisna domain we wanted to test if the Fisna domain is sufficient for interaction between the proteins. Furthermore, taken into consideration that the NACHT domain is believed to be responsible for NLR oligomerization (Proell, Riedl et al. 2008) we also tested interaction properties between the NACHT domains of the four groups. In addition to that we checked if the combination of Fisna and NACHT domain altered the results as obtained from the isolated domains.

3.3.3.1 Interaction between Fisna and NACHT domains using yeast two hybrid

Figure 3.12 summarizes the results from comprehensive interaction studies in which binding between domains of all groups were tested with each other. For every group representative members were taken that had as many common sequence features within its group as possible. On the left side interaction studies between Fisna domains are depicted. The middle part shows interaction between the NACHT domains and on the right side the results of binding tests between combined FisnaNACHT is summarized. The yeast cells expressing domains of identical groups are applied diagonal from top left to bottom right.

Yeast cells that expressed Fisna domains from the identical groups grew up to a dilution of 0.001OD on strong selective media, indicating a strong protein binding. The

same results were obtained when interaction between combined FisnaNACHT domains either of group I or II was tested. Yeast cells expressing FisnaNACHT III or IV only grew up to a dilution of 0.01OD. Interaction tests between NACHT domains from the same groups also resulted in cell growth up to 0.01OD. In contrast to the results that were observed from the same groups, tests of protein binding between foreign groups resulted in less cell growth which means weaker protein binding. Interaction between Fisna I and Fisna II and III showed yeast cell growth up to 0.1OD. This was also observed for FisnaNACHT III and IV.

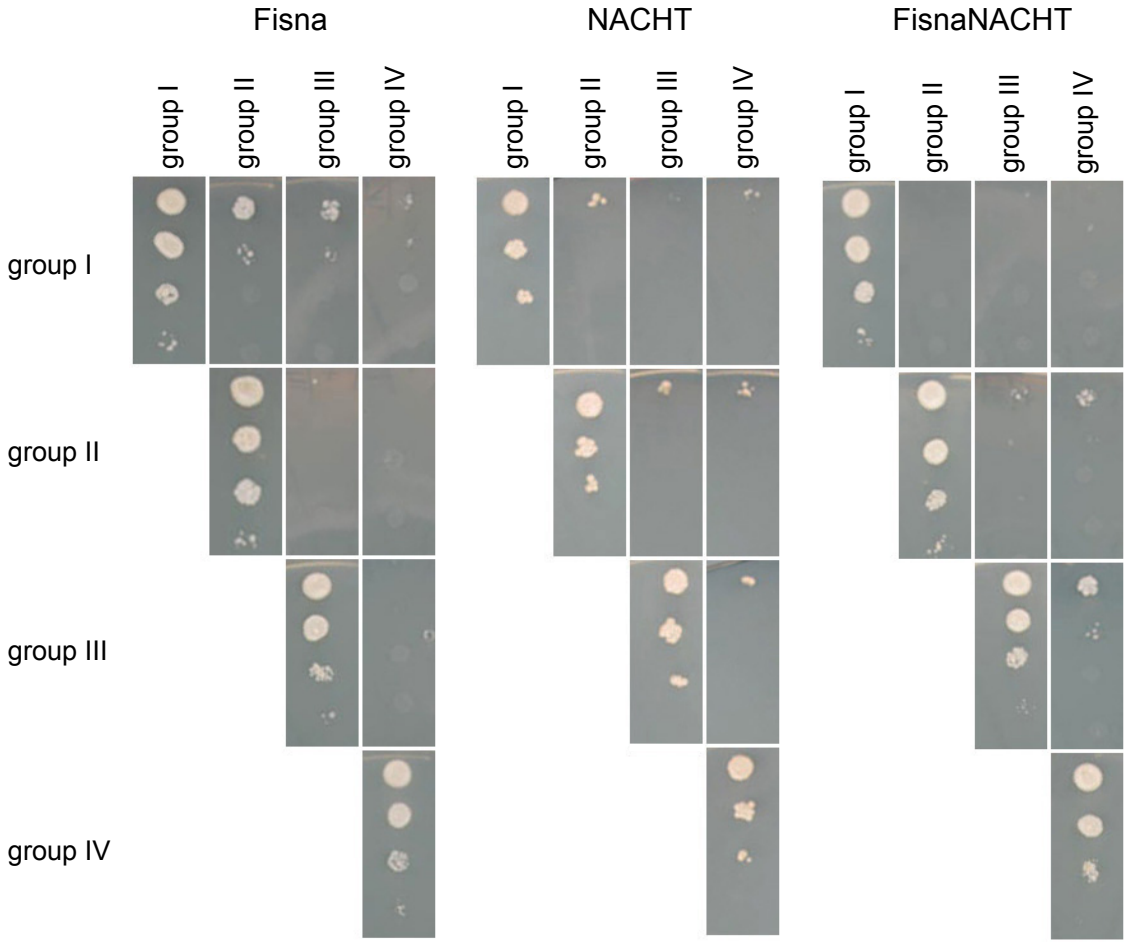

Figure 3.12: Comparative interaction studies on Fisna, NACHT and FisnaNACHT domains of all four NLR groups using yeast two hybrid. Reporter strain AH109 was cotransformed with the respective BD- and AD-fusion constructs expressing the domains of interest. Yeast cells were ad on SD- LEU/-TRP/-ADE in a dilution series starting at the top with 1OD and ending at the bottom with 0.001OD.

The remaining tests between domains of different NLR groups resulted in no cell growth at all or very few at 1OD. From these results it was concluded that protein

binding between domains of the same groups was always stronger compared to interaction between foreign groups.

3.3.3.2 Interaction between identical FisnaNACHT domains using co-immuniprecipitations

To validate the results for protein interaction co-immunoprecipitations (co-IPs) were performed. By this means the possibility of false-positive results from the yeast two hybrid experiments could be excluded. As the yeast two hybrid results did not show significant differences of binding between separated Fisna and NACHT domains or both domains combined, co-IPs were performed for the combined FisnaNACHT domains. The BD-fusion constructs and AD-fusion constructs that were used for the yeast two hybrid experiments expressed a myc-tag or a HA-tag at their N-terminus, respectively. Therefore, co-IP experiments could be performed in yeast cells using an anti-HA and an anti-myc antibody to test binding between two FisnaNACHT constructs. Interactions between domains from same groups and between different NLR groups were tested in the same combinations as during the yeast two hybrid experiments. The results are presented in figures 3.13 to 3.17, parts (A) to (D).

Parts (A) and (B) or (C) and (D) always show corresponding immunoprecipitation results of one protein pair. Parts (A) and (C) show the results of the IP performed with an anti-myc antibody and protein detection by two Western blot analyses using an anti-HA antibody (the upper row) and an anti-myc antibody (the lower row). (B) an (D) show the IP and Western blot results performed vice versa. The first columns on the left show the detection of the tagged FisnaNACHT domains by the used antibodies during Western blots, referred to as input. The second column shows the precipitated protein from the IP experiment. Column three shows a negative control for the IP (see below) and columns four and five indicate protein abundance in the supernantants of the IP or the negative control. Compared to the IP sample only one fifth of the lysates were taken for the input sample and for the supernatants. This caused an about five times stronger signal for the IP samples (second row from the left). Compared to the protein marker the bands appeared at the expected size of 50kDa.

Figures 3.13 and 3.14 show the co-IP results of FisnaNACHT constructs originating from identical NLR groups. In 3.13 binding between two of each FisnaNACHT I and II domains is depicted. Figure 3.14 shows the results for interaction between constructs of groups III and IV, respectively. The IPs performed with FisnaNACHT constructs originating from the same NLR groups co-precipitated in all the experiments. Western blot analyses on the IP samples (second columns) detected both tagged versions of

the FisnaNACHT constructs in every case. For the negative control experiments yeast cells expressing only one plasmid were used. In the case of performing an IP with an anti-myc antibody ((A) and (C) in figures 3.13 and 3.14), the control cells were only transformed with the plasmid carrying the HA-tagged version of the FisnaNACHT construct. For a HA-based IP the myc-tagged version of the construct was used for the negative control (see figures 3.13 and 3.14 (B) and (D)). The negativ control showed no signal at the size of 50kDa in Western blot analyses indicating no unspecific binding of the antibody to the different tagged FisnaNACHT construct during the IP.

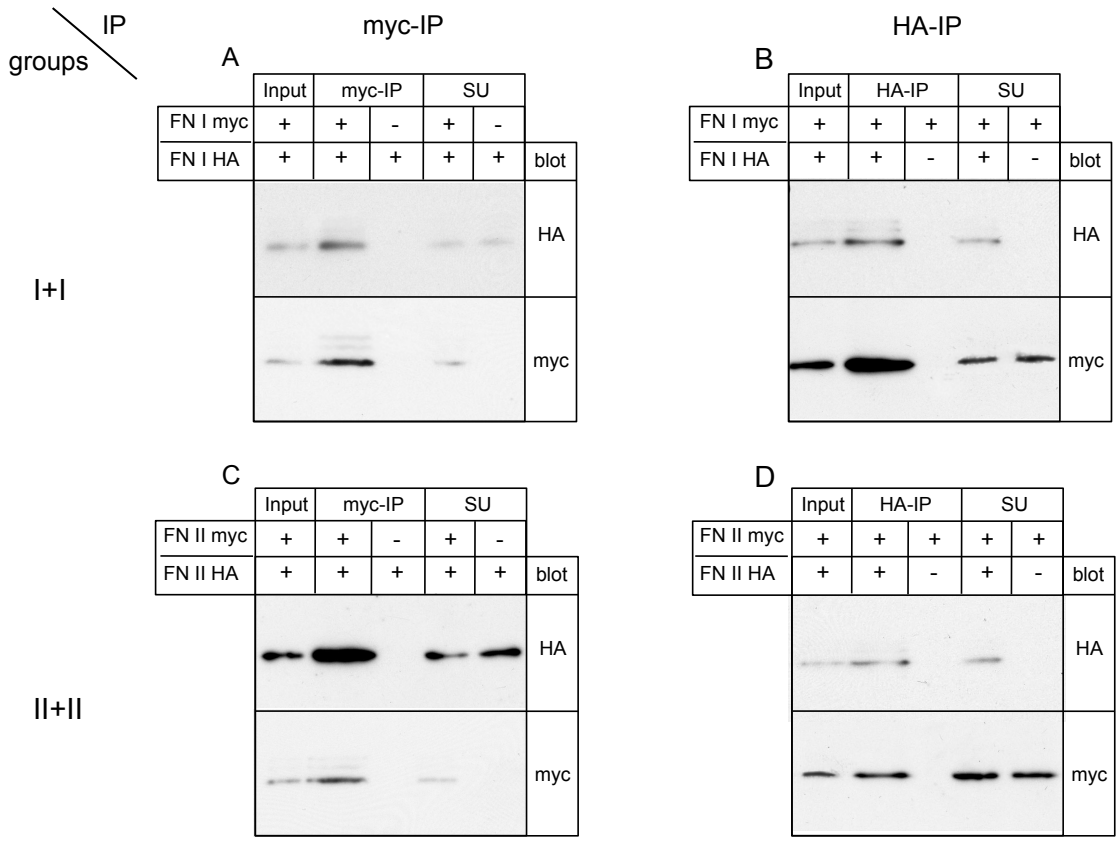

Figure 3.13: Co-IP experiments for FisnaNACHT I and FisnaNACHT II. (A) Co-IP for FisnaNACHT I myc and FisnaNACHT I HA constructs performed with an anti-myc antibody. (B) Co-IP for FisnaNACHT I myc and FisnaNACHT I HA constructs performed with an anti-HA antibody. (C) and (D) show the corresponding experiments with FisnaNACHT II myc and FisnaNACHT II HA constructs. Input, IPs and SU samples were analyzed using western blot analyses. The labels below 'blot' indicates the antibody used for the Western blots. + and - indicates if the yeast cells contained the constructs depicted left. SU is supernatant.

