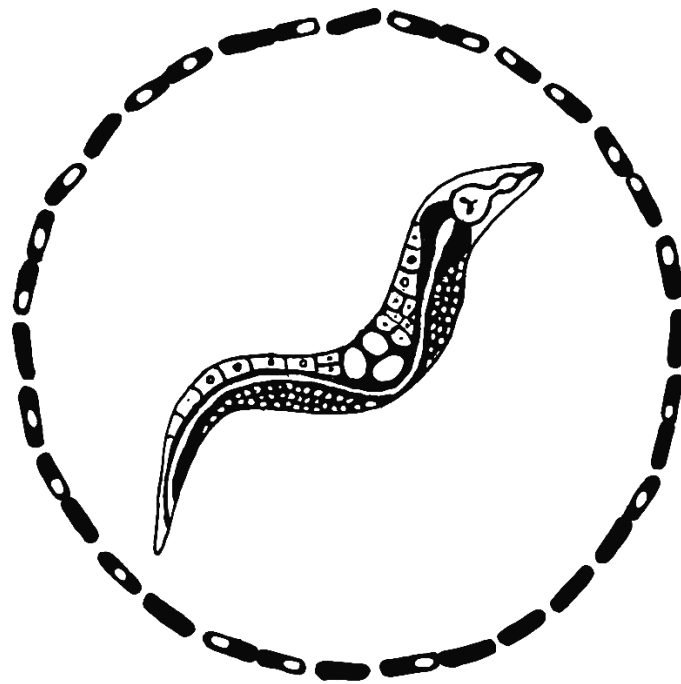


THE ROLE OF C-TYPE LECTIN-LIKE DOMAIN GENES
IN *C. ELEGANS* IMMUNITY



Dissertation

in fulfillment of the requirements for the degree

Doctor rerum naturalium

of the Faculty of Mathematics and Natural Sciences

at the University of Kiel

Submitted by Barbara Pees

Department of Evolutionary Ecology and Genetics

Zoological Institute, Kiel University

Kiel, 2015

First referee: Prof. Dr. Hinrich Schulenburg

Second referee: Prof. Dr. Matthias Leippe

Date of oral examination: 08.02. 2016

Approved for publication: 08.02. 2016

Signature: _____

Contents

Declaration	1
Authors' Contributions.....	3
Summary	5
Zusammenfassung.....	6
Introduction	7
Chapter I.....	27
High Innate Immune Specificity through Diversified C-type Lectin-like Domain Proteins in Invertebrates.	
Chapter II.....	47
The Involvement of C-type Lectin-like Domain Genes in <i>C. elegans</i> Behavioral Immune Defenses.	
Chapter III.....	63
Put to the Test: Immune Function of <i>C. elegans</i> C-type Lectin-like Domain Gene <i>clec-4</i> and its Paralogs.	
Additional Results	83
List of Abbreviations.....	105
Graphical Acknowledgement.....	109

Declaration

I, **Barbara Pees**, declare that:

Apart from my supervisor's guidance the content and design of the thesis is all my own work;

Specific aspects of my thesis were supported by colleagues; their contribution is specified in detail in the following section "Authors' Contributions";

The thesis has not already been submitted neither partially nor wholly as part of a doctoral degree to another examining body. Apart from the included published paper (Chapter I, High Innate Immune Specificity through Diversified C-type Lectin-like Domain Proteins in Invertebrates. *Journal of Innate Immunity*) no other part of the thesis has been published nor submitted for publishing;

The thesis has been prepared subject to the Rules of Good Scientific Practice of the German Research Foundation (DFG).

Signature: _____

Authors' Contributions

This PhD thesis consists of three chapters, each represented by a publication, or unpublished manuscripts. Barbara Pees developed original ideas and wrote the manuscripts with major contribution for chapter II and III, and on collaborative basis for chapter I.

Chapter I

High Innate Immune Specificity through Diversified C-type Lectin-like Domain Proteins in Invertebrates. *J Innate Immun*, 2015.

Barbara Pees, Wentao Yang, Alejandra Zárata-Potes, Hinrich Schulenburg, Katja Dierking

BP, WY, AZP, HS, and KD jointly conceived of the idea, reviewed the relevant literature and wrote the manuscript.

WY performed all meta-analyses.

Chapter II

The Involvement of C-type Lectin-like Domain Genes in *C. elegans* Behavioral Immune Defenses. *Unpublished manuscript*.

Barbara Pees, Anke Kloock, Camilo Barbosa, Rania Nakad, Hinrich Schulenburg, Katja Dierking

BP developed original ideas, supervised AK, planned and performed the experiments, analyzed and interpreted the data, and wrote the manuscript.

AK performed part of the assays (avoidance assay) and provided technical support for other assays (mutant survival screen, validation of screen).

CB analyzed part of the data (mutant survival screen).

RN helped performing part of the assays (mutant screen) and discussed the project.

HS supervised and discussed the project.

KD conceived of the original ideas, supervised and discussed the project, interpreted the data, and wrote the manuscript.

Chapter III

Put to the Test: Immune Function of *C. elegans* C-type Lectin-like Domain Gene *clec-4* and its Paralogs. *Unpublished manuscript.*

Barbara Pees, Anke Kloock, Wentao Yang, Hinrich Schulenburg, Katja Dierking

BP developed original ideas, supervised AK, planned and performed the experiments, analyzed and interpreted the data, and wrote the manuscript.

AK performed part of the assays (avoidance assay).

WY performed the meta-analysis and created the heat map.

HS supervised and discussed the project.

KD conceived of the original ideas, supervised and discussed the project, interpreted the data, and wrote the manuscript.

Hiermit bestätige ich als Betreuer die oben stehenden Angaben.

Summary

C-type lectin-like domain (CTLD) proteins, are mainly known for the crucial roles they play in vertebrate immunity. They function primarily as pattern recognition receptors (PRRs) to identify components of pathogenic intruders, but also as antimicrobial molecules to eliminate pathogens. In invertebrates the exact function of CTLD proteins in immunity is less described despite of the abundance of their encoding genes in many metazoan genomes. One of the most studied model organism – *Caenorhabditis elegans* – particular stands out for its enormous set of highly diversified CTLD proteins, which has the potential to generate immune specificity. Surprisingly, their exact function and contribution to *C. elegans* immunity still remains elusive.

During my PhD I systematically screened the CTLD gene (*clec*) repertoire of *C. elegans* to identify genes that are involved in the defense against the two bacterial pathogens *Pseudomonas aeruginosa* (PA) and *Bacillus thuringiensis* (BT). Unexpectedly, in the tested mutants *clec*-knock-out led equally to susceptibility and resistance to pathogen infection, indicating that *clec* genes mediate both immune activation and suppression. Resistance of *clec* mutants on BT could partially be explained by enhanced pathogen avoidance behavior. We could confirm for the *clec* gene *C54G4.4* that it negatively regulates behavioral immune responses and thus demonstrated for the first time that beside a potential contribution to the physiological immune system, *clec* genes might also function in behavioral immunity.

In parallel to identifying *clec* immune response genes by a candidate mutant screen, we analyzed the expression of *clec* genes upon infection and stress using a database for *C. elegans* expression data. We identified *clec-4* as another interesting candidate, due to its highly up-regulated profile upon exposure to many pathogens and stresses. Surprisingly, the *clec-4* mutant showed resistance towards infection with BT, which suggests a function in negative regulation of the immune response. We further examined the role of one co- and one contra-expressed *clec-4* paralog, *clec-41* and *clec-10*, respectively, which were required for defense against BT infection and both suppressed the resistance phenotype of the *clec-4* mutant. These findings give hint to a complex regulation network of interacting *clec* genes. Moreover, we started to analyze the function of the recombinant CLEC-4 protein in pathogen binding assays.

Taken together, our work is the first step towards a more detailed analysis of the exact *clec*/CLEC function in immunity at the gene and protein level.

Zusammenfassung

Proteine mit einer C-Typ Lektin-ähnlichen Domäne (CTLD) sind vor allem für ihren essentiellen Beitrag im Immunsystem der Vertebraten bekannt. Dort agieren sie hauptsächlich als Mustererkennungsrezeptoren (PRRs), indem sie pathogene Struktur motive identifizieren, oder aber als antimikrobielle Moleküle, die Pathogene eliminieren. Im Immunsystem von Invertebraten ist die Aufgabe der CTLD-Proteine weitaus weniger erforscht, obwohl deren kodierende Gene in vielen metazoischen Genomen zahlreich vertreten sind. Besonders *Caenorhabditis elegans* – einer der meist untersuchten Modellorganismen – sticht durch seine enorme Menge an stark diversifizierten CTLD-Proteinen hervor, die das Potenzial haben, Immunspezifität zu erzeugen. Dennoch weiß man kaum etwas über die genaue Funktion und ihren Beitrag zur Immunabwehr bei *C. elegans*.

Während meiner Doktorarbeit habe ich das Repertoire der CTLD-Gene (*clecs*) von *C. elegans* systematisch untersucht, um Gene, die bei der Immunabwehr gegen zwei bakterielle Pathogene – *Pseudomonas aeruginosa* (PA) und *Bacillus thuringiensis* (BT) – beteiligt sind, zu identifizieren. Die getesteten Knockout-Mutanten zeigten gleichermaßen Resistenz sowie Anfälligkeit gegenüber der Pathogeninfektion. Dies weist darauf hin, dass *clec*-Gene das Immunsystem sowohl aktivieren als auch unterdrücken können. Die Resistenz der *clec*-Mutanten auf BT konnte teilweise durch das erhöhte Pathogenvermeidungsverhalten erklärt werden. Zudem bestätigten wir, dass das *clec*-Gen *C54G4.4* das Vermeidungsverhalten negativ reguliert. Neben der potentiellen Beteiligung an der physiologischen Immunantwort konnten wir somit für *clec*-Gene erstmals einen Einfluss auf das Immunverhalten nachweisen.

Um weitere interessante Kandidaten zu ermitteln, untersuchten wir gleichzeitig mit Hilfe einer *C. elegans*-Transkriptomdatenbank die Expression von *clec*-Genen nach Infektion und Stress. Die Wahl fiel auf *clec-4*, da dieses *clec*-Gen nach Infektion mit einer Vielzahl von Pathogenen und Stress ein hochreguliertes Profil aufwies. Die *clec-4*-Mutante zeigte eine unerwartete Resistenz gegenüber Infektion mit BT, was möglicherweise mit einer negativen Regulation der Immunantwort erklärt werden kann. Des Weiteren untersuchten wir zwei *clec-4* Paraloge, *clec-41* und *clec-10*. Trotz Co-Regulation bzw. Contra-Regulation mit *clec-4* wurden beide für die Immunabwehr gegen BT benötigt und unterdrückten zudem den Resistenzphänotyp von *clec-4*. Diese Ergebnisse deuten auf ein komplexes Regulationsnetzwerk hin, in dem *clec*-Gene miteinander interagieren. Darüber hinaus haben wir mit Hilfe von mikrobiellen Bindungsexperimenten begonnen, die Funktion des rekombinanten CLEC-4-Proteins zu analysieren.

Zusammenfassend ist unsere Arbeit ein erster Schritt zu einer detaillierten Analyse der genauen *clec*/CLEC-Funktion im Immunsystem, sowohl auf genetischer als auch auf Proteinebene.

Introduction

In this PhD project I addressed the immune defenses of an invertebrate host, *Caenorhabditis elegans*, against two bacterial pathogens, Gram-positive *Bacillus thuringiensis* and Gram-negative *Pseudomonas aeruginosa*. I was particularly interested in the involvement of C-type lectin-like domain genes and proteins questioning their exact functions in immune regulation and pathogen recognition or clearance. The model system *C. elegans* with its immune system, the pathogenic antagonists as well as the group of C-type lectin-like domain genes/proteins will be in detail discussed on the following pages.

Caenorhabditis elegans – the Flagship Model of Invertebrates

Caenorhabditis elegans (Figure 1) is a small, free-living nematode, which occurs in nature primarily in rotting material^{1,2}. It mainly exists as a self-fertilizing hermaphrodite, yet males appear infrequently in a low percentage of the hermaphrodite's progeny (e.g. 0.1-0.2% in the commonly used strain N2) due to a non-disjunction of the X chromosome or in approximately 50% after mating with males. The nematode passes four larval stages (L1-L4), which are defined by a molting event at the end of each stage leading to reproducible adulthood. When exposed to stress *C. elegans* may undergo an alternative cycle where the L2 larvae enter an almost indefinite dauer stage until favorable conditions induce the transition to the L4 stage³.

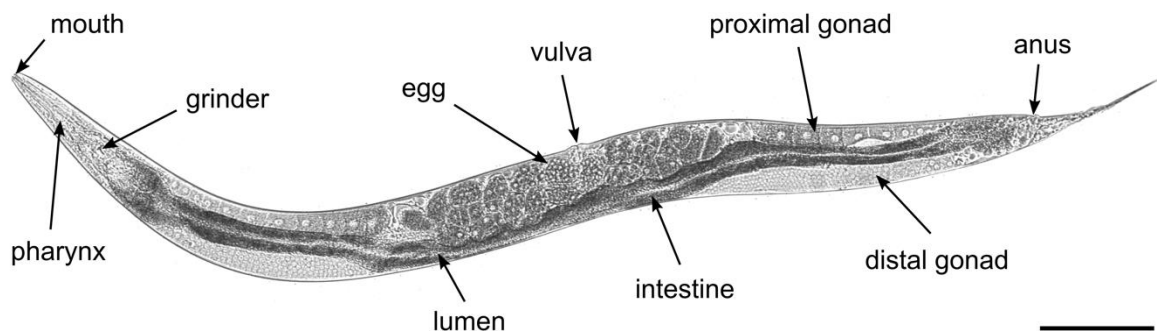


Figure 1. Microscopic picture of *C. elegans* hermaphrodite wild type adult. Photo was taken using Zeiss Axio Observer Z.1 (Carl Zeiss AG, Jena, Germany) with 40 x magnification. Single pictures were stitched and edited with image editing software GIMP 2.8.14 (<http://www.gimp.org/>). The scale bar represents 0.1 mm.

When Sydney Brenner first introduced *C. elegans* as model organism⁴ he laid the foundation for a plethora of significant discoveries in life sciences, some even awarded with Nobel prizes in different disciplines⁵⁻⁷. The reasons for the worm's crusade through biological and medical research are numerous. Such a small animal – as adult it reaches about 1 mm in length – is easy and cheap to maintain

as it can be held in large numbers on Petri dishes, fed with bacteria, and one life cycle from egg to adult only takes 2.5 days at room temperature. The transparency facilitates to observe cellular processes under the microscope *in vivo* (Figure 1) which led to a description of every cell's fate already decades ago⁸, not to mention the constant cell number of 959 in hermaphrodites and 1031 in males. However, there is one major advantage researchers working with other multicellular model organisms can only dream of: *C. elegans* can be frozen unlimitedly at -80 °C and revived within minutes after thawing. Furthermore, besides the fully sequenced genome⁹ a variety of different genetic tools are available ranging from RNA interference¹⁰, comprehensive mutant libraries (e.g. Caenorhabditis Genetics Center, National BioResource Project), and GFP tagging¹¹ (e.g. Figure 2B) to novel technologies like CRISPR-Cas9-mediated gene editing¹². Taking these facts together it is not surprising that *C. elegans* nowadays is one of the leading multicellular models in research laboratories worldwide with a constantly growing “worm community”.

***C. elegans* as a Model for Studying Immunity**

As discussed above *C. elegans* provides remarkable advantages for research but what makes it so special for studying immunity?

In general, the simplicity of model organisms enables scientists to focus on certain aspects postponing the complex processes to higher organisms. As invertebrate *C. elegans* lacks an adaptive immune response engaging professional immune cells (e.g. T and B lymphocytes), which facilitates research purely on the innate immune system. The fact that common immune regulators (e.g. NF- κ B, MyD88) are absent or do not play a crucial role in immunity as in vertebrates (e.g. TOL-1¹³), opens up the possibility to decipher alternative mechanisms of how a host fights pathogenic intruders. Simultaneously, key questions on the evolution of immune systems can be addressed as defensive strategies of vertebrates and invertebrates share impressive similarities¹⁴.

However, despite of its “basic” immune response the nematode exhibits exceptional strategies to sustain in a microbe-rich environment.

Behavioral Defenses

The first, yet in classical immunology widely overlooked, strategy to prevent pathogenic infection is the behavioral avoidance of encountering or/and ingesting harmful microbes.

C. elegans is capable of discriminating harmful microbes from food bacteria and responds accordingly by physical avoidance before mounting a physiological response. The latter is not only costly¹⁵ but might also lead to collateral host damage since reactive oxygen species (ROS) production can be a response towards a pathogen¹⁶. The exact microbial cues which are sensed by the worm can be volatile or water-soluble compounds mostly deriving from bacterial metabolism by-products¹⁷. Examples for bacteria-originated repellents are the surfactant serrawettin produced by *Serratia marcescens*¹⁸ or the secondary

metabolites phenazine-1-carboxamide and pyochelin secreted by *Pseudomonas aeruginosa*¹⁹. But there is also indication for an indirect recognition by cellular damage being a supplementary trigger of aversive behavior as non-pathogenic bacteria do not initiate immune behavior while native non-pathogenic bacteria expressing toxins do²⁰.

The key players in chemosensory regulation of pathogen avoidance are the eleven pairs of ciliated amphid neurons in the head¹⁷. Some of them have been shown to be implicated in two neural circuits: the AWB and AWC neurons mediate the innate olfactory preference for bacteria via expression of neurotransmitters and neuropeptides¹⁸, whereas the ADF neurons are rather responsible for the aversive learning response, a behavior of wild type worms studied on *P. aeruginosa*, where initial attraction to the bacterial lawn turns after a learning phase into avoidance^{21,22}. The aversive learning is based on serotonin-signaling through the tryptophan hydroxylase TPH-1 as the limiting factor in serotonin synthesis²³. The recognition of environmental cues on the molecular level might be modulated by a superfamily of putative chemoreceptors with more than 1000 predicted members, namely the G protein-coupled receptors (GPCRs), since a reasonable number is expressed in the amphid neurons²⁴. Other sensory neurons likewise contribute to the modulation of immune behavior, e.g. HECT domain-containing E3 ubiquitin ligase encoding *hecw-1* is expressed in the OLL mechanosensory neurons to repress *npr-1*²⁵, which in turn is required to promote oxygen-dependent avoidance²⁶.

Additionally, as mentioned above, *C. elegans* might react to pathogen exposure with feeding cessation which is mediated by the insulin-like receptor (ILR) pathway²⁷.

Physical Barriers

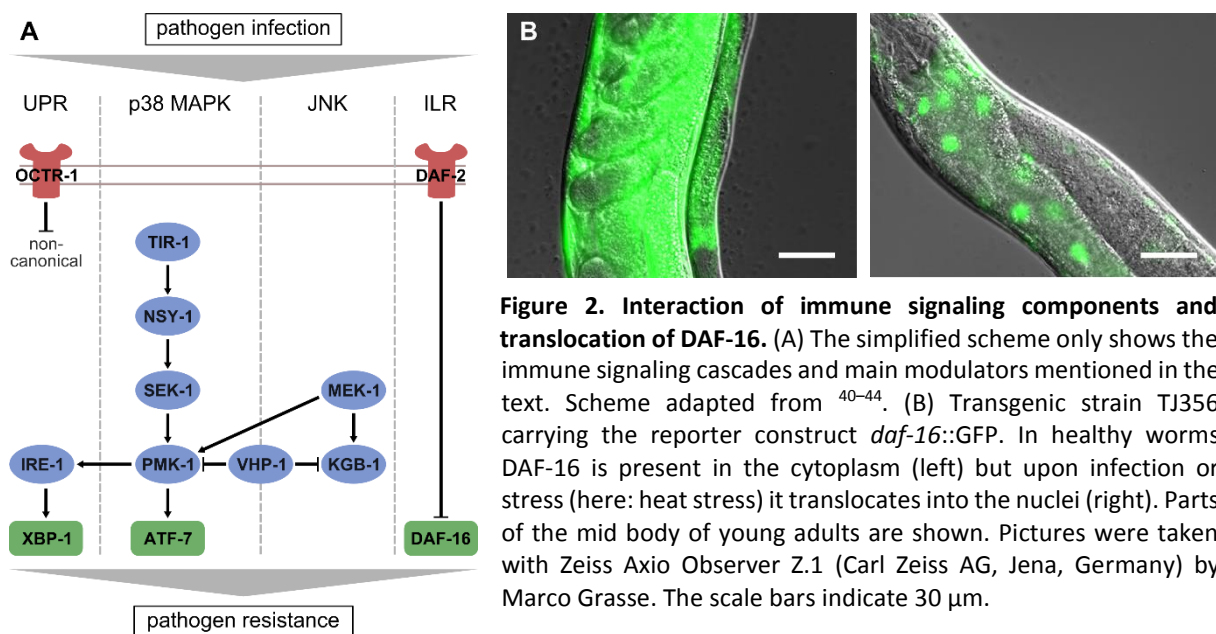
If host and pathogen eventually encounter each other entering the host is hampered by physical barriers. The cuticle as most apparent obstacle provides a thick, impermeable surface, which should prevent attachment of pathogens that attack from the outside, such as *Drechmeria coniospora*²⁸ or *Microbacterium nematophilum*²⁹. A more common route for pathogens to establish intestinal infections in the host is the mouth. Therefore, a mechanic structure in the second bulb of the pharynx, called grinder (Figure 1), protects this opening by destroying sucked in bacteria as first line of digestion but also defense³⁰.

Physiological Response – Recognition, Processing and Elimination

The final defense mechanism is the induction of physiological responses to fight off and clear pathogen attack. The inducible immune response comprises three main steps: (i) the recognition of microbe-associated molecular patterns (MAMPs) or damage-associated molecular patterns (DAMPs) by pattern recognition receptors (PRRs), (ii) consequent signaling forwarding the received information to transcription factors, which mediate (iii) expression of effector molecules, that eventually eliminate the intruders³¹. These levels will be discussed in the following one by one.

PRRs in *C. elegans* have long been a subject of immunological studies since the worm lacks prominent immune receptors that fulfill the task of pattern recognition in higher organisms' innate immunity. For example, the only Toll-like receptor (TLR) orthologue TOL-1 does not hold the crucial function of initiating antimicrobial peptide (AMP) expression upon pathogen perception as it does in *Drosophila* and many vertebrates¹³, and other PRRs like NOD-like receptors (NLRs) of vertebrates do not exist at all³². C-type lectin-like domain (CTLD) proteins, which will be discussed in depth throughout the thesis, as well as the aforementioned chemosensory GPCRs have been proposed as potential PRRs^{33,34}. An example for the latter is FSHR-1, which is required for defense against several pathogens and oxidative stresses but also pathogen avoidance, and acts independently of known immune pathways in the intestine^{35,36}. More detailed findings for a function of a GPCR in pathogen recognition could be demonstrated by Zugasti and colleagues, who identified the GPCR DCAR-1 with its endogenous ligand HPLA (4-hydroxyphenyllactic acid), a DAMP generated upon disruption of the epidermis by *D. coniospora* or wounding. DCAR-1 acts in the epidermis to modulate the expression of AMPs through the conserved p38 mitogen-activated protein kinase (MAPK) cascade (see below)³⁷. Together with the RIG-I (retinoic acid inducible gene)-like helicase DRH-1 (dicer related helicase), which is a receptor for virus-derived RNA³⁸ these two receptors are heretofore the only PRRs described in *C. elegans*.

Despite the knowledge gap on how *C. elegans* recognizes pathogenic attack the several intermediate signaling pathways activating the expression of effector molecules have been described and repeatedly reviewed^{31,39}. With regards to the model system used in this study I will focus on the most relevant pathways implicated in the intestinal host response to *Bacillus thuringiensis* and *P. aeruginosa* (see "Pathogens and their Interaction with *C. elegans*" and Figure 2A).



The principal and most ancient immune cascade is the conserved p38 MAPK pathway with its key kinase regulators NSY-1, SEK-1, and PMK-1. First described in 2002 upon infection with *P. aeruginosa*⁴¹ it shortly was shown to be implicated in the defense against a variety of bacterial and fungal pathogens, e.g.⁴⁵⁻⁴⁷. Upstream of the p38 cascade acts Toll-interleukin 1 receptor TIR-1, which contains multiple leucine-rich repeats, a well-known feature of TLRs in plants and animals⁴⁸. One of the downstream targets of PMK-1 was identified as the transcription factor ATF-7, which expresses infection-induced genes upon phosphorylation by PMK-1⁴⁶. The p38 MAPK requires the activation by a second MAPK pathway, the c-Jun N-terminal kinase (JNK) cascade previously shown to be involved in the stress response to heavy metal. MEK-1 and VHP-1 cross-regulate the activity of PMK-1 in pathogen defense^{40,49}. In the immune response against pore-forming toxins (PFTs), as produced by *B. thuringiensis*, JNK-like MAPK KGB-1 even seems to be the central regulator for transcriptional induction of defense⁵⁰ (Figure 2A).

The ILR pathway including the receptor tyrosine kinase DAF-2 and the FoxO transcription factor DAF-16 has been discovered due to its role in longevity and dauer regulation⁵¹. Interestingly, *daf-2* mutants do not only live twice as long as the wild type⁵², but also exhibit resistance against multiple stresses (e.g. heavy metal, heat, or osmotic stress⁵³⁻⁵⁵) and pathogen exposure^{42,56}. The activation of DAF-2 leads to the inactivation of DAF-16 in the cytoplasm whereas *daf-2* deficiency results in an active nuclear DAF-16 inducing gene expression⁵⁷. In how far DAF-16 contributes to resistance against pathogens or rather generally to stress is not yet clear. Pleading for the latter is the observation that the expression of PMK-1-dependent genes upon infection is repressed by DAF-16 but induced during stress^{58,59}. Nonetheless, in a study from our group it was shown that DAF-16 directly translocates to the nucleus following *B. thuringiensis* exposure⁶⁰ (Figure 2B).

A further signaling cascade that interacts with the pivotal p38 MAPK pathway belongs as well to the compartment of stress response management, the unfolded protein response (UPR). Upon pathogen attack PMK-1 mediates in addition to expression of immune effector genes the activation of the UPR in order to handle cellular accumulation of un- and misfolded proteins. The UPR is thought to be the protective force against self-damage caused by AMPs of the immune defense⁶¹. The p38 MAPK-dependent activation of the conserved IRE-1-XBP-1 branch of the UPR was shown to act in the defense against *P. aeruginosa*⁶¹ and *B. thuringiensis*' PFTs⁴⁴. Sun *et al.*, moreover, presented a GPCR in the amphid neurons, OCTR-1 (octopamine receptor), to suppress the non-canonical UPR in non-neuronal tissue independent of behavioral pathogen avoidance⁴³, suggesting an intervention of the nervous system in the immune response (Figure 2A). The mitochondrial UPR (UPR^{mt}) additionally seems to promote resistance by protecting from mitochondrial damage caused by bacteria that target mitochondrial processes, like *P. aeruginosa*⁶².

After a successful signal transmission from PRRs via the signaling cascades to transcription factors, the expression of effector molecules is triggered. I will here focus on AMP expression and give some examples. *C. elegans* AMPs have been identified mostly by their up-regulation after pathogenic infection coupled with homology searches for known metazoan antimicrobial peptides and proteins. Based on that, several families are considered to hold antimicrobial activity, but the exact function has only been demonstrated for few proteins^{63,64}. For example, the defensin-like peptide ABF-2 – identified through homology with *Ascaris suum* antibacterial factors (ASABFs)⁶⁵, and consequently named Ce-ABF-2 or only ABF-2 – affects a broad range of pathogens but more strongly Gram-positive bacteria⁶⁶. Caenopores or saposin-like proteins (SPPs), share similarities with amoebapores of the protozoan parasite *Entamoeba histolytica*. Several SPPs have been reported to display antimicrobial activity against a variety of bacteria, including *P. aeruginosa* and *B. thuringiensis* used in the present study, their expression – primarily in the intestine – is partially regulated via the ILR pathway^{67–70}. Furthermore, genes encoding neuropeptide like peptides (NLPs) and the closely related caenacins (CNCs) are transcriptionally induced by infection with *D. coniospora*, the *nlp*s additionally by *S. marcescens* infection in both intestinal and epidermal epithelium^{71,72}. Yet, only NLP-31 has been reported to show antifungal activity at the protein level *in vitro*⁷².

Despite the mentioned examples of research on AMP characterization *C. elegans* AMPs remain a poorly studied group of proteins due to the lack of functional studies on both the gene and protein level.

Pathogens and their Interaction with *C. elegans*

Another benefit of the model organism *C. elegans* is the large variety of infection agents, which can be applied in immunity studies. Since *C. elegans* lives in decaying organic matter, in other words microbe-rich environments^{1,2}, and as bacteriovore feeds on the very same, experimentally introducing intestinal pathogens is rather uncomplicated. Also many human pathogens are able to establish infections in the worm, which allows the investigation of clinically relevant host-pathogen interaction *in vivo*. A considerable diversity of microbes has been used to infect *C. elegans*; Gram-positive, Gram-negative, fungal, and microsporidial parasites with intestinal, epidermal, anal, and vulval routes of infection, applying a plethora of killing mechanisms^{31,73}. In our study we employed the likely natural *B. thuringiensis* and human opportunistic pathogen *P. aeruginosa*, which will be introduced in more detail in the following.