Supernatants (SU) from the IPs (fourth columns), including the negative control (fifth columns), were also used for western blot analyses. The signals obtained from the supernatants show the amount of proteins left after the IP and thus the IP effeciency.

The more protein is pulled out from the input the less remains in the supernatant and the higher is the IP efficiency.

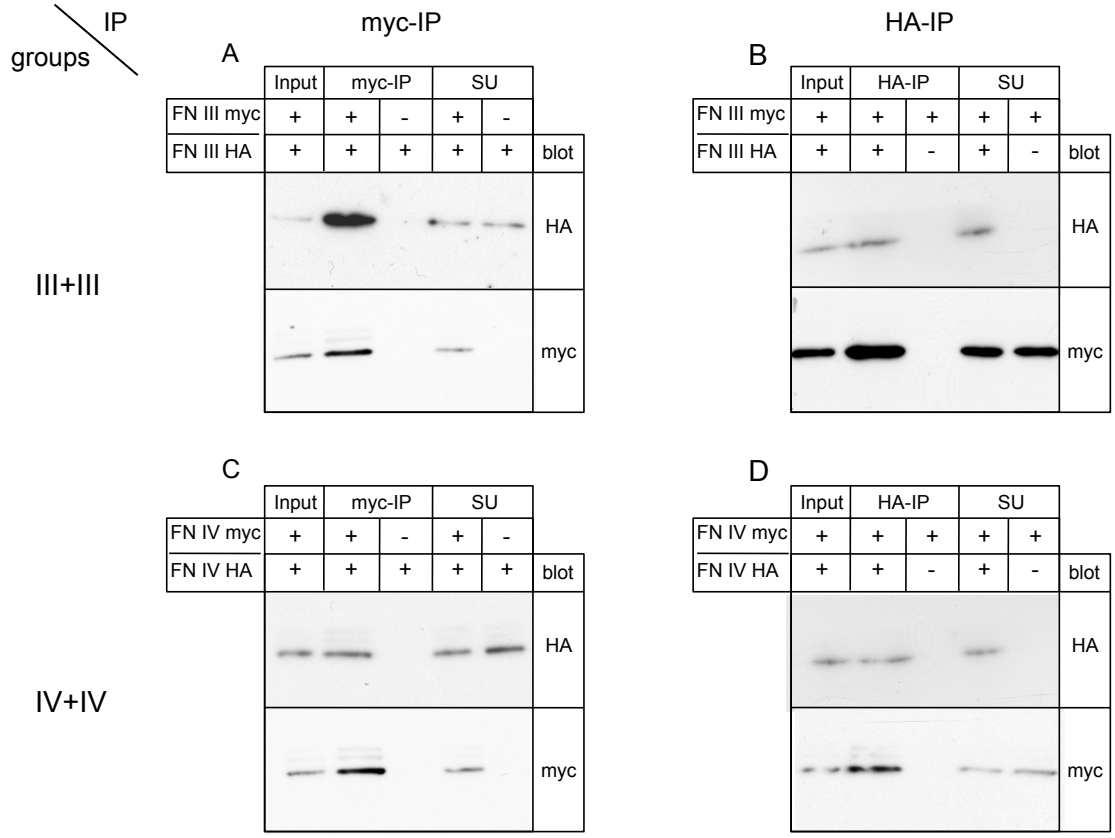

Figure 3.14: Co-IP experiments for FisnaNACHT III and FisnaNACHT IV. (A) Co-IP for FisnaNACHT III myc and FisnaNACHT III HA constructs performed with an anti-myc antibody. (B) Co-IP for FisnaNACHT III myc and FisnaNACHT III HA constructs performed with an anti-HA antibody. (C) and (D) show the corresponding experiments with FisnaNACHT IV myc and FisnaNACHT IV HA constructs. Input, IPs and SU samples were analyzed with Western blot analyses. For labelling compare to figure 3.13.

The results obtained from co-IP experiments testing interactions between FisnaNACHT constructs originating from the same NLR groups showed binding between all tested candidates. By this means the outcome regarding strong protein interactions from yeast two hybrid experiments were confirmed: NLRs bind group-specifically through their Fisna and NACHT domains.

3.3.3.3 Interaction between different FisnaNACHT domains using co-immuniprecipitations

Aside from testing protein binding between identical groups, FisnaNACHT constructs were also tested for group-unspecific binding meaning interactions between all groups

were tested by co-IPs. Figures 3.15 to 3.17 show co-IPs for every combinations between different NLR groups. The order and way of representing the results is the same as in figures 3.13 and 3.14.

The antibody signals in the input in figures 3.15, 3.16 and 3.17 indicated expression for all transformed FisnaNACHT constructs. While the yeast two hybrid experiments indicated weak interaction between FisnaNACHT II and IV and between FisnaNACHT III and IV (see above) the co-IPs showed no interaction between FisnaNACHT domains from different NLR groups. The negative controls indicated no unspecific binding of the antibodies. The input signal in every experiment was similar to the signal of the supernatant indicating that the same amount of protein was present after the IP. This also shows that none of the proteins was pulled out during the IP.

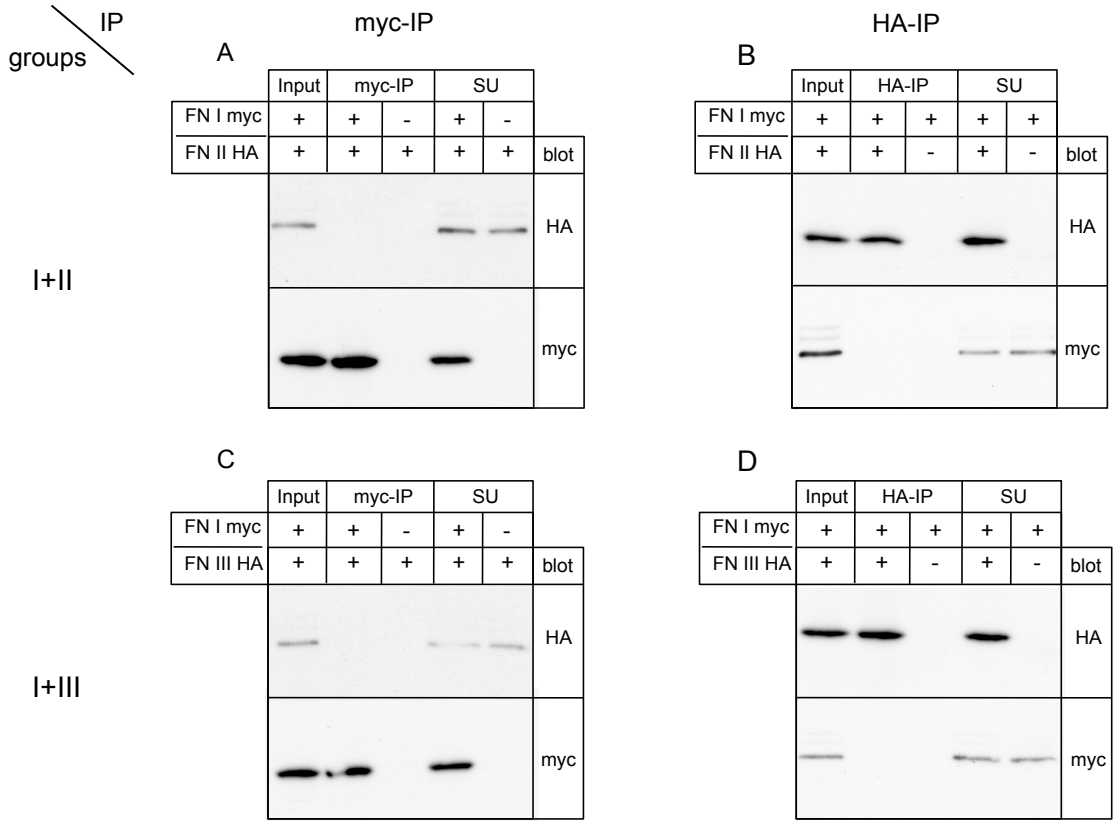

Figure 3.15: Co-IP experiments for FisnaNACHT I with II and FisnaNACHT I with III. (A) Co-IP for FisnaNACHT I myc and FisnaNACHT II HA constructs performed with an anti-myc antibody. (B) Co-IP for FisnaNACHT I myc and FisnaNACHT II HA constructs performed with an anti-HA antibody. (C) and (D) show the corresponding experiments with FisnaNACHT I myc and FisnaNACHT III HA constructs. Input, IPs and SU samples were analyzed using Western blot analyses. For labelling compare to figure 3.13.

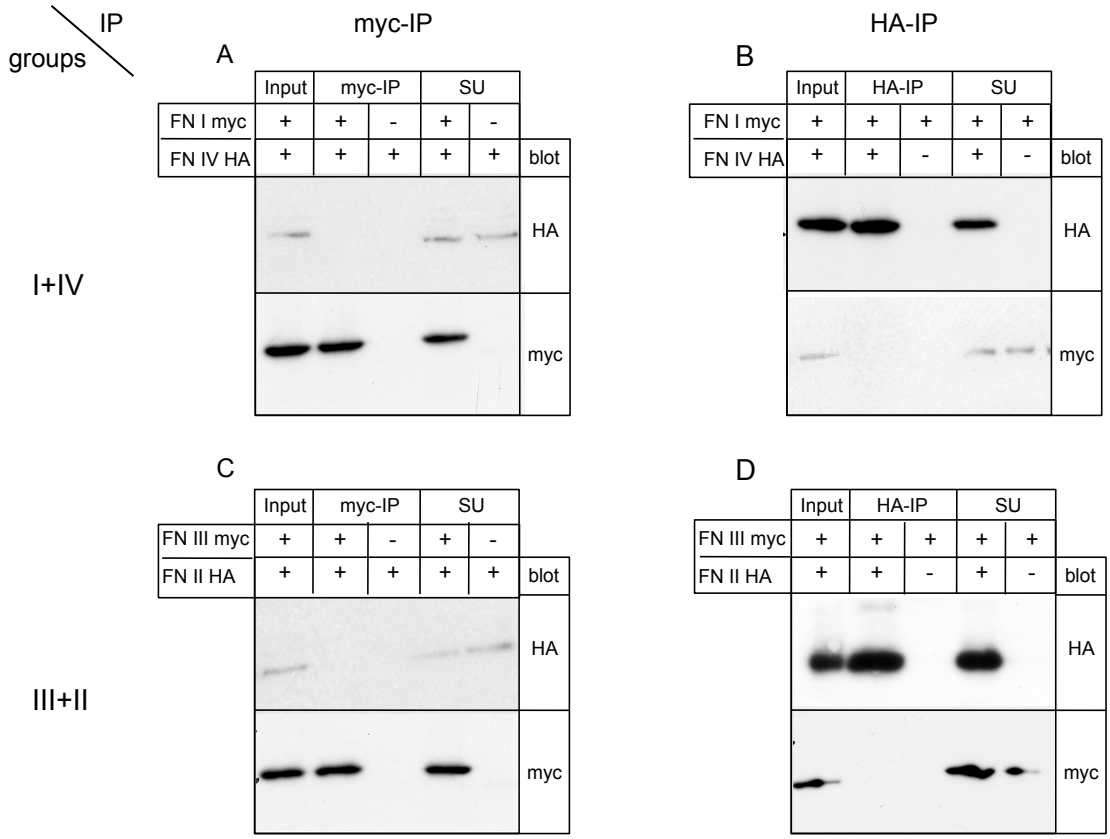

Figure 3.16: Co-IP experiments for FisnaNACHT I with IV and FisnaNACHT III with II. (A) Co-IP for FisnaNACHT I myc and FisnaNACHT IV HA constructs performed with an anti-myc antibody. (B) Co-IP for FisnaNACHT I myc and FisnaNACHT IV HA constructs performed with an anti-HA antibody. (C) and (D) shows the corresponding experiments with FisnaNACHT III myc and FisnaNACHT II HA constructs. Input, IPs and SU samples were analyzed using Western blot analyses. For labelling compare to figure 3.13.

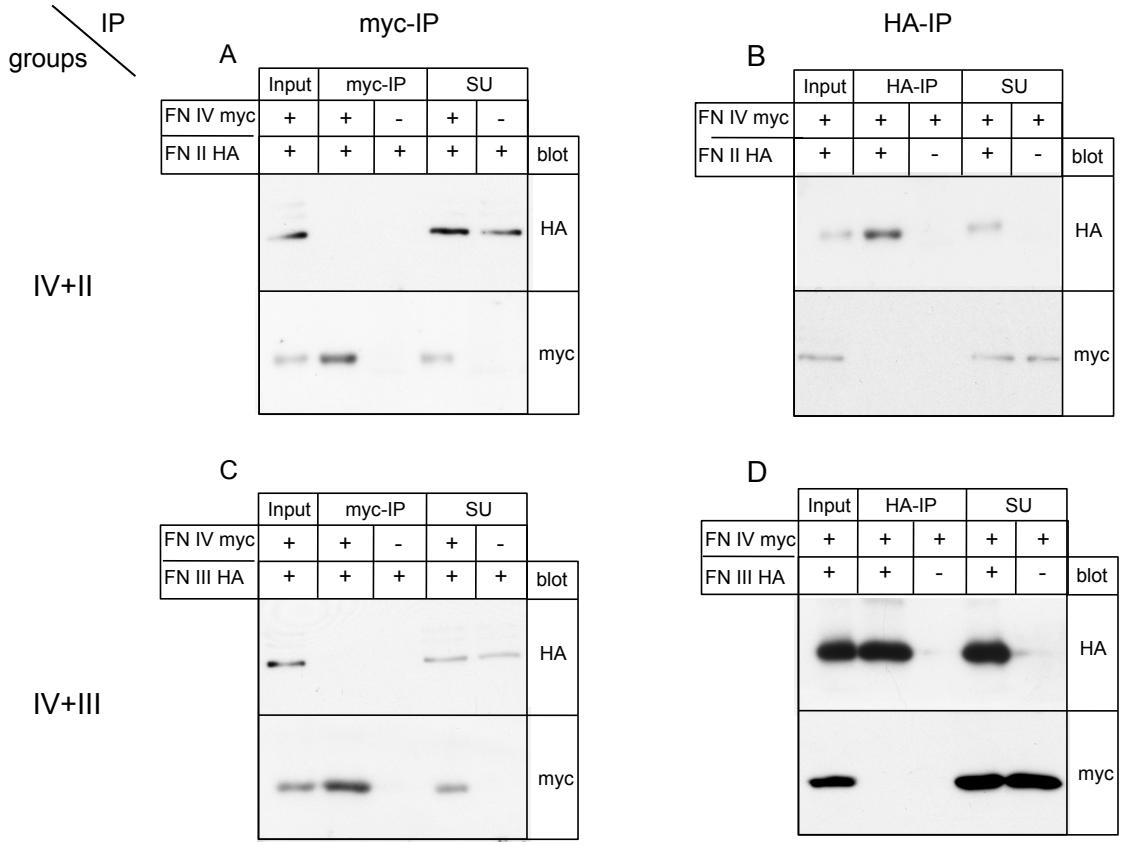

Figure 3.17: Co-IP experiments for FisnaNACHT IV with II and FisnaNACHT IV with III. (A) Co-IP for FisnaNACHT IV myc and FisnaNACHT II HA constructs performed with an anti-myc antibody. (B) Co-IP for FisnaNACHT IV myc and FisnaNACHT II HA constructs performed with an anti-HA antibody. (C) and (D) show the corresponding experiments with FisnaNACHT IV myc and FisnaNACHT IIV HA constructs. Input, IPs and SU samples were analyzed using western blot analyses. For labelling compare to figure 3.13.

To sum up results of all interaction studies, Fisna domains as well as NACHT domains are able to interact with the respective domains of identical NLR groups. This fact is not changed by the combination of both domains, here referred to as FisnaNACHT. Furthermore, the protein interaction was specific between identical NLR groups. For a summary of all tested interactions also see figure 3.18. Three “+“ indicate strong interaction between the tested fragments obtained by yeast two hybrid experiments. One ″+“ stands for a weak interaction according to the tests in yeast two hybrid and ″-“ means no interaction observed. Fields with grey backgrounds indicate that these interactions were verified by co-IP experiments. Only interactions that were classified as strong during yeast two hybrid have been validated.

These results raised the question how the group-specificities of NLR interactions were mediated.

		HA GAL4 transcription activation domain fusion											
		Fisna I	Fisna II	Fisna III	Fisna IV	NACHT I	NACHT II	NACHT III	NACHT IV	FisnaNACHT I	FisnaNACHT II	FisnaNACHT III	FisnaNACHT IV
myc GAL4 DNA-binding domain fusion	Fisna I	+++	+	+	+	n.t.				++	-	-	+
	Fisna II	+	+++	-	-					-	++	-	-
	Fisna III	+	-	+++	-					-	-	++	-
	Fisna IV	+	-	-	+++					+	-	-	++
	NACHT I	n.t.				+++	+	-	+	++	+	+	-
	NACHT II					+	+++	+	-	+	++	-	-
	NACHT III					-	+	+++	+	+	-	+++	-
	NACHT IV					+	-	+	+++	-	-	-	+++
	FisnaNACHT I	n.t.				n.t.				+++	-	-	-
	FisnaNACHT II									-	+++	+	+
	FisnaNACHT III									-	+	+++	+
	FisnaNACHT IV									-	+	+	+++

Figure 3.18: Summary of protein interactions between NLR domains of all four groups. - indicates that no protein interaction was observed, + means that cell growth was observed for 1OD and +++ means that yeast cells grew at least until the dilution of 0.01OD. Cells with a grey coloured background indicate that protein binding was verified by co-IP experiments. n.t means not tested.

3.3.4 Searching for NLR binding sites

To narrow down sites within tested domains that are responsible for protein interaction we first gathered information about interaction studies obtained from conserved NLRs that could be applicable to our NLRs.

A few years ago the group of R. Schwarzenbacher published structure and sequence analyses on conserved NLR proteins (Proell, Riedl et al. 2008). These studies revealed that most NLRs and their pioneer, APAF1, share many secondary structure features (Proell, Riedl et al. 2008). Mechanistic data that were available for APAF1 was linked to conserved features of NLRs and homologous interface residues were identified that most likely are responsible for intermolecular interaction between NLRs during oligomerization. They identified 6 sequence stretches that, according to the predicted

structure, either contributes to interactions with the left partner (No. 1 and 3) or with the right partner (No. 2, 4, 5 and 6) of an NLR-oligomer.

We compared these features to the fish NLRs and identified interface-residue candidates that may play a role in protein interactions (see figure 3.19). In figure 3.19 only interface residues (No. 2 to 6) located within the NACHT domains I to IV were included. The idea was to examine if these four interaction-site candidates mediate the binding between two interaction partners or rather are responsible for group-specific interactions between fish-specific NLRs. In figure 3.19 it is shown that residues 3 and 4 are located upstream of the Walker B motif and no. 5 and 6 are located more towards the C-termini of the domains, downstream of the Walker B motif. For residue groups no. 5 and 6 a 33.33% to 50% identity applies throughout the four NACHT domains while 3 and 4 show no identities.

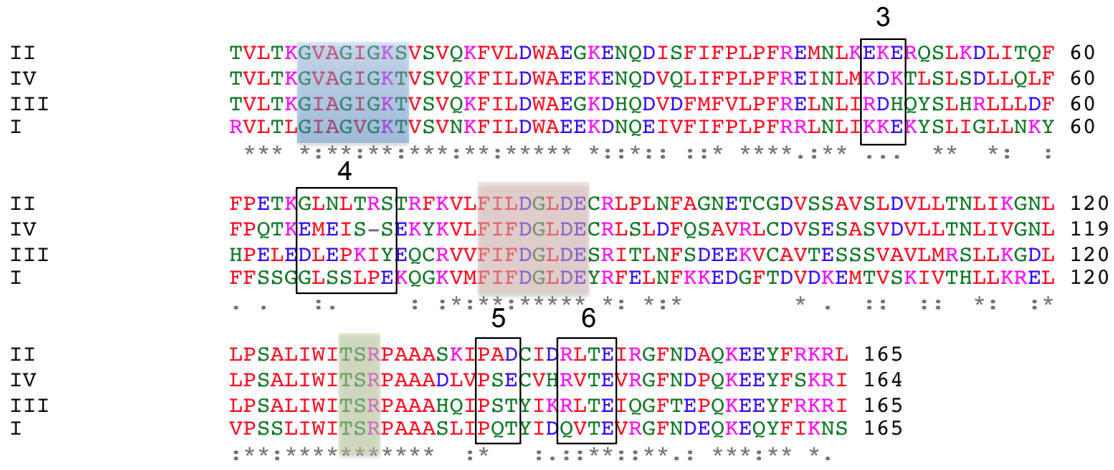

Figure 3.19: Protein sequence alignment of NACHT domains of NLR groups I to IV. Walker A motif, Walker B motif and sensor 1 are indicated by blue, rose and green boxes, respectively. Numbers on the right side show the number and position of amino acids, asterisks mean 'identical' and ':' indicate the conserved substitutions. Black frames 3 to 6 mark interaction-site candidates.

3.3.4.1 Generating hybrid NACHT constructs to narrow down binding sites

To analyze the impact of sites no. 3, 4 and 5, 6 24 hybrid NACHT constructs were generated in a way that the part upstream of the Walker B motif belonged to a different NLR group than the part downstream of Walker B. The cloning strategy for two hybrid NACHT domains is depicted in figure 3.20. In the first step the N-terminal part of one NACHT domain and the C-terminal part of a different native NACHT domain was amplified using PCR. For amplification of the N-terminal half the reverse primer included the Walker B motif and had an additional sequence of a second NACHT domain included. For the C-terminal half the forward primer added a 5' overhang

sequence that belonged to the first NACHT domain. In a second step the non-native overhang was filled up with complementary nucleotides during the PCR. The third step included an additional PCR using the product of the initial PCR as template so much that the new 5' and 3' overhangs were filled up with nucleotides and the hybrid NACHT domain was completed.

Hybrid NACHT domains were designed and generated in every possible combination for all four groups (see figures 3.20 and 3.21). This way binding properties between native NACHT domains and both halfs of all four NACHT domains could be tested independently. Binding of hybrid NACHT constructs to native NACHT domains was checked using yeast two hybrid.