Bacillus thuringiensis – the Natural Pathogen

Gram-positive *B. thuringiensis* (BT) is known especially for its application as biological pesticide, as alternative to chemical pesticides, and tool in genetic engineering to generate transgenic resistant crop in agriculture^{74,75}. This utilization it owes to its ability to infect and kill a wide range of insect taxa such as lepidopterans, dipterans, hymenopterans, and coleopterans, serving the purpose as pest control^{76,77}.

BT kills through PFTs called Cry toxins due to their crystalline structure (Figure 3A) produced during sporulation. Secreted by bacteria PFTs bind to host receptors, multimerize and pierce a hole into the plasma membrane eventually causing cell lysis and ion leakage. The solubilization of inactive pro-toxins by the gut's milieu, the activation through proteolytic cleavage as well as the specificity of the receptor are inevitable parts of the infection process provided by the host after ingestion^{78,79}(Figure 3B).

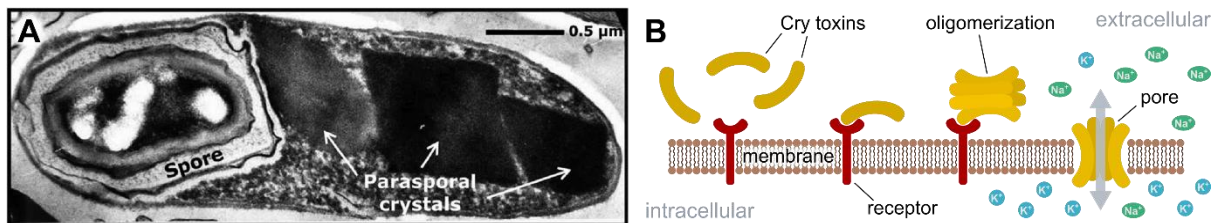


Figure 3. Pathogenic *B. thuringiensis* with spores and crystals, and PFTs' mode of infection. (A) Transmission electron microscopic picture of single sporulated BT cell taken from⁸⁰. (B) Scheme of infection by PFTs adapted from^{79,81}.

The host spectrum of each BT strain is determined by its repertoire of Cry toxin encoding genes and is based almost exclusively on the presence of the specific toxin receptor^{81,82}. For instance, Cry4Ba is specifically active against dipterans such as *Anopheles* and *Aedes*, Cry1Aa against Lepidopterans such as *Diatraea* and *Chilo*⁸³⁻⁸⁸, whereas synergistic combinations might expand the host spectrum^{89,90}.

The BT strains used in the following study are strain BT18247 which produces Cry6Ba⁹¹ and strain BT407 Cry⁻ as non-pathogenic counterpart.

B. thuringiensis vs. *C. elegans*

More than a decade ago BT was introduced as a pathogen of model organism *C. elegans* in order to study its toxicity on parasitic nematodes, improve the understanding of the toxins' mode of action^{92,93}, and – most importantly for the *C. elegans* community – to learn more about the nematode's immune response towards this potentially natural pathogen^{30,34}. Commonly, the effect of BT toxins towards the worm is tested by applying toxin expressing *E. coli* instead of the complete bacterium or spore. Like this it has been demonstrated that *C. elegans* is susceptible to several Cry toxin variants⁹⁴. But only Cry5B was used to discover Cry toxin-resistant mutants, which were resistant to Cry5B⁹² and cross-resistant to Cry14A⁹⁵. The underlying mutations were identified in the *bre* (Bacillus-toxin resistant) genes, which encode glycosyltransferases being involved in the synthesis of glycolipid receptors⁹⁶. Thus, the ability of Cry5B and Cry14A expressing BT strains to infect *C. elegans* is modulated by these specific gut receptors. Alternatively, the glycolipid receptor can be actively blocked by *C. elegans* immune defenses promoting resistance due to competitive inhibition: Galectin LEC-8 is expressed in

the intestine upon Cry5B exposure and is able to bind to the very same receptors BT requires for infection, consequently preventing a successful docking of the virulence factor to the host tissue⁹⁷.

Resistance to exposure with Cry5B further requires components of the two aforementioned MAPK pathways, p38 and JNK-like cascade⁹⁸, where the former activates UPR signaling⁴⁴, which in turn interacts with the hypoxia pathway⁹⁹. The ILR pathway, moreover, is mainly involved in the response to Cry6B produced by BT18247 in both physiological and behavioral immunity^{27,60}.

Pseudomonas aeruginosa – the Human Pathogen

Pseudomonads are ubiquitously occurring Gram-negative bacteria, occupying a variety of different niches, with *P. aeruginosa* (PA) being the most extensively studied genus member^{100,101}. First described in 1894¹⁰², PA quickly became the center of attention since it is a major threat of injured, burned and immuno-compromised patients in hospitals and might cause chronic pulmonary infections in cystic fibrosis patients^{103,104}. Thus, it is not surprising that the most prominent PA strains research is done on, PA01 and PA14, originally arose from clinical isolates. Both strains were fully sequenced^{105,106}, but the more virulent PA14 is nowadays predominantly subject of infection studies in a broad range of host organisms, such as plants, insects, other non-mammalian hosts^{107–111}, and also antibiotic resistance research¹¹².

The ability to infect organisms of such evolutionary divergence might be associated with the multiplicity of virulence factors PA employs¹¹³. Key PA virulence factors are pyoverdinin siderophores (also the reason for its characteristic green color) which sequester iron from the host, an embattled resource for redox enzymes among both pro- and eukaryotes¹¹⁴. The most toxic secreted virulence factor, exotoxin A, however, is known to inhibit the protein synthesis of eukaryotic host cells, eventually leading to cell death^{113,115}. The whole PA toxin variety has been reviewed in e.g.^{116,117}.

P. aeruginosa vs. *C. elegans*

PA is the most commonly applied bacterial pathogen in *C. elegans* immunity studies¹¹⁸. One major convenience of using the nematode as a model is that different PA toxins can be examined according to the set-up of infection. If worms are fed on high-osmolarity solid media, referred to as “fast killing”, diffusible toxins, such as phenazine pyocyanin, kill *C. elegans* within few hours. “Slow killing” happens, on the other hand, on low-nutrient media, where PA proliferates in the gut and causes typical symptoms of infection (e.g. paralysis, distended intestinal lumen) eventually leading to death after several days^{119,120}. A modification of the latter method in liquid media identified pyoverdinin as major infectious agent independent of the other previous mentioned factors¹²¹.

Similar to the immune response to BT’s Cry toxins, PA infections activate the ancient p38 MAPK cascade in cross-talk with the JNK-like pathway⁴¹, and additional involvement of the transcription factor ZIP-2 and the already mentioned GPCR FSHR-1, particularly against exotoxin A¹²². Both slow and fast

killing can induce the UPR^{mt} as a consequence of mitochondrial dysfunction⁶², whereas the UPR is activated in a microRNA-¹²³ but also p38 MAPK-dependent manner⁶¹. Interestingly, PA is also able to actively manipulate the host's immune response by suppressing the expression of immune effectors via the ILR pathway⁵⁶. PA further elicits an aversive learning response as mentioned in section “Behavioral Defenses”.

The Crucial Role of C-type Lectins in Vertebrate Innate Immunity

C-type lectins are calcium-dependent (C-type) carbohydrate-binding lectins which share a characteristic carbohydrate-recognition domain (CRD)¹²⁴. The CRD motif, however, does not necessarily confer sugar- nor calcium-binding activity which is why a more general term, the C-type lectin-like domain (CTLD) protein, was introduced¹²⁵. CTLD proteins are mainly found throughout metazoan taxa but also – although more rarely – present in plants, bacteria, and viruses¹²⁶. In innate immunity they are of crucial importance being involved in both, initial pathogen recognition and final pathogen elimination (cf. “Physiological Response – Recognition, Processing and Elimination”).

C-type Lectin Receptors and their Immune Signaling

C-type lectin receptors (CLRs) recognize MAMPs coming from diverse pathogens as well as DAMPs released by injured cells. Recognition of fungal infections is the best studied function of vertebrate CLRs. Dectin-1 and Dectin-2, for example confer protection against *Candida albicans* (among others), the causing agent of candidiasis in mucosal membranes¹²⁷, by binding to different cell wall components, β -glucan or α -mannose respectively^{128,129}. The binding affinity of Mincle and the mannose receptor (MR) has been shown to be directed against fungi, too, but also against few bacterial species, and regarding MR even viruses and protozoan parasites^{130,131}. CLRs further have been associated with viral recognition, either by promoting (e.g. CLEC5A and dengue virus¹³²) or counteracting (e.g. DNGR-1 and vaccinia virus¹³³) viral infections. The detection of DAMPs, such as F-actin and uric acid crystals, has been attributed to DNGR-1^{134,135} and Clec12a¹³⁶, respectively (Figure 4A). Nevertheless, in contrast to their contribution to fighting fungal pathogenesis, classical CLRs acting in defense against bacterial infections are less described.

For some transmembrane C-type lectins the cytoplasmic domain determines the direction of signaling: the immunoreceptor tyrosine-based activation motif (ITAM) or ITAM-like motif promote subsequent signaling (e.g. of Dectin-1, DNGR-1) while the immunoreceptor tyrosine-based inhibitory motif (ITIM) prevents inappropriate immune responses (e.g. of MICAL). ITAM-containing CLRs are coupled to Syk (spleen tyrosine kinase)-signaling, resulting in activation of different immune responses that might involve MAPK pathways such as the p38, ERK, and JNK cascades, or NF- κ B. These immune activations can be negatively regulated by ITIM-containing CLRs either directly at the ITAM site of other receptors or at the course of the activating signaling¹³⁷. Interestingly, CLRs might synergistically

collaborate with each other (e.g. Dectin-1 and -2) but also with other types of PRRs (e.g. Dectin-1 and TLR2)¹³¹ (Figure 4A).

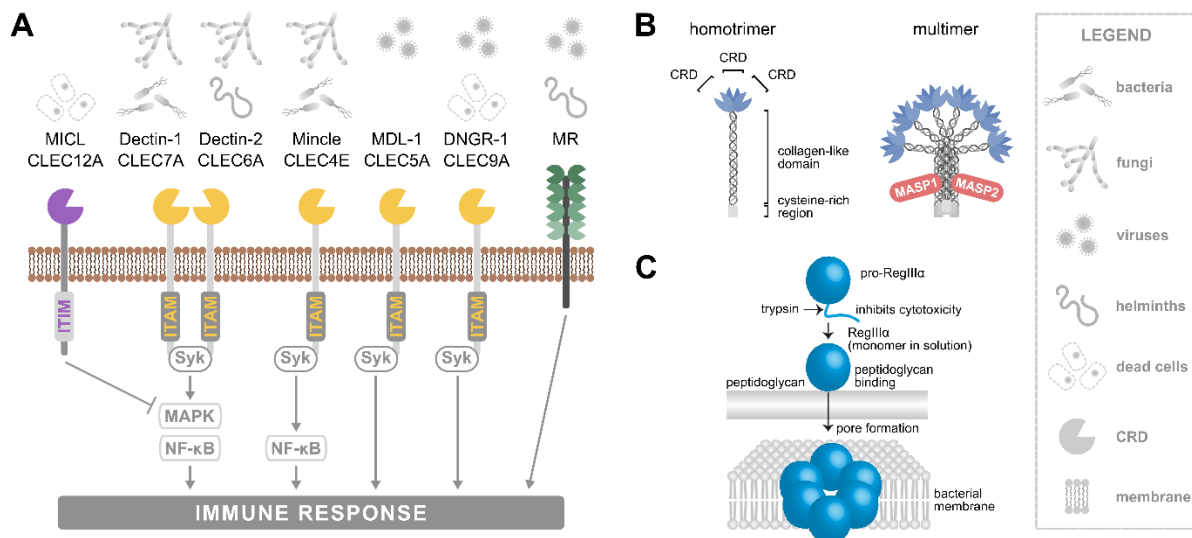


Figure 4. Examples of C-type lectins in vertebrate immunity. (A) Simplified scheme of CLR, which were mentioned in the text. Shown are the pathogens or endogenous constituents which they are able to recognize, their C-type lectin name as well as an alternative name, the domain architecture, and some major signaling components that lead to activation of the immune response. Adapted from ¹³⁷⁻¹³⁹. (B) Domain architecture and multimerization of MBL. Adapted from ^{140,141}. (C) Model for pore-forming process of human RegIIIα. After activation of soluble pro-RegIIIα by trypsin, mature RegIIIα binds to peptidoglycan of Gram-positive bacteria and final oligomerization of monomers leads to formation of pore. Taken from ¹⁴².

The Mannose-binding Lectin as Inducer of the Complement System

Other important C-type lectins within vertebrate immunity are the secreted collectins (collagen-containing C-type lectins). Here, I focus on the mannose-binding lectin (MBL) since it is the major regulator of the lectin pathway, which is next to the classical and the alternative pathway the evolutionary most ancient pathway to activate the complement system^{143,144}. MBL circulates in the blood serum as a multimer of homotrimers which facilitates binding of more complex ligands (Figure 4B). The range of infectious agents being recognized by MBL comprises all principal pathogen groups with specific binding of mannose, fucose, and GlcNAc (N-acetyl-glucosamine) (summarized in ^{145,146}). Furthermore, MBL is not only able to distinguish between self, non-self but also altered-self in the shape of necrotic, apoptotic or cancer cells^{147,148}. Upon recognition of these MAMPs or DAMPs, MBL forms a complex with MBL-associated serine proteases (MASP1 and MASP2) at the collagenous region and is then able to induce/trigger complement activation. Thus, in interacting with the classical pathway MBL triggers inflammation, lysis and opsonization¹⁴⁹. The crucial role of MBL has been repeatedly demonstrated as MBL-deficient patients suffer from increased infection and susceptibility to sepsis potentially leading to death¹⁵⁰.

C-type Lectins as Effector Proteins

Last but not least, vertebrate C-type lectins have also been shown to be directly involved in downstream processes of immunity, such as elimination of pathogens. The best examples for such C-type lectin effectors are members of the RegIII family. RegIII γ in mice (RegIII α in humans), for instance, is strongly expressed in the intestinal epithelium in a MyD88-dependent manner triggered by bacterial gut colonization and infection^{151,152}. It interacts preferably with Gram-positive bacteria, as peptidoglycan is a ligand, but on the contrary is inhibited by lipopolysaccharide (LPS) in the cell wall of Gram-negative bacteria. Secreted into the gut lumen RegIII γ attacks bacteria through oligomerization and pore formation and thereby limits the colonization of the epithelial surface^{142,151} (Figure 4C). The close relative RegIII β acts similar to RegIII γ , but targets to a greater extent Gram-negative bacteria due to binding of Lipid A of LPS^{153,154}. Thus, these C-type lectins function as antimicrobial proteins and are thought to represent an evolutionary primitive innate immune response distinct from lectin-mediated complement activation, but critical for maintenance of the microbiome homeostasis in the gut^{142,151,152}.

C-type Lectins in *C. elegans* and other Invertebrates

As above exemplarily summarized for vertebrates (specifically mice and humans), the multifaceted contribution of C-type lectins in immunity emphasizes the significance of this protein family in both the humoral and cellular arm of the innate immune defense and as supplement to adaptive defenses. But CTLD proteins also exist in various invertebrate taxa (e.g. ¹⁵⁵⁻¹⁵⁷), yet, their role in immunity is not as well explored. Only single C-type lectins have been implicated in promoting immune processes encompassing phagocytosis, encapsulation, or melanization (e.g. ^{158,159}) and also directly in pathogen clearance (e.g. ¹⁶⁰⁻¹⁶²). Nevertheless, the indispensability of CTLD proteins in immune defenses like in vertebrates has not been demonstrated yet. That might not be surprising given the variety of defense strategies found in invertebrate taxa already in different species¹⁶³.

C. elegans has an intriguingly high number of genes encoding CTLD proteins, which are predicted to be very diverse¹⁶⁴. The majority of the CTLD protein encoding genes (*clecs*) are organized in genomic clusters, which hints to a rapidly evolving gene family under strong selective pressure by co-evolving pathogens¹⁶⁵. Together with their up-regulation upon various pathogenic threats and stresses CTLD proteins have always been considered to be involved in *C. elegans* immunity. But still we lack fundamental knowledge about the exact role they might play in defending the host organism.

A more detailed and thorough discussion about CTLD proteins in invertebrate and particularly in *C. elegans* immunity is provided in the first chapter.

PhD Thesis Content

The current PhD thesis addresses the role of *clec* genes and their encoded proteins as potential regulators of immunity in the model nematode *C. elegans*. The aim of the study was (i) to question the inevitable label given to *clec* genes as “immune genes” serving crucial functions in *C. elegans* immune defense, (ii) to systematically screen the *clec* gene repertoire in order to extract potential candidate immune genes, (iii) and ultimately to characterize the selected genes for their specific function in immunity.

Chapter I represents a review on “High Innate Immune Specificity through Diversified C-type Lectin-like Domain Proteins in Invertebrates”. It gives a detailed introduction into the current knowledge of *clec* gene and protein functions in invertebrate immunity. It deals with the issue of immune specificity in invertebrate organisms and suggests CTLD proteins as promising candidates for generating immune specificity. First compiling the latest findings in representative taxa (insects and crustaceans) the review then focuses on nematodes, which show extreme CTLD gene diversification. Uncovering the largely unknown functions of invertebrate CTLD proteins requires novel and comprehensive research approaches, which are discussed in the concluding remarks of the review.

Chapter II is a manuscript that is now ready for submission. We here demonstrate that *C. elegans clec* genes might function in behavioral defense responses. A survival screen of *clec* gene knock-out mutants exposed to bacterial pathogens revealed resistant as well as susceptible phenotypes. The resistance phenotype correlated with the number of escaped worms. In a thorough analysis of the *clec* gene *C54G4.4* we confirmed that CTLD proteins are involved in *C. elegans* immune behaviors.

Another in depth analysis of one *clec* gene is described in **Chapter III**, which is a preliminary manuscript. Infection with several *C. elegans* pathogens leads to the strong up-regulation of *clec-4*. The knock-out unexpectedly showed a resistance survival phenotype on *B. thuringiensis*, which is suppressed by additional RNAi knock-down of the two *clec-4* paralogs *clec-41* and *clec-10*. This indicates that different *clec* genes mediate immune responses by interacting with each other. Furthermore, analysis of the recombinant protein provides indication that CLEC-4 binds certain bacterial pathogens. These results represent an exciting basis for untangling genetic networks within the *clec* superfamily and discovering potential antagonists of pathogenic intruders.

The chapter **Additional Results** comprises obtained data on additional *clec* genes, which were insufficient to include in a manuscript. It contains the initial efforts to find interesting *clec* gene candidates by systematic RNAi screens of the majority of the *clec* genes repertoire and also more detailed analyses of selected *clec* genes. Furthermore, the *in vivo* expression pattern of some *clec* genes is shown and all generated transgenic reporter strains are compiled in a list that also gives information about the used transgene concentrations and primers.

References – Introduction

1. Petersen, C., Dirksen, P., Prah, S., Strathmann, E. A. & Schulenburg, H. The prevalence of *Caenorhabditis elegans* across 1.5 years in selected North German locations: the importance of substrate type, abiotic parameters, and *Caenorhabditis* competitors. *BMC Ecol.* **14**, 4 (2014).
2. Félix, M.-A. & Braendle, C. The natural history of *Caenorhabditis elegans*. *Curr. Biol.* **20**, R965–R969 (2010).
3. Corsi, A. K. A transparent window into biology: A primer on *Caenorhabditis elegans*. *WormBook* 1–31 (2015). doi:10.1895/wormbook.1.177.1
4. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
5. Zamore, P. D. RNA interference: Big applause for silencing in Stockholm. *Cell* **127**, 1083–1086 (2006).
6. Marx, J. Nobels run the gamut from cells to the cosmos. *Science* **298**, 526 (2002).
7. Zimmer, M. GFP: From jellyfish to the Nobel prize and beyond. *Chem. Soc. Rev.* **38**, 2823 (2009).
8. Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64–119 (1983).
9. The *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* **282**, 2012–2018 (1998).
10. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
11. Chalfie, M., Yuan, T., Euskirchen, G., Ward, W. W. & Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805 (1994).
12. Friedland, A. E. *et al.* Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat. Methods* **10**, 741–743 (2013).
13. Pujol, N. *et al.* A reverse genetic analysis of components of the Toll signaling pathway in *Caenorhabditis elegans*. *Curr. Biol.* **11**, 809–821 (2001).
14. Kimbrell, D. A. & Beutler, B. The evolution and genetics of innate immunity. *Nat. Genet.* **2**, 256–267 (2001).
15. Schmid-Hempel, P. & Ebert, D. On the evolutionary ecology of specific immune defence. *Trends Ecol. Evol.* **18**, 27–32 (2003).
16. Chavez, V., Mohri-Shiomi, A., Maadani, A., Vega, L. A. & Garsin, D. A. Oxidative stress enzymes are required for DAF-16-mediated immunity due to generation of reactive oxygen species by *Caenorhabditis elegans*. *Genetics* **176**, 1567–1577 (2007).
17. Bargmann, C. I. Chemosensation in *C. elegans*. *WormBook* (2006). doi:10.1895/wormbook.1.123.1
18. Pradel, E. *et al.* Detection and avoidance of a natural product from the pathogenic bacterium *Serratia marcescens* by *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* **104**, 2295–2300 (2007).
19. Meisel, J. D., Panda, O., Mahanti, P., Schroeder, F. C. & Kim, D. H. Chemosensation of bacterial secondary metabolites modulates neuroendocrine signaling and behavior of *C. elegans*. *Cell* **159**, 267–280 (2014).
20. Meisel, J. D. & Kim, D. H. Behavioral avoidance of pathogenic bacteria by *Caenorhabditis elegans*. *Trends Immunol.* **35**, 465–470 (2014).
21. Zhang, Y., Lu, H. & Bargmann, C. I. Pathogenic bacteria induce aversive olfactory learning in *Caenorhabditis elegans*. *Nature* **438**, 179–184 (2005).
22. Ha, H. *et al.* Functional organization of a neural network for aversive olfactory learning in *Caenorhabditis elegans*. *Neuron* **68**, 1173–1186 (2010).
23. Shivers, R. P., Kooistra, T., Chu, S. W., Pagano, D. J. & Kim, D. H. Tissue-specific activities of an immune signaling module regulate physiological responses to pathogenic and nutritional bacteria in *C. elegans*. *Cell Host Microbe* **6**, 321–330 (2009).

24. Robertson, H. The putative chemoreceptor families of *C. elegans*. *WormBook* (2006). doi:10.1895/wormbook.1.66.1
25. Chang, H. C., Paek, J. & Kim, D. H. Natural polymorphisms in *C. elegans* HECW-1 E3 ligase affect pathogen avoidance behaviour. *Nature* (2011). doi:10.1038/nature10643
26. Reddy, K. C., Andersen, E. C., Kruglyak, L. & Kim, D. H. A polymorphism in *npr-1* is a behavioral determinant of pathogen susceptibility in *C. elegans*. *Science* **323**, 382–384 (2009).
27. Hasshoff, M., Böhnisch, C., Tonn, D., Hasert, B. & Schulenburg, H. The role of *Caenorhabditis elegans* insulin-like signaling in the behavioral avoidance of pathogenic *Bacillus thuringiensis*. *FASEB J.* **21**, 1801–1812 (2007).
28. Jansson, H.-B. Adhesion of conidia of *Drechmeria coniospora* to *Caenorhabditis elegans* wild type and mutants. *J. Nematol.* **26**, 430 (1994).
29. Hodgkin, J., Kuwabara, P. E. & Corneliussen, B. A novel bacterial pathogen, *Microbacterium nematophilum*, induces morphological change in the nematode *C. elegans*. *Curr. Biol.* **10**, 1615–1618 (2000).
30. Gravato-Nobre, M. J. & Hodgkin, J. *Caenorhabditis elegans* as a model for innate immunity to pathogens. *Cell. Microbiol.* **7**, 741–751 (2005).
31. Engelmann, I. & Pujol, N. in *Invertebrate Immunity* 105–121 (Springer, 2010).
32. Bryant, C. E. & Monie, T. P. Mice, men and the relatives: Cross-species studies underpin innate immunity. *Open Biol.* **2**, 120015–120015 (2012).
33. Nicholas, H. R. & Hodgkin, J. Responses to infection and possible recognition strategies in the innate immune system of *Caenorhabditis elegans*. *Mol. Immunol.* **41**, 479–493 (2004).
34. Schulenburg, H., Kurz, C. L. & Ewbank, J. J. Evolution of the innate immune system: The worm perspective. *Immunol. Rev.* **198**, 36–58 (2004).
35. Powell, J. R., Kim, D. H. & Ausubel, F. M. The G protein-coupled receptor FSHR-1 is required for the *Caenorhabditis elegans* innate immune response. *Proc. Natl. Acad. Sci.* **106**, 2782–2787 (2009).
36. Miller, E. V., Grandi, L. N., Giannini, J. A., Robinson, J. D. & Powell, J. R. The conserved G-protein coupled receptor FSHR-1 regulates protective host responses to infection and oxidative stress. *PLOS ONE* **10**, e0137403 (2015).
37. Zugasti, O. *et al.* Activation of a G protein-coupled receptor by its endogenous ligand triggers the innate immune response of *Caenorhabditis elegans*. *Nat. Immunol.* **15**, 833–838 (2014).
38. Ashe, A. *et al.* A deletion polymorphism in the *Caenorhabditis elegans* RIG-I homolog disables viral RNA dicing and antiviral immunity. *Elife* **2**, e00994 (2013).
39. Ewbank, J. Signaling in the immune response. *WormBook* (2006). doi:10.1895/wormbook.1.83.1
40. Kim, D. H. *et al.* Integration of *Caenorhabditis elegans* MAPK pathways mediating immunity and stress resistance by MEK-1 MAPK kinase and VHP-1 MAPK phosphatase. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 10990–10994 (2004).
41. Kim, D. H. *et al.* A conserved p38 MAP kinase pathway in *Caenorhabditis elegans* innate immunity. *Science* **297**, 623–626 (2002).
42. Garsin, D. A. Long-lived *C. elegans daf-2* mutants are resistant to bacterial pathogens. *Science* **300**, 1921–1921 (2003).
43. Sun, J., Singh, V., Kajino-Sakamoto, R. & Aballay, A. Neuronal GPCR controls innate immunity by regulating noncanonical unfolded protein response genes. *Science* **332**, 729–732 (2011).
44. Bischof, L. J. *et al.* Activation of the unfolded protein response is required for defenses against bacterial pore-forming toxin *in vivo*. *PLoS Pathog.* **4**, e1000176 (2008).
45. Bolz, D. D., Tenor, J. L. & Aballay, A. A conserved PMK-1/p38 MAPK is required in *Caenorhabditis elegans* tissue-specific immune response to *Yersinia pestis* infection. *J. Biol. Chem.* **285**, 10832–10840 (2010).