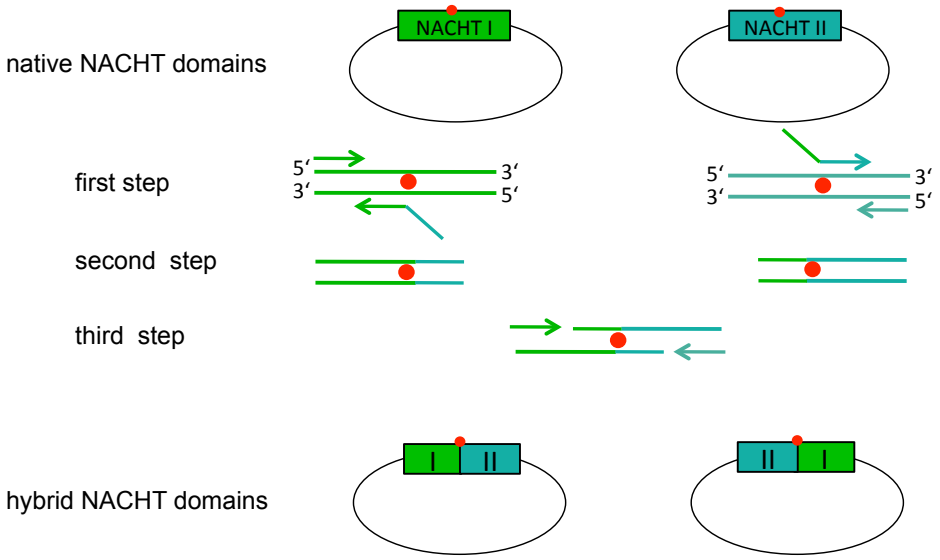

Figure 3.20: Cloning scheme for hybrid NACHT constructs using NACHT domains of NLR groups I and II as an example. The upper part shows native NACHT I and II clones as they were used for previous yeast two hybrid experiments and here served as templates for designing hybrid NACHT domains. First, second and third step indicate the two PCR procedures as described in the text above. The red circles indicate the Walker B motif which is used as boundary between the N-terminal an C-terminal part of the hybrids. The arrows indicate primers and the colours represent the affiliation of the sequences: light green depicts NACHT I and dark green marks NACHT II. The black cirkles associated to the NACHT domains indicate the plasmids that were used for yeast two hybrid experiments.

3.3.4.2 Yeast two hybrid experiments with hybrid NACHT constructs

To investigate the impact of binding site candidates no. 3 to 6 (see above) protein binding was tested between the four native NACHT construct and every hybrid NACHT construct. The names for the hybrid constructs are composed of two roman numerals such as the first one indicates the origin of the N-terminal half and the second one the C-terminal half. Native NACHT I, II, III and IV in the prey constructs was here referred

to as I-I, II-II, III-III and IV-IV and binding of each native NACHT domain to itself was tested again for comparison to the binding of native NACHT domains with the respective hybrid constructs. The co-transformed yeast cells were plated on strong selection medium (SD-LEU/-TRP/-ADE) in a serial dilution from 1OD to 0.001OD (from left to right) as it was done in previous yeast two hybrid experiments.

Figure 3.21 (A) shows the results of yeast two hybrid experiments between native NACHT I and the six NACHT I containing hybrids. The prey constructs expressing NACHT hybrids with the NACHT I N-terminal half showed strong protein binding to the native NACHT I domain since the yeast cells were growing from 1OD to 0.01OD.

In contrast to the N-terminal half of NACHT I the C-terminal half mediated only weak protein binding to native NACHT I as it was reflected by yeast cell growth only in the low dilutions: While yeast cells expressing the NACHT hybrid domain with NACHT I in the N-terminal half grew up to a dilution of 0.01OD, the cells expressing the C-terminal NACHT I hybrid showed reduced growth that was restricted to 1OD, or in some cases to 0.1OD.

In figure 3.21 (B), (C) and (D) the results of corresponding experiments using native NACHT II, III and IV, respectively, were shown. Binding of every native NACHT domain was tested to NACHT hybrids that included one half of the native NACHT domain. In addition, yeast two hybrid experiments with the native NACHT domain as a bait as well as a prey was performed in parallel to have a direct comparisons of yeast cell growth.

In general, the results that were obtained from native NACHT I with NACHT I containing hybrids could also be applied to interaction with NACHT II, III and IV.

In cases with equal N-terminal halfs for the native NACHT domain and the hybrid NACHT domain, protein interaction was much stronger than with identical C-terminal parts. In the first mentioned case cell growth was observed up to the highest dilution of 0.001OD, as it was observed for NACHT II and II-I hybrid (II x II-I) and II x II-III. Furthermore, it was also true for III x III-I, III x III-II, IV x IV-I and IV x IV-II (see figure 3.21 (B)-(D)).

Comparison between interaction of two native proteins with interaction between a native NACHT domain and an the NACHT hybrid, including the identical N-terminal half as in the bait, cell growth was about the same indicating an equal strength of protein binding. The results suggest that the N-terminal half of the NACHT domain mediates group-specific interaction. If binding site candidates no. 3 and 4 are responsible for the group affinitiy has to be analyzed further.

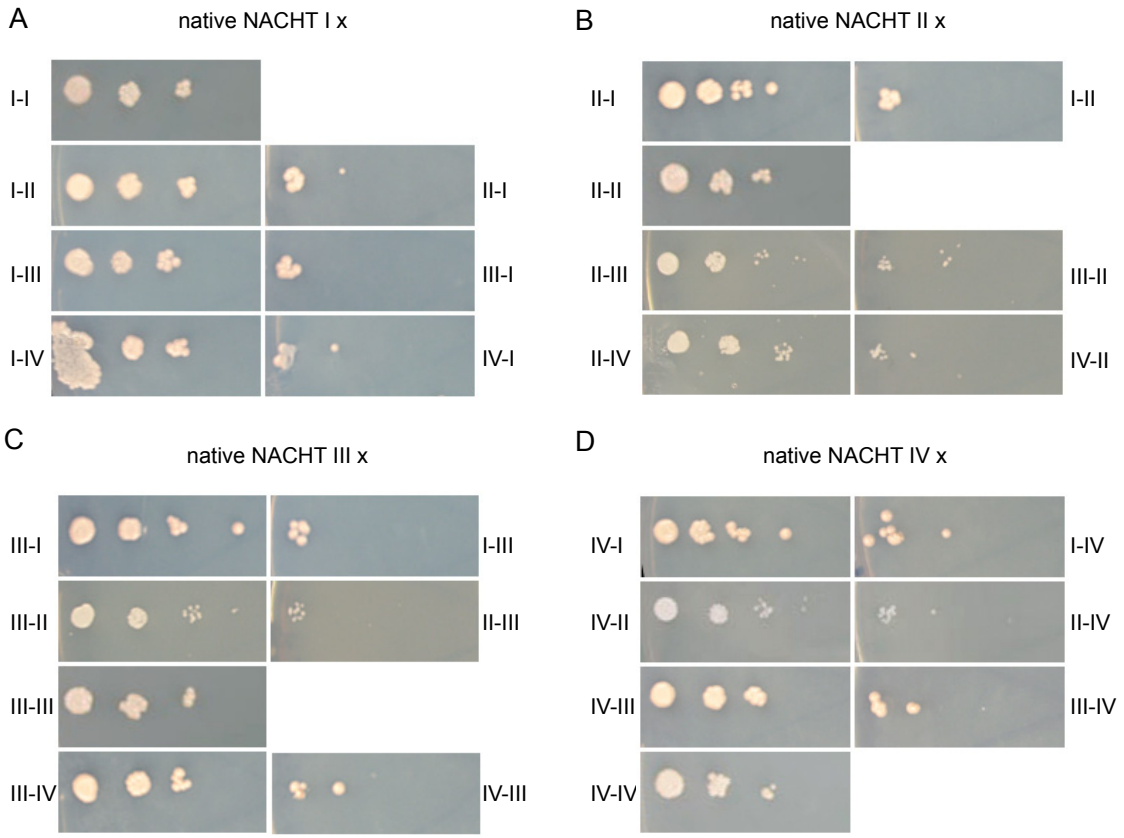

Figure 3.21: Yeast two hybrid results from testing protein binding between native NACHT domains and hybrid NACHT domains. The composition of each hybrid NACHT domain is reflected by its nomenclature: The first roman numeral indicates the origin of the N-terminal half of the hybrid and the second numeral indicates the C-terminal half. (A) Protein binding tested between the native NACHT I domain and hybrid NACHT domains, containing either the N-terminal ort he C-terminal part of the NACHT I domain. (B) Protein binding tested between the native NACHT II domain and hybrid NACHT domains, containing either the N-terminal or the C-terminal part of the NACHT II domain. (C) Protein binding tested between the native NACHT III domain and hybrid NACHT domains, containing either the N-terminal ort he C-terminal part of the NACHT III domain and (D) protein binding tested between the native NACHT IV domain and hybrid NACHT domains, containing either the N-terminal ort he C-terminal part of the NACHT IV domain. Yeast two hybrid experiments were performed by cotransformation of the bait and prey constructs into AH109 reporter strain. The bait included the native NACHT domain and the prey contained the hybrid NACHT domain. Cotransformed yeast cells were plated on strong selection medium from the left side to the right side a serial dilution from 1OD to 0.001OD. Redundant pictures were left out.

4. Discussion

The zebrafish embryo became a popular model organism to study components of the innate immune system. In contrast to mammalian systems the zebrafish offers the possibility to investigate the innate immune system in the absence of the adaptive immunity. Many components and pathways that have been studied in mammals are conserved in the zebrafish and are present at early developmental stages in the embryo (Chluba, Jault et al. 2004; Stein, Caccamo et al. 2007; Matsuo, Oshiumi et al. 2008; Stein 2010). Very recently a group from New Zeeland has demonstrated an antibacterial role for zebrafish NOD1 and NOD2 (Oehlers, Flores et al. 2011).
Yet, in some cases functional differences have been uncovered (Sepulcre, Alcaraz-Perez et al. 2009) and for many factors that are associated to mammalian innate immune functions a contribution in defence strategies has not been found or has not been investigated in zebrafish.

4.1 NLRs: a divergent protein family

The presence of NACHT domain containing genes has been shown for many organisms, including plants, echinoderms, chordates and vertebrates (Huang, Yuan et al. 2008; Maekawa, Kufer et al. 2011). Yet, the divergence of the *nlr* gene family is reflected by lineage-specific expansions (LSEs) as reported for *Danio rerio*, pufferfish (Stein, Caccamo et al. 2007), purple sea urchin and amphiouxus (Huang, Wang et al. 2011). Studies have shown that especially molecules that directly interact with pathogens show higher degrees of divergence in contrast to highly conserved pathways of signal transduction (Waterhouse, Kriventseva et al. 2007). One explanation for this phenomenon is a coevolution of pathogens and receptors. The C-terminal LRRs that are associated to pathogen recognition exhibit a great variation in their numbers and arrangements (Waterhouse, Kriventseva et al. 2007) which can be due to the affinity to different ligands. LSEs have also been reported for other innate immune related factors in zebrafish as, for example, the fish novel tripartite motif (finTRIM) family (van der Aa, Levraud et al. 2009).
A great divergence in LRR composition was observed in the zebrafish expansion of the novel NLRs. Number and positions within their C-termini vary among the different family members.
LSEs in novel fish NLRs are mainly characterized by the presence of a Fisna domain and the absence of a typical effector binding domain. Yet, the N-termini of fish NLRs show great differences in their sequences. In contrast, the fish NLRs exhibit an

enormous degree of sequence similarity within their Fisna, NACHT and helical domains which are all located on one exon. This similarity of the family can only be exceeded by the resemblance within the four subgroups. Both, the great number of *nlr* genes in zebrafish as well as the high degree of conservation makes it difficult to study their gene expression.

4.2 Expression of fish *nlr* genes

One of the aims of this study was to investigate the expression of *nlr* genes during embryonic and larval development and to compare that to their expressional responses after challenging the innate immune system. However, as mentioned above, the size and degree of similarity of the gene family made it rather impossible to draw up an expression profile of every single *nlr* gene by PCR or Northern blot analyses. In addition to that the family of fish-specific NLRs is not completely annotated in the databases.

For these reasons RT-PCRs and Northern blot analyses were performed by targeting a whole group of genes. Yet, some of the results were inconclusive: semiquantitative RT-PCRs showed the strongest expression for *nlr I* and *II*. In contrast, qRT-PCR results stated that *nlr III* showed the highest expression level, followed by *nlr II* and groups *I* and *IV* were equally low expressed.