46. Shivers, R. P. *et al.* Phosphorylation of the conserved transcription factor ATF-7 by PMK-1 p38 MAPK regulates innate immunity in *Caenorhabditis elegans*. *PLoS Genet.* **6**, e1000892 (2010).
47. Pukkila-Worley, R., Ausubel, F. M. & Mylonakis, E. *Candida albicans* infection of *Caenorhabditis elegans* induces antifungal immune defenses. *PLoS Pathog* **7**, e1002074 (2011).
48. Boraschi, D. & Tagliabue, A. The interleukin-1 receptor family. *Semin. Immunol.* **25**, 394–407 (2013).
49. Koga, M., Zwaal, R., Guan, K.-L., Avery, L. & Ohshima, Y. A *Caenorhabditis elegans* MAP kinase kinase, MEK-1, is involved in stress response. *EMBO J.* **19**, 5148–5156 (2000).
50. Kao, C.-Y. *et al.* Global functional analyses of cellular responses to pore-forming toxins. *PLoS Pathog.* **7**, e1001314 (2011).
51. Murphy, C. T. Insulin/insulin-like growth factor signaling in *C. elegans*. *WormBook* 1–43 (2013). doi:10.1895/wormbook.1.164.1
52. Kenyon, C., Chang, J., Gensch, E., Rudner, A. & Tabtlang, R. A *C. elegans* mutant that lives twice as long as wild type. *Nature* **366**, 461–464 (1993).
53. Barseyte, D., Lovejoy, D. A. & Lithgow, G. J. Longevity and heavy metal resistance in *daf-2* and *age-1* long-lived mutants of *Caenorhabditis elegans*. *FASEB J.* **15**, 627–634 (2001).
54. Lithgow, G. J., White, T. M., Melov, S. & Johnson, T. E. Thermotolerance and extended life-span conferred by single-gene mutations and induced by thermal stress. *Proc. Natl. Acad. Sci.* **92**, 7540–7544 (1995).
55. Lamitina, S. T. Transcriptional targets of DAF-16 insulin signaling pathway protect *C. elegans* from extreme hypertonic stress. *AJP Cell Physiol.* **288**, C467–C474 (2005).
56. Evans, E. A., Kawli, T. & Tan, M.-W. *Pseudomonas aeruginosa* suppresses host immunity by activating the DAF-2 insulin-like signaling pathway in *Caenorhabditis elegans*. *PLoS Pathog.* **4**, e1000175 (2008).
57. Murphy, C. T. *et al.* Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*. *Nature* **424**, 277–283 (2003).
58. Shapira, M. *et al.* A conserved role for a GATA transcription factor in regulating epithelial innate immune responses. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 14086–14091 (2006).
59. Troemel, E. R. *et al.* p38 MAPK regulates expression of immune response genes and contributes to longevity in *C. elegans*. *PLoS Genet.* **2**, e183 (2006).
60. Wang, J., Nakad, R. & Schulenburg, H. Activation of the *Caenorhabditis elegans* FOXO family transcription factor DAF-16 by pathogenic *Bacillus thuringiensis*. *Dev. Comp. Immunol.* **37**, 193–201 (2012).
61. Richardson, C. E., Kooistra, T. & Kim, D. H. An essential role for XBP-1 in host protection against immune activation in *C. elegans*. *Nature* **463**, 1092–1095 (2010).
62. Pellegrino, M. W. *et al.* Mitochondrial UPR-regulated innate immunity provides resistance to pathogen infection. *Nature* (2014). doi:10.1038/nature13818
63. Tarr, D. E. K. Nematode antimicrobial peptides. *ISJ-Invertebr. Surviv. J.* **9**, 122–133 (2012).
64. Bogaerts, A., Beets, I., Schoofs, L. & Verleyen, P. Antimicrobial peptides in *Caenorhabditis elegans*. *Invertebr. Surviv. J.* **7**, 45–52 (2010).
65. Kato, Y. & Komatsu, S. ASABF, a novel cysteine-rich antibacterial peptide isolated from the nematode *Ascaris suum* Purification, primary structure, and molecular cloning of cDNA. *J. Biol. Chem.* **271**, 30493–30498 (1996).
66. Kato, Y. *et al.* *abf-1* and *abf-2*, ASABF-type antimicrobial peptide genes in *Caenorhabditis elegans*. *Biochem J* **361**, 221–230 (2002).
67. Roeder, T. *et al.* Caenopores are antimicrobial peptides in the nematode *Caenorhabditis elegans* instrumental in nutrition and immunity. *Dev. Comp. Immunol.* **34**, 203–209 (2010).
68. Hoekendorf, A. & Leippe, M. SPP-3, a saposin-like protein of *Caenorhabditis elegans*, displays antimicrobial and pore-forming activity and is located in the intestine and in one head neuron. *Dev. Comp. Immunol.* **38**, 181–186 (2012).

69. Hoekendorf, A., Stanisak, M. & Leippe, M. The saposin-like protein SPP-12 is an antimicrobial polypeptide in pharyngeal neurons of *Caenorhabditis elegans* and participates in defence against a natural bacterial pathogen. *Biochem. J.* (2012). doi:10.1042/BJ20112102
70. Bányai, L. & Patthy, L. Amoebapore homologs of *Caenorhabditis elegans*. *Biochim. Biophys. Acta BBA-Protein Struct. Mol. Enzymol.* **1429**, 259–264 (1998).
71. Mallo, G. V. *et al.* Inducible antibacterial defense system in *C. elegans*. *Curr. Biol.* **12**, 1209–1214 (2002).
72. Couillault, C. *et al.* TLR-independent control of innate immunity in *Caenorhabditis elegans* by the TIR domain adaptor protein TIR-1, an ortholog of human SARM. *Nat. Immunol.* **5**, 488–494 (2004).
73. Darby, C. Interactions with microbial pathogens. *WormBook* (2005). doi:10.1895/wormbook.1.21.1
74. Schnepf, E. *et al.* *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol. Mol. Biol. Rev.* **62**, 775–806 (1998).
75. Romeis, J., Meissle, M. & Bigler, F. Transgenic crops expressing *Bacillus thuringiensis* toxins and biological control. *Nat. Biotechnol.* **24**, 63–71 (2006).
76. Aronson, A. I., Beckman, W. & Dunn, P. *Bacillus thuringiensis* and related insect pathogens. *Microbiol. Rev.* **50**, 1–24 (1986).
77. Whiteley, H. R. & Schnepf, H. E. The molecular biology of parasporal crystal body formation in *Bacillus thuringiensis*. *Annu. Rev. Microbiol.* **40**, 549–576 (1986).
78. Los, F. C. O., Randis, T. M., Aroian, R. V. & Ratner, A. J. Role of pore-forming toxins in bacterial infectious diseases. *Microbiol. Mol. Biol. Rev.* **77**, 173–207 (2013).
79. Parker, M. W. & Feil, S. C. Pore-forming protein toxins: from structure to function. *Prog. Biophys. Mol. Biol.* **88**, 91–142 (2005).
80. Ibrahim, M. A., Griko, N., Junker, M. & Bulla, L. A. *Bacillus thuringiensis*: A genomics and proteomics perspective. *Bioeng. Bugs* **1**, 31–50 (2010).
81. Griffiths, J. S. & Aroian, R. V. Many roads to resistance: how invertebrates adapt to Bt toxins. *BioEssays* **27**, 614–624 (2005).
82. Van Rie, J., Jansens, S., Höfte, H., Degheele, D. & Van Mellaert, H. Receptors on the brush border membrane of the insect midgut as determinants of the specificity of *Bacillus thuringiensis* delta-endotoxins. *Appl. Environ. Microbiol.* **56**, 1378–1385 (1990).
83. Pigott, C. R. & Ellar, D. J. Role of receptors in *Bacillus thuringiensis* crystal toxin activity. *Microbiol. Mol. Biol. Rev.* **71**, 255–281 (2007).
84. de Barros Moreira Beltrão, H. & Silva-Filha, M. H. N. L. Interaction of *Bacillus thuringiensis* svar. *israelensis* Cry toxins with binding sites from *Aedes aegypti* (Diptera: Culicidae) larvae midgut. *FEMS Microbiol. Lett.* **266**, 163–169 (2007).
85. Hua, G., Zhang, R., Abdullah, M. A. F. & Adang, M. J. *Anopheles gambiae* cadherin AgCad1 binds the Cry4Ba Toxin of *Bacillus thuringiensis israelensis* and a Fragment of AgCad1 synergizes toxicity. *Biochemistry (Mosc.)* **47**, 5101–5110 (2008).
86. Park, Y., Hua, G., Abdullah, M. A. F., Rahman, K. & Adang, M. J. Cadherin fragments from *Anopheles gambiae* synergize *Bacillus thuringiensis* Cry4Ba's toxicity against *Aedes aegypti* larvae. *Appl. Environ. Microbiol.* **75**, 7280–7282 (2009).
87. Fiuza, L., Nielsen-Leroux, C., Goze, E., Frutos, R. & Charles, J. Binding of *Bacillus thuringiensis* Cry1 toxins to the midgut brush border membrane vesicles of *Chilo suppressalis* (Lepidoptera: Pyralidae): Evidence of shared binding sites. *Appl. Environ. Microbiol.* **62**, 1544–1549 (1996).
88. Davolos, C. C. *et al.* Binding analysis of *Bacillus thuringiensis* Cry1 proteins in the sugarcane borer, *Diatraea saccharalis* (Lepidoptera: Crambidae). *J. Invertebr. Pathol.* **127**, 32–34 (2015).
89. Poncet, S., Delécluse, A., Klier, A. & Rapoport, G. Evaluation of synergistic interactions among the CryIVA, CryIVB, and CryIVD toxic components of *B. thuringiensis* subsp. *israelensis* crystals. *J. Invertebr. Pathol.* 131–135 (1995).

90. Monnerat, R. *et al.* Synergistic activity of *Bacillus thuringiensis* toxins against *Simulium* spp. larvae. *J. Invertebr. Pathol.* **121**, 70–73 (2014).
91. Masri, L. *et al.* Host–pathogen coevolution: The selective advantage of *Bacillus thuringiensis* virulence and its cry toxin genes. *PLOS Biol.* **13**, e1002169 (2015).
92. Marroquin, L. D., Elyassnia, D., Griffiths, J. S., Feitelson, J. S. & Aroian, R. V. *Bacillus thuringiensis* (Bt) toxin susceptibility and isolation of resistance mutants in the nematode *Caenorhabditis elegans*. *Genetics* **155**, 1693–1699 (2000).
93. Devidas, P. & Rehberger, L. A. The effects of exotoxin (thuringiensin) from *Bacillus thuringiensis* on *Meloidogyne incognita* and *Caenorhabditis elegans*. *Plant Soil* **145**, 115–120 (1992).
94. Wei, J.-Z. *et al.* *Bacillus thuringiensis* crystal proteins that target nematodes. *Proc. Natl. Acad. Sci.* **100**, 2760–2765 (2003).
95. Griffiths, J. S., Whitacre, J. L., Stevens, D. E. & Aroian, R. V. Bt toxin resistance from loss of a putative carbohydrate-modifying enzyme. *Science* **293**, 860–864 (2001).
96. Griffiths, J. S. *et al.* Glycolipids as receptors for *Bacillus thuringiensis* crystal toxin. *Science* **307**, 922–925 (2005).
97. Ideo, H. *et al.* A *Caenorhabditis elegans* glycolipid-binding galectin functions in host defense against bacterial infection. *J. Biol. Chem.* **284**, 26493–26501 (2009).
98. Huffman, D. L. *et al.* Mitogen-activated protein kinase pathways defend against bacterial pore-forming toxins. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 10995–11000 (2004).
99. Bellier, A., Chen, C.-S., Kao, C.-Y., Cinar, H. N. & Aroian, R. V. Hypoxia and the hypoxic response pathway protect against pore-forming toxins in *C. elegans*. *PLoS Pathog.* **5**, e1000689 (2009).
100. Hardalo, C. & Edberg, S. C. *Pseudomonas aeruginosa*: Assessment of risk from drinking water. *Crit. Rev. Microbiol.* **23**, 47–75 (1997).
101. Botzenhart, K. & Döring, G. in *Pseudomonas aeruginosa as an opportunistic pathogen* (eds. Campa, M., Bendinelli, M. & Friedman, H.) 1–18 (Springer US, 1993).
102. Palleroni, N. J. The *Pseudomonas* story: Editorial. *Environ. Microbiol.* **12**, 1377–1383 (2010).
103. Govan, J. R. & Deretic, V. Microbial pathogenesis in cystic fibrosis: Mucoicid *Pseudomonas aeruginosa* and *Burkholderia cepacia*. *Microbiol. Rev.* **60**, 539–574 (1996).
104. Kerr, K. G. & Snelling, A. M. *Pseudomonas aeruginosa*: A formidable and ever-present adversary. *J. Hosp. Infect.* **73**, 338–344 (2009).
105. Stover, C. K. *et al.* Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature* **406**, 959–964 (2000).
106. Lee, D. G. *et al.* Genomic analysis reveals that *Pseudomonas aeruginosa* virulence is combinatorial. *Genome Biol.* **7**, R90 (2006).
107. Jander, G., Rahme, L. G. & Ausubel, F. M. Positive correlation between virulence of *Pseudomonas aeruginosa* mutants in mice and insects. *J. Bacteriol.* **182**, 3843–3845 (2000).
108. D’Argenio, D. A., Gallagher, L. A., Berg, C. A. & Manoil, C. *Drosophila* as a model host for *Pseudomonas aeruginosa* infection. *J. Bacteriol.* **183**, 1466–1471 (2001).
109. Rahme, L. G. *et al.* Common virulence factors for bacterial pathogenicity in plants and animals. *Science* **268**, 1899–1902 (1995).
110. Mahajan-Miklos, S., Rahme, L. G. & Ausubel, F. M. Elucidating the molecular mechanisms of bacterial virulence using non-mammalian hosts. *Mol. Microbiol.* **37**, 981–988 (2000).
111. Clatworthy, A. E. *et al.* *Pseudomonas aeruginosa* infection of zebrafish involves both host and pathogen determinants. *Infect. Immun.* **77**, 1293–1303 (2009).
112. McCarthy, K. *Pseudomonas aeruginosa*: Evolution of antimicrobial resistance and implications for therapy. *Semin. Respir. Crit. Care Med.* **36**, 044–055 (2015).

113. Cao, H., Baldini, R. L. & Rahme, L. G. Common mechanisms for pathogens of plants and animals. *Annu. Rev. Phytopathol.* **39**, 259–284 (2001).
114. Cassat, J. E. & Skaar, E. P. Iron in infection and immunity. *Cell Host Microbe* **13**, 509–519 (2013).
115. Wick, M. J., Hamood, A. N. & Iglewski, B. H. Analysis of the structure-function relationship of *Pseudomonas aeruginosa* exotoxin A. *Mol. Microbiol.* **4**, 527–535 (1990).
116. Strateva, T. & Mitov, I. Contribution of an arsenal of virulence factors to pathogenesis of *Pseudomonas aeruginosa* infections. *Ann. Microbiol.* **61**, 717–732 (2011).
117. Sawa, T. The molecular mechanism of acute lung injury caused by *Pseudomonas aeruginosa*: From bacterial pathogenesis to host response. *J Intensive Care* **2**, (2014).
118. Kirienko, N. V., Cezairliyan, B. O., Ausubel, F. M. & Powell, J. R. in *Pseudomonas Methods and Protocols* (eds. Filloux, A. & Ramos, J.-L.) **1149**, 653–669 (Springer New York, 2014).
119. Tan, M.-W., Mahajan-Miklos, S. & Ausubel, F. M. Killing of *Caenorhabditis elegans* by *Pseudomonas aeruginosa* used to model mammalian bacterial pathogenesis. *Proc. Natl. Acad. Sci.* **96**, 715–720 (1999).
120. Mahajan-Miklos, S., Tan, M.-W., Rahme, L. G. & Ausubel, F. M. Molecular mechanisms of bacterial virulence elucidated using a *Pseudomonas aeruginosa*-*Caenorhabditis elegans* pathogenesis model. *Cell* **96**, 47–56 (1999).
121. Kirienko, N. V. *et al.* *Pseudomonas aeruginosa* disrupts *Caenorhabditis elegans* iron homeostasis, causing a hypoxic response and death. *Cell Host Microbe* **13**, 406–416 (2013).
122. McEwan, D. L., Kirienko, N. V. & Ausubel, F. M. Host translational inhibition by *Pseudomonas aeruginosa* exotoxin A triggers an immune response in *Caenorhabditis elegans*. *Cell Host Microbe* **11**, 364–374 (2012).
123. Dai, L.-L., Gao, J.-X., Zou, C.-G., Ma, Y.-C. & Zhang, K.-Q. *mir-233* modulates the unfolded protein response in *C. elegans* during *Pseudomonas aeruginosa* infection. *PLoS Pathog.* **11**, e1004606 (2015).
124. Drickamer, K. Two distinct classes of carbohydrate-recognition domains in animal lectins. *J. Biol. Chem.* **263**, 9557–9560 (1988).
125. Weis, W. I., Taylor, M. E. & Drickamer, K. The C-type lectin superfamily in the immune system. *Immunol. Rev.* **163**, 19–34 (1998).
126. Zelensky, A. N. & Gready, J. E. The C-type lectin-like domain superfamily. *FEBS J.* **272**, 6179–6217 (2005).
127. Pfaller, M. A. & Diekema, D. J. Epidemiology of invasive candidiasis: A persistent public health problem. *Clin. Microbiol. Rev.* **20**, 133–163 (2007).
128. Brown, G. D. & Gordon, S. A new receptor for β -glucans. *Nature* **413**, 36–37 (2001).
129. Saijo, S. *et al.* Dectin-2 recognition of α -mannans and induction of Th17 cell differentiation is essential for host defense against *Candida albicans*. *Immunity* **32**, 681–691 (2010).
130. Gazi, U. & Martinez-Pomares, L. Influence of the mannose receptor in host immune responses. *Immunobiology* **214**, 554–561 (2009).
131. Drummond, R. A., Saijo, S., Iwakura, Y. & Brown, G. D. The role of Syk/CARD9 coupled C-type lectins in antifungal immunity. *Eur. J. Immunol.* **41**, 276–281 (2011).
132. Chen, S.-T. *et al.* CLEC5A is critical for dengue-virus-induced lethal disease. *Nature* **453**, 672–676 (2008).
133. Iborra, S. *et al.* The DC receptor DNGR-1 mediates cross-priming of CTLs during vaccinia virus infection in mice. *J. Clin. Invest.* **122**, 1628–1643 (2012).
134. Zhang, J.-G. *et al.* The dendritic cell receptor Clec9A binds damaged cells via exposed actin filaments. *Immunity* **36**, 646–657 (2012).
135. Ahrens, S. *et al.* F-actin is an evolutionarily conserved damage-associated molecular pattern recognized by DNGR-1, a receptor for dead cells. *Immunity* **36**, 635–645 (2012).
136. Neumann, K. *et al.* Clec12a is an inhibitory receptor for uric acid crystals that regulates inflammation in response to cell death. *Immunity* **40**, 389–399 (2014).

137. Sancho, D. & Reis e Sousa, C. Signaling by myeloid C-type lectin receptors in immunity and homeostasis. *Annu. Rev. Immunol.* **30**, 491–529 (2012).
138. Hoving, J. C., Wilson, G. J. & Brown, G. D. Signalling C-Type lectin receptors, microbial recognition and immunity: C-type lectins in immunity. *Cell. Microbiol.* **16**, 185–194 (2014).
139. Dambuza, I. M. & Brown, G. D. C-type lectins in immunity: recent developments. *Curr. Opin. Immunol.* **32**, 21–27 (2015).
140. Takahashi, K., Ip, W. E., Michelow, I. C. & Ezekowitz, R. A. B. The mannose-binding lectin: A prototypic pattern recognition molecule. *Curr. Opin. Immunol.* **18**, 16–23 (2006).
141. Fujita, T. Evolution of the lectin-complement pathway and its role in innate immunity. *Nat. Rev. Immunol.* **2**, 346–353 (2002).
142. Mukherjee, S. *et al.* Antibacterial membrane attack by a pore-forming intestinal C-type lectin. *Nature* **505**, 103–107 (2013).
143. Holmskov, U., Thiel, S. & Jensenius, J. C. Collectins and ficolins: Humoral lectins of innate immune defense. *Annu. Rev. Immunol.* **21**, 547–578 (2003).
144. Sekine, H. *et al.* An ancient lectin-dependent complement system in an ascidian: novel lectin isolated from the plasma of the solitary ascidian, *Halocynthia roretzi*. *J. Immunol.* **167**, 4504–4510 (2001).
145. Fraser, I. P., Koziel, H. & Ezekowitz, R. A. B. The serum mannose-binding protein and the macrophage mannose receptor are pattern recognition molecules that link innate and adaptive immunity. in *Seminars in Immunology* **10**, 363–372 (Elsevier, 1998).
146. Weis, W. I., Drickamer, K. & Hendrickson, W. A. Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature* **360**, 127–134 (1992).
147. Nauta, A. J. *et al.* Mannose-binding lectin engagement with late apoptotic and necrotic cells. *Eur. J. Immunol.* **33**, 2853–2863 (2003).
148. Ma, Y. *et al.* Antitumor activity of mannan-binding protein in vivo as revealed by a virus expression system: mannan-binding protein-independent cell-mediated cytotoxicity. *Proc. Natl. Acad. Sci.* **96**, 371–375 (1999).
149. Dunkelberger, J. R. & Song, W.-C. Complement and its role in innate and adaptive immune responses. *Cell Res.* **20**, 34–50 (2010).
150. Keizer, M. P., Wouters, D., Schlapbach, L. J. & Kuijpers, T. W. Restoration of MBL-deficiency: Redefining the safety, efficacy and viability of MBL-substitution therapy. *Mol. Immunol.* **61**, 174–184 (2014).
151. Vaishnava, S. *et al.* The antibacterial lectin RegIII γ promotes the spatial segregation of microbiota and host in the intestine. *Science* **334**, 255–258 (2011).
152. Cash, H. L., Whitham, C. V., Behrendt, C. L. & Hooper, L. V. Symbiotic bacteria direct expression of an intestinal bactericidal lectin. *Science* **313**, 1126–1130 (2006).
153. Miki, T., Holst, O. & Hardt, W.-D. The bactericidal activity of the C-type lectin RegIII β against Gram-negative bacteria involves binding to lipid A. *J. Biol. Chem.* **287**, 34844–34855 (2012).
154. Miki, T. & Hardt, W.-D. Outer membrane permeabilization is an essential step in the killing of Gram-negative bacteria by the lectin RegIII β . *PLoS ONE* **8**, e69901 (2013).
155. Dheilly, N. M. *et al.* A family of variable immunoglobulin and lectin domain containing molecules in the snail *Biomphalaria glabrata*. *Dev. Comp. Immunol.* **48**, 234–243 (2015).
156. Li, R. *et al.* Jellyfish venomomics and venom gland transcriptomics analysis of *Stomolophus meleagris* to reveal the toxins associated with sting. *J. Proteomics* **106**, 17–29 (2014).
157. Wang, X.-W. & Wang, J.-X. Diversity and multiple functions of lectins in shrimp immunity. *Dev. Comp. Immunol.* **39**, 27–38 (2013).
158. Yang, J. *et al.* CfLec-3 from scallop: an entrance to non-self recognition mechanism of invertebrate C-type lectin. *Sci. Rep.* **5**, 10068 (2015).

159. Wang, J.-L. *et al.* Involvement of a pattern recognition receptor C-type lectin 7 in enhancing cellular encapsulation and melanization due to its carboxyl-terminal CRD domain in the cotton bollworm, *Helicoverpa armigera*. *Dev. Comp. Immunol.* **44**, 21–29 (2014).
160. Zhang, X.-W. *et al.* Overexpression of a C-type lectin enhances bacterial resistance in red swamp crayfish *Procambarus clarkii*. *Fish Shellfish Immunol.* **34**, 1112–1118 (2013).
161. Wang, X.-W., Zhang, X.-W., Xu, W.-T., Zhao, X.-F. & Wang, J.-X. A novel C-type lectin (FcLec4) facilitates the clearance of *Vibrio anguillarum* *in vivo* in Chinese white shrimp. *Dev. Comp. Immunol.* **33**, 1039–1047 (2009).
162. Yu, X.-Q. & Kanost, M. R. *Manduca sexta* lipopolysaccharide-specific immulectin-2 protects larvae from bacterial infection. *Dev. Comp. Immunol.* **27**, 189–196 (2003).
163. Loker, E. S., Adema, C. M., Zhang, S.-M. & Kepler, T. B. Invertebrate immune systems – not homogeneous, not simple, not well understood. *Immunol. Rev.* **198**, 10–24 (2004).
164. Schulenburg, H., Hoepfner, M. P., Weiner, J. & Bornberg-Bauer, E. Specificity of the innate immune system and diversity of C-type lectin domain (CTLD) proteins in the nematode *Caenorhabditis elegans*. *Immunobiology* **213**, 237–250 (2008).
165. Thomas, J. H. Analysis of homologous gene clusters in *Caenorhabditis elegans* reveals striking regional cluster domains. *Genetics* **172**, 127–143 (2006).

Chapter I

High Innate Immune Specificity through Diversified C-type Lectin-like Domain Proteins in Invertebrates

Barbara Pees¹, Wentao Yang¹, Alejandra Zárata-Potes¹, Hinrich Schulenburg¹, Katja Dierking^{1†}

¹ Department of Evolutionary Ecology and Genetics, Zoological Institute, Christian Albrechts University of Kiel, Kiel, Germany

† Corresponding author: kdierking@zoologie.uni-kiel.de

High Innate Immune Specificity through Diversified C-Type Lectin-Like Domain Proteins in Invertebrates

Barbara Pees Wentao Yang Alejandra Zárata-Potes Hinrich Schulenburg
Katja Dierking

Department of Evolutionary Ecology and Genetics, Zoological Institute, Christian Albrechts University of Kiel, Kiel, Germany

Key Words

C-type lectin-like domain proteins · Immune specificity · Innate immunity · Insects · Crustaceans · *Caenorhabditis elegans*

Abstract

A key question in current immunity research is how the innate immune system can generate high levels of specificity. Evidence is accumulating that invertebrates, which exclusively rely on innate defense mechanisms, can differentiate between pathogens on the species and even strain level. In this review, we identify and discuss the particular potential of C-type lectin-like domain (CTLD) proteins to generate high immune specificity. Whilst several CTLD proteins are known to act as pattern recognition receptors in the vertebrate innate immune system, the exact role of CTLD proteins in invertebrate immunity is much less understood. We show that CTLD genes are highly abundant in most metazoan genomes and summarize the current state of knowledge on CTLD protein function in insect, crustacean and nematode immune systems. We then demonstrate extreme CTLD gene diversification in the genomes of *Caenorhabditis* nematodes and provide an update of data from CTLD gene function studies in *C. elegans*, which indicate that the diversity of

CTLD genes could contribute to immune specificity. In spite of recent achievements, the exact functions of the diversified invertebrate CTLD genes are still largely unknown. Our review therefore specifically discusses promising research approaches to rectify this knowledge gap.

© 2015 S. Karger AG, Basel

Introduction

Animals depend on the specificity of their immune reaction to efficiently and economically protect themselves against pathogens. For a long time, the generation of a highly specific immune response was considered a hallmark of adaptive immunity, based on its immense diversity of T and B cell receptors generated by somatic recombination and hypermutation. There is, however, accumulating evidence that invertebrates, which solely rely on innate defense mechanisms and lack T and B cells, can differentiate between pathogens on the species and even strain level [1–3]. Whether and how such high levels of specificity can be generated by the innate immune system is one of the main questions in current immunity research.

A prerequisite for the generation of a pathogen-specific immune reaction is a repertoire of diversified receptors, which enables the organism to specifically recognize and respond to different immune challenges. In vertebrates the well-characterized pattern recognition receptors (PRRs) of the innate system, such as the Toll-like receptors (TLRs), retinoic acid-inducible gene-I-like receptors (RLRs), and NOD-like receptors (NLRs), recognize structures that are conserved among microbial species and mediate a broader pattern of specificity (i.e. on the level of a particular pathogen group, such as Gram-negative or Gram-positive bacteria). The diversity of ligand binding can be expanded by the formation of receptor homo- or heterodimers and/or the interaction with other PRR coreceptors, as is well described for mammalian TLRs. However, in invertebrates, high levels of specificity (on the species and strain level) are unlikely mediated by TLRs, NLRs or RLRs, as these are too few in number in most, although not all (see below), animal species.

There are few examples of the diversification of alternative putative immune receptors in invertebrates. The hypervariable protein Down syndrome cell adhesion molecule (Dscam), which is generated by alternative somatic splicing, was shown to be implicated in the immune response of insects and crustaceans [4, 5], and suggested to act as a pathogen-specific recognition molecule. Further experimental evidence to support this hypothesis is, however, still scarce [6]. A high level of genetic polymorphism was observed in the immunoglobulin (Ig)-type variable (V) region-containing chitin-binding proteins (VCBPs) of the cephalochordate *Branchiostoma floridae* [7], and in the Ig domain of fibrinogen-related proteins (FREPs) of the snail *Biomphalaria glabrata* [8]. However, the molecular mechanisms underlying the diversification of VCBPs and FREPs and their functional significance for immune specificity are not well understood. Yet another example of diversification of putative innate immune system proteins is the Sp185/333 system in echinoderms, such as the purple sea urchin *Strongylocentrotus purpuratus*. The members of this highly diverse, novel gene family are not homologous to any other known genes, they are highly upregulated after immune challenge and expressed in the phagocyte class of sea urchin coelomocytes. The function of Sp185/333 proteins is unknown. The high diversification of Sp185/333 genes and proteins is achieved through different mechanisms, including maintenance of high polymorphism, gene recombination and modifications at the RNA and protein level [9]. Similar multigene expansions, usually based on repeated gene duplications that potentially produce high-sequence di-

versification, were observed for TLR and NLR gene families in the genomes of the sea urchin *S. purpuratus* [10] and the lancelet *B. floridae* [11], but the functional consequence of this diversified repertoire of potential receptors remains elusive. Here, we argue that another group of known, yet neglected, PRRs – the C-type lectins – show surprisingly high diversity in most metazoan genomes and have the potential to generate immune specificity.

C-type lectins are a family of both soluble and membrane-bound proteins with a characteristic carbohydrate-recognition domain (CRD) that mediates ligand binding. The CRD has a distinct protein fold determined by highly conserved residues that form a characteristic sequence motif [12]. It was first described in Ca²⁺-dependent ('C-type') carbohydrate binding animal proteins ('lectins'), but has subsequently been found in many other proteins that do not bind sugar and/or calcium. Thus, we here use the term C-type lectin-like domain (CTLD) in its broadest sense, as proposed previously [13], and use sequence similarity to identify genes containing CTLDs in different metazoan genomes. CTLD proteins have a wide range of functions and bind a wide variety of ligands, such as sugars, proteins, lipids and inorganic compounds [13].