One reason for these varieties can be a different *nlr* expression in the embryos that were used for RNA isolation. Another reason can be sequence variation within primer binding sites among the embryos. A high degree of polymorphism is a feature of zebrafish. Even within one zebrafish strain innate immunity related genes showing polymorphisms within internal splice sites have been identified (Stein 2010). The probability of polymorphisms within a group of more than 300 genes is therefore very likely. The samples used for the expression analyses in this study were isolated from different animals, or in the case of embryos from a pool of animals, which can be an explanation for the different results. Polymorphisms can not be avoided in zebrafish strains since establishing inbreds leads to infertility after a few generations. Experiments with samples originating from one batch of embryos or from the same parents minimize this problem.

A further explanation could be a contamination with genomic DNA. Although the RNA samples have been treated with DNAse this possibility can not be completely excluded. If genomic DNA was still present in some of the cDNA samples both kinds of nucleic acids would compete for primer recruitment and some primers were kept from amplifying cDNA.

It is also possible that the different embryo batches were exposed to different unscheduled factors that activated the immune system as e. g. temperature variations or pathogen contamination in the embryo medium.

Further ambigous results were obtained by Northern blot analyses. The two Northern blots presented in this study showed different results for expression levels of *nlr I* and *IV* which was elevated in one experiment (supp 1) compared to the levels in a second experiment (fig. 3.5) Furthermore, the first blot only showed signals at 48hpf and for adult fish while in the second Northern blot for *nlr III* additional signals at 2wpf and 4wpf were detected. Another difference was the expression level at 48hpf for *nlr II* which was also higher in the first experiment than in the second. Differences between the two results might state the obvious that the analyzed embryos differed in their *nlr* expression. In both Northern blots strong signals were detected at around 8kb which is in contrary to the annotated gene size with a maximum of about 3,7kb. The denaturating conditions that were applied to perform the gel electrophoresis should normally inhibit the formation of secondary structures. However, RNA molecules tend to form those structures if they cool down before they are loaded on a gel. The topology (i.e., circularity) can affect migration, making RNAs appear longer on the gel than they actually are. In addition to that it is not clear if the signals are caused by the several identical *nlr* RNA molecules or if they are different. For nlr II and III multiple bands were detected which suggests different transcripts per group since *nlr* gene differ in their sizes (see introduction).

Nevertheless, the outcome from the qRT-PCRs and the Northern blot analyses coincided in that *nlr III* showed the highest expression level of the four groups and *nlr I* and *IV* the lowest level. Based on the latest gene annotations this distribution also fits to the sizes of the four *nlr* groups. However, as mentioned above the dataset is not yet complete.

The Southern blot analysis that was applied to test the specificity of each probe that was used for the Northen blots demonstrated group-specific hybridizations to constructs originating from the same *nlr* group as the probe. Except for probe FN III which hybridized similar strong to *nlr* group *I* and *III*. However, this affinity was not observed vice versa when probe FN I hybridization was tested on *nlr III* constructs. A reason for these hybridization properties is not clear but since this affinity was only observed in one direction it can be considered as artefact. Moreover, the hybridizations of FN I and FN II I during Northern blot analyses led to different and non-overlapping signals.

To further investigate the expression of fish-specific NLRs we want to sequence *nlrs* from different zebrafish. The project involves deep sequencing of genomic DNA and

cDNA of adult zebrafish and their offspring. The embryos would also be challenged with different pathogens to analyze gene-specific expression responses. By these means it would be possible to analyse the presence of *nlr* genes in the genome, the expression of *nlr* genes and their expression response upon pathogens. Furthermore, the annotation of *nlr* genes that are not yet covered by the databases has to be completed. Both projects are pursued in colaboration with the group of Dr. Howe from the Sanger Institute in Hinxton and the group of Prof. Dr. Wiehe in Cologne.

Visualizing *nlr* expression either by antibody staining or whole mount in situ hybridization would be difficult for such a large family. Antibody stainings would require at least four different antibodies that would be able to distinguish between the NLRs. So far there is no working antibody available. For in situ hybridization it would be possible to design probes that distinguish between the different *nlr* groups but one would expect a staining of the complete embryonic or larval body that would not lead to an insight of *nlr* expression that was not covered by the experiments presented in this study.

4.2.1 *nlr* response on pathogen treatment

To compare *nlr* expression before and after pathogen treatment embryos of one batch were used to inject the different pathogens and isolate RNA in parallel. Semiquantitative and qRT-PCRs showed a response on poly IC of *nlr I* expression 3hpi and 30hpf (equivalent to 24hpi). Both analyses indicated an upregulation of *nlr I* expression which was five fold compared to non-infected embryos according to qRT-PCR results. Poly IC is a synthetic RNA analog that is often used to mimic viral infections. A study on a different immune-related gene family, the finTRIMs, showed a virus-induced expression in rainbow trout leucocytes after viral hemorrhagic septicemia virus or poly IC infection (van der Aa, Levraud et al. 2009). finTRIMs also harbour a B30.2 domain as it was shown for zebrafish NLRs I, II and III. Yet, if the virus induced upregulation of finTRIM expression was induced by their own B30.2 domains ist not clear. An analysis of B30.2 domain-specific function or a contribution in virus-recognition has not been done for zebrafish.

For *nlr III* a strong induction of RNA expression upon *E. coli* infection has been observed 3hpi in qRT-PCR experiments. The semiquantitative RT-PCR results also showed a higher expression level of *nlr III* in *E. coli*-infected embryos but the upregulation was not that abundant as it was observed by qRT-PCR. *E. coli* are no natural fish pathogens in contrast to *Y. ruckeri* which caused no induction of *nlr* expression. The only expression response for *E. coli* infection in zebrafish embryos

reported so far was an induced *tnfα* expression *(Sieger, Stein et al. 2009).* One possible explanation is that the expression of receptors that recognize rather exotic PAMPs is induced while the receptors for more frequent appearing ligands are constantly expressed. To get further insight into NLR-ligand combinations it would be interesting to determine the exact genes which get amplified upon *E. coli* or poly IC infections in contrast to the analysis of a group-specific response. The idea of an inducible gene expression caused by pathogen infections seems to be rather controversy to the presence of a large and diverse receptor family that shows at least a basal expression under normal conditions. Yet, by our results it can not be distinguished between expression of one single gene or a group of genes under normal conditions or after pathogen injection. Moreover, the unchanged gene expressions of all four nlr groups after *Y. ruckeri* does not necessarily mean that they did not recognize this kind of pathogen.

4.3 Protein-binding between NLR proteins

Typical effector-binding domains of NLR proteins feature homotypic binding to downstream signalling molecules and result in inflammasome, nodosome or apoptosome activation. It would be interesting to investigate an interaction between different N-terminal ends of fish NLRs and the conserved downstream effector molecules as for example caspases, ASC or RIP2. By this means it could be figured out if novel fish NLRs contribute to the typical pathways as known from other NLRs.
To unreval putative pathways the fish NLRs are involved in binding partners of for the N-terminus of an NLR III protein were searched for using yeast two hybrid.

4.3.1 Evaluation of the zebrafish cDNA library

The cDNA library was generated from mRNA isolated from one adult zebrafish. It was also considered to use mRNA from an infected zebrafish to activate immunity components but the risk of isolating bacterial mRNA that would contaminate the zebrafish library was too high. When the yeast two hybrid screen was about to be started no zebrafish cDNA library was commercially available.
The insert sizes of our generated library sizes were relatively small, only 7 % was between 0.8kb and 1kb the other inserts were smaller. The addition of all cDNA inserts led to a library size of 107 million bases compared to about 700 million so far annotated bases according to the Vega genome browser from ENSEMBL. Yet, cDNA inserts from the library could be overlapping transcripts. Information about the entire zebrafish transcriptome are not available that a statement can be made about the part that was covered by the cDNA library.

4.3.2 Yeast two hybrid screen results

The N-terminus of a NLR III protein bound three times to NLR proteins under medium strong selection conditions and to 27 NLR proteins under low stringency conditions. The entire set of NLR III binding partners that were identified during the screen were classified as NLR I group members. These results were in contrast to further performed yeast two hybrid experiments that indicated a clear group-specific interaction between identical NLR groups. In fact an interaction between NLR I and III proteins could be verified but it was much weaker compared to a protein interaction between two NLR group III proteins. Regarding this it was rather suprising that the bait protein did not pull out any NLR III proteins during the screen. An explanation for this might be the quality and incompleteness of the cDNA library. Since expression of *nlr III* in adult zebrafish was confirmed by former experiments it was rather unlikely that NLR III proteins were not included in the library because they were not part of the transcriptome. A more likely possibility was that the transcriptome was not entirely represented in the cDNA library and nlr III cDNA was missing which is why the bait protein bound to NLR I proteins instead.

The reason for choosing yeast cells was that they offered a eucaryotic, NLR-free system that was advantageous to test NLR binding. It would have been difficult to test interaction between NLRs in a fish cell line or zebrafish embryos if the targeted NLR constructs were amongst hundreds of similar proteins including many possible interaction partners. Yet, it remains to be discussed if the interaction between NLR proteins is direct or mediated by a protein which belongs to the *Sacharomyces cerevisiae* transcriptome. This possibility can be rouled out by a pulldown assay with a recombinant GST-tagged construct and e.g. a radioactive labelled in vitro translated version of the tested constructs.

Yeast two hybrid experiments were often discussed to be imprecise and lead to false-positive results due to biased transcriptional responses (Serebriiskii, Estojak et al. 2000). Therefore binding candidates that were identified during a yeast two hybrid screen required further confirmation by other methods which is why co-IP experiments were performed with combined FisnaNACHT domains. The results obtained by co-IPs were in agreement with the ones from the yeast two hybrid experiments regarding the group-specific interactions. However, interactions between foreign NLR groups were not shown at all whereas the yeast two hybrid experiments indicated weak binding between some of the different fish NLRs.

4.3.3 Group-specific protein binding between NLR proteins

Comprehensive yeast two hybrid experiments revealed a clear affinity between NLR proteins of identical groups compared to binding with other NLRs. The results also indicate that the NACHT and Fisna domains are sufficent for group recognition and strong protein interactions. Studies on dominant negative conserved NLRs have demonstrated that the NACHT domain is essential for oligomerization or dimerization of the proteins. Yet, it was unclear if this was due to NTP binding and hydrolysis accomplished by the NACHT domain or a physical interaction between two of the domains (Zurek, Proell et al. 2011). In this study here it has been shown that NLR proteins physically interact with each other through these domains. Moreover, it was demonstrated that binding sites located in the N-terminal half of the NACHT domain were responsible for an interaction to a domain originating from the same NLR group. Residues that are hypothetical binding sites in APAF1 have been identified by in silico analyses (Proell, Riedl et al. 2008). The authors identified six sites of binding residues that were exposed to the outer surface of the protein. Sequence comparisons between APAF1 and fish NLRs identified four candidates for those sites (three to six) that were located on the NACHT domains of fish NLRs. Sites three and four were upstream of the Walker B motif and sites five and six were downstream of it. Candidates for binding sites one and two were upstream of the NACHT domain, located in the Fisna domain. Putative contribution of the first two sites were not analyzed in this study but are planned in future experiments.

Experiments on protein binding with hybrid NACHT domains revealed an important contribution for the N-terminal part of the NACHT domain to group-specificity. Binding site candidates that were located in the N-terminal half of the NACHT domains showed no sequence similarities between the four NLRs whereas the sites within the half downstream of the Walker B motif were more similar to each other. This might indicate a function for group-specific recognition and binding for the N-terminal half. However, it has not been pinpointed to the suggested binding sites no. 3 and 4 but to a sequence of 78 aa. In order to narrow down specific binding sites it would be usefull to look into structure predictions for fish NLRs by computational analyses. By this means the positions for the putative binding sites could be localized to find out if they were accessible for intramolecular interactions. Furthermore, mutational analyses would also shed light upon this issue.

The Fisna and NACHT domains are independently able to bind other NLRs in a group-specific way. Co-IPs did not show any interactions between fish NLRs from different groups which strongly suggests that a group-specific interaction fullfills a functional purpose. It can be hypothesized that fish NLRs of one group build a signalling platform

that induce pathways distinct to the three other groups. To further analyze a contribution to downstream pathways signalling molecules have to be found and identified. What speaks against this hypothesis is the diversity in their N-terminus which would suggest binding to distinct molecules.

Abstract

NLRs (NACHT domain and leucine-rich repeat containing proteins) comprise a family of intracellular pathogen recognition receptors that can activate the innate immune system, e.g. via inflammasomes or by activating the NFκB signalling pathway. The highly diverse family of *nlr* genes features a lineage-specific expansion in *Danio rerio* with at least 366 members. The predicted protein domain structure resembles the architecture of more intense studied mammalian NLRs as they all share a middle NACHT domain and leucine-rich repeats (LRRs) in their C-terminus. A structural hallmark of the fish NLRs is the fish-specific NACHT associated (Fisna) domain upstream of the NACHT domain which is why the proteins were called fish-specific NLRs. According to sequence similarities within the Fisna and NACHT domains they were subdivided into four groups. Since annotations of *nlrs* are not completed yet the numbers of *nlr* group members are only preliminary but based on current data *nlr* group *III* has the most members, followed by group *II*. *nlr I* and *IV* have similar few members.

In this study data from different RT-PCRs and Northern blot experiments are presented that mainly confirmed the sizes of the four *nlr* groups. It also showed an expression of *nlrs* already at maternal stages that is maintained during embryonic and larval development. Moreover, it was shown that poly IC injection induced nlr I expression 3hpi. Similar results were observed for *nlr III* genes after the embryonic innate immune system was challenged with *E. coli* at 3hpi and 24hpi.

Furthermore, in this study a group-specific interaction between NLRs has been demonstrated. A yeast two hybrid screen first revealed interaction between NLRs which was further specified and validated by yeast two hybrid experiments and co-IPs. All experiments consistently showed a strong protein binding between the Fisna and the NACHT domains in a group-specific way. To further narrow down the responsible recognition and binding sites domain swapping experiments were performed. Hybrid NACHT domains were generated that incorporated two different NLR groups. These were tested in comprehensive binding experiments that could restrict the binding side to the 78 aa of the N-terminal half of the NACHT domains.

Zusammenfassung

NLRs (NACHT domain and leucine-rich repeat containing proteins) bilden eine Familie von intrazellulären Rezeptoren, welche in die Zelle eingedrungende Pathogene erkennen und daraufhin das angeborene Immunsystem stimulieren. Dies geschieht zum Beispiel durch die Aktivierung und Bildung von Inflammasomen oder des NFκB Signalweges. Die *nlr* Genfamilie ist extrem divers und weist spezifisch für die Spezies Zebrafish eine phylogenetische Expansion von mindestens 366 Genen auf.

Die vorausgesagte Proteinstruktur dieser neuen Familie ähnelt jener von den besser untersuchten Säugetier NLRs. Beide Gruppen weisen eine zentrale NACHT Domäne und sogenannte LRRs (leucine-rich repeats) in ihren C-terminalen Enden auf. Was NLRs in Fischen von denen in Säugern unterscheidet und gleichzeitig ein Kennzeichen für erstere bildet, ist die Fisna (fish-specific NACHT associated) Domäne. Diese Domäne ist der NACHT Domäne direkt vorgelagert. Anhand von Sequenzähnlichkeiten, die sowohl in der Fisna als auch der NACHT Domäne zu finden sind, lassen sich die Fisch NLRs in vier Gruppen unterteilen. Da die Annotationen der nlr Gene noch nicht vollständig abgeschlossen sind, wird die Anzahl der jeweiligen Gruppenmitglieder als vorläufig betrachtet. Basierend auf den aktuellen Daten lässt sich jedoch bereits festhalten, dass Gruppe III die meisten Mitglieder umfasst, gefolgt von Gruppe II, während Gruppen I und IV gleichermaßen die wenigsten Mitglieder aufweisen.

In dieser Arbeit konnte durch RT-PCRs und Northern Blots diese Mengenverteilung auf die vier Gruppen größtenteils bestätigt werden. Darüber hinaus wurde mit Hilfe derselben Methoden, dass die Expression der *nlrs* bereits maternal vorhanden ist und sich über die Embryonalentwicklung und im Larvenstatium fortsetzt. Es konnte hier ebenfalls gezeigt werden, dass die Injektion von poly IC nach drei Stunden eine Erhöhung der *nlr I* Expression bewirkt. Auf ähnliche Weise reagiert die *nlr III* Untergruppe auf *E. coli* Bakterien. Drei bzw. 24 Stunden nachdem die Erreger injiziert wurden, konnte ein erhöhtes Level der *nlr III* Genexpression verzeichnet werden.

Im weiteren Verlauf wurde eine gruppenspezifische Interaktion zwischen NLR Proteinen beobachtet. Zum ersten Mal wurde eine Bindung zwischen Fisch NLR Proteinen während eines Hefe Zwei Hybrid „Screens“ entdeckt, was in weiteren Hefe Zwei Hybrid Experimenten und Co-Immunoprezipitationen bestätigt werden konnte. Jedes dieser Experimente zeigte übereinstimmend die Interaktionen zwischen den

Fisna und NACHT Domänen, welche aus denselben NLR Gruppen stammten. Um die Sequenz, welche für Erkennung und Interaktion verantwortlich war, einzugrenzen, wurden hybride NACHT Konstrukte generiert. Diese hybriden NACHT Domänen vereinten zwei NLR Gruppen in sich. Die N-terminale Hälfte gehörte zu einer anderen NLR Untergruppe als der C-terminale Teil. Die hybriden NACHT Konstrukte wurden in umfassenden Interaktionsexperimenten eingesetzt, welche die Bindungsstelle auf die 78 Aminosäuren lange N-terminale Häfte der NACHT Domänen eingrenzen konnten.

References

Abi Rached, L., McDermott, M. F. , and Pontarotti, P. (1999). "The MHC big bang." Immonol Rev **167**: 33-44.

Acehan, D., X. Jiang, et al. (2002). "Three-dimensional structure of the apoptosome: implications for assembly, procaspase-9 binding, and activation." Mol Cell **9**(2): 423-432.

Aggad, D., C. Stein, et al. (2010). "In vivo analysis of Ifn-gamma1 and Ifn-gamma2 signaling in zebrafish." J Immunol **185**(11): 6774-6782.

Akashi, K., D. Traver, et al. (2000). "A clonogenic common myeloid progenitor that gives rise to all myeloid lineages." Nature **404**(6774): 193-197.

Albrecht, M., F. S. Domingues, et al. (2003). "Structural localization of disease-associated sequence variations in the NACHT and LRR domains of PYPAF1 and NOD2." FEBS Lett **554**(3): 520-528.

Alder, M. N., I. B. Rogozin, et al. (2005). "Diversity and function of adaptive immune receptors in a jawless vertebrate." Science **310**(5756): 1970-1973.

Alexander, J. J., A. J. Anderson, et al. (2008). "The complement cascade: Yin-Yang in neuroinflammation--neuro-protection and -degeneration." J Neurochem **107**(5): 1169-1187.

Azumi, K., R. De Santis, et al. (2003). "Genomic analysis of immunity in a Urochordate and the emergence of the vertebrate immune system: "waiting for Godot"." Immunogenetics **55**(8): 570-581.

Babior, B. M., J. T. Curnutte, et al. (1976). "The particulate superoxide-forming system from human neutrophils. Properties of the system and further evidence supporting its participation in the respiratory burst." J Clin Invest **58**(4): 989-996.

Berman, J. N., J. P. Kanki, et al. (2005). "Zebrafish as a model for myelopoiesis during embryogenesis." Exp Hematol **33**(9): 997-1006.

Bertrand, J. Y., A. D. Kim, et al. (2007). "Definitive hematopoiesis initiates through a committed erythromyeloid progenitor in the zebrafish embryo." Development **134**(23): 4147-4156.

Biacchesi, S., M. LeBerre, et al. (2009). "Mitochondrial antiviral signaling protein plays a major role in induction of the fish innate immune response against RNA and DNA viruses." J Virol **83**(16): 7815-7827.

Bilodeau-Bourgeois, L., B. G. Bosworth, et al. (2008). "Differences in mortality, growth, lysozyme, and toll-like receptor gene expression among genetic groups of catfish exposed to virulent Edwardsiella ictaluri." Fish Shellfish Immunol **24**(1): 82-89.

Bitter, W., A. M. van der Sar, et al. (2003). "Zebrafish embryos as a model host for the real time analysis of Salmonella typhimurium infections." Cellular Microbiology **5**(9): 601-611.

Bonardi, V., S. Tang, et al. (2011). "Expanded functions for a family of plant intracellular immune receptors beyond specific recognition of pathogen effectors." Proc Natl Acad Sci U S A.

Castro, R., D. Bernard, et al. (2011). "T cell diversity and TcR repertoires in teleost fish." Fish Shellfish Immunol **31**(5): 644-654.

Cerqueira, B. A., W. V. Boas, et al. (2011). "Increased concentrations of IL-18 and uric acid in sickle cell anemia: Contribution of hemolysis, endothelial activation and the inflammasome." Cytokine **56**(2): 471-476.

Cerretti, D. P., C. J. Kozlosky, et al. (1992). "Molecular cloning of the interleukin-1 beta converting enzyme." Science **256**(5053): 97-100.

Chang, M., B. Collet, et al. (2011). "Expression and functional characterization of the RIG-I-like receptors MDA5 and LGP2 in Rainbow trout (Oncorhynchus mykiss)." J Virol **85**(16): 8403-8412.

Chluba, J., C. Jault, et al. (2004). "Toll-like receptor gene family and TIR-domain adapters in Danio rerio." Molecular Immunology **40**(11): 759-771.

Danilova, N., J. Bussmann, et al. (2005). "The immunoglobulin heavy-chain locus in zebrafish: identification and expression of a previously unknown isotype, immunoglobulin Z." Nat Immunol **6**(3): 295-302.

Danot, O., E. Marquenet, et al. (2009). "Wheel of Life, Wheel of Death: A Mechanistic Insight into Signaling by STAND Proteins." Structure **17**(2): 172-182.

Davidson, A. J. and L. I. Zon (2004). "The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis." Oncogene **23**(43): 7233-7246.

Deretic, V. (2011). "Autophagy in immunity and cell-autonomous defense against intracellular microbes." Immunol Rev **240**(1): 92-104.

Dios, S., A. Romero, et al. (2010). "Effect of the temperature during antiviral immune response ontogeny in teleosts." Fish Shellfish Immunol **29**(6): 1019-1027.

Dunne, A. (2011). "Inflammasome activation: from inflammatory disease to infection." Biochem Soc Trans **39**(2): 669-673.

Fang, Y., C. Xu, et al. (1998). "Expression of complement receptors 1 and 2 on follicular dendritic cells is necessary for the generation of a strong antigen-specific IgG response." J Immunol **160**(11): 5273-5279.

Feldmeyer, L., M. Keller, et al. (2007). "The inflammasome mediates UVB-induced activation and secretion of interleukin-1beta by keratinocytes." Curr Biol **17**(13): 1140-1145.

Fitzgerald, K. A., D. C. Rowe, et al. (2003). "LPS-TLR4 signaling to IRF-3/7 and NF-kappaB involves the toll adapters TRAM and TRIF." J Exp Med **198**(7): 1043-1055.

Flajnik, M. F. and L. Du Pasquier (2004). "Evolution of innate and adaptive immunity: can we draw a line?" Trends Immunol **25**(12): 640-644.