In vertebrates, CTLD proteins are known to fulfill important tasks in the immune system by acting as PRRs [for a review, see 14], as well as effector proteins with bactericidal activity [15]. As we will demonstrate further below, CTLD genes are highly abundant in many metazoan genomes. Yet, our knowledge on CTLD protein function in immunity mainly derives from the few well-studied vertebrate CTLD proteins, while the function of the majority of CTLD proteins is unknown. In invertebrates, it has repeatedly been suggested that the diversified CTLD proteins contribute to immunity, yet exact functional genetic information is largely absent and the few exceptions mainly focus on the assessment of single CTLD proteins. Moreover, the particular role of CTLD proteins in generating immune specificity is completely unexplored. Therefore, we require a systematic research program that aims at elucidating the role of diversified CTLD proteins in immune specificity.

Here, we highlight that CTLD genes constitute expanded gene families in many metazoan taxa and summarize the available evidence for a role of CTLD proteins in invertebrate immunity, mainly inferred from studies on crustaceans and insects. We then turn to the nematode genus *Caenorhabditis*, in which several complete genomes are available, to demonstrate extreme CTLD gene diversification, followed by a summary of data from *C. elegans* that strongly indicate a role of the diversified CTLD proteins in immune specificity. Finally, we empha-

Table 1. Number of CTLD genes (proteins) in different metazoan genomes

Species	Taxon	Genome size, Mb	Total coding genes	Total InterPro ¹ genes (proteins), n	Total Pfam ¹ genes (proteins), n	Immune function ²	Ref.	Genome assembly
<i>H. sapiens</i>	Mammalia	3,209.29	20,805	100 (277)	94 (268)	e.g. DC-SIGN, DECTIN-1, DECTIN-2, MINCLE, Langerin (CD207), BDCA-2, MBL2, COLEC11, MRC-1, MRC-2, SELP, SELE, SELL, CLEC5A, CLECSF8, DCIR, CLEC9A, DEC-205, and others	reviewed in 14, 61, 62, 73	GRCh37
<i>M. musculus</i>	Mammalia	2,798.79	23,148	132 (355)	129 (339)	e.g. SIGNR1, SIGNR3, SIGNR5, Dectin-1, Dectin-2, Mincle, Langerin (Cd207), MBL-A, MBL-C, Colec11, Mrc-1, Mrc-2, Selp, Sele, Sell, Clec5a, Clecsf8, Dcir, Dcir2, DNNGR-1 (Clec9a), DEC-205, and others	reviewed in 14, 61, 62, 73	GRCm38
<i>D. rerio</i>	Teleostei	1,412.46	26,459	176 (381)	169 (360)			Zv9
<i>B. floridae</i>	Cephalochordata	521.90	30,601 ³	692 (694)	646 (647)	AmphiCTL1	74	v1.0
<i>S. purpuratus</i>	Echinodermata	936.58	28,987	293 (303)	241 (244)	SpEchinoidin	75, 76	Spur_3.1
<i>A. gambiae</i>	Insecta	265.03	12,810	37 (99)	33 (93)	CTL4, CTLMA2	34	AgamP3
<i>B. mori</i>	Insecta	481.82	14,623	21 (32)	17 (25)	BmLBP, BmMBP	21, 26, 77	ASM15162v1
<i>D. melanogaster</i>	Insecta	139.49	13,937	56 (96)	48 (84)	DL1, DL2, DL3	18–20	BDGP 5
<i>C. brenneri</i>	Nematoda	190.37	30,667	319 (320)	159 (159)			WS227
<i>C. briggsae</i>	Nematoda	108.48	21,936	183 (183)	111 (111)			WS224
<i>C. elegans</i>	Nematoda	100.29	20,541	283 (288)	166 (185)	CLEC-39, CLEC-49, CLEC-79 + 9 <i>clec</i> genes (see table 2)	see table 2	WS235
<i>C. japonica</i>	Nematoda	166.26	29,964	111 (139)	58 (72)			WS227
<i>C. remanei</i>	Nematoda	145.44	31,444	292 (292)	158 (158)			WS185
<i>N. vectensis</i>	Cnidaria	356.61	24,773	67 (67)	62 (62)			ASM20922v1
<i>A. queenslandica</i>	Porifera	166.70	30,060	2 (2)	1 (1)			Aqu1
<i>T. adhaerens</i>	Placozoa	105.63	11,520	45 (45)	40 (40)			ASM15027v1

The total number of CTLD genes are listed for each species.

¹ The number of genes (proteins) with at least one CTLD were obtained by automatic annotation from the InterPro (v53.0) [16] and Pfam (v28.0) databases [17]. The number of CTLD proteins is higher than the CTLD gene number as it includes different isoforms encoded by the same gene. Proteins with multiple CTLDs are counted only once.

² CTLD proteins for which a function in immune defense is supported by experimental evidence. The list is nonexhaustive for human and mouse CTLD genes, for which gene name synonyms are shown that are commonly used in the literature.

³ Noncurated annotation.

size the current lack of exact understanding of the immune functions of the highly diverse CTLD proteins across metazoan taxa and define promising avenues for future research that integrate functional genetic with protein analysis approaches to rectify this knowledge gap.

CTLD Genes Are Abundant in Metazoan Genomes and Mediate Immune Responses in a Variety of Invertebrate Taxa

While CTLD proteins have been shown to fulfill diverse and fundamental tasks in the vertebrate immune system, their function in invertebrate immunity is much

less understood. This lack of data on CTLD function in invertebrates is surprising given the fact that genes encoding CTLD proteins are present in all invertebrate genomes sequenced to date and in some cases in very large numbers. In table 1 we list the available genome sequence data for some selected species as representatives of the major invertebrate groups and show the number of CTLD genes and proteins as provided by the InterPro v53.0 [16] and Pfam v28.0 databases [17]. CTLD genes are highly abundant in many metazoan genomes, including, for example, 132 members in *Mus musculus* (the 30th most abundant gene family in *M. musculus*) or 283 gene family members in the nematode *C. elegans* (the 7th most numerous gene family in *C. elegans*). The CTLD encod-

ing gene repertoire in the analyzed invertebrate species varies from two in the sponge *Amphimedon queenslandica* to more than 600 in the cephalochordate *B. floridae* [11]. We also looked for the total number of CTLD genes in the genomes of nine plant species (*Arabidopsis thaliana*, *Glycine max*, *Zea mays*, *Oryza sativa japonica* subsp., *Populus trichocarpa*, *Medicago truncatula*, *Triticum aestivum*, *Hordeum vulgare* and *Sorghum bicolor*). Interestingly, the plant genomes encode only very few (1–4) CTLD genes, which further emphasizes that the expansion of CTLD genes in metazoans is remarkable. Experimental evidence of an immune function of most invertebrate CTLD-encoding genes is missing, although their role in immunity is usually assumed. Within the existing information on invertebrate CTLD proteins, the functional data is biased towards certain insect, crustacean and nematode taxa. In the following paragraphs, we will summarize the evidence of CTLD proteins functioning in the immune defenses of insects and crustaceans, before focusing on the nematode genus *Caenorhabditis*.

Insects

Although genes encoding CTLD proteins are numerous in many insect genomes (table 1), research on insect CTLD proteins has as yet been focused on assessment of single CTLD proteins from different species. For example, only 3 of the more than 50 CTLD-encoding genes in *Drosophila melanogaster* were characterized in more detail [18–20]. Several insect CTLD proteins were suggested to function as PRRs, including those found to mediate important insect immune defense mechanisms, such as hemocyte nodule formation, encapsulation, melanization and activation of phagocytosis [18, 21, 22]. Most insect CTLD proteins were first isolated from hemolymph either due to their ability to agglutinate red blood cells [19, 22], to bind to bacterial cells [21], or were selectively purified from larval plasma by carbohydrate affinity chromatography after bacterial challenge [23]. The purified CTLD proteins were subsequently characterized biochemically in vitro, for example by testing their binding properties and hemagglutination activities. These protein-level studies demonstrated CTLD proteins to possess both broad and more specific binding abilities.

While most studied insect proteins only have one CTLD, some characterized proteins from lepidopteran species, such as *Bombyx mori* [24], *Manduca sexta* [23], and the cotton bollworm *Helicoverpa armigera* [25], have two CTLDs in tandem. The *B. mori* multibinding protein (BmMBP) is an example of an insect CTLD protein with a broad binding capacity. Each of its two divergent tan-

dem CTLDs has distinct binding specificities, enabling the molecule to bind to a large variety of microorganisms ranging from Gram-positive to Gram-negative bacteria, as well as yeasts. After binding to its target, BmMBP is able to trigger cellular immune responses, such as aggregation of hemocytes and nodule formation [26]. The expression of the four characterized *M. sexta* CTLD proteins immuectin-1 to -4 is induced upon bacterial exposure in the fat body. Immuectin-1 and -2 function in activation of the prophenoloxidase [27] and immuectin-3 and -4 act in enhancing encapsulation and melanization [28, 29]. They bind to microbe-associated molecular patterns (MAMPs) [30] with different affinities and specificities in vitro. While immuectin-1, -3 and -4 bind to a broad range of MAMPs, immuectin-2 specifically binds to lipopolysaccharide [28, 29, 31].

These and other studies provide important information on the potential immune functions of insect CTLD proteins. However, there is hardly any experimental evidence from functional genetic studies on the role of CTLD proteins in insect immune defenses in vivo. The three exceptions tested the effect of CTLD gene expression knockdown or knockout in a lepidopteran, the fruit fly and a mosquito. RNAi knockdown of the *M. sexta* CTLD gene immuectin-2 renders animals more susceptible to infection with the insect pathogen *Photorhabdus luminescens* [32]. The galactose-specific *D. melanogaster* CTLD protein DL1, whose expression is induced by injury [19] and which specifically binds to *Escherichia coli* and *Erwinia chrysanthemi*, but not to the Gram-positive *Staphylococcus aureus* or the yeast *Saccharomyces cerevisiae*, agglutinates *E. coli* and promotes the association of a *Drosophila* hemocytes-derived cell line with *E. coli* [20]. A DL1 deletion mutant did not show any defects in antimicrobial peptide expression after pricking or infection with *E. coli*, indicating that DL1 does not participate in the humoral immune response in vivo, but might still be involved in cellular immune responses such as phagocytosis [20]. In a functional genetic screen of about 100 immune candidate genes using RNAi in the mosquito *Anopheles gambiae* (Diptera) two CTLD proteins, CTL4 and CTLMA2, were discovered to affect *Plasmodium berghei* development by acting as negative regulators of the vector's melanization response to *P. berghei* ookinetes [33]. In a subsequent genetic analysis [34], CTL4 or CTLMA2 silencing by RNAi significantly reduced the survival of adult female mosquitoes infected with *E. coli*, *Enterobacter cloacae* and Pseudomonadaceae H2.26, but not with *S. aureus*, *Micrococcus luteus* and *Enterococcus faecalis*, suggesting a specific role in defense against Gram-negative

bacteria. CTL4 and CTLMA2 expression is induced by *E. coli* and *S. aureus* infection and they both produce a heterodimer that is secreted into the mosquito hemolymph [34].

In summary, while some of the characterized insect CTLD proteins have the ability to bind to a broad range of microorganisms or MAMPs, others exhibit more specific binding properties. Insect CTLD proteins are involved in the activation of cellular immune responses rather than the humoral immune response, but also act as opsonization molecules. A direct effector function such as antimicrobial activity has not been described and their target signaling pathways are unknown. Overall, the potential immune functions of the high variety of insect CTLD proteins (table 1) are far from being understood.

Crustaceans

The second most comprehensive data set on invertebrate CTLD protein functions comes from research on crustacean species that are of economic importance, such as the Chinese white shrimp *Fenneropenaeus chinensis* or the Chinese mitten crab *Eriocheir sinensis*. CTLD protein research in these species mainly aims at the development of antiparasitic drugs, as parasitic threats are a particular concern in crustacean aquaculture. For example, the white spot syndrome virus infects most of the commercially important shrimp species and can cause major production losses [35]. Since CTLD proteins are known to participate in the immune response of other invertebrates and since many expressed sequence tags encoding lectins were identified in cDNA libraries from both healthy and immune-challenged crustaceans, characterization of immune-relevant CTLD proteins in these economically important species has been a particular focus of the current research activities. Comprehensive data sets are thus available on the molecular cloning, isolation and in vitro biochemical characterization of purified crustacean CTLD proteins (either the recombinant or the native protein), especially from the economically important shrimp species, as reviewed in depth previously [36]. Like insect CTLD proteins, crustacean CTLD proteins have repeatedly been suggested to act as PRRs in various species. They can enhance encapsulation, promote phagocytosis and induce the production of reactive oxygen intermediates in hemocytes in vitro [36]. In addition, several crustacean CTLD proteins contribute to pathogen clearance in vivo [37–39].

Interestingly, some crustacean CTLD proteins exhibit in vitro antibacterial or antiviral activity and are thus assumed to act as antimicrobials. For example, Fc-hsL from

F. chinensis inhibits the growth of Gram-positive as well as Gram-negative bacteria and fungi, with the highest activity against Gram-positive bacteria [40]. Similarly, the two *E. sinensis* CTLD proteins EsLecA and EsLecG produce antimicrobial activity against both Gram-positive and Gram-negative bacteria in in vitro growth-inhibitory activity assays [41]. Most characterized crustacean CTLD proteins are expressed in the gills, intestine, hemocytes and hepatopancreas (the counterpart of the fat body in insects), and are likely to be secreted into the hemolymph upon immune challenge. Indeed, almost all crustacean CTLD proteins that have been studied so far show an inducible expression by either bacterial and/or viral pathogens. The majority of characterized crustacean CTLD proteins have one CTLD, the remaining have two [36]. Surprisingly, all studied crustacean CTLD proteins possess broad binding properties active towards all the bacterial strains tested. One exception might be the *Penaeus monodon* PmLec, which binds lipopolysaccharide and mainly agglutinates Gram-negative bacteria in vitro [42].

Since the genomes of these economically important crustacean species have not been fully sequenced yet, and shrimps and crabs have (as yet) not been amenable to genetic manipulation, genomic analyses and functional genetic studies on the crustacean CTLD genes are almost completely missing. The exceptions are three recent studies in which expression of crustacean CTLD proteins was genetically manipulated in vivo. Overexpression of the CTLD protein PcLec4 from the red swamp crayfish *Procambarus clarkii*, which is induced by infection with *Vibrio anguillarum*, facilitated bacterial clearance and increased survival of crayfish infected with *V. anguillarum* [39]. Knockdown of the expression of the CTLD protein MjHeCL from the kuruma shrimp *Marsupenaeus japonicus* by RNAi caused increased bacterial proliferation in the hemolymph and led to a change in expression of several antimicrobial peptides in hemocytes both in vivo and in vitro. It was thus suggested that MjHeCL plays a role in restricting the growth of the hemolymph microbiota by regulating antimicrobial peptide expression [43]. Similarly, the role of another *M. japonicus* CTLD protein, hFcLec4, and of FcLec4 from *F. chinensis* in promoting phagocytosis by acting as opsonins in hemocytes was supported by the finding that clearance of the pathogenic *V. anguillarum* from the hemolymph was delayed by RNAi knockdown of hFcLec4 and FcLec4 [44].

Interestingly, FcLec4 and hFcLec4 interact with the transmembrane protein β -integrin via their N-terminal domains and this interaction is required for their opsonic activity [44]. β -Integrins are known to participate in

phagocytosis in other invertebrates, such as flies [45]. Shrimp FcLec4 and hFcLec4 may thus act as PRRs recognizing bacteria via their CTLD and their binding to β -integrin in the membrane of hemocytes leads to cytoskeletal reorganization, which induces phagosome formation to ingest the bacteria [44]. This study is one of only two studies providing the first insight into interacting and signaling partners of invertebrate CTLD proteins. In the other study, a conserved binding motif for the transcription factor NF- κ B was identified in the flanking promoter sequence of the CTLD gene LvCTL3 from *Litopenaeus vannamei*. LvCTL3 contributed to opsonization and thus marking bacteria and cells for phagocytosis. Deletion of the NF- κ B binding site led to the abolishment of LvCTL3 expression, suggesting its regulation by NF- κ B. This finding represents a first step towards the incorporation of CTLD proteins into the poorly understood immune signaling network in crustaceans [46].

Taken together, the functional role of several crustacean CTLD proteins was inferred at the protein level by isolating the respective proteins and biochemically analyzing their physical interaction with microorganisms or MAMPs in vitro. Almost all crustacean CTLD proteins are able to bind to a broad range of microorganism and/or MAMPs. They seem to be involved in the activation of cellular immune responses, such as encapsulation and phagocytosis, as well as in the direct elimination of pathogens by exhibiting bactericidal activity. However, the actual number of genes encoding CTLD proteins, their genomic and structural organization, and the underlying impact on the specificity of the interactions with microorganisms remains unknown due to the lack of genome sequences and further experimental analysis. Particularly, studies of the protein function in the context of the whole organism inferred from altered gene products, such as mutant or overexpression phenotypes, are rare.

CTLD Genes Are Extremely Diversified in *Caenorhabditis* Genomes

The nematode genus *Caenorhabditis* provides an impressive example of extreme CTLD gene diversification. This was previously noted for the model nematode *C. elegans* [47], which possesses 283 CTLD genes (*clec* genes) in its genome (online suppl. table S1; for all online suppl. material, see www.karger.com/doi/10.1159/000441475; note that this is an updated number and thus different from the previous report due to further improvements of

the whole genome sequence). Of the 283 *clec* genes, 81% are predicted to have a signal peptide [48] and are thus likely secreted proteins, highlighting a particular role of the CTLD proteins in the pseudocoel, in the intestinal lumen and/or as components of the secreted cuticle. In contrast, 21% of the 283 *clec* genes possess membrane-spanning properties, suggesting that these function as components of cellular membranes. *clec* gene evolution is likely highly dynamic and subject to repeated duplication events, as indicated by their genomic localization in clusters (125 out of 283, equivalent to 44%; online suppl. table S1) and the high phylogenetic similarity of the genes within a cluster [47]. It is possible that the repeated duplications are favored by pathogen-mediated selection if *clec* gene duplicates allow the worm to recognize and/or eliminate a larger variety of pathogen strains [47].

In recent years, complete genome sequences became available for additional *Caenorhabditis* species, allowing cross-species comparison of *clec* gene diversification. Next to *C. elegans*, four additional species have comparatively well-annotated genomes, namely *C. briggsae*, *C. remanei*, *C. brenneri* and *C. japonica*. These five *Caenorhabditis* genomes contain between 111 (*C. japonica*) and 319 (*C. brenneri*) *clec* genes (online suppl. table S2). Overall, the largest group of *Caenorhabditis clec* genes are those with a single CTLD, making up half of the CTLD genes in three of the five analyzed *Caenorhabditis* species and approximately one third in the remaining two species. With one exception (*C. japonica*), at least one third of the *Caenorhabditis clec* genes also have additional domains. These include the complement C1r/C1s Uegf Bmp1 (CUB) domain, the conserved cysteine and tryptophan residues (CW, also called PAN-3) domain, or the von Willebrand factor type A (VWA) domain. The joint presence of one or multiple CTLDs with other domains in one protein could indicate possible functions in immune signaling. Moreover, several *clec* genes (27% of all *clec* genes for *C. remanei*, 24% for *C. elegans*, 23% for *C. briggsae*, 19% for *C. brenneri* and 10% for *C. japonica*) encode proteins with more than one CTLD within the same protein. Such internal domain duplications might increase protein versatility; BmMBP from *B. mori* is one example for a tandem repetition of CTLDs that leads to extended binding abilities [26].

The majority of the *Caenorhabditis* CTLD proteins are predicted to be secreted. Besides having a putative cell-autonomous recognition function, these secreted CTLD proteins might act as highly specific immune effectors eliminating pathogens in analogy to the vertebrate lectins that are part of the complement system [49].

In summary, *clec* genes show a remarkable diversity in *Caenorhabditis* species, encoding secreted and membrane-bound proteins with one or multiple CTLDs, as well as multidomain proteins. The binding capabilities of CTLD proteins are potentially enlarged by repetitions of CTLDs within one protein. An additional functional diversification is conceivable at the protein level through dimerization or oligomerization of the CTLD proteins, as previously reported for CTLD proteins from other organisms such as CTLD protein homodimers, heterodimers and oligomers in humans and snakes [13], or the heterodimer CTL4-CTLMA2 in *A. gambiae* [34]. CTLD protein-protein interactions might have synergistic effects on CTLD binding specificity (i.e. two or multiple CTLD proteins are able to generate a higher degree of specificity if they interact with each other and with other proteins than if they act alone) and can potentially further increase the diversity available at the genetic level. Such synergistic interactions between CTLD proteins are, however, currently unexplored for the *Caenorhabditis* taxon and for invertebrates in general. Taken together, the *Caenorhabditis* CTLD proteins are particularly diverse and may thus be able to mediate immune specificity at both the level of pathogen recognition and elimination.

CTLD Proteins in *C. elegans* Show Potential for Generating High Immune Specificity

The most comprehensive data on *clec* gene function is available from the model nematode *C. elegans*. Some of the data strongly suggests that *clec* gene diversity could indeed contribute to immune specificity.

C. elegans clec Genes Are Differentially Regulated following Pathogen Infection

A large number of previous transcriptome studies in *C. elegans* demonstrate that putative immune recognition and effector gene families, such as the *clec*, lysozyme (*lys* and *ilys*) and caenopore (spp.) families, show pathogen-dependent activation upon infection [47]. Because of their extreme diversification, the *clec* genes have the highest potential to generate immune specificity. Of all *C. elegans clec* genes, 84% (237 out of 283) are differentially expressed following an infection with pathogens; 104 *clec* genes are only upregulated, 103 *clec* genes are both, up- and downregulated (online suppl. table S1), and 30 *clec* genes are only downregulated upon pathogen exposure. Importantly, some *C. elegans clec* genes are only differentially expressed in response to a specific pathogen, sug-

gesting their highly specific regulation. In contrast, other *clec* genes are differentially expressed after infection with several pathogens. For these genes, gene induction does not appear to be influenced by general pathogen characteristics, such as the route or site of infection (e.g. the epidermis or intestine) or general pathogen surface properties (e.g. no obvious difference between Gram-negative or Gram-positive bacteria). Instead of these broader patterns of specificity expected to be common for innate immunity, a highly specific pattern of regulation is displayed (fig. 1a). For instance, *clec-67* shows upregulation upon a relatively broad range of pathogens, including both Gram-positive and Gram-negative bacteria, whereas *clec-165* is induced only by the Gram-negative *P. luminescens*. Other *clec* genes are specifically downregulated, such as *clec-273* by infection with the Gram-negative *Xenorhabdus nematophila* (fig. 1a). Interestingly, *clec* genes within a genomic cluster are not coexpressed, but seem to undergo regulation at the individual level. For example, *clec-58*, *clec-59*, *clec-60*, *clec-61* and *clec-143* lie within a cluster on chromosome II and have the same domain architecture. However, *clec-59*, *clec-60* and *clec-61* are differentially expressed after infection with several pathogens, while *clec-58* has so far not been shown to be differentially regulated upon pathogen exposure, and expression of *clec-143* is only induced following infection with *P. aeruginosa*. Different expression of *clec* genes within a genomic cluster might further indicate functional diversification.

C. elegans clec Genes Function in Pathogen Resistance

Our knowledge of the actual function of CLEC proteins in *C. elegans* is still in its infancy and it is as yet unknown why so many *clec* genes are down- or upregulated after exposure to pathogens. Few studies have tried to further elucidate the exact involvement of *clec* genes in *C. elegans* immunity by using functional genetic approaches and as yet only one study could confirm a role for CLEC proteins in pathogen binding (table 2).

The most common approach, which is taken to further analyze the function of *clec* genes, is survival analysis of either the respective *clec* mutant or *clec* RNAi-treated worms [49, 50]. In some studies, survival analysis was combined with the generation of transgenic *C. elegans* strains carrying a transcriptional *clec* gene reporter and the subsequent analysis of spatial as well as infection-induced expression patterns in vivo [51–53]. These studies demonstrated that most of the investigated *clec* genes are expressed in the intestine (fig. 1c), where most bacterial infections take place (fig. 1b), and that some *clec* genes are indeed required for resistance against infection (table 2).

Table 2. Functional analyses of CTLD genes/proteins in *C. elegans*

Gene/protein	Approach	Phenotypic measurement	Phenotype/result	Pathogen	Ref.
<i>clec-17</i>	RNAi knockdown	infection	hyperdeformed anal region (Dar) response, severe constipation	MN	53
<i>clec-39</i> <i>clec-49</i>	knockout mutant	survival egg laying	higher susceptibility lower egg laying		
CLEC-39 CLEC-49	ELISA-based assay with recombinant proteins	binding assays	binding to live/dead pathogen	SM	55
<i>clec-52</i>	promoter fusion reporter	fluorescence induction	expression in intestine	SA	52
	promoter fusion reporter	fluorescence induction	expression in intestine		
<i>clec-60</i>	RNAi knockdown	infection	hyperdeformed anal region (Dar) response, severe constipation	MN	53
	promoter fusion reporter	fluorescence induction	expression in intestine	SA	52
<i>clec-60/61</i> cluster	overexpression	survival	higher susceptibility	PA	52
<i>clec-65</i>	RNAi knockdown	survival	higher susceptibility	EC	51
<i>clec-67</i>	promoter fusion reporter	fluorescence induction	expression in intestine	SE	56
<i>clec-70</i>	RNAi knockdown	survival	higher susceptibility	SA	52
	promoter fusion reporter	fluorescence induction	expression in intestine		
<i>clec-70/71</i> cluster	overexpression	survival	higher resistance higher susceptibility	SA PA	52
CLEC-79	glycoconjugate microarray with recombinant protein	binding assay	binding to Galb1-3GalNAc		54
<i>clec-85</i>	promoter fusion reporter	fluorescence intensity	dependence on immune components <i>tir-1, nsy-1, dbl-1, daf-16</i>	SA PA	78
<i>clec-86</i>	RNAi knockdown	infection	hyperdeformed anal region (Dar) response, severe constipation	MN	53
<i>clec-174</i>	RNAi knockdown	survival	higher susceptibility	VC	50

EC = *E. coli* strain LF82; SE = *S. enterica*. For other abbreviations, see the legend to figure 1.

Fig. 1. The highly specific pattern of *clec* gene regulation through pathogen exposure, intestinal expression of *C. elegans clec-66*, and analysis of protein sequence similarity between *C. elegans* CTLD proteins and CTLD proteins from 10 other taxa. **a** Hierarchical clustering of differential expressions of randomly selected *C. elegans clec* genes (rows) after exposure to different microbes (columns), grouped in Gram-positive bacteria (blue), Gram-negative bacteria (pink) and fungi (green). Each column represents *clec* up- (red) or downregulation (blue) according to published transcriptome studies. DB27 and I-3690 are strains of *B. thuringiensis* and *Lactobacillus rhamnosus*, respectively. *clec* expression after infection with *B. thuringiensis* [63] represents merged data from two *B. thuringiensis* concentrations and two time points. BM = *B. megaterium*; BT = *B. thuringiensis*; DC = *D. coniospora*; EC = *E. carotovora*; EF = *E. faecalis*; HA = *Harposporium* spp.; LR = *L. rhamnosus*; ML = *M. luteus*; MN = *M. nematophilum*; PA = *P. aeruginosa*;

PL = *P. luminescens*; SA = *S. aureus*; SM = *S. marcescens*; VC = *V. cholerae*; XN = *X. nematophilus*; YP = *Y. pestis*. **b** GFP-labeled spores of the pathogenic *B. thuringiensis* strain MYBT18247 accumulate in the *C. elegans* anterior intestine at an early time point after infection. **c** Animals carrying a *pclec-66::GFP* reporter that is expressed throughout the intestine with highest expression in the anterior and posterior intestine. The coinjection marker *pttx-3::RFP* (red) is expressed in head neurons. **d** The number of *C. elegans* CTLD proteins showing different levels of sequence similarity to the CTLD protein repertoire of other vertebrates and invertebrates. Sequence similarity was assessed with BLAST and a comparison of the 283 *C. elegans* CTLD proteins to those of 10 selected vertebrate and invertebrate taxa. The matrix shows the number of proteins falling into different percentage similarity categories. The observed variation is further highlighted by different shades of gray (dark gray for large numbers, light gray for small numbers).

There have been only two studies investigating the function of CLEC proteins on the protein level. One of these provided experimental evidence for the sugar-binding qualities of CLEC-79 [54], though not in the context of an immune response to infection, while the other demonstrated that CLEC-39 and CLEC-49 were able to bind live and dead *Serratia marcescens* in a Ca^{2+} -independent manner [55]. Both *clec-39* and *clec-49* seem to be required for nematode fitness in the course of an encounter with *S. marcescens* since knockout mutants produced lower reproductive and survival rates. The authors reasoned that due to a lack of bactericidal activity these CLEC proteins could act as PRRs, inducing downstream immune pathways after recognition of the pathogenic bacteria.

There is thus accumulating evidence that *clec* genes function in *C. elegans* resistance to pathogen infection. However, the exact role of the vast majority of CTLD proteins in *C. elegans* immunity remains unclear.

Expression of C. elegans clec Genes Is Regulated by Stress and Immunity Pathways

In *C. elegans* several highly conserved signaling pathways (e.g. p38 MAPK, insulin signaling and TGF β) are required for the induction of immune effector molecules and, ultimately, host survival. The expression of numerous *clec* genes was shown to be under the control of these *C. elegans* immunity pathways and other known regulators of immunity, such as the intestine-specific GATA transcription factor ELT-2 [56] or the E-box transcription factor HLH-30 [57].

Taken together, CTLD proteins are highly diversified in the *C. elegans* genome with a potential further extension of binding capabilities through oligomerization and internal domain duplications. Experimental evidence demonstrates a *clec* gene function in *C. elegans* innate immunity, their regulation by known immunity pathways, and their highly specific response to different pathogens. Especially the latter emphasizes the particular potential of *C. elegans* CTLD proteins to generate immune specificity.

Similarity of *C. elegans* CTLD Proteins to Those from Other Taxa

Although CTLD proteins from the various taxa are related through the presence of the conserved CTL domain, it is not clear how similar the complete protein sequences are to each other. To address this point, we used the *C. elegans* CTLD protein repertoire as an example and determined their sequence similarity to the CTLD proteins

from a selected range of 10 invertebrate and vertebrate taxa, including *Homo sapiens*, *M. musculus*, *D. melanogaster* and *Nematostella vectensis*. This analysis revealed that similarities above 70% are only found for comparisons within the same genus (fig. 1d). In comparisons to distinct taxonomic groups, similarity scores for the large majority of proteins are below 50% and none of the comparisons revealed a value of more than 60%. Interestingly, even in the comparisons to the other *Caenorhabditis* taxa, approximately half of the proteins produce similarity scores of less than 50%.

These results highlight that: (i) CTLD proteins are primarily related through the presence of the CTL domain and a likely function in immunity (at least indicated for numerous CTLD proteins), but otherwise seem to show high diversification in the various phylogenetic lineages, (ii) CTLD proteins thus appear to comprise a highly dynamic, fast-evolving gene family, possibly in response to selective pressure from pathogens, and (iii) information on the function of a certain CTLD protein from one taxon, for example *C. elegans*, cannot be directly transferred onto a specific CTLD protein from a different taxon, for example humans, even though general functional similarities can be deduced through a comparative approach. These findings reinforce our notion that CTLD proteins represent a gene family that is of particular interest to our understanding of the evolution of innate and invertebrate immune systems, and especially their dynamic nature within different phylogenetic lineages.

Current Limitations in Our Understanding of Invertebrate CTLD Proteins

The recently described, surprisingly high degree of specificity of innate immune systems must be based on as yet unknown underlying molecular mechanisms. One of the possibilities may rely on genomically diversified gene families, such as those of the CTLD genes. These genes are numerous in many invertebrate species (table 1) and possess several additional characteristics that support a role in high immune specificity. In the model nematode *C. elegans*, the expression of many *clec* genes is regulated through pathogen exposure in a highly specific form (fig. 1a) [47]. CTLD proteins play an important role in the innate immune system of various invertebrate taxa by activating cellular, and in one case humoral [43], immune responses. Several insect CTLD proteins were found to specifically bind to Gram-negative [18, 20, 31, 34] or Gram-positive bacteria [25]. On the protein level, further

functional diversity of CTLD proteins might be achieved through oligomerization and internal domain duplications.

To date, however, the exact role of invertebrate CTLD proteins in generating immune specificity is completely unexplored. In most cases, only single CTLD proteins from a particular invertebrate species have been studied, but not a larger set of the gene family. On the protein level, some invertebrate CTLD proteins were found to bind to either a broad range of microorganisms (both Gram-positive and Gram-negative bacteria) or alternatively only to a particular pathogen group (Gram-negative or Gram-positive bacteria), already indicating a certain level of immune specificity. At the moment, none of the studied CTLD proteins specifically bind to a single pathogen species. The reasons for this might be that most, if not all, binding analyses on CTLD proteins are performed *in vitro* and might not reflect what is happening during pathogen infection *in vivo*. Moreover, they usually do not consider the pathogen variants, which coexist and coevolve with the host in nature and against which high immune specificity may thus have evolved across time. In the future, a larger variety of CTLD proteins should be studied for a particular species. For these studies, it is important that CTLD gene function is analyzed both *in vitro* and *in vivo*. The latter is particularly important (yet also challenging) because it can reveal to what extent interactions with other proteins, dimerization and oligomerization might influence the binding properties of CTLD proteins. In addition, the identification of highly specific host-pathogen interactions requires a study system in which the host is infected in a natural way with a natural pathogen and obviously with different strains of the same pathogen species, a setup that is so far not available for studies in *C. elegans*, crustacean and insect models.

Moreover, at the moment, the general functions of the majority of invertebrate CTLD proteins have also not been characterized. While the studies on single CTLD proteins in invertebrates provide important evidence on their role in innate immunity, an understanding of their exact properties and overall functions is far from complete. For example, we lack detailed information regarding how the expression of CTLD proteins is regulated and controlled, or their exact binding properties and targets. In *C. elegans*, CTLD proteins have repeatedly been suggested to be involved in immune defense by acting as antimicrobial effectors, and signaling pathways regulating CTLD gene expression have been identified, whereas in insects and crustaceans CTLD proteins are mainly as-

sumed to act as PRRs activating cellular immune responses. This discrepancy might be due to the fact that in *C. elegans* comprehensive genomic and transcriptomic data sets are available, which revealed that the majority of *C. elegans* CTLD proteins are induced upon pathogen infection and also contain a signal peptide and are thus likely to be secreted. In addition, several *C. elegans* CTLD gene knockout or knockdown animals have been phenotypically characterized, supporting the view that CTLD proteins are important for host resistance to infection. Nevertheless, for this nematode, we currently lack data on CTLD protein interactions with other proteins, the possible downstream molecular signaling pathways, and also data from biochemical assays to infer CTLD binding properties on the protein level.

In contrast, such analyses at the protein level are available for different crustaceans and insects. In particular, *in vitro* studies using recombinant or native proteins revealed the binding characteristics of several CTLD proteins and provided information on the involvement of CTLD proteins in activating cellular immune defenses. However, in these taxa information from functional genetic studies to confirm a role of CTLD proteins in immune defense *in vivo* are extremely scarce. Such functional genetic analyses may represent a particular challenge for such large gene families because of functional redundancy among paralogs. Thus, a single gene knockout or knockdown may not necessarily result in a visible phenotype even if the gene is of functional importance. Two approaches have been successfully used to target functionally redundant genes and may prove useful in the future: the simultaneous introduction of mutations in several or all members of a gene family using the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein 9 (Cas9) system [58, 59] and the simultaneous knockdown of several genes using RNAi-based approaches [60].

To conclude, it remains unclear for most invertebrate CTLD proteins if they either act as PRRs, directly recognizing pathogens by binding to MAMPs, as signaling mediators or as antimicrobial effectors. To identify mechanisms of CTLD protein action biochemical analyses on the purified native or recombinant proteins to identify binding capabilities and interacting proteins need to be combined with assays on bactericidal effects and also *in vivo* functional genetic analyses in the respective host taxa. The study by Wang et al. [44] represents a first example of such an integrated approach.

Even though our review focused on invertebrates, we would like to note that similar CTLD protein diversifica-

tion within and across species is also found in vertebrates. The potential role of CTLD proteins in contributing to immune specificity may thus also be of relevance in vertebrate taxa. Consistent with this idea, Drickamer and Taylor [61] proposed in a recent review that mammalian CTLD proteins, which bind to endogenous or self-glycans, are generally conserved, while those which bind to MAMPs are generally divergent and vary in number among species. Examples of such species-specific differences are: (i) two SIGN molecules expressed in humans (DC-SIGN and DC-SIGNR) compared to eight SIGN molecules expressed in mice, (ii) two loci coding for functional mannose-binding lectin in mice compared to only one in humans [61], (iii) one BDCA-2 gene in humans with no reported homologue in mice and (iv) one DCIR-coding locus in humans compared to four (*Dcir1–4*) in mice [62]. The CTLD proteins in mice and humans do

not only differ in number but also in their binding specificities, suggesting that they could indeed contribute to specific responses to pathogens and that their diversification may be a consequence of an arms race with specific coevolving pathogens enabling the host to meet the particular challenges imposed by rapidly evolving pathogens.

Acknowledgements

We thank Andrei Papkou for providing the picture of *C. elegans* infected with GFP-labeled *B. thuringiensis* spores. B.P., W.Y., H.S. and K.D. are supported by grants from the German Science foundation (Grant DI 1687/1-1 to K.D. and SCHU 1415/9-2 to H.S.). A.Z.-P. and W.Y. are members of the IMPRS for Evolutionary Biology. K.D. is additionally supported by institutional funding from Kiel University.

References






- Little TJ, O'Connor B, Colegrave N, Watt K, Read AF: Maternal transfer of strain-specific immunity in an invertebrate. *Curr Biol* 2003; 13:489–492.
- Roth O, Sadd BM, Schmid-Hempel P, Kurtz J: Strain-specific priming of resistance in the red flour beetle, *Tribolium castaneum*. *Proc R Soc B Biol Sci* 2009;276:145–151.
- Sadd BM, Schmid-Hempel P: Insect immunity shows specificity in protection upon secondary pathogen exposure. *Curr Biol* 2006; 16:1206–1210.
- Watson FL: Extensive diversity of Ig-superfamily proteins in the immune system of insects. *Science* 2005;309:1874–1878.
- Watthanasurorot A, Jiravanichpaisal P, Liu H, Söderhäll I, Söderhäll K: Bacteria-induced Dscam isoforms of the Crustacean, *Pacifastacus leniusculus*. *PLoS Pathog* 2011;7:e1002062.
- Armitage SAO, Peuss R, Kurtz J: Dscam and pancrustacean immune memory – a review of the evidence. *Dev Comp Immunol* 2014;48: 315–323.
- Cannon JP, Haire RN, Schnitker N, Mueller MG, Litman GW: Individual protochordates have unique immune-type receptor repertoires. *Curr Biol* 2004;14:R465–R466.
- Zhang S-M, Adema CM, Kepler TB, Loker ES: Diversification of Ig superfamily genes in an invertebrate. *Science* 2004;305:251–254.
- Ghosh J, Buckley KM, Nair SV, Raftos DA, Miller C, Majeske AJ, et al: Sp185/333: a novel family of genes and proteins involved in the purple sea urchin immune response. *Dev Comp Immunol* 2010;34:235–245.
- Hibino T, Loza-Coll M, Messier C, Majeske AJ, Cohen AH, Terwilliger DP, et al: The immune gene repertoire encoded in the purple sea urchin genome. *Dev Biol* 2006;300:349–365.
- Huang S, Yuan S, Guo L, Yu Y, Li J, Wu T, et al: Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. *Genome Res* 2008;18:1112–1126.
- Weis WI, Taylor ME, Drickamer K: The C-type lectin superfamily in the immune system. *Immunol Rev* 1998;163:19–34.
- Zelensky AN, Gready JE: The C-type lectin-like domain superfamily. *FEBS J* 2005;272: 6179–6217.
- Hoving JC, Wilson GJ, Brown GD: Signalling C-type lectin receptors, microbial recognition and immunity: C-type lectins in immunity. *Cell Microbiol* 2014;16:185–194.
- Vaishnav S, Yamamoto M, Severson KM, Ruhn KA, Yu X, Koren O, et al: The antibacterial lectin RegIII γ promotes the spatial segregation of microbiota and host in the intestine. *Science* 2011;334:255–258.
- Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, et al: InterPro: the integrative protein signature database. *Nucleic Acids Res* 2009;37:D211–D215.
- Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al: Pfam: the protein families database. *Nucleic Acids Res* 2014;42:D222–D230.
- Ao J, Ling E, Yu X-Q: *Drosophila* C-type lectins enhance cellular encapsulation. *Mol Immunol* 2007;44:2541–2548.
- Haq S, Kubo T, Kurata S, Kobayashi A, Natori S: Purification, characterization, and cDNA cloning of a galactose-specific C-type lectin from *Drosophila melanogaster*. *J Biol Chem* 1996;271:20213–20218.
- Tanji T, Ohashi-Kobayashi A, Natori S: Participation of a galactose-specific C-type lectin in *Drosophila* immunity. *Biochem J* 2006;396: 127–138.
- Koizumi N, Morozumi A, Imamura M, Tanaka E, Iwahana H, Sato R: Lipopolysaccharide-binding proteins and their involvement in the bacterial clearance from the hemolymph of the silkworm *Bombyx mori*. *Eur J Biochem* 1997;248:217–224.
- Wilson R, Chen C, Ratcliffe NA: Innate immunity in insects: the role of multiple, endogenous serum lectins in the recognition of foreign invaders in the cockroach, *Blattella discoidalis*. *J Immunol* 1999;162:1590–1596.
- Yu X-Q, Zhu Y-F, Ma C, Fabrick JA, Kanost MR: Pattern recognition proteins in *Manduca sexta* plasma. *Insect Biochem Mol Biol* 2002; 32:1287–1293.
- Koizumi N, Imai Y, Morozumi A, Imamura M, Kadotani T, Yaoi K, et al: Lipopolysaccharide-binding protein of *Bombyx mori* participates in a hemocyte-mediated defense reaction against Gram-negative bacteria. *J Insect Physiol* 1999;45:853–859.
- Chai L-Q, Tian Y-Y, Yang D-T, Wang J-X, Zhao X-F: Molecular cloning and characterization of a C-type lectin from the cotton bollworm, *Helicoverpa armigera*. *Dev Comp Immunol* 2008;32:71–83.
- Watanabe A, Miyazawa S, Kitami M, Tabunoki H, Ueda K, Sato R: Characterization of a novel C-type lectin, *Bombyx mori* multibinding protein, from the *B. mori* hemolymph: mechanism of wide-range microorganism recognition and role in immunity. *J Immunol* 2006;177:4594–4604.


- 27 Kanost MR, Jiang H, Yu X-Q: Innate immune responses of a lepidopteran insect, *Manduca sexta*. *Immunol Rev* 2004;198:97–105.
- 28 Yu X-Q, Tracy ME, Ling E, Scholz FR, Trenczek T: A novel C-type lectin-3 from *Manduca sexta* is translocated from hemolymph into the cytoplasm of hemocytes. *Insect Biochem Mol Biol* 2005;35:285–295.
- 29 Yu X-Q, Ling E, Tracy ME, Zhu Y: Immulectin-4 from the tobacco hornworm *Manduca sexta* binds to lipopolysaccharide and lipoteichoic acid. *Insect Mol Biol* 2006;15:119–128.
- 30 Ausubel FM: Are innate immune signaling pathways in plants and animals conserved? *Nat Immunol* 2005;6:973–979.
- 31 Yu X-Q, Kanost MR: Immulectin-2, a lipopolysaccharide-specific lectin from an insect, *Manduca sexta*, is induced in response to Gram-negative bacteria. *J Biol Chem* 2000;275:37373–37381.
- 32 Eleftherianos I, Millichap PJ, Ffrench-Constant RH, Reynolds SE: RNAi suppression of recognition protein mediated immune responses in the tobacco hornworm *Manduca sexta* causes increased susceptibility to the insect pathogen *Photorhabdus*. *Dev Comp Immunol* 2006;30:1099–1107.
- 33 Osta MA, Christophides GK, Kafatos FC: Effects of mosquito genes on plasmodium development. *Science* 2004;303:2030–2032.
- 34 Schnitger AKD, Yassine H, Kafatos FC, Osta MA: Two C-type lectins cooperate to defend *Anopheles gambiae* against Gram-negative bacteria. *J Biol Chem* 2009;284:17616–17624.
- 35 Sánchez-Paz A: White spot syndrome virus: an overview on an emergent concern. *Vet Res* 2010;41:43.
- 36 Wang X-W, Wang J-X: Diversity and multiple functions of lectins in shrimp immunity. *Dev Comp Immunol* 2013;39:27–38.
- 37 Huang Y, Huang X, Wang Z, Tan J-M, Hui K-M, Wang W, et al: Function of two novel single-CRD containing C-type lectins in innate immunity from *Eriocheir sinensis*. *Fish Shellfish Immunol* 2014;37:313–321.
- 38 Wang X-W, Zhang X-W, Xu W-T, Zhao X-F, Wang J-X: A novel C-type lectin (FcLec4) facilitates the clearance of *Vibrio anguillarum* in vivo in Chinese white shrimp. *Dev Comp Immunol* 2009;33:1039–1047.
- 39 Zhang X-W, Liu Y-Y, Mu Y, Ren Q, Zhao X-F, Wang J-X: Overexpression of a C-type lectin enhances bacterial resistance in red swamp crayfish *Procambarus clarkii*. *Fish Shellfish Immunol* 2013;34:1112–1118.
- 40 Sun Y-D, Fu L-D, Jia Y-P, Du X-J, Wang Q, Wang Y-H, et al: A hepatopancreas-specific C-type lectin from the Chinese shrimp *Fenneropenaeus chinensis* exhibits antimicrobial activity. *Mol Immunol* 2008;45:348–361.
- 41 Jin X-K, Li S, Guo X-N, Cheng L, Wu M-H, Tan S-J, et al: Two antibacterial C-type lectins from crustacean, *Eriocheir sinensis*, stimulated cellular encapsulation in vitro. *Dev Comp Immunol* 2013;41:544–552.
- 42 Luo T, Yang H, Li F, Zhang X, Xu X: Purification, characterization and cDNA cloning of a novel lipopolysaccharide-binding lectin from the shrimp *Penaeus monodon*. *Dev Comp Immunol* 2006;30:607–617.
- 43 Wang X-W, Xu J-D, Zhao X-F, Vasta GR, Wang J-X: A shrimp C-type lectin inhibits proliferation of the hemolymph microbiota by maintaining the expression of antimicrobial peptides. *J Biol Chem* 2014;289:11779–11790.
- 44 Wang X-W, Zhao X-F, Wang J-X: C-type lectin binds to β -integrin to promote hemocytic phagocytosis in an invertebrate. *J Biol Chem* 2014;289:2405–2414.
- 45 Shiratsuchi A, Mori T, Sakurai K, Nagaosa K, Sekimizu K, Lee BL, et al: Independent recognition of *Staphylococcus aureus* by two receptors for phagocytosis in *Drosophila*. *J Biol Chem* 2012;287:21663–21672.
- 46 Li M, Li C, Ma C, Li H, Zuo H, Weng S, et al: Identification of a C-type lectin with antiviral and antibacterial activity from pacific white shrimp *Litopenaeus vannamei*. *Dev Comp Immunol* 2014;46:231–240.
- 47 Schulenburg H, Hoepfner MP, Weiner J, Bornberg-Bauer E: Specificity of the innate immune system and diversity of C-type lectin domain (CTLN) proteins in the nematode *Caenorhabditis elegans*. *Immunobiology* 2008;213:237–250.
- 48 Petersen TN, Brunak S, von Heijne G, Nielsen H: SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 2011;8:785–786.
- 49 Schulenburg H, Kurz CL, Ewbank JJ: Evolution of the innate immune system: the worm perspective. *Immunol Rev* 2004;198:36–58.
- 50 Sahu SN, Lewis J, Patel I, Bozdog S, Lee JH, LeClerc JE, et al: Genomic analysis of immune response against *Vibrio cholerae* hemolysin in *Caenorhabditis elegans*. *PLoS One* 2012;7:e38200.
- 51 Simonsen KT, Møller-Jensen J, Kristensen AR, Andersen JS, Riddle DL, Kallipolitis BH: Quantitative proteomics identifies ferritin in the innate immune response of *C. elegans*. *Virulence* 2011;2:120–130.
- 52 Irazoqui JE, Troemel ER, Feinbaum RL, Luhachack LG, Cezairliyan BO, Ausubel FM: Distinct pathogenesis and host responses during infection of *C. elegans* by *P. aeruginosa* and *S. aureus*. *PLoS Pathog* 2010;6:e1000982.
- 53 O'Rourke D, Baban D, Demidova M, Mott R, Hodgkin J: Genomic clusters, putative pathogen recognition molecules, and antimicrobial genes are induced by infection of *C. elegans* with *M. nematophilum*. *Genome Res* 2006;16:1005–1016.
- 54 Takeuchi T, Sennari R, Sugiura K, Tateno H, Hirabayashi J, Kasai K: A C-type lectin of *Caenorhabditis elegans*: its sugar-binding property revealed by glycoconjugate microarray analysis. *Biochem Biophys Res Commun* 2008;377:303–306.
- 55 Miltsch SM, Seeberger PH, Lepenies B: The C-type lectin-like domain containing proteins Clec-39 and Clec-49 are crucial for *Caenorhabditis elegans* immunity against *Serratia marcescens* infection. *Dev Comp Immunol* 2014;45:67–73.
- 56 Kerry S, TeKippe M, Gaddis NC, Aballay A: GATA transcription factor required for immunity to bacterial and fungal pathogens. *PLoS One* 2006;1:e77.
- 57 Visvikis O, Ihuegbu N, Labea SA, Luhachack LG, Alves A-MF, Wollenberg AC, et al: Innate host defense requires TFEB-mediated transcription of cytoprotective and antimicrobial genes. *Immunity* 2014;40:896–909.
- 58 Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, et al: One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 2013;153:910–918.
- 59 Peng D, Kurup SP, Yao PY, Minning TA, Tarleton RL: CRISPR-Cas9-mediated single-gene and gene family disruption in *Trypanosoma cruzi*. *mBio* 2014;6:e02097-14.
- 60 Kim J, Badaloni A, Willert T, Zimmer-Strobl U, Kühn R, Wurst W, et al: An RNAi-based approach to down-regulate a gene family in vivo. *PLoS One* 2013;8:e80312.
- 61 Drickamer K, Taylor ME: Recent insights into structures and functions of C-type lectins in the immune system. *Curr Opin Struct Biol* 2015;34:26–34.
- 62 Dambuzza IM, Brown GD: C-type lectins in immunity: recent developments. *Curr Opin Immunol* 2015;32:21–27.
- 63 Yang W, Dierking K, Esser D, Tholey A, Leippe M, Rosenstiel P, et al: Overlapping and unique signatures in the proteomic and transcriptomic responses of the nematode *Caenorhabditis elegans* toward pathogenic *Bacillus thuringiensis*. *Dev Comp Immunol* 2015;51:1–9.
- 64 Coolon JD, Jones KL, Todd TC, Carr BC, Herman MA: *Caenorhabditis elegans* genomic response to soil bacteria predicts environment-specific genetic effects on life history traits. *PLoS Genet* 2009;5:e1000503.
- 65 Wong D, Bazopoulou D, Pujol N, Tavernarakis N, Ewbank JJ: Genome-wide investigation reveals pathogen-specific and shared signatures in the response of *Caenorhabditis elegans* to infection. *Genome Biol* 2007;8:R194.
- 66 Boehnisch C, Wong D, Habig M, Isermann K, Michiels NK, Roeder T, et al: Protist-type lysozymes of the nematode *Caenorhabditis elegans* contribute to resistance against pathogenic *Bacillus thuringiensis*. *PLoS One* 2011;6:e24619.
- 67 Treitz C, Cassidy L, Höckendorf A, Leippe M, Tholey A: Quantitative proteome analysis of *Caenorhabditis elegans* upon exposure to nematocidal *Bacillus thuringiensis*. *J Proteomics* 2015;113:337–350.

- 68 Sinha A, Rae R, Iatsenko I, Sommer RJ: System wide analysis of the evolution of innate immunity in the nematode model species *Caenorhabditis elegans* and *Pristionchus pacificus*. *PLoS One* 2012;7:e44255.
- 69 Grompone G, Martorell P, Llopis S, González N, Genovés S, Mulet AP, et al: Anti-inflammatory *Lactobacillus rhamnosus* CNCM I-3690 strain protects against oxidative stress and increases lifespan in *Caenorhabditis elegans*. *PLoS One* 2012;7:e52493.
- 70 Bolz DD, Tenor JL, Aballay A: A conserved PMK-1/p38 MAPK is required in *Caenorhabditis elegans* tissue-specific immune response to *Yersinia pestis* infection. *J Biol Chem* 2010; 285:10832–10840.
- 71 Shapira M, Hamlin BJ, Rong J, Chen K, Ronen M, Tan M-W: A conserved role for a GATA transcription factor in regulating epithelial innate immune responses. *Proc Natl Acad Sci USA* 2006;103:14086–14091.
- 72 Engelmann I, Griffon A, Tichit L, Montaña-Sanchis F, Wang G, Reinke V, et al: A comprehensive analysis of gene expression changes provoked by bacterial and fungal infection in *C. elegans*. *PLoS One* 2011;6:e19055.
- 73 Drummond RA, Brown GD: Signalling C-type lectins in antimicrobial immunity. *PLoS Pathog* 2013;9:e1003417.
- 74 Yu Y, Yu Y, Huang H, Feng K, Pan M, Yuan S, et al: A short-form C-type lectin from *Amphioxus* acts as a direct microbial killing protein via interaction with peptidoglycan and glucan. *J Immunol* 2007;179:8425–8434.
- 75 Multerer KA, Smith LC: Two cDNAs from the purple sea urchin, *Strongylocentrotus purpuratus*, encoding mosaic proteins with domains found in factor H, factor I, and complement components C6 and C7. *Immunogenetics* 2004;56:89–106.
- 76 Terwilliger DP, Clow LA, Gross PS, Smith LC: Constitutive expression and alternative splicing of the exons encoding SCRs in Sp152, the sea urchin homologue of complement factor B: implications on the evolution of the Bf/C2 gene family. *Immunogenetics* 2004;56:531–543.
- 77 Koizumi N, Imamura M, Kadotani T, Yaoi K, Iwahana H, Sato R: The lipopolysaccharide-binding protein participating in hemocyte nodule formation in the silkworm *Bombyx mori* is a novel member of the C-type lectin superfamily with two different tandem carbohydrate-recognition domains. *FEBS Lett* 1999;443:139–143.
- 78 Alper S, McBride SJ, Lackford B, Freedman JH, Schwartz DA: Specificity and complexity of the *Caenorhabditis elegans* innate immune response. *Mol Cell Biol* 2007;27:5544–5553.
- 79 Troemel ER, Chu SW, Reinke V, Lee SS, Ausubel FM, Kim DH: p38 MAPK regulates expression of immune response genes and contributes to longevity in *C. elegans*. *PLoS Genet* 2006;2:e183.
- 80 Mallo GV, Kurz CL, Couillault C, Pujol N, Granjeaud S, Kohara Y, et al: Inducible antibacterial defense system in *C. elegans*. *Curr Biol* 2002;12:1209–1214.
- 81 Huffman DL, Abrami L, Sasik R, Corbeil J, van der Goot FG, Aroian RV: Mitogen-activated protein kinase pathways defend against bacterial pore-forming toxins. *Proc Natl Acad Sci USA* 2004;101:10995–11000.
- 82 Pujol N, Zugasti O, Wong D, Couillault C, Kurz CL, Schulenburg H, et al: Anti-fungal innate immunity in *C. elegans* is enhanced by evolutionary diversification of antimicrobial peptides. *PLoS Pathog* 2008;4:e1000105.

Supplementary Tables

Table S1. Diversity of CTLD proteins in the *C. elegans* genome.

Class	Protein domains ^a	Total ^b #	Secreted/membr.-bound ^c			Genomic clusters ^d #	Cluster names	Pathogen-induction ^e	
			# S	# M	# S + M			#	Pathogens
I	1 CTLD 	144	115	8	13	65	I.1, I.2, I.3, II.2, II.3, II.4, III.1, IV.1, IV.2, IV.3, IV.4, IV.5, V.1, V.3, V.6, V.7, V.8	99	BM, BT, LR, SM, PL, EC, HA, EF, DC, ML, VC, SA, MN, YP, PA, XN
II	2-3 CTLD 	44	23	17	1	26	II.2, II.3, IV.4, IV.5, V.4, V.5	40	SM, EF, PL, HA, EC, ML, SA, DC, BT, PA
III	1-3 CTLD + 1-3 CUB 	31	22	1	7	9	II.1, V.2, V.6	25	MN, PL, PA, DC, EF, SM, HA, BT, XN, SA
IV	1-2 CTLD + 1-2 CW 	35	24	4	0	14	I.2, I.4, II.3, IV.4, V.8, V.9	20	HA, EF, SM, PL, DC, PA
V	1 CTLD + 1-2 VWA 	14	14	0	0	10	II.5, II.6	12	DC, EF, SM, PL, PA, HA, SA, MN, YP, BT, EC

Class	Protein domains ^a	Total ^b #	Secreted/membr.-bound ^c			Genomic clusters ^d		Pathogen-induction ^e	
			# S	# M	# S + M	#	Cluster names	#	Pathogens
VI	Complex structure 	15	4	2	8	1	I.3	11	EF, PL, SM, DC, HA, YP, BM
Sum		283	202	32	29	125		207	

^a Number of protein domains present. Domain abbreviations: CTLD, C-type lectin-like domain; CUB, Complement C1r/C1s Uegf Bmp1 domain; CW, conserved cysteine and tryptophan residues, Caenorhabditis-specific domain; VWA, Von Willebrand factor type A domain; Complex structure, complex domain structure, where CTLDs might not determine the actual function of the protein.