Fritz, J. H., L. Le Bourhis, et al. (2007). "Nod1-mediated innate immune recognition of peptidoglycan contributes to the onset of adaptive immunity." Immunity **26**(4): 445-459.

Gack, M. U., Y. C. Shin, et al. (2007). "TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity." Nature **446**(7138): 916-920.

Giacomin, P. R., D. L. Gordon, et al. (2008). "The role of complement in innate, adaptive and eosinophil-dependent immunity to the nematode Nippostrongylus brasiliensis." Mol Immunol **45**(2): 446-455.

Gongora, R., F. Figueroa, et al. (1998). "Independent duplications of Bf and C3 complement genes in the zebrafish." Scand J Immunol **48**(6): 651-658.

Hall, C., M. V. Flores, et al. (2009). "Transgenic zebrafish reporter lines reveal conserved Toll-like receptor signaling potential in embryonic myeloid leukocytes and adult immune cell lineages." J Leukoc Biol **85**(5): 751-765.

Hoshino, K., O. Takeuchi, et al. (1999). "Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the Lps gene product." J Immunol **162**(7): 3749-3752.

Hu, Y., M. A. Benedict, et al. (1999). "Role of cytochrome c and dATP/ATP hydrolysis in Apaf-1-mediated caspase-9 activation and apoptosis." EMBO J **18**(13): 3586-3595.

Huang, J., V. Canadien, et al. (2009). "Activation of antibacterial autophagy by NADPH oxidases." Proc Natl Acad Sci U S A **106**(15): 6226-6231.

Huang, S., X. Wang, et al. (2011). "The evolution and regulation of the mucosal immune complexity in the basal chordate amphioxus." J Immunol **186**(4): 2042-2055.

Huang, S., S. Yuan, et al. (2008). "Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity." Genome Res **18**(7): 1112-1126.

Inohara, Chamaillard, et al. (2005). "NOD-LRR proteins: role in host-microbial interactions and inflammatory disease." Annu Rev Biochem **74**: 355-383.

Inohara, N., Y. Ogura, et al. (2001). "Human Nod1 confers responsiveness to bacterial lipopolysaccharides." J Biol Chem **276**(4): 2551-2554.

Janeway, C. A., Jr. (2001). "How the immune system works to protect the host from infection: a personal view." Proc Natl Acad Sci U S A **98**(13): 7461-7468.

Jin, M. S., S. E. Kim, et al. (2007). "Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide." Cell **130**(6): 1071-1082.

Jing, L. and L. I. Zon (2011). "Zebrafish as a model for normal and malignant hematopoiesis." Dis Model Mech **4**(4): 433-438.

Kajava, A. V. (1998). "Structural diversity of leucine-rich repeat proteins." J Mol Biol **277**(3): 519-527.

Kimmel, C. B., W. W. Ballard, et al. (1995). "Stages of embryonic development of the zebrafish." Dev Dyn **203**(3): 253-310.

Klein, J. and N. Nikolaidis (2005). "The descent of the antibody-based immune system by gradual evolution." Proceedings of the National Academy of Sciences of the United States of America **102**(1): 169-174.

Kobayashi, K. S., M. Chamaillard, et al. (2005). "Nod2-dependent regulation of innate and adaptive immunity in the intestinal tract." Science **307**(5710): 731-734.

Kobe, B. and A. V. Kajava (2001). "The leucine-rich repeat as a protein recognition motif." Current Opinion in Structural Biology **11**(6): 725-732.

Kufer, T. A. (2008). "Signal transduction pathways used by NLR-type innate immune receptors." Molecular Biosystems **4**(5): 380-386.

Kumar, H., T. Kawai, et al. (2009). "Pathogen recognition in the innate immune response." Biochem J **420**(1): 1-16.

Kvarnhammar, A. M., T. Petterson, et al. (2011). "NOD-like receptors and RIG-I-like receptors in human eosinophils: activation by NOD1 and NOD2 agonists." Immunology **134**(3): 314-325.

Lam, S. H., H. L. Chua, et al. (2004). "Development and maturation of the immune system in zebrafish, Danio rerio: a gene expression profiling, in situ hybridization and immunological study." Dev Comp Immunol **28**(1): 9-28.

Laroui, H., Y. Yan, et al. (2011). "L-Ala-{gamma}-D-Glu-meso-diaminopimelic Acid (DAP) Interacts Directly with Leucine-rich Region Domain of Nucleotide-binding Oligomerization Domain 1, Increasing Phosphorylation Activity of Receptor-interacting Serine/Threonine-protein Kinase 2 and Its Interaction with Nucleotide-binding Oligomerization Domain 1." J Biol Chem **286**(35): 31003-31013.

Levin, T. C., K. E. Wickliffe, et al. (2008). "Heat shock inhibits caspase-1 activity while also preventing its inflammasome-mediated activation by anthrax lethal toxin." Cellular Microbiology **10**(12): 2434-2446.

Levraud, J. P., P. Boudinot, et al. (2007). "Identification of the zebrafish IFN receptor: implications for the origin of the vertebrate IFN system." J Immunol **178**(7): 4385-4394.

Lieschke, G. J., A. C. Oates, et al. (2001). "Morphologic and functional characterization of granulocytes and macrophages in embryonic and adult zebrafish." Blood **98**(10): 3087-3096.

Link, V. (2006). "Proteomics of early zebrafish embryo." BMC Developmental Biology.

Maekawa, T., T. A. Kufer, et al. (2011). "NLR functions in plant and animal immune systems: so far and yet so close." Nat Immunol **12**(9): 817-826.

Mariathasan, S., K. Newton, et al. (2004). "Differential activation of the inflammasome by caspase-1 adaptors ASC and Ipaf." Nature **430**(6996): 213-218.

Mariathasan, S., D. S. Weiss, et al. (2006). "Cryopyrin activates the inflammasome in response to toxins and ATP." Nature **440**(7081): 228-232.

Matsuo, A., H. Oshiumi, et al. (2008). "Teleost TLR22 recognizes RNA duplex to induce IFN and protect cells from birnaviruses." J Immunol **181**(5): 3474-3485.

Meijer, A. H., S. F. G. Krens, et al. (2004). "Expression analysis of the Toll-like receptor and TIR domain adaptor families of zebrafish." Molecular Immunology **40**(11): 773-783.

Meylan, E., J. Tschopp, et al. (2006). "Intracellular pattern recognition receptors in the host response." Nature **442**(7098): 39-44.

Miao, E. A., J. V. Rajan, et al. (2011). "Caspase-1-induced pyroptotic cell death." Immunol Rev **243**(1): 206-214.

Motani, K., H. Kushiyama, et al. (2011). "Caspase-1 Protein Induces Apoptosis-associated Speck-like Protein Containing a Caspase Recruitment Domain (ASC)-mediated Necrosis Independently of Its Catalytic Activity." J Biol Chem **286**(39): 33963-33972.

Neal, M. D., C. Leaphart, et al. (2006). "Enterocyte TLR4 mediates phagocytosis and translocation of bacteria across the intestinal barrier." J Immunol **176**(5): 3070-3079.

Oehlers, S. H., M. V. Flores, et al. (2011). "The inflammatory bowel disease (IBD) susceptibility genes NOD1 and NOD2 have conserved anti-bacterial roles in zebrafish." Dis Model Mech.

Ohtani, M., J. Hikima, et al. (2010). "Evolutional conservation of molecular structure and antiviral function of a viral RNA receptor, LGP2, in Japanese flounder, Paralichthys olivaceus." J Immunol **185**(12): 7507-7517.

Paik, E. J. and L. I. Zon (2010). "Hematopoietic development in the zebrafish." Int J Dev Biol **54**(6-7): 1127-1137.

Palti, Y. (2011). "Toll-like receptors in bony fish: From genomics to function." Dev Comp Immunol.

Pancer, Z., C. T. Amemiya, et al. (2004). "Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey." Nature **430**(6996): 174-180.

Parish, C. R. and E. R. O'Neill (1997). "Dependence of the adaptive immune response on innate immunity: some questions answered but new paradoxes emerge." Immunol Cell Biol **75**(6): 523-527.

Perrot, I., F. Deauvieau, et al. (2010). "TLR3 and Rig-like receptor on myeloid dendritic cells and Rig-like receptor on human NK cells are both mandatory for production of IFN-gamma in response to double-stranded RNA." J Immunol **185**(4): 2080-2088.

Phelan, P. E., M. T. Mellon, et al. (2005). "Functional characterization of full-length TLR3, IRAK-4, and TRAF6 in zebrafish (Danio rerio)." Mol Immunol **42**(9): 1057-1071.

Proell, M., S. J. Riedl, et al. (2008). "The Nod-like receptor (NLR) family: a tale of similarities and differences." PLoS One **3**(4): e2119.

Qin, H., S. M. Srinivasula, et al. (1999). "Structural basis of procaspase-9 recruitment by the apoptotic protease-activating factor 1." Nature **399**(6736): 549-557.

Reubold, T. F., S. Wohlgemuth, et al. (2011). "Crystal structure of full-length Apaf-1: how the death signal is relayed in the mitochondrial pathway of apoptosis." Structure **19**(8): 1074-1083.

Ribeiro, C. M., T. Hermsen, et al. (2010). "Evolution of recognition of ligands from Gram-positive bacteria: similarities and differences in the TLR2-mediated response between mammalian vertebrates and teleost fish." J Immunol **184**(5): 2355-2368.

Richardson, W. M., C. P. Sodhi, et al. (2010). "Nucleotide-binding oligomerization domain-2 inhibits toll-like receptor-4 signaling in the intestinal epithelium." Gastroenterology **139**(3): 904-917, 917 e901-906.

Riedl, S. J., W. Li, et al. (2005). "Structure of the apoptotic protease-activating factor 1 bound to ADP." Nature **434**(7035): 926-933.

Rosenstiel, P. and S. Schreiber (2009). "NOD-Like Receptors-Pivotal Guardians of the Immunological Integrity of Barrier Organs." Target Pattern Recognition in Innate Immunity **653**: 35-47.

Rosenstiel, P., A. Till, et al. (2007). "NOD-like receptors and human diseases." Microbes Infect **9**(5): 648-657.

Sepulcre, M. P., F. Alcaraz-Perez, et al. (2009). "Evolution of lipopolysaccharide (LPS) recognition and signaling: fish TLR4 does not recognize LPS and negatively regulates NF-kappaB activation." J Immunol **182**(4): 1836-1845.

Serebriiskii, I., J. Estojak, et al. (2000). "Approaches to detecting false positives in yeast two-hybrid systems." Biotechniques **28**(2): 328-330, 332-326.

Shintani, S., J. Terzic, et al. (2000). "Do lampreys have lymphocytes? The Spi evidence." Proc Natl Acad Sci U S A **97**(13): 7417-7422.

Sieger, D., C. Stein, et al. (2009). "The role of gamma interferon in innate immunity in the zebrafish embryo." Dis Model Mech **2**(11-12): 571-581.

Song, H. D., X. J. Sun, et al. (2004). "Hematopoietic gene expression profile in zebrafish kidney marrow." Proc Natl Acad Sci U S A **101**(46): 16240-16245.

Stein, C. (2010). Innate immune genes in the zebrafish, danio rerio. Biology. Cologne, University of Cologne.

Stein, C., M. Caccamo, et al. (2007). "Conservation and divergence of gene families encoding components of innate immune response systems in zebrafish." Genome Biol **8**(11): R251.

Sullivan, C., J. Charette, et al. (2009). "The gene history of zebrafish tlr4a and tlr4b is predictive of their divergent functions." J Immunol **183**(9): 5896-5908.

Tattoli, I., L. H. Travassos, et al. (2007). "The Nodosome: Nod1 and Nod2 control bacterial infections and inflammation." Semin Immunopathol **29**(3): 289-301.