^b Total number (#) of CTLD proteins per class.

^c Number (#) of CTLD proteins with either signal sequence (S, secreted), transmembrane-spanning region (M, membr.-bound) or both according to SignalP 4.0 [48].

^d Number (#) of clec genes within genomic clusters and their respective names. A genomic clusters is defined as at least three genes within a continuous 20 kb fragment. Roman and Arabic numbers of the cluster names refer to the chromosome and cluster position within each chromosome, respectively.

^e Number of microorganism-induced clec genes per CTLD class taken from previous gene expression studies [50,52,53,64–66,68–72,79–82]. Abbreviations for microorganisms as explained in figure 1

Table S2. Diversity of CTLD genes in *Caenorhabditis* genomes.

Species	<i>C. remanei</i>				<i>C. japonica</i>				<i>C. briggsae</i>				<i>C. brenneri</i>				
	Protein domains ^a	Total ^b	Secreted/membr.-bound ^c			Total ^b	Secreted/membr.-bound ^c			Total ^b	Secreted/membr.-bound ^c			Total ^b	Secreted/membr.-bound ^c		
			#	# S	# M		# S + M	#	# S		# M	# S + M	#		# S	# M	# S + M
1 CTLD	107	79	5	7	71	35	3	7	85	54	6	6	166	98	7	15	
2-3 CTLD	48	30	8	5	9	4	1	1	21	15	2	0	26	16	3	2	
1-3 CTLD + 1-3 CUB	39	25	1	5	9	6	0	1	28	22	0	4	44	32	0	6	
1-2 CTLD + 1-2 CW	52	26	2	3	14	8	1	1	19	10	1	1	56	26	1	2	
1 CTLD + 1-2 VWA	33	28	1	0	4	1	0	1	18	13	0	1	15	12	0	1	
Complex Structure	13	6	1	3	6	4	0	0	11	5	2	2	11	6	3	2	
Sum	292	194	18	23	113	58	5	11	182	119	11	14	318	190	14	28	

^a Number of protein domains present. Domain abbreviations as explained in table S1.

^b Total number (#) of CTLD proteins per class.

^c Number (#) of CTLD proteins with either signal sequence (S, secreted), transmembrane-spanning region (M, membr.-bound) or both according to SignalP 4.0 [50].

Chapter II

The Involvement of C-type Lectin-like Domain Genes in *C. elegans* Behavioral Immune Defense

Barbara Pees¹, Anke Kloock¹, Camilo Barbosa¹, Rania Nakad², Hinrich Schulenburg¹, Katja Dierking^{1†}

¹ Department of Evolutionary Ecology and Genetics, Zoological Institute, Christian Albrechts University of Kiel, Kiel, Germany

² Cologne Excellence Cluster for Cellular Stress Responses in Aging-Associated Diseases (CECAD), Cologne, Germany

† Corresponding author: kdierking@zoologie.uni-kiel.de

Abstract

C-type lectin-like domain (CTLCD) proteins occupy crucial functions in the immune system of vertebrates, but their role in invertebrate immunity is much less understood. The nematode *Caenorhabditis elegans* possesses a highly diverse CTLCD protein encoding (*clec*) gene repertoire. A role of *C. elegans clec* genes in the immune response is always assumed, yet, experimental evidence for *clec* immune function is rare. To systematically test the contribution of *clec* genes to the *C. elegans* defense against pathogens, we screened 39 *clec* mutants for survival on the Gram-positive pathogen *Bacillus thuringiensis* (BT) and the Gram-negative pathogen *Pseudomonas aeruginosa* (PA14). We found that 41% of the tested *clec* mutants showed altered survival after infection on BT and 27% on PA14, exhibiting both increased susceptibility and resistance. As resistance to pathogen infection was positively correlated with escape behavior, we then investigated the role of *C. elegans clec* genes in pathogen avoidance behavior. We show that the highly resistant *C54G4.4(ok2110)* mutant exhibits a particularly strong avoidance phenotype combined with prolonged feeding cessation when exposed to BT. Together, our data suggest that *C. elegans clec* genes mediate both resistance and susceptibility to infection and further provide evidence for a role of *C54G4.4* in regulating behavioral immune responses.

Introduction

Metazoan C-type lectins are known to be involved in many cellular processes such as cell-cell adhesion or endocytosis¹, but fulfill their most distinctive tasks in the context of immunity². C-type lectins are characterized by their ability to recognize and bind carbohydrates (“lectins”) in a calcium-dependent (“C-type”) fashion. The sugar-protein interaction is mediated by a common sequence motif, the carbohydrate recognition domain (CRD)³. Although C-type lectins share primary and secondary structural homology in their CRDs, not all are capable of binding sugar nor calcium, hence, the more general term C-type lectin-like domain (CTLCD) was introduced^{4,5}.

The function of CTLCD proteins is best characterized in the vertebrate immune system, where CTLCD proteins recognize pathogens as pattern recognition receptors (PRR) and modulate consequent immune responses (reviewed in ⁶). In invertebrates, the exact contribution of CTLCD proteins to immunity is much less understood despite the genomic abundance of CTLCD encoding genes in several taxa, e.g. insects⁷⁻¹⁰, echinodermata^{11,12}, or nematodes¹³. There are, however, few examples of invertebrate CTLCD proteins, for which a function in promoting phagocytosis, encapsulation, and melanization, or in initiating the prophenoloxidase-activating system could experimentally be demonstrated (e.g. ^{7,14-16}). The nematode *Caenorhabditis elegans* has an exceptionally diversified repertoire of CTLCD encoding (*clec*) genes composed of 283 members. In every gene expression study done in *C. elegans* after pathogen exposure several *clec* genes were found to be up-regulated, and that in a pathogen-specific manner (reviewed in ¹³). It is thus generally assumed that *C. elegans clec* genes are involved in the immune response during pathogen infection. The experimental evidence for such a contribution to *C. elegans* immunity is, however, still restricted to a very limited number of *clec* genes: Few studies describe a susceptibility phenotype for *clec* mutants following an infection (e.g. ^{17,18}) or demonstrate an intestinal expression of *clec* genes upon infection (e.g. ¹⁹). Only two studies examined CLEC function on the protein level^{20,21}.

In our study, we aimed at systematically testing the contribution of *clec* genes to the *C. elegans* defense against pathogens and screened 39 *clec* mutants for survival on the Gram-positive *Bacillus thuringiensis* (BT) and the Gram-negative *Pseudomonas aeruginosa* (PA14). 41% of the tested *clec* mutants showed altered susceptibility to infection on BT, 27% on PA14, surprisingly exhibiting both higher susceptibility and higher resistance. We noticed a positive correlation between resistance and pathogen avoidance and could confirm a role for the CTLCD encoding gene *C54G4.4* in avoidance behavior and feeding cessation. Taken together, we here provide evidence for a role of *C. elegans clec* genes in behavioral immune defenses.

Materials and Methods

Strains and Culture Conditions

Worms were grown and maintained on nematode growth medium (NGM) agar plates seeded with *Escherichia coli* OP50 as previously described²². Wild type strain N2 (Bristol) and all *clec* mutant strains were obtained from the Caenorhabditis Genetics Center (CGC, Minnesota, USA) or the National BioResource Project (NBRP, Tokyo, Japan)²³. The complete list of worm strains including their names, genotype, WormBase ID (Version WB250), and the source they were received from can be found in “Supplementary Tables and Figures” Table S1.

The BT strains MY-BT18247 (in the following BT18247) and BT407 were originally obtained from the Agriculture Research Service Patent Culture Collection of Microorganisms and Cell Cultures (United States Department of Agriculture, Peoria, Illinois, USA). BT spore-toxin mixtures were generated according previous protocols^{24–26}. The spore-toxin mixtures were frozen at -20 °C in 250 µl aliquots with a spore concentration ranging from 3*10⁹ to 8*10⁹ particles/ml for BT18247 and 1.1*10¹⁰ particles/ml for BT407 and immediately used for *C. elegans* infection. The *Pseudomonas aeruginosa* strain PA14 (provided by Dennis Kim) was first grown on LB plates and then cultured overnight at 37 °C in LB broth prior to inoculation of assay plates.

Survival, Avoidance, and Ingestion Assays

Peptone free medium (PFM) agar plates (6 cm diameter) were inoculated with 75 µl of a mixture of BT spore-toxin mixture and *E. coli* OP50 in PBS (adjusted to OD 5) at different concentrations (e.g. 1:50 and 1:100 in the initial screen). Plates were kept at 20 °C overnight. PA14 requires enriched NGM (0.35% instead of 0.25% peptone) for efficient “slow killing” of worms²⁷. NGM plates seeded with a mixture of PA14 (OD 1) and *E. coli* OP50 (OD 5) at a concentration of 1:3 were incubated for 24 h at 37 °C, followed by another incubation for 24 h at 25 °C²⁸. The non-pathogenic BT407 and *E. coli* OP50 were used as controls. Thirty synchronized worms at the fourth larval (L4) stage were picked onto each plate and survival was scored after 24 h for BT and 72 h for PA14 infection experiments. Worms were considered to be dead if they did not respond to light touch.

For assays of pathogen avoidance behavior plates were prepared as above with the exception that only 30 µl of the bacterial mixture were inoculated in the middle of the plate (9 cm diameter). Low BT concentrations were chosen in order to challenge, but not to kill the worms. Ten synchronized L4 hermaphrodites were picked onto each bacterial treatment at time point 0 hpi (hours post infection) and the worms residing on that spot were scored every second hour. The avoidance index was calculated as follows: (initial total number of worms - worms on bacterial spot)/initial total number of worms. Dead individuals were excluded from the total number of worms per plate.

For measuring the bacterial uptake, synchronized L4 hermaphrodites were exposed to mild concentration as described above and the pumping of the grinder was measured in individual worms within 20 s after 1, 6, and 24 hpi.

Statistical Analyses

Statistical analyses were done with RStudio (Version 0.98.977), graphs created with its package ggplot2 (Version 0.9.3.1) and edited with Inkscape (Version 0.48).

Results and Discussion

C. elegans clec mutants are both more resistant and more susceptible to infection with pathogenic BT and PA14.

We aimed at systematically testing the contribution of *clec* genes to the *C. elegans* defense against infection with the Gram-positive pathogen BT and the Gram-negative pathogen PA14. We screened 39 *clec* knock-out mutants that were available from the CGC and NBRP (“Supplementary Tables and Figures” Table S 1) for their survival phenotype on both pathogens. While nine *clec* mutant strains exhibited a susceptibility phenotype on BT and four *clec* mutants on PA14, seven *clec* mutant strains showed a resistance phenotype on BT and six *clec* mutants on PA14 (Figure 1; Table S 2, Table S 3). Except for *clec-90(tm3402)*, none of the *clec* mutant strains showed a significant survival phenotype on the non-pathogenic treatment controls, *E. coli* OP50 or BT407 (data not shown). We could confirm the results of the first screen for most of the *clec* mutants, which had shown significantly altered susceptibility to infection, in a secondary screen with an increased number of replicates on PA14 (data not shown) and an additional concentration for BT (Figure S 1A; Table S 2, Table S 3). Considering the common assumption that *C. elegans* CTLD proteins are key players in the immune defense it is somewhat surprising that 18% of the here tested *clec* mutants on BT and 16% of the *clec* mutants on PA14 showed an increased resistance phenotype. The genetic manipulation of *clec* genes by RNAi knock-down has been previously shown to affect worm survival after pathogen exposure¹⁷⁻²⁰. Using knock-out mutants Miltsch and colleagues found a higher susceptibility on *Serratia marcescens* for *clec-49(ok2416)* and *clec-39(ok2271)*²⁰, whereas the latter has shown to be more resistant on PA14 in our screen (Figure 1B). Given the fact that *S. marcescens* and PA14 are both Gram-negative bacteria the contrary survival phenotypes of the *clec-39* mutant might hint to a species-specific role in regulating the immune response, as previously proposed¹³. However, to the best of our knowledge it has never been reported before that *C. elegans clec* mutants are resistant to infection. Our results suggest that *C. elegans* CTLD proteins can mediate suppression of the immune response. For vertebrate C-type lectin receptors it is known that they play activatory as well as inhibitory roles in regulating the immune responses, one example is the inhibitory receptor Clec12a^{29,30}.

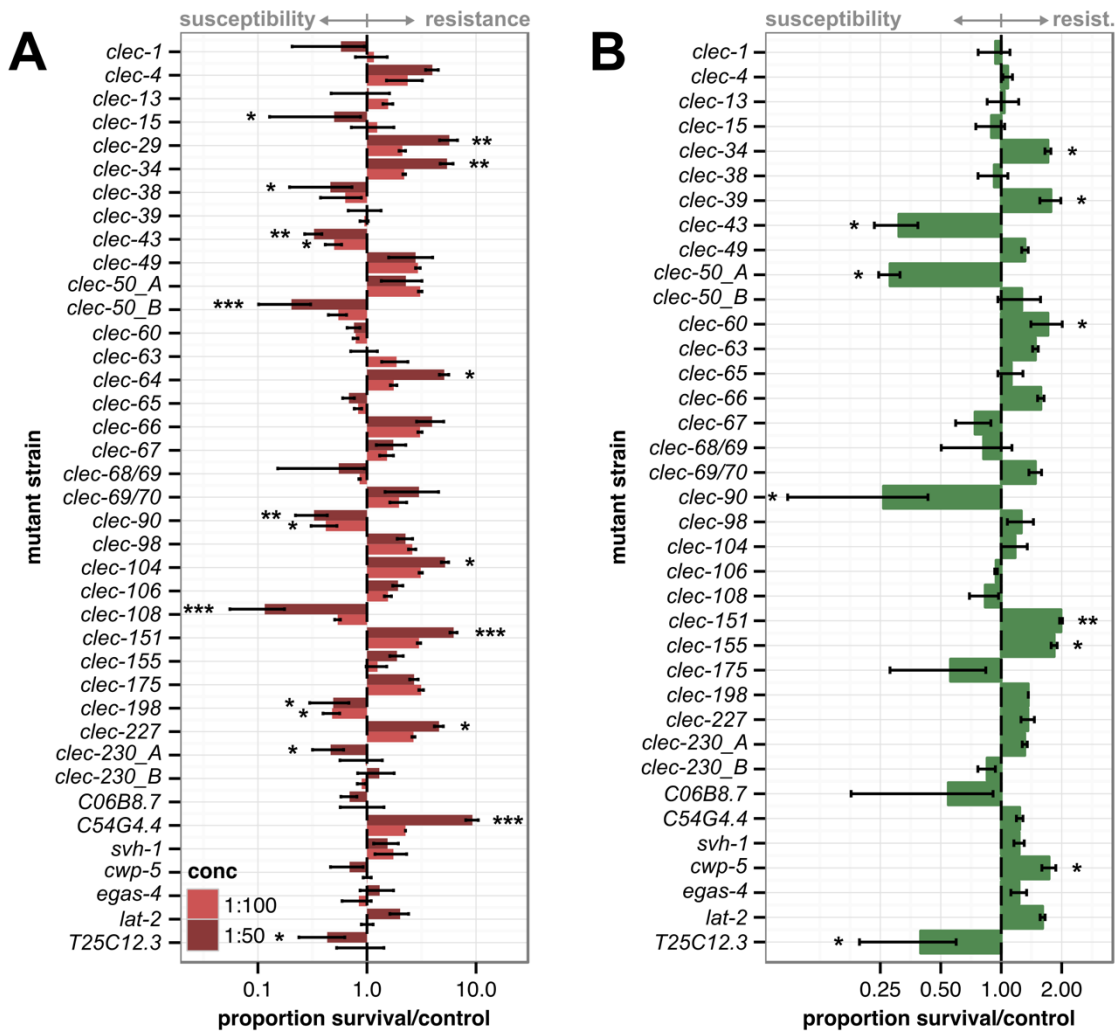


Figure 1: Survival screen of *C. elegans clec* mutants upon exposure to pathogenic (A) BT18247 and (B) PA14. The two different shades of red represent different concentrations (conc) of BT mixed with *E. coli* OP50. The proportional survival was scored (A) 24 hpi or (B) 72 hpi and normalized by the wild type control N2 ($x = 1$, vertical dashed line). Values < 1 indicate more susceptible, values > 1 more resistant strains. Means of three replicates are shown; the error bars are the standard error (SE) of the mean. Significant differences were determined with a randomization test⁴⁹, corrected for multiple comparisons with false discovery rate (FDR)⁵⁰ and denoted by asterisks ($*P < 0.025$, $**P < 0.005$, $***P < 0.0005$). Note that the x-axes are log₁₀ and log₂ transformed, respectively.

Increased resistance of *clec* mutants correlates with escape behavior.

During our initial survival screen of the 39 *clec* mutants, we noticed for some of the *clec* mutants that a noticeable fraction was missing on BT at 24 hpi, the time point of scoring. As BT is known to induce behavioral pathogen avoidance responses in *C. elegans*³¹ and as increased pathogen avoidance behavior can lead to increased resistance to infection³², we examined if we could observe this phenomenon in the *clec* mutants of our screen. Therefore, we plotted the survival rate against the number of missing worms (“leaving index”) and indeed, increased survival was positively correlated with increased leaving

(Figure S 1B). Thus, worms, which showed avoidance towards the pathogen, enhanced their chances of coping with the pathogenic threat.

Pathogen avoidance behavior as natural strategy to escape infection has been described already for many other pathogens in *C. elegans* (e.g. ³³⁻³⁶). In an evolutionary context, evading potential harm is a highly efficient defense strategy found across the animal kingdom³⁷, which is not as costly as mounting a physiological immune response³⁸. Further, avoidance of contact with pathogens and thus infection also prevents tissue damage, which might arise not only from virulence mechanisms of the pathogen, but also from the host's own immune effector mechanisms, such as the production of reactive oxygen species (ROS)³⁹.

***C54G4.4* negatively regulates avoidance behavior towards pathogenic BT.**

Based on our observation that increased resistance was correlated with increased escape in our initial screen, we tested if the *clec* mutant with the strongest resistant phenotype on BT, *C54G4.4(ok2110)* (Figure 1A), showed high pathogen avoidance. The exposure to mild concentrations of pathogenic BT18247 caused the previously described pathogen avoidance behavior in both strains, the wild type and *C54G4.4(ok2110)* (for all comparisons $P < 0.001$ according to a generalized linear mixed model (GLMM)⁴⁰, Bonferroni corrected⁴¹). Moreover, the *C54G4.4(ok2110)* mutant exhibited a much stronger behavioral avoidance response on the pathogen compared to the wild type (Figure 2A). These results thus confirmed our initial observation that the resistant *C54G4.4(ok2110)* mutant shows high avoidance behavior on BT. In addition, we validated the increased survival phenotype of the *C54G4.4(ok2110)* mutant on BT (Figure 2B) which conforms to the observed correlation between increased resistance and escape behavior in the screened *clec* mutants (Figure S 1B). In addition to strong avoidance behavior, an enhanced physiological immune response could underlie the high resistance of the *C54G4.4* mutant on BT. To rule out the latter, we conducted another survival assay, in which all plates were fully covered with the bacterial lawn. This is a common approach to enforce an encounter of worms and pathogens and thus to analyze the contribution of the physiological immune response to resistance while worms cannot escape (e.g. ³²). However, the *C54G4.4(ok2110)* mutant still managed to escape and most worms desiccated at the rim of the Petri dish, so that the results were not reliable (data not shown). Thus, we cannot entirely exclude the involvement of physiological immune mechanisms in *C54G4.4*'s resistance to BT, but our data strongly supports the importance of the behavioral response.

We here present for the first time evidence for a role of a CTLD encoding gene, *C54G4.4*, in *C. elegans* pathogen avoidance behavior. The high avoidance behavior of the *C54G4.4* mutant on BT indicates that *C54G4.4* negatively regulates *C. elegans* behavioral immunity. This phenomenon is reminiscent of the microbial aversion behavior triggered in worms, in which cellular core processes are disrupted by RNAi⁴². However, this scenario can be ruled out here since the worms showed elevated avoidance only on the pathogen but not on the control (Figure 2A).

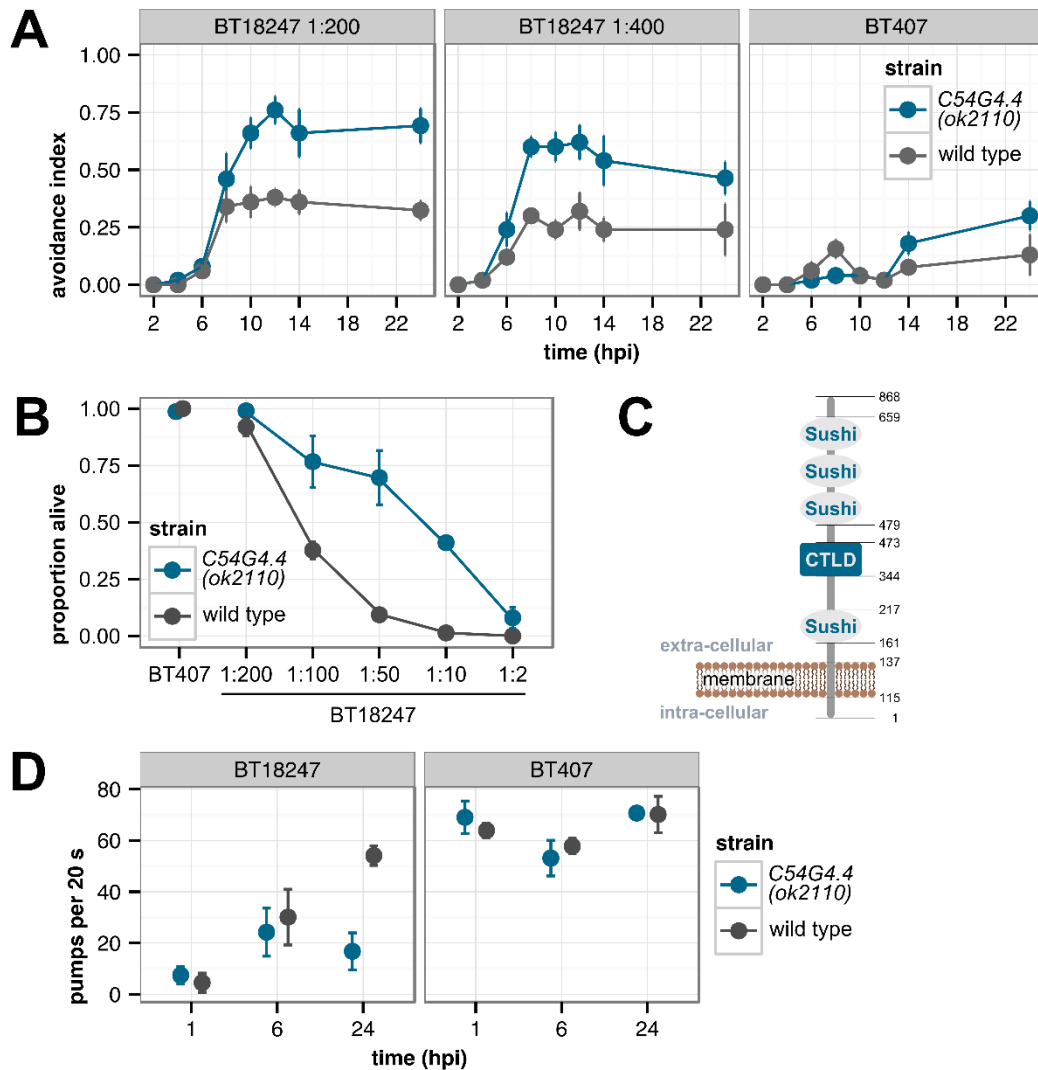


Figure 2. Immune behavior, survival phenotype and domain architecture of *C54G4.4(ok2110)* knock-out strain. (A) Avoidance behavior of *C54G4.4(ok2110)* compared to wild type over time. Each panel shows the applied bacterial treatment. Means \pm SE of five replicates are shown. Comparisons between both strains, $P < 0.001$ for each BT18247 treatment, generalized linear mixed model (GLMM)⁴⁰, Bonferroni corrected⁴¹. (B) Survival of *C54G4.4(ok2110)* and wild type on a 1:2 dilution of BT407 and serial dilutions of BT18247 24 hpi. Means \pm SE of three replicates are shown. Comparison between both strains, $P < 0.001$, generalized linear model (GLM)⁴⁸, Bonferroni corrected⁴¹. (C) Domain architecture adapted from SMART (<http://smart.embl-heidelberg.de/>) taking UniProt as source database. Numbers represent the amino acid position. (D) Pumping activity of *C54G4.4(ok2110)* and wild type on 1:100 dilutions of BT407 and BT18247 each mixed with *E. coli* OP50. Grinder movements (“pumps”) within 20 s of three individuals per replicate were scored. Mean \pm SE of three replicates are shown. Comparisons between both strains, $P = 0.006$ in BT18247 treatment at 24 hpi, GLM⁴⁸, Bonferroni corrected⁴¹.

C54G4.4 encodes a transmembrane CTLD protein and has, in addition to its CTLD, four extracellular Sushi domains (Figure 2C). Sushi domains are important protein-protein interaction domains occurring in proteins of the complement system, hence the alias complement control protein (CCP) module⁴³, and *C54G4.4* is orthologous to human complement factor H and complement component 4 binding protein.

According to expression data on WormBase (Version WB250), *C54G4.4* is expressed in neurons, where it could act to mediate behavioral immune responses. We generated transgenic reporter strains carrying transcriptional *C54G4.4::GFP* fusion constructs to analyze *C54G4.4* expression *in vivo*, but could unfortunately not detect any GFP signal, possibly due to low expression levels of *C54G4.4*. However, the *C. elegans* neuropeptide receptor gene *npr-1* is an example of a gene expressed in neurons, which mediates resistance to PA14 through oxygen-dependent behavioral avoidance of the pathogen³².

Which exact role *C54G4.4* plays in mediating pathogen susceptibility and regulating avoidance behavior to BT and in which immune cascade its action is implemented remains to be elucidated.

***C54G4.4* negatively regulates feeding cessation on pathogenic BT.**

Another behavioral defense strategy in *C. elegans* is feeding cessation. Harmful microbes evoke a reduction in food uptake as reported for both BT^{24,31} and PA14²⁷, and also other pathogens (e.g. ^{35,44,45}). Exposure of worms to *E. coli* expressing one of the main virulence factors of BT, the Cry toxin Cry6Aa2, is sufficient to trigger feeding cessation⁴⁶. A common approach to assess food up-take in *C. elegans* is to measure the movement of the grinder (“pumps”)⁴⁷. We therefore observed the pumping activity of *C54G4.4(ok2110)* on mild concentrations of BT over time. As expected, in both, the *C54G4.4(ok2110)* mutant and wild type worms, the pumping activity was lower on pathogenic BT18247 compared to the non-pathogenic control ($P < 0.001$ according to a GLM⁴⁸, Bonferroni corrected⁴¹). The wild type stopped feeding at an early time point after pathogen exposure (1 hpi), but resumed feeding at later time points (6 and 24 hpi). Interestingly, *C54G4.4(ok2110)* behaved similarly at 1 and 6 hpi, but then remained at a low feeding rate at 24 hpi (Figure 2D). This result suggests that not only increased avoidance behavior, but also prolonged feeding cessation contributes to the highly resistant survival phenotype of *C54G4.4(ok2110)* mutants.

Conclusions

With a systematic screen of 39 *clec* mutants for survival on the two pathogens BT18247 and PA14, we wanted to test the common assumption that *clec* genes function in *C. elegans* immunity. While our results indeed provide evidence for a role of *clec* genes in resistance, our data also suggest that *clec* genes function as negative regulators of immunity (Figure 1). We found that the resistance phenotype was positively correlated with the occurrence of missing worms, i.e. pathogen avoidance behavior (Figure S 1B). For one *clec* gene, *C54G4.4* (Figure 2), we could confirm that the high resistance of the *C54G4.4(ok2110)* mutant is likely due to a high avoidance behavior on BT and also a prolonged reduction in feeding on the pathogen. We thus provide for the first time evidence for a role of a *C. elegans clec* gene in regulating immune behaviors.

Acknowledgement

We thank Sabrina Butze and Christian Pawlitzki for technical assistance; Andrei Papkou for help with statistical analyses; the Evolutionary Ecology Genetics Department for valuable feedback; and Dennis Kim for providing PA14. All knock-out strains were provided by either the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440) or by the National Bioresource Project coordinated by S. Mitani, as indicated in “Supplementary Tables and Figures” Table S 1. BP, HS, and KD are supported by grants from the German Science foundation (Grant DI 1687/1-1 to K.D. and SCHU 1415/9-2 to H.S.). CB is member of the IMPRS for Evolutionary Biology. KD is additionally supported by institutional funding from Kiel University.

References – Chapter II

1. Weis, W. I. & Drickamer, K. Structural basis of lectin-carbohydrate recognition. *Annu. Rev. Biochem.* **65**, 441–473 (1996).
2. Weis, W. I., Taylor, M. E. & Drickamer, K. The C-type lectin superfamily in the immune system. *Immunol. Rev.* **163**, 19–34 (1998).
3. Drickamer, K. Two distinct classes of carbohydrate-recognition domains in animal lectins. *J. Biol. Chem.* **263**, 9557–9560 (1988).
4. Cummings, R. D. & McEver, R. P. in *Essentials of Glycobiology* (eds. Varki, A. et al.) (Cold Spring Harbor Laboratory Press, 2009).
5. Drickamer, K. C-type lectin-like domains. *Curr. Opin. Struct. Biol.* **9**, 585–590 (1999).
6. Dambuza, I. M. & Brown, G. D. C-type lectins in immunity: recent developments. *Curr. Opin. Immunol.* **32**, 21–27 (2015).
7. Ao, J., Ling, E. & Yu, X.-Q. *Drosophila* C-type lectins enhance cellular encapsulation. *Mol. Immunol.* **44**, 2541–2548 (2007).
8. Arensburger, P. et al. Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. *Science* **330**, 86–88 (2010).
9. Haq, S., Kubo, T., Kurata, S., Kobayashi, A. & Natori, S. Purification, characterization, and cDNA cloning of a galactose-specific C-type lectin from *Drosophila melanogaster*. *J. Biol. Chem.* **271**, 20213–20218 (1996).
10. Tanji, T., Ohashi-Kobayashi, A. & Natori, S. Participation of a galactose-specific C-type lectin in *Drosophila* immunity. *Biochem. J.* **396**, 127–138 (2006).
11. Multerer, K. A. & Smith, L. C. Two cDNAs from the purple sea urchin, *Strongylocentrotus purpuratus*, encoding mosaic proteins with domains found in factor H, factor I, and complement components C6 and C7. *Immunogenetics* **56**, 89–106 (2004).
12. Terwilliger, D. P., Clow, L. A., Gross, P. S. & Smith, L. C. Constitutive expression and alternative splicing of the exons encoding SCRs in Sp152, the sea urchin homologue of complement factor B. Implications on the evolution of the Bf/C2 gene family. *Immunogenetics* **56**, 531–543 (2004).
13. Pees, B., Yang, W., Zárate-Potes, A., Schulenburg, H. & Dierking, K. High innate immune specificity through diversified C-type lectin-like domain proteins in invertebrates. *J. Innate Immun.* (2015). doi:10.1159/000441475
14. Wang, X.-W. & Wang, J.-X. Diversity and multiple functions of lectins in shrimp immunity. *Dev. Comp. Immunol.* **39**, 27–38 (2013).
15. Wilson, R., Chen, C. & Ratcliffe, N. A. Innate immunity in insects: the role of multiple, endogenous serum lectins in the recognition of foreign invaders in the cockroach, *Blaberus discoidalis*. *J. Immunol.* **162**, 1590–1596 (1999).
16. Yu, X.-Q., Gan, H. & Kanost, M. Immulectin, an inducible C-type lectin from an insect, *Manduca sexta*, stimulates activation of plasma prophenol oxidase. *Insect Biochem. Mol. Biol.* **29**, 585–597 (1999).
17. Sahu, S. N. et al. Genomic analysis of immune response against *Vibrio cholerae* hemolysin in *Caenorhabditis elegans*. *PLoS ONE* **7**, e38200 (2012).
18. Simonsen, K. T. et al. Quantitative proteomics identifies ferritin in the innate immune response of *C. elegans*. *Virulence* **2**, 120–130 (2011).
19. Irazoqui, J. E. et al. Distinct pathogenesis and host responses during infection of *C. elegans* by *P. aeruginosa* and *S. aureus*. *PLoS Pathog.* **6**, e1000982 (2010).
20. Miltsch, S. M., Seeberger, P. H. & Lepenies, B. The C-type lectin-like domain containing proteins Clec-39 and Clec-49 are crucial for *Caenorhabditis elegans* immunity against *Serratia marcescens* infection. *Dev. Comp. Immunol.* **45**, 67–73 (2014).

21. Takeuchi, T. *et al.* A C-type lectin of *Caenorhabditis elegans*: Its sugar-binding property revealed by glycoconjugate microarray analysis. *Biochem. Biophys. Res. Commun.* **377**, 303–306 (2008).
22. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
23. The *C. elegans* Deletion Mutant Consortium. Large-scale screening for targeted knockouts in the *Caenorhabditis elegans* genome. *G3* **2**, 1415–1425 (2012).
24. Hasshoff, M., Böhnisch, C., Tonn, D., Hasert, B. & Schulenburg, H. The role of *Caenorhabditis elegans* insulin-like signaling in the behavioral avoidance of pathogenic *Bacillus thuringiensis*. *FASEB J.* **21**, 1801–1812 (2007).
25. Leyns, F., Borgonie, G., Arnaut, G. & Waele, D. D. Nematicidal activity of *Bacillus thuringiensis* isolates. *Fundam. Appl. Nematol.* **18**, 211–218 (1995).
26. Schulte, R. D., Makus, C., Hasert, B., Michiels, N. K. & Schulenburg, H. Multiple reciprocal adaptations and rapid genetic change upon experimental coevolution of an animal host and its microbial parasite. *Proc. Natl. Acad. Sci.* **107**, 7359–7364 (2010).
27. Tan, M.-W., Mahajan-Miklos, S. & Ausubel, F. M. Killing of *Caenorhabditis elegans* by *Pseudomonas aeruginosa* used to model mammalian bacterial pathogenesis. *Proc. Natl. Acad. Sci.* **96**, 715–720 (1999).
28. Powell, J. R. & Ausubel, F. M. in *Innate Immunity* 403–427 (Springer, 2008).
29. Marshall, A. S. J. *et al.* Identification and characterization of a novel human myeloid inhibitory C-type lectin-like receptor (MICL) that is predominantly expressed on granulocytes and monocytes. *J. Biol. Chem.* **279**, 14792–14802 (2004).
30. Neumann, K. *et al.* Clec12a is an inhibitory receptor for uric acid crystals that regulates inflammation in response to cell death. *Immunity* **40**, 389–399 (2014).
31. Schulenburg, H. & Müller, S. Natural variation in the response of *Caenorhabditis elegans* towards *Bacillus thuringiensis*. *Parasitology* **128**, 433–443 (2004).
32. Reddy, K. C., Andersen, E. C., Kruglyak, L. & Kim, D. H. A polymorphism in *npr-1* is a behavioral determinant of pathogen susceptibility in *C. elegans*. *Science* **323**, 382–384 (2009).
33. Pujol, N. *et al.* A reverse genetic analysis of components of the Toll signaling pathway in *Caenorhabditis elegans*. *Curr. Biol.* **11**, 809–821 (2001).
34. Schulenburg, H. & Ewbank, J. J. The genetics of pathogen avoidance in *Caenorhabditis elegans*. *Mol. Microbiol.* **66**, 563–570 (2007).
35. Sicard, M., Hering, S., Schulte, R., Gaudriault, S. & Schulenburg, H. The effect of *Photorhabdus luminescens* (Enterobacteriaceae) on the survival, development, reproduction and behaviour of *Caenorhabditis elegans* (Nematoda: Rhabditidae). *Environ. Microbiol.* **9**, 12–25 (2007).
36. Zhang, Y., Lu, H. & Bargmann, C. I. Pathogenic bacteria induce aversive olfactory learning in *Caenorhabditis elegans*. *Nature* **438**, 179–184 (2005).
37. Curtis, V., de Barra, M. & Aunger, R. Disgust as an adaptive system for disease avoidance behaviour. *Philos. Trans. R. Soc. B Biol. Sci.* **366**, 389–401 (2011).
38. Schmid-Hempel, P. & Ebert, D. On the evolutionary ecology of specific immune defence. *Trends Ecol. Evol.* **18**, 27–32 (2003).
39. Chavez, V., Mohri-Shiomi, A., Maadani, A., Vega, L. A. & Garsin, D. A. Oxidative stress enzymes are required for DAF-16-mediated immunity due to generation of reactive oxygen species by *Caenorhabditis elegans*. *Genetics* **176**, 1567–1577 (2007).
40. Bolker, B. M. *et al.* Generalized linear mixed models: A practical guide for ecology and evolution. *Trends Ecol. Evol.* **24**, 127–135 (2009).
41. Dunn, O. J. Multiple comparisons among means. *J. Am. Stat. Assoc.* **56**, 52 (1961).
42. Melo, J. A. & Ruvkun, G. Inactivation of conserved *C. elegans* genes engages pathogen- and xenobiotic-associated defenses. *Cell* **149**, 452–466 (2012).

43. Norman, D. G. *et al.* Three-dimensional structure of a complement control protein module in solution. *J. Mol. Biol.* **219**, 717–725 (1991).
44. Aballay, A., Yorgey, P. & Ausubel, F. M. *Salmonella typhimurium* proliferates and establishes a persistent infection in the intestine of *Caenorhabditis elegans*. *Curr. Biol.* **10**, 1539–1542 (2000).
45. O’Quinn, A. L., Wiegand, E. M. & Jeddloh, J. A. *Burkholderia pseudomallei* kills the nematode *Caenorhabditis elegans* using an endotoxin-mediated paralysis. *Cell. Microbiol.* **3**, 381–393 (2001).
46. Luo, H., Xiong, J., Zhou, Q., Xia, L. & Yu, Z. The effects of *Bacillus thuringiensis* Cry6A on the survival, growth, reproduction, locomotion, and behavioral response of *Caenorhabditis elegans*. *Appl. Microbiol. Biotechnol.* **97**, 10135–10142 (2013).
47. Raizen, D. Methods for measuring pharyngeal behaviors. *WormBook* 1–13 (2012). doi:10.1895/wormbook.1.154.1
48. Nelder, J. A. & Wedderburn, R. W. M. Generalized linear models. *J. R. Stat. Soc. Ser. Gen.* **135**, 370 (1972).
49. Whitlock, M. C. & Schluter, D. *The analysis of biological data*. (Roberts and Company Publishers, 2009).
50. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**, 289–300 (1995).
51. Wilcoxon, F. Individual comparisons by ranking methods. *Biom. Bull.* **1**, 80 (1945).

Supplementary Tables and Figures

Table S 1. Knock-out *clec* gene mutants provided by CGC^a and NBRP^b.

strain	affected gene	genotype	WormBase ID	source
	<i>clec-1</i>	<i>clec-1(tm1291)</i>	WBGene00017772	NBRP
RB1660	<i>clec-4</i>	<i>clec-4(ok2050)</i> II	WBGene00012583	CGC
	<i>clec-13</i>	<i>clec-13(tm4129)</i>	WBGene00010399	NBRP
	<i>clec-15</i>	<i>clec-15(tm3833)</i>	WBGene00011584	NBRP
RB2345	<i>clec-29</i>	<i>clec-29(ok3181)</i> V	WBGene00012024	CGC
RB1701	<i>clec-34</i>	<i>clec-34(ok2120)</i> V	WBGene00012023	CGC
	<i>clec-38</i>	<i>clec-38(tm2035)</i>	WBGene00012025	NBRP
RB1767	<i>clec-39</i>	<i>clec-39(ok2271)</i> V	WBGene00012022	CGC
RB2350	<i>clec-43</i>	<i>clec-43(ok3188)</i> II	WBGene00019917	CGC
RB1870	<i>clec-49</i>	<i>clec-49(ok2416)</i> V	WBGene00012251	CGC
RB1897	<i>clec-50^c</i>	<i>clec-50(ok2455)</i> V	WBGene00012253	CGC
RB1898	<i>clec-50^d</i>	<i>clec-50(ok2456)</i> V	WBGene00012253	CGC
	<i>clec-60</i>	<i>clec-60(tm2319)</i>	WBGene00014046	NBRP
RB2004	<i>clec-63</i>	<i>clec-63(ok2656)</i> II	WBGene00009394	CGC
RB2121	<i>clec-64</i>	<i>F35C5.7(ok2806)</i> II	WBGene00009395	CGC
RB1803	<i>clec-65</i>	<i>clec-65(ok2337)</i> II	WBGene00009396	CGC
RB1745	<i>clec-66</i>	<i>clec-66(ok2230)</i> II	WBGene00009397	CGC
RB2095	<i>clec-67</i>	<i>F56D6.2(ok2770)</i> IV	WBGene00018971	CGC
VC1954	<i>clec-68/clec-69</i>	<i>clec-68;</i> <i>clec-69(ok2549)</i> IV	WBGene00018970/ WBGene00044642	CGC
RB1663	<i>clec-69/clec-70</i>	<i>clec-69;</i> <i>clec-70(ok2061)</i> IV	WBGene00044642/ WBGene00021581	CGC
	<i>clec-90</i>	<i>clec-90(tm3402)</i>	WBGene00019009	NBRP
VC1600	<i>clec-98</i>	<i>ZK39.7(ok2085)</i> I	WBGene00013932	CGC
RB1403	<i>clec-104</i>	<i>Y18D10A.10(ok1596)</i> I	WBGene00012480	CGC
RB2300	<i>clec-106</i>	<i>Y18D10A.12(ok3123)</i> I	WBGene00012482	CGC
VC2345	<i>clec-106</i>	<i>clec-106;</i> <i>Y18D10A.23(ok3090)</i> I	WBGene00012482	CGC
	<i>clec-108</i>	<i>clec-108(tm1622)</i>	WBGene00012504	NBRP
RB1761	<i>clec-151</i>	<i>F10F2.7(ok2264)</i> III	WBGene00008659	CGC
VC1818	<i>clec-155</i>	<i>T04A8.3;</i> <i>tag-243(ok1855)</i> III	WBGene00011405	CGC
RB2217	<i>clec-175</i>	<i>R08C7.6(ok2999)</i> IV	WBGene00019950	CGC
VC2919	<i>clec-198</i>	<i>clec-198(gk3149)</i> IV; <i>W03G11.3(gk1251)</i> X	WBGene00008203	CGC
RB1984	<i>clec-227</i>	<i>clec-227(ok2618)</i> V	WBGene00008593	CGC
RB2308	<i>clec-230^e</i>	<i>C29F3.5(ok3131)</i> V	WBGene00007806	CGC
RB2326	<i>clec-230^f</i>	<i>clec-230(ok3158)</i> V	WBGene00007806	CGC
VC2185	<i>C06B8.7</i>	<i>C06B8.7(ok2814)</i> V	WBGene00007372	CGC
RB1697	<i>C54G4.4</i>	<i>C54G4.4(ok2110)</i> I	WBGene00008314	CGC
VC2172	<i>svh-1</i>	<i>C07G1.2(ok2531)</i> IV	WBGene00006620	CGC
UR116	<i>cwp-5</i>	<i>him-5(e1490)</i> V; <i>cwp-5(tm1893)</i> X	WBGene00009844	CGC
	<i>egas-4</i>	<i>F55G1.13(tm4826)</i>	WBGene00018906	NBRP
VC158	<i>lat-2</i>	<i>lat-2(ok301)</i> II	WBGene00002252	CGC
	<i>T25C12.3</i>	<i>T25C12.3(tm4691)</i>	WBGene00012018	NBRP

^a Caenorhabditis Genetics Center, Minnesota, USA.

^b National BioResource Project, Tokyo, Japan.

^c Denoted as *clec-50_A* in Figure 1.

^d Denoted as *clec-50_B* in Figure 1.

^e Denoted as *clec-230_A* in Figure 1.

^f Denoted as *clec-230_B* in Figure 1.

Table S 2. *P* values of *clec* gene mutant screen on BT18247 according to randomization test (Figure 1 A).

mutant strain	BT conc ^a	<i>P</i>	corrected <i>P</i>	validated ^b	remark
<i>clec-15</i>	1:50	0.0027	0.0222	yes	
<i>clec-29</i>	1:50	< 0.0005	0.0022	yes	confirmed after out-crossing
<i>clec-34</i>	1:50	< 0.0005	0.0039	yes	confirmed after out-crossing
<i>clec-38</i>	1:50	0.0015	0.0156	yes	
<i>clec-43</i>	1:50	< 0.0005	0.0022	yes	lost phenotype after out-crossing
	1:100	0.0027	0.0222		
<i>clec-50_B</i>	1:50	< 0.0005	< 0.0005	no	
<i>clec-64</i>	1:50	< 0.0005	0.0062	yes	
<i>clec-90</i>	1:50	< 0.0005	0.0022	yes	
	1:100	0.0006	0.0085		
<i>clec-104</i>	1:50	< 0.0005	0.0052	yes	
<i>clec-108</i>	1:50	< 0.0005	< 0.0005	yes	lost phenotype after out-crossing
<i>clec-151</i>	1:50	< 0.0005	< 0.0005	yes	confirmed after out-crossing
<i>clec-198</i>	1:50	0.0021	0.0193	yes	
	1:100	0.0018	0.0176		
<i>clec-227</i>	1:50	0.0011	0.0132	yes	lost phenotype after out-crossing
<i>clec-230_A</i>	1:50	0.0015	0.0156	yes	
<i>C54G4.4</i>	1:50	< 0.0005	< 0.0005	yes	validated in Figure 2B
<i>T25C12.3</i>	1:50	0.0010	0.0130	yes	lost phenotype after out-crossing

^a conc, concentration.

^b Mutant strains, which included in validation of initial screen (Figure S 1A).

Table S 3. *P* values of *clec* gene mutant screen on PA14 according to randomization test (Figure 1B).

mutant strain	<i>P</i>	corrected <i>P</i>	validated ^a	remark
<i>clec-34</i>	0.0023	0.0167	no	
<i>clec-39</i>	0.0010	0.0123	yes	lost phenotype after out-crossing
<i>clec-43</i>	< 0.0005	0.0064	yes	lost phenotype after out-crossing
<i>clec-50_A</i>	< 0.0005	0.0063	no	
<i>clec-60</i>	0.0023	0.0167	no	
<i>clec-90</i>	< 0.0005	0.0063	yes	susceptibility also on OP50 control
<i>clec-151</i>	< 0.0005	0.0022	yes	
<i>clec-155</i>	< 0.0005	0.0063	yes	
<i>cwp-5</i>	0.0018	0.0163	no	
<i>T25C12.3</i>	0.0018	0.0163	no	

^a Mutant strains, which were included in validation of initial screen (Figure S 1A).

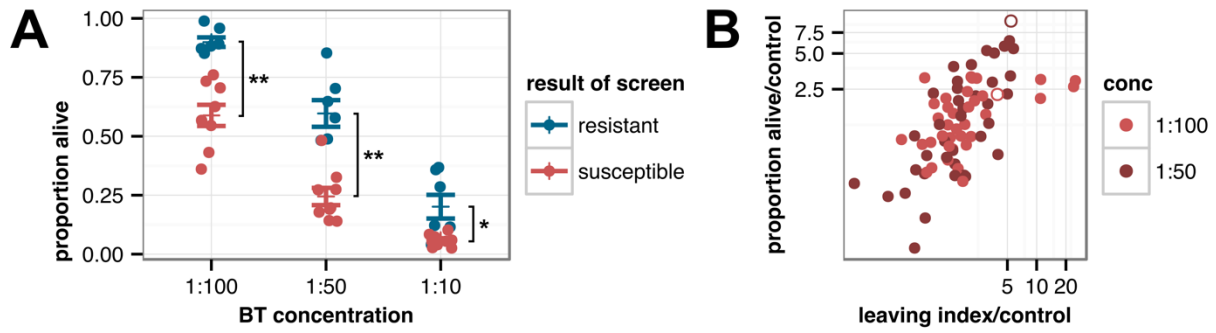


Figure S 1. Validation of mutant screen on BT and positive Pearson's correlation between survival and avoidance behavior. (A) *clec* mutants that showed significant alterations in survival – in the first screen on BT (Figure 1A) were retested on three different concentrations of BT mixed with *E. coli* OP50. Each dot represents the resulting mean of five replicates colored according to their survival phenotype in the first screen. The error bars represent the means of group “resistant” or “susceptible” \pm SE per BT concentration. Wilcoxon rank sum test⁵¹ between survival phenotype groups per BT concentration, FDR corrected⁵⁰; $P = 0.001$ for 1:100, $P = 0.003$ for 1:50, $P = 0.01$ for 1:10. (B) Correlated are the means of proportion alive as shown in Figure 1 and the means of the leaving index ((initial total number of worms - present worms 24 hpi)/initial total number of worms), both normalized by the control values of N2. The two different shades of red represent different concentrations (conc) of BT mixed with *E. coli* OP50; the open circles represent data of the *C54G4.4(ok2110)* mutant. Note that both axes are log2 transformed. Spearman's rank correlation⁴⁹; $\rho = 0.507$ for 1:100, $\rho = 0.726$ for 1:50; $P < 0.001$ for each concentration.

Chapter III

Put to the Test: Immune Function of *C. elegans* C-type Lectin-like Domain Gene *clec-4* and its Paralogs

Barbara Pees¹, Anke Kloock¹, Wentao Yang¹, Hinrich Schulenburg¹, Katja Dierking^{1†}

¹ Department of Evolutionary Ecology and Genetics, Zoological Institute, Christian Albrechts University of Kiel, Kiel, Germany

† Corresponding author: kdierking@zoologie.uni-kiel.de

Abstract

Despite the increasing understanding of the innate immune system of the nematode *Caenorhabditis elegans* certain aspects of its immunity, such as pathogen recognition and clearance still need to be determined. The study of the highly diverse *C. elegans* gene family encoding C-type lectin-like domain (CTLN) proteins (*clec* genes) has the potential to close these knowledge gaps, as *clec* genes might function in both pathogen recognition and clearance. A role of *C. elegans* *clec* genes in immunity has indeed been repeatedly suggested and in some cases even been taken for granted, although their exact function is completely unknown. Using WormExp, a recently established database for *C. elegans* gene expression analysis, we analyzed microbe- and stress-induced expression of *clec* genes and identified *clec-4* as being highly responsive to infection with different pathogens and exposure to several stressors. We selected the *clec-4* gene for subsequent functional analyses at the gene and protein level. Contrary to our expectations, *clec-4(ok2050)* mutants were more resistant to infection with pathogenic *Bacillus thuringiensis* than wild type worms and did not show a survival phenotype on *Pseudomonas aeruginosa*. By looking at expression of *clec-4* paralogs, we identified *clec-41*, which is mainly co-expressed with *clec-4*, and *clec-10*, which is differentially regulated in the opposite direction from *clec-4*. Silencing *clec-41* and *clec-10* expression by RNAi led to increased susceptibility to BT infection in wild type worms and suppression of the resistance phenotype in *clec-4* mutants. On the protein level, we tested binding of the recombinantly expressed CLEC-4 protein to different bacterial species and one yeast species and found that CLEC-4 binds to a pathogenic and non-pathogenic BT as well as yeast in a calcium-dependent manner. Together, our data suggest that *clec-4* functions in the negative regulation of the *C. elegans* immune response to BT infection and genetically interacts with its paralogs *clec-41* and *clec-10*. The binding activity of CLEC-4 towards BT reinforces the notion of *clec-4* playing a role in host-microbe interactions.

Introduction

Like all invertebrates the nematode *Caenorhabditis elegans* solely relies on innate immunity. Yet, highly sophisticated physiological mechanisms aid the worm to persist in its microbe-rich environment, which is full of potential pathogens¹. These mechanisms can generally be divided into three modules: (i) Direct and indirect recognition of pathogens and initiation of downstream signaling is mediated by pattern recognition receptors (PRRs), of which, however, only two have been discovered in *C. elegans* – the RIG-1-like helicase DHR-1 recognizes viral RNA and triggers the RNAi cascade², and the G protein-coupled receptor (GPCR) DCAR-1 binds the endogenous, putative damage-associated molecular pattern (DAMP) 4-hydroxyphenyllactic acid (HPLA) and activates expression of the antimicrobial peptide *nlp-29* in the epidermis³. (ii) Transmission of the immune response by evolutionary conserved signaling pathways, with the p38 MAPK cascade being one of the central pathways, acting in both the epidermis and the intestine^{4,5}. (iii) Finally the expression of antimicrobial peptides such as lysozymes (*lys*, *ilys*), caenopores (*spp*), or members of the neuropeptide-like protein (*nlp*) family (reviewed in ^{6,7}) and elimination of the pathogen. Nevertheless, one of the major knowledge gaps in our understanding of *C. elegans* immunity concerns the question of how the host discriminates between bacterial pathogens and commensals or food bacteria and what are the exact triggers inducing the immune response.

The study of one protein family in *C. elegans* could help to shed light on these constantly recurring questions – the C-type lectin-like domain (CTLTD) proteins. Due to their essential, well described role in recognition and activation of the vertebrate immune response to fungal infection, *C. elegans* CTLTD proteins have repeatedly been considered to be involved in immunity, although experimental evidence supporting this idea remains scarce. The CTLTD proteins are characterized by a conserved carbohydrate-recognition domain (CRD), with which sugar can be bound in a calcium-dependent (C-type) manner. But not all of the proteins carrying a CRD bind glycans and/or require calcium for binding, which is why we will use the more general term CTLTD proteins in the following, as previously suggested⁸. In vertebrates CTLTD proteins occupy important immune functions, acting mainly as classical PRRs binding ligands derived from fungi, bacteria or viruses, and as dead and cancerous cell sensors enhancing tumor killing activities of natural killer cells (reviewed in ⁹). Some CTLTD proteins that seem to play a role in immune defense have also been characterized in invertebrates, especially in economically important crustaceans and in insects. Yet, their exact immune functions are far from being completely understood (reviewed in ¹⁰).

C. elegans stands out for its highly diversified CTLTD encoding gene (*clec*) repertoire (reviewed in ¹⁰). 283 CTLTD proteins are encoded in the nematode's genome with half of them carrying an additional CTLTD or other domains such as CUB (Complement C1r/C1s Uegf Bmp1), CW (conserved cysteine and tryptophan residues), or VWA (Von Willebrand factor type A)⁷. Several *C. elegans* transcriptome analyses revealed that the majority of the *clec* genes are highly up-regulated upon pathogen exposure

and this (e.g. ¹¹⁻¹³) in a highly specific pattern⁷. By reverse genetic analyses some *clec* genes were shown to be required for defense against infection with certain pathogens, e.g. *clec-17*, *-60*, and *-86* for *Microbacterium nematophilum*¹⁴, *clec-70* for *Staphylococcus aureus*¹⁵, *clec-65* for the pathogenic *Escherichia coli* strain LF82¹⁶, or *clec-174* for *Vibrio cholerae*¹⁷. On the protein level, only two CLEC proteins were demonstrated to bind bacteria *in vitro*: CLEC-39 and CLEC-40 bind to dead and alive *Serratia marcescens*¹⁸.

In this study we thus aimed at understanding in how far pathogen-responsive *clec* genes really function in *C. elegans* immunity. Thus, we employed a recently established database for *C. elegans* gene expression analysis, WormExp¹⁹, and analyzed the expression of *clec* genes after exposure to pathogens and stress. We found one *clec* gene, *clec-4*, to be highly up-regulated upon infection with many pathogens and exposure to several stressors. Here we show that the *clec-4(ok2050)* mutant is resistant to infection with the Gram-positive *Bacillus thuringiensis* (BT), suggesting a role of *clec-4* in immune suppression. Further, we attempt to disentangle the much discussed issue of functional redundancy within the *clec* family by analyzing *clec-4* interaction with two of its paralogs, *clec-10* and *clec-41*. Knocking-down *clec-41* or *clec-10* expression by RNAi results in higher susceptibility to BT in wild type worms and suppresses the resistance phenotype in *clec-4* mutants, suggesting opposing roles in modulation of immunity. Moreover and most importantly, we started to examine the function of CLEC-4 on the protein level by testing the binding activity of the recombinant protein *in vitro*.

To summarize, these results suggest a negative regulatory role of *clec-4* in immunity against BT and give a glimpse at a complex regulation network involving different *clec* genes. Further studies are needed for a more detailed understanding of *clec-4*/CLEC-4 function on both the gene and protein level.

Materials and Methods

Meta-Analysis of Transcriptional Response

The information on *clec-4* paralogs were downloaded from WormBase Version WS250 (<http://www.wormbase.org>)²⁰, gene expression data sets from the two categories “Microbes” (pathogens) and “Chemicals/stress” were analyzed using WormExp (<http://wormexp.zoologie.uni-kiel.de/wormexp/>)¹⁹. The heat map only shows the data sets, in which *clec-4* was differentially expressed (Figure 1A).

Strains and Culture Conditions

Worms were grown and maintained on NGM agar plates seeded with *Escherichia coli* OP50 as previously described²¹. The wild type strain N2 (Bristol) and the mutant strain RB1660 carrying the

clec-4(2050) allele were ordered at the Caenorhabditis Genetics Center (CGC, Minnesota, USA), the latter back-crossed 3x with N2 prior to use.

Bacillus thuringiensis strains MY-BT18247 (in the following BT18247) and BT407 were originally obtained from the Agriculture Research Service Patent Culture Collection of Microorganisms and Cell Cultures (United States Department of Agriculture, Peoria, Illinois, USA). Spore-toxin mixtures were generated following previous protocols²²⁻²⁴ and frozen at -20 °C in aliquots with a spore concentration ranging from 3×10^9 to 8×10^9 particles/ml, depending on the culture, for BT18247 and 1.1×10^{10} particles/ml for BT407. Stocks were thawed and then immediately applied in survival/avoidance assays. *Pseudomonas aeruginosa* strain PA14 (provided by Dennis Kim) was grown first on LB plates and then in LB broth at 37 °C overnight prior to inoculation of assay plates.