Thornberry, N. A., H. G. Bull, et al. (1992). "A novel heterodimeric cysteine protease is required for interleukin-1 beta processing in monocytes." Nature **356**(6372): 768-774.

Tschopp, J., F. Martinon, et al. (2003). "NALPs: a novel protein family involved in inflammation." Nat Rev Mol Cell Biol **4**(2): 95-104.

Uchil, P. D., B. D. Quinlan, et al. (2008). "TRIM E3 ligases interfere with early and late stages of the retroviral life cycle." Plos Pathogens **4**(2).

van der Aa, L. M., J. P. Levraud, et al. (2009). "A large new subset of TRIM genes highly diversified by duplication and positive selection in teleost fish." BMC Biol **7**: 7.

Wagner, R. N., M. Proell, et al. (2009). "Evaluation of Nod-like receptor (NLR) effector domain interactions." PLoS One **4**(4): e4931.

Wang, Z., S. Zhang, et al. (2008). "Complement activity in the egg cytosol of zebrafish Danio rerio: evidence for the defense role of maternal complement components." PLoS One **3**(1): e1463.

Waterhouse, R. M., E. V. Kriventseva, et al. (2007). "Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes." Science **316**(5832): 1738-1743.

Willett, C. E., A. Cortes, et al. (1999). "Early hematopoiesis and developing lymphoid organs in the zebrafish." Dev Dyn **214**(4): 323-336.

Wolf, D. and S. P. Goff (2007). "TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells." Cell **131**(1): 46-57.

Ye, Z., J. D. Lich, et al. (2008). "ATP binding by monarch-1/NLRP12 is critical for its inhibitory function." Mol Cell Biol **28**(5): 1841-1850.

Yu, X., D. Acehan, et al. (2005). "A structure of the human apoptosome at 12.8 A resolution provides insights into this cell death platform." Structure **13**(11): 1725-1735.

Yu, X., L. Wang, et al. (2006). "Three-dimensional structure of a double apoptosome formed by the Drosophila Apaf-1 related killer." J Mol Biol **355**(3): 577-589.

Zarkadis, I. K., D. Mastellos, et al. (2001). "Phylogenetic aspects of the complement system." Dev Comp Immunol **25**(8-9): 745-762.

Zhang, Y. B., F. Sun, et al. (2011). "Fish MITA Serves as a Mediator for Distinct Fish IFN Gene Activation Dependent on IRF3 or IRF7." Journal of Immunology **187**(5): 2531-2539.

Zou, H., W. J. Henzel, et al. (1997). "Apaf-1, a human protein homologous to C. elegans CED-4, participates in cytochrome c-dependent activation of caspase-3." Cell **90**(3): 405-413.

Zou, J., M. Chang, et al. (2009). "Origin and evolution of the RIG-I like RNA helicase gene family." BMC Evol Biol **9**: 85.

Zurek, B., M. Proell, et al. (2011). "Mutational analysis of human NOD1 and NOD2 NACHT domains reveals different modes of activation." Innate Immun.

Supplementary data

Table supp.1: Positive Interaction clones with low stringency obtained from yeast two hybrid screen

NCBI Ref Seq or other source	numbers	longest clone (bp)	gene
ENSDARG00000079877	27	384	NACHT NTPases
ENSDARG00000025391	29	179	prefoldin subunit 2
ENSDARG00000058105	26	176	60S ribosomal protein L36a
ENSDARG00000035860	33	258	40S ribosomal protein S28
ENSDARG00000020850	25	380	Ef1a
ENSDARG00000092423	18	289	c-x-c chemokine 11-like
ENSDARG00000053462	18	57	Complement C3
ENSDARG00000023151	11	95	uncoupling protein 1
unknown	11	279	Aldo/Ketoreduktase
unknown	7	110	si:ch211-217k17.10
ENSDARG00000024478	3	178	brain protein 44
unknown	4	154	1. Novel protein, unknown function
unknown	3	126	2. Novel protein, unknown function
unknown	4	371	3. Novel protein, unknown function
ENSDARG00000042996	6	59	Cox 16
ENSDARG00000069296	3	505	monooxygenase, DBH-like 1, like
ENSDARG00000035327	6	216	muscle creatine kinase a
ENSDARG00000090428	3	215	chymotrypsin B1
ENSDARG00000007018	1	508	membrane-spanning 4-domains subfamily A member 4A-like
ENSDARG00000036893	2	186	F13a1
ENSDARG00000023055	3	238	slc25a43
ENSDARG00000043493	2	231	clathrin, heavy

			polypeptide a
ENSDARG00000079018	4	372	FKBP3
ENSDARG00000078022	3	388	FYVE
ENSDARG00000095577	3	93	cytochrome b r1
ENSDARG00000012972	1	512	non-muscle cofilin 1
ENSDARG00000091751	1	92	hydroxysteroid (17-beta) dehydrogenase 7
ENSDARG00000015822	2	315	sestrin-3
ENSDARG00000043102	1	98	Latexin
ENSDARG00000015065	2	103	Complement C4-B
ENSDARG00000040856	1	465	lactate dehydrogenase A4
ENSDARG00000045487	2	257	40S ribosomal protein S16
ENSDARG00000019181	1	225	40S ribosomal protein SA
ENSDARG00000021838	5	373	40S ribosomal protein S23
ENSDARG00000014867	1	364	60S ribosomal protein L8
ENSDARG00000051783	1	260	60S acidic ribosomal protein P0
ENSDARG00000017339	1	227	glutamine synthetase b
ENSDARG00000025073	1	417	60S ribosomal protein L18a
ENSDARG00000035871	1	535	60S ribosomal protein L30
ENSDARG00000092609	3	247	ancient ubiquitous protein 1
ENSDARG00000076892	3	89	nucleoside diphosphate kinase B
ENSDARG00000017339	2	509	glutamine synthetase b
ENSDARG00000068385	2	413	tropomyosin beta chain
ENSDARG00000037555	1	460	protein atonal homolog 8
ENSDARG00000089356	4	254	Titin-like
ENSDARG00000039034	2	324	MARCKS-like 1
ENSDARG00000021257	2	338	euk. Translation initiation

			factor 3, subunit D
ENSDARG00000002344	1	38	tubulin, beta 2c
ENSDARG00000077126	1	313	U1 small nuclear ribonucleoprotein 70 kD
unknown	1	307	hypothetical protein 1
unknown	1	498	hypothetical protein 2
ENSDARG00000062788	2	296	immunoresponsive gene 1, like
unknown	2	123	hypothetical protein 4
unknown	1	112	hypothetical protein 5
ENSDARG00000004687	1	481	3-ketoacyl-CoA thiolase B
ENSDARG00000016649	2	451	aflatoxin B1 aldehyde reductase member 2

Table supp. 1 summarizes interaction partners that were identified during the yeast two hybrid screen under low interaction conditions. The clones were identified by sequencing the isolated plasmids from positive yeast cells. The number of clones from one gene were indicated and the longest clone was used for blast searches on ENSEMBL. ENSDARG numbers identify the genes in the database.

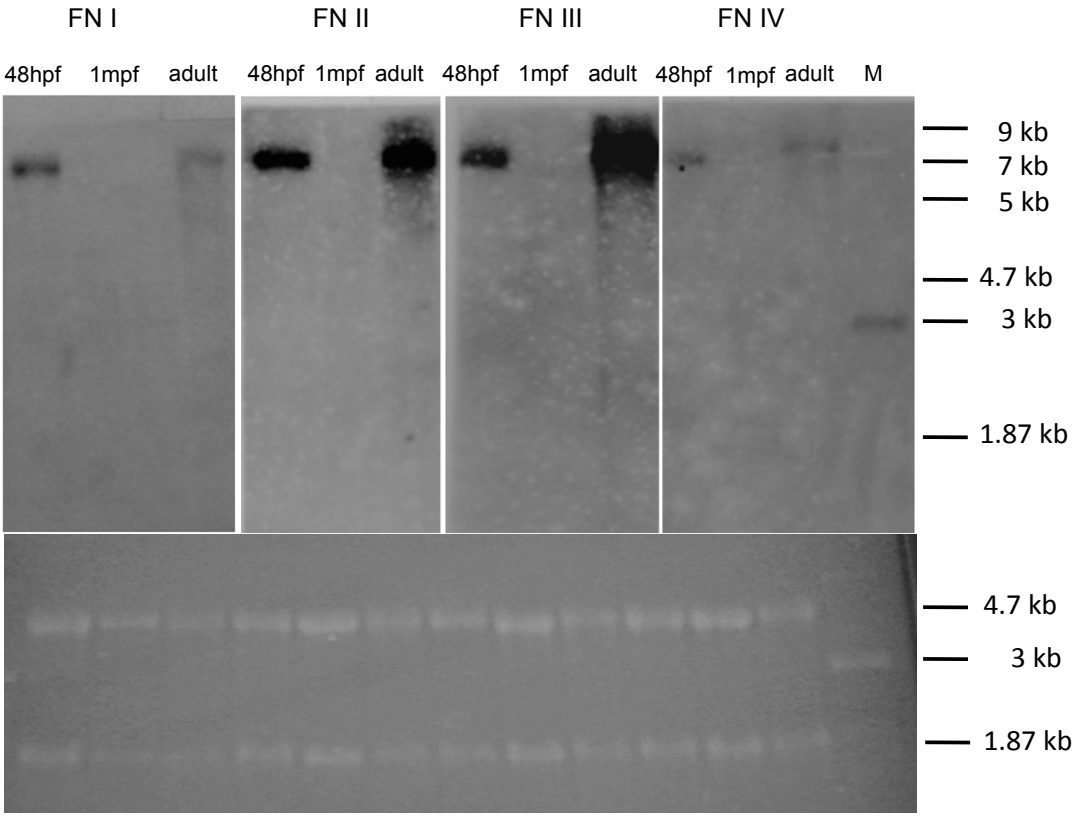

Figure supp.1: Northern blot analysis on zebrafish embryos and adult fish. Northern blot analyses using probes FN I to FN IV were performed on total RNA samples isolated from embryos 48hpf, 1mpf and from adult fish. M indicates a RNA ladder that shows a band at 3kb. The part below shows the different RNA samples run on a agarose gel to identify the amount of RNA and to exclude degradation. The visible bandscontain ribosomal RNA, the large subunit running at 4.7kb and the small subunit at 1.87kb.

Danksagung

Danke

Im Besonderen an Maria Leptin für die Möglichkeit meine Arbeit in ihrer Arbeitsgruppe schreiben zu dürfen, für die interessante Themenstellung, die Hilfsbereits- und Diskussionsbereitschaft, die sie mir entgegengebracht hat.

An Herrn PD Dr. Niels Gehring und Herrn Prof. Dr. Siegfried Roth für die Begutachtung der Arbeit und die Übernahme des Prüfungsvorsitzes.

An die gesamte Arbeitsgruppe Leptin in Köln und in Heidelberg für die Zusammenarbeit, Hilfe und reibungslose Abläufe und für den Spaß, was mir alles zusammen ermöglicht hat, meine Arbeit durchführen zu können.

An meine ehemaligen Arbeitskollegen Cornelia Stein und Dirk Sieger für ihre Hilfestellungen, Antworten und ihren Beitrag an Spaß und Motivation im Labor.

An meine Familie und meine Freunde für ihre liebevolle Unterstützung und für die aufmunternden und motivierenden Worte.

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzen Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen -, die in anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlicht worden ist sowie, dass eine solche Veröffentlichung vor Abschluss des Promotionsverfahren nicht vornehmen werde.
Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Maria Leptin betreut worden.
Teilpublikationen liegen nicht vor.

Köln, den 30.11.2011

Teilpublikationen

Es liegen keine Teilpublikationen vor.