Generation of Transgenic Strains

The *pclec-4::GFP* reporter construct was generated by PCR fusion as previously described²⁵. The *clec-4* promoter (~1.2 kb upstream of start codon) was amplified from genomic DNA by PCR with primer A (5'-ACCAGAAGCCGAGAATCCAC-3') and primer B (5'-**AGTCGACCTGCAGGCATGCAAGCTGGCTCTTTTATT** CGCAATTTCCAT-3', with bold letters being the overlapping sequence to the *gfp* vector). The *gfp* coding sequence plus the 3'-UTR of *unc-54* was amplified from the Fire vector pPD95.75 with primer C (5'-AGCTTGCATGCCTGCAGGTCGACT-3') and D (5'-AAGGGCCCGTACGGCCGACTAGTAGG-3'). The transgenic construct *pclec-4::GFP* was finally synthesized using PCR fusion with primer A* (5'-TTTGAAATTCCACCACTGTTCCAC-3') and D* (GGAAACAGTTATGTTTGGTATATTGGG) and directly injected at a concentration of 10 ng/μl. The co-injection marker *pttx-3::RFP*, which is expressed in the AIY interneuron pair of successfully transformed animals, was used at a concentration of 40 ng/μl. Of four generated transgenic lines, MY1031 *yaEx26* was further analyzed by microscopy. All pictures were taken with the confocal microscope LSM 700 by Zeiss (Carl Zeiss AG, Jena, Germany).

Survival and Avoidance Assays

PFM agar plates (6 cm diameter) were inoculated with 75 μl of a mixture of BT and *E. coli* OP50 at OD 5 at different concentrations. The survival plates were kept at 20 °C overnight. PA14 required enriched NGM (0.35% instead of 0.25 % peptone) for efficient killing of worms. Seeded with a mixture of PA14 (OD 1) and *E. coli* OP50 (OD 5) at a concentration of 1:3 plates were incubated for 24 h at 37 °C and another 24 h at 25 °C, the temperature, at which the PA14 survival assay was conducted. Plates seeded solely with *E. coli* OP50 or non-pathogenic BT407 were used as controls in BT survival assays, and *E. coli* OP50 plates in survival assays with PA14. 30 synchronized L4 larvae were picked onto each plate and worm survival was scored once after 24 h (BT) or every day until all worms on the pathogen were dead (PA14). Worms were considered to be dead if they did not respond to light touch.

For scoring the avoidance behavior the appropriate media plates (9 cm diameter) were prepared as described above for the survival assays but with a 30 μ l spot of the bacterial mixture in the middle. For BT exposure mild concentrations were chosen in order to challenge but not to kill the worms. Ten previously synchronized L4 hermaphrodites were picked onto each bacterial spot in the middle of the plate at time point 0 hpi (hours post infection) and the worms residing on that spot were scored every second hour. The leaving index was calculated as follows: (total number of worms - worms on bacterial spot)/total number of worms. Dead individuals were excluded from the total number of worms per plate.

Protein Binding Assays

Protein CLEC-4 with an N-terminal His tag was commercially obtained from GenScript (<http://www.genscript.com/>, Piscataway, New Jersey, USA); it was expressed in *E. coli* BL21 (DE3) with vector E3, purified, refolded and confirmed by SDS-PAGE and Western blot using Mouse-anti-His mAb. The recombinant protein was stored at -80 °C in TBS buffer (50 mM Tris-HCl, 150 mM NaCl, pH 7.6) with 10% glycerol. Bacteria were grown at 37 °C overnight in LB (yeast in YPD), pelleted, washed, and resuspended in TBS buffer with or without CaCl₂ (10 mM final concentration) at OD 2. The bacterial solution with 10 μ g of the recombinant CLEC-4 was incubated in 50 μ l of total volume for 30 min at room temperature. After taking out 16 μ l “total” sample (T) the reaction mixtures were centrifuged at 6000 g at 4 °C for 5 min, the “supernatant” sample (S) was taken off, and the remaining “pellet” sample (P) resuspended in 16 μ l fresh TBS buffer. All samples were boiled for 5 min mixed with Laemmli sample buffer²⁶ and subjected to 12% SDS-PAGE including subsequent analysis by Coomassie Brilliant Blue (CBB) staining.

Statistical Analyses

Statistical analyses were done with RStudio (Version 0.98.977), graphs created with its package ggplot2 (Version 0.9.3.1) and edited with Inkscape (Version 0.48).

Results and Discussion

Expression of *clec-4* is highly up-regulated upon pathogen exposure and stress.

One of the main reasons why *C. elegans clec* genes have repeatedly been suggested to be involved in immune responses is that *clec* genes are always among the genes that are highly up-regulated by pathogen infection and also stress. The repertoire of induced *clec* genes differs from pathogen to pathogen, suggesting a highly specific regulation (reviewed in ¹⁰). However, the expression of few *clec* genes is activated by infection with several pathogens, indicating a more general role in *C. elegans*

immunity. As we were interested in testing the potential immune function of *C. elegans clec* genes, we performed a transcriptome meta-analysis using WormExp, a recently established database for *C. elegans* gene expression analyses¹⁹. We looked at the pathogen (microbe)- and stress-dependent expression pattern of *clec* genes and identified *clec-4* as one of the *clec* genes, which is highly up-regulated upon infection with a broad array of pathogens and several stressors (Figure 1A).

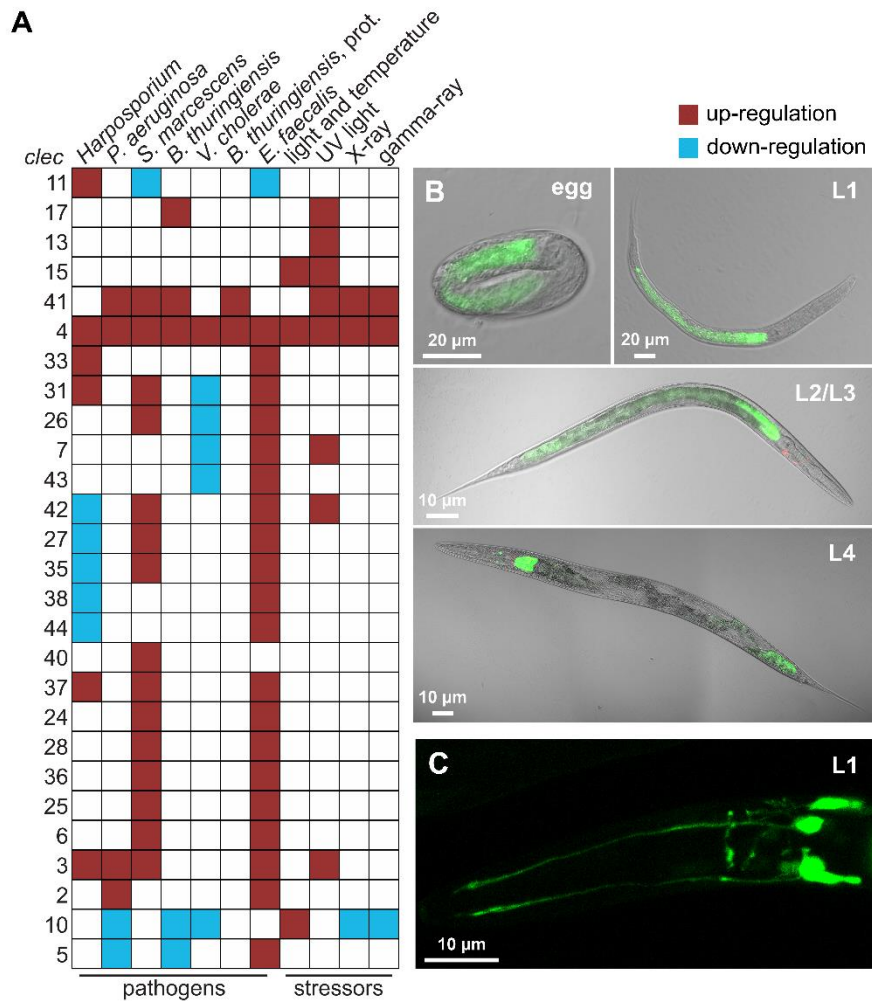


Figure 1. Pathogen- and stress-dependent expression pattern of *clec-4* and a subset of its paralogs and *in vivo* expression of a *pclec-4::GFP* reporter. (A) Heat map showing the differential expression of *clec-4* and its paralogs in worms exposed to different pathogens and stressors. Note that one BT column is based on proteome data (prot.). Twelve *clec-4* paralogs, which did not exhibit a differential expression have been excluded. Genes are organized by hierarchical clustering from Cluster 3.0³⁰. Red and blue colors indicate up- and down-regulation, respectively. Transcriptomic data taken from ^{12,17,27,28,31–34} were analyzed using WormExp¹⁹. (B) The *pclec-4::GFP* reporter is constitutively expressed throughout the intestine in the embryo and at all larval stages and most strongly in int1 and the posterior intestine in adults. The co-injection marker *pttx-3::RFP* is expressed in the AIY interneuron pair. (C) Expression of *pclec-4::GFP* in the amphid neurons and amphid nerves in the head of a L1 larva.

Interestingly, in the case of infection with the Gram-positive *Bacillus thuringiensis* Treitz *et al.* could demonstrate in a quantitative proteome analysis that the abundance of the CLEC-4 protein indeed highly increased (highest observed protein ratio)^{27,28}. In addition to that, *clec-4* up-regulation after infection with PA14 was confirmed by qRT-PCR²⁹. The broad transcriptional response to a variety of pathogens and stressors suggests that *clec-4* is part of a more general defense response. Taken all these observations together, *clec-4* represents a promising *clec* candidate gene to analyze in more detail.

A *pclec-4::GFP* reporter is constitutively expressed in the intestine.

In order to visualize the expression pattern of *clec-4* *in vivo* we generated a transcriptional reporter fusing the *clec-4* promoter to the green fluorescent protein (GFP). The reporter showed highly constitutive expression of *pclec-4::GFP* in the intestine throughout all stages. Expression seems to be decreased in adults, in which a strong GFP signal could only be observed in the most anterior and posterior cells of the intestine (Figure 1B). Intestinal expression of several other *C. elegans* *clec* genes has been reported before (^{14,15,35} and unpublished data from our group). Given the fact that most bacterial pathogens are taken up orally by the worm and establish intestinal infections, the intestinal expression of *clec-4* is in line with its potential immune function. In contrast to the analyzed transcriptome data, an infection with pathogenic BT or PA14 did not enhance the expression of our transcriptional reporter. This might be due to the fact that the constitutive expression of *pclec-4::GFP* is already so strong that a further up-regulation of the fluorescent reporter is not discernable by eye. An analysis of the available *C. elegans* transcriptome data indeed showed that *clec-4* is expressed also on non-pathogenic bacteria (data not shown).

***pclec-4::GFP* is expressed in amphid neurons, but *clec-4* is not required for pathogen avoidance behavior.**

Intriguingly, *clec-4* is also constitutively expressed in the amphid neurons from the first larval stage to the adult (Figure 1C). Neuronal expression of putative antimicrobial genes in the head has been reported in *C. elegans* before, e.g. for lysozymes *lys-1*, *lys-2* and *lys-8*³⁶, as well as for genes encoding antimicrobially active caenopores, *spp-3* and *spp-12*^{37,38}. These are interesting observations, as the *C. elegans* amphid neurons are chemosensory and thermosensory neurons with openings to the exterior that are generally known to be implemented in detecting microbial cues (important for food foraging and lawn-leaving behavior) or mating behavior, but also pathogen avoidance behavior³⁹. It is thus intriguing to speculate that *clec-4* functions in amphid neurons to bind bacterial components and mediate avoidance behavior. Consequently, we tested the *clec-4(ok2050)* mutant for its avoidance behavior on BT, but could not detect any difference to the behavior of wild type worms (Figure S 1). Still, *clec-4* might mediate the physiological immune response via the neuronal network (see below).

clec-4 knock-out confers resistance to BT.

Expression of *clec-4* is highly up-regulated by infection with Gram-negative PA14 and Gram-positive BT. For the latter this up-regulation on the transcriptional level was confirmed with higher CLEC-4 abundance on the protein level²⁸. We thus chose these two pathogens and did functional genetic analyses to test if *clec-4* is indeed required for *C. elegans* resistance to infection. While the *clec-4(ok2050)* mutant was more resistant than wild type worms to BT infection (Figure 2A), there was no difference in survival on PA14 (Figure S 2). These findings were unexpected, since the strong induction of *clec-4* gene expression upon pathogen exposure indicates an important role in pathogen defense, as has been recurrently suggested or taken for granted in the literature.

The unexpected resistance phenotype of *clec-4(ok2050)* might have four alternative explanations: (i) *clec-4* might be required for cellular core processes that are strictly monitored. If these processes are disturbed, i.e. lack of one component as in the knock-out mutant, surveillance pathways up-regulate general physiological and behavioral immune responses⁴⁰. (ii) *clec-4* might mediate susceptibility to infection through negatively regulating pathogen avoidance behavior, as was shown for another *C. elegans* *clec* gene, *C54G4.4* (Pees *et al.*, unpublished)⁴¹. (iii) *clec-4* might encode a host factor, which is required for BT to successfully infect. For instance, binding of certain BT toxins to intestinal cells depends on *C. elegans* glycolipid receptors⁴². (iv) *clec-4* might modulate immune defenses negatively, e.g. through down-regulation of immune genes. We can exclude the first two explanations as the *clec-4(ok2050)* mutants (i) exhibited normal (and not enhanced) survival and avoidance behavior on non-pathogenic BT407 compared to wild type worms, and (ii) did not show increased pathogen avoidance behavior on pathogenic BT (Figure S 1). Although we cannot exclude that *clec-4* encodes a host factor required for successful infection, we hypothesize that *clec-4* plays a role in suppressing the immune system. As *clec-4* is also expressed in amphid neurons, it could play a role in the neuronal regulation of the *C. elegans* immune response, as it was shown for the G protein-coupled receptor gene *octr-1* or the insulin-like peptide INS-7. The octopamine receptor OCTR-1 is expressed in the amphid neurons and while it is not required for pathogen avoidance behavior, it suppresses the physiological immune defense towards PA14 in non-neuronal tissues⁴³. The insulin-like peptide INS-7 is expressed in neurons, but modulates the immune response to infection with PA14 via activation of the insulin/insulin-like growth factor receptor (ILR) DAF-2 in the intestine⁴⁴. In mammals, one CTLD protein that act as immune repressor has also been reported: Clec12a binds uric acid crystals and suppresses the respiratory burst thus preventing hyper-inflammation in mice⁴⁵. This inhibitory CTLD protein is a transmembrane receptor with an intracellular ITIM (immunoreceptor tyrosine-based inhibitory motif) domain that mediates negative regulation of immune signaling and thereby counteracts activatory receptors. *clec-4*, however, lacks such a motif and contains a signal peptide and is therefore predicted to be secreted,

suggesting a different mechanism in negative regulation of immunity. The exact role and mechanism of *clec-4* action in this context remains to be further elucidated.

The lack of a survival phenotype of the *clec-4* mutant on PA14 is somewhat less surprising as Miltsch *et al.* showed that the *clec-4(ok2050)* mutant was as resistant as wild type worms on *Serratia marcescens*¹⁸. Absence of a phenotype might be due to genetic or functional redundancy between the numerous *clec* genes in *C. elegans* (see further below).

***clec-41* and *clec-10* RNAi suppress *clec-4* resistance phenotype.**

The *C. elegans* genome consists of a remarkable repertoire of protein-encoding gene families situated in clusters of duplicated genes, one family being the *clec* genes⁴⁶. The duplication rate of *C. elegans* exceeds the one of comparable model organisms, such as *Drosophila melanogaster* tenfold⁴⁷. Consequently, almost one third of the genome encodes genes with one or more paralogs. These clustered genes or genetic duplicates might function analogously to the original gene⁴⁸.

As *clec-4* is located directly upstream of its paralog *clec-8*, we wanted to test if there is functional redundancy between the two genes. We thus knocked down *clec-8* expression by RNAi in the loss-of-function mutant *clec-4(ok2050)* and again tested survival on BT. While the *clec-4* mutant was clearly more resistant to BT infection than the wild type, hence confirming our previous results, silencing *clec-8* expression did not have an effect on survival, neither in the *clec-4* mutant, nor in wild type worms. (Figure 2B). This indicates that *clec-8* is neither required for, nor does it function redundantly to *clec-4* in the defense against BT.

In addition to *clec-8*, *clec-4* has 38 paralogs in *C. elegans* (WormBase Version WS250). Since paralogs of genes are considered to possess overlapping functions⁴⁹ and since co-expressed genes are predicted to have similar function or to be involved in the same cellular process, we examined the pathogen- and stress-induced expression of these 39 *clec-4* paralogs to identify co-expressed genes (Figure 1A). The expression of the gene *clec-41* was induced like *clec-4* in seven of eleven data sets, in case of BT infection at both the transcriptome and proteome level^{27,28}. In contrast, the expression of *clec-10* was regulated in the opposite way to *clec-4* in five of the eleven data sets.

To analyze genetic interaction between *clec-4* and *clec-41* and *clec-10*, we did a BT survival assay using RNAi to silence *clec-41* or *clec-10* expression in the *clec-4* mutant and wild type worms. In contrast to the resistance phenotype of the *clec-4(ok2050)* mutant, *clec-41* RNAi-treated worms were highly susceptible to BT infection. Unexpectedly, silencing *clec-41* expression in *clec-4(ok2050)* mutants suppressed the resistance phenotype of *clec-4(ok2050)* to the level of wild type worm survival (Figure 2C). This result shows that *clec-4* and *clec-41* interact genetically with opposing effects on the immune response to BT infection. Comparable results were reported e.g. for the p38 MAPK/JNK pathway cross-talk. The MAPKK MEK-1 is required for a full activation of the p38 MAPK PMK-1,

while the MAPK phosphatase VHP-1 negatively regulates PMK-1 leading to susceptibility of *mek-1(ks54)* mutants and resistance of *vhp-1* RNAi-treated worms on PA14. *mek-1(ks54);vhp-1*(RNAi) worms exhibit the same survival phenotype as wild type worms⁵⁰. Considering this as an example for a scenario, in which two genes play opposing roles in regulating the immune response and taking the transcriptome/proteome data into account, we hypothesize that *clec-41* acts as a positive, while *clec-4* functions as a negative regulator of the response to BT infection.

In the context of BT infection, *clec-4* and *clec-41* thus do not function redundantly, since *clec-41* RNAi caused an opposite phenotype to *clec-4(ok2050)*⁵¹.

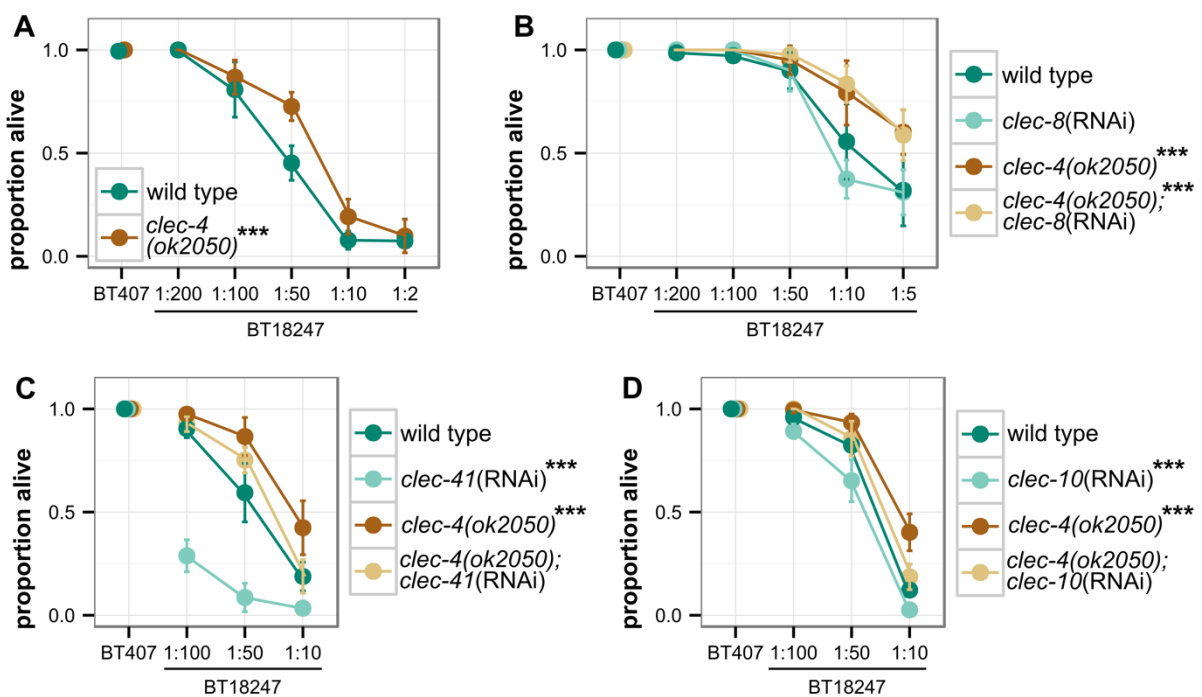


Figure 2. (A) Survival of *clec-4(ok2050)* mutant and (B-D) effect of *clec-8*, *clec-41*, and *clec-10* RNAi on survival of *clec-4(ok2050)* and wild type worms on BT. Worm strains were exposed to non-pathogenic BT407 and serial dilutions of pathogenic BT18247. Means \pm SD of five replicates are shown. Significant differences compared against wild type worms on BT18247 were determined by a generalized linear model (GLM)⁵², *P* values corrected for multiple comparisons with Bonferroni⁵³ and denoted with ****P* < 0.001 in the figure legend.

Surprisingly, silencing the expression of *clec-10*, which showed an opposite regulation to *clec-4*, also resulted in higher susceptibility to BT infection (Figure 1A, Figure 2D), similarly to *clec-41* RNAi, albeit the susceptibility phenotype was not as prominent. The down-regulation of genes during an infection usually suggests their participation in non-immune stress responses like metabolic processes¹⁴. Therefore, *clec-10*, whose expression is down-regulated on pathogenic BT and also on PA14¹³ might act in digestion, yet, at the same time be required for a proper recovery during and after pathogenic

attack. Understanding the exact role of *clec-10* in the response to BT and its interaction with *clec-4* requires further investigation.

CLEC-4 binds to BT and yeast, but not to *E. coli* OP50 and PA14.

Most research on *clec* gene function in *C. elegans* is based on genetic manipulation and subsequent phenotypic analyses, while only two studies tackled functional analyses on the protein level. (i) Takeuchi *et al.* analyzed the sugar-binding activity for CLEC-79⁵⁴ and (ii) Miltsh and colleagues demonstrated binding of CLEC-39 and CLEC-49 to *S. marcescens*¹⁸ *in vitro*. CTLD proteins are commonly considered to function as PRRs since the characterizing CRD is predestinated to recognize and bind glycans found on eukaryotic and prokaryotic cell membranes⁵⁵. Soluble mannose-binding lectin (MBL), for example, is activated by microbe associated molecular patterns (MAMPs) such as mannose, fucose, and N-acetylglucosamine, subsequently initiating the complement system in mammals^{56,57}. CLEC-4 has two CTLDs and one CUB domain (Figure 3A) and might thus be involved in MAMP recognition by the CTLDs and immune system activation mediated by the CUB domain, which occurs for example in MBL-associated serine proteases (MASPs)^{58,59}. Alternatively, secreted CTLD proteins might have antimicrobial effector protein functions like demonstrated for the antimicrobial members of the RegIII family^{60,61} or AmphiCTL1 in *Amphioxus*⁶².

Hence, to investigate the function of CLEC-4 in *C. elegans*-pathogen interactions on the protein level, we performed bacterial binding assays with recombinantly expressed CLEC-4 and the bacterial species used in this study (*E. coli* OP50, BT407, BT18247, PA14) as well as one wild yeast isolate *Pichia* sp. (provided by Philipp Dirksen). The preliminary results of the binding assays indicate that CLEC-4 binds to non-pathogenic BT407, pathogenic BT18247, and yeast in a calcium-dependent fashion, but does not bind the Gram-negative food bacterium *E. coli* OP50 or PA14 (Figure 3B-F).

Please note that these results are only preliminary as the binding assay requires further optimization. Hence, we can as yet only speculate about CLEC-4's function in digestion, pathogen recognition or pathogen clearance.

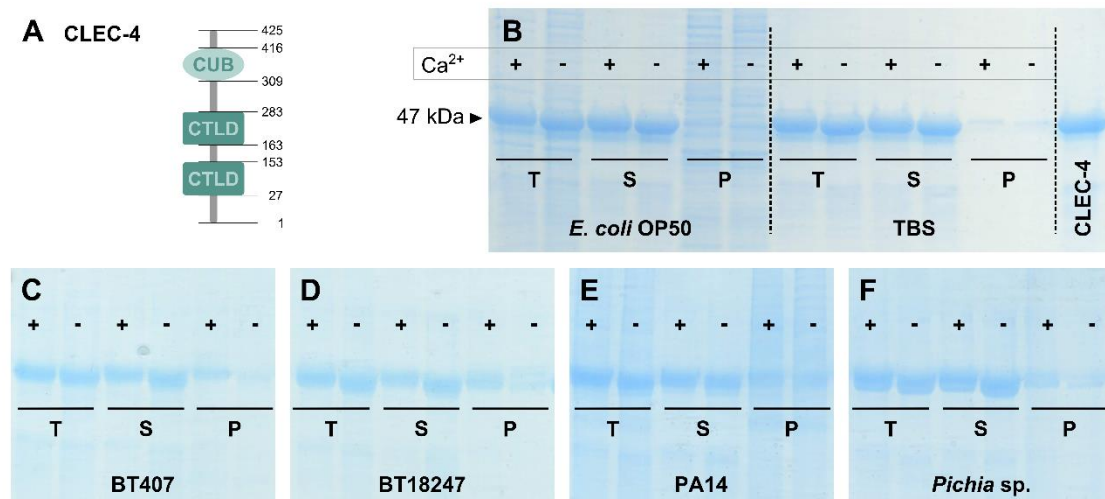


Figure 3. Domain architecture of CLEC-4 and SDS-PAGE of bacterial binding assays. (A) Domain architecture adapted from SMART (<http://smart.embl-heidelberg.de/>) taking UniProt as source database. Numbers represent the amino acid position. (B-F) SDS-PAGE showing bands of CLEC-4 upon incubation with different microbes in different fractions (T, total; S, supernatant; P, pellet) and with or without supplementation of Ca^{2+} . Reaction buffer TBS without microorganisms was used for technical control of the assay (B right panels). While CLEC-4 did not precipitate in the pellet with *E. coli* bacteria (B left panels) or PA14 (E) after centrifugation, it did precipitate in the pellet with both BT strains, the non-pathogenic BT407 (C), the pathogenic BT18247 (D), and yeast (F). For more information see “Materials and Methods” section.

Conclusion

Here, we take first steps into the functional analysis of the CTLD protein encoding gene *clec-4*, which is highly up-regulated upon pathogen infection, on both the gene and protein level. We show that *clec-4* is expressed in the *C. elegans* intestine and amphid neurons. Phenotypic analyses of the *clec-4(ok2050)* mutant revealed that *clec-4* negatively regulates the immune response to BT. We further identified the co- and contra-expressed *clec-4* paralogs *clec-41* and *clec-10*, respectively, which are both required for the immune defense against BT and interact with *clec-4* genetically. Preliminary results of binding studies with the recombinantly expressed protein indicate that CLEC-4 binds to BT and yeast, but not to *E. coli* or PA. These first results raise many interesting research questions we would like to tackle in the near future (see Box 1).

Heretofore, we conclude that it is not possible to infer a gene’s immune function purely from transcriptional data showing up-regulation upon pathogen exposure. CTLD gene expression in *C. elegans* shows an impressive response to pathogenic attack on the transcriptional level, however, experimental evidence of their involvement in immunity is scarce. Nevertheless, CTLD proteins remain an interesting family for the study of the evolution of innate immunity and the genetic interplay of an expanded gene family.

Box 1. Outstanding questions and future work.

What is the evolutionary cost for the *clec-4* mutants for being more resistant to BT?

The fact that the *clec-4* mutant exhibits higher survival after BT infection indicates that it negatively regulates the immune response. This raises the question whether the seemingly beneficial deletion comes with an evolutionary trade-off. A classic example for this in *C. elegans* is the long-lived knock-out mutant of the insulin-like receptor gene *daf-2*, which negatively regulates the FoxO transcription factor DAF-16⁶³. The activation of DAF-16 due to the *daf-2* knock-out leads to an increased resistance to pathogens and stresses^{64,65}, but also causes slower development and a decrease in reproduction in non-challenging environments⁶⁶. *clec-4(ok2050)* might therefore also show decreased evolutionary fitness, which can be tested with corresponding assays (e.g. brood size, reproduction rate, lifespan).

Does *clec-4* mediate susceptibility or resistance to infection with other pathogens? Can the missing survival phenotype of *clec-4* on certain pathogens be explained by redundant paralogs?

Our and previous results present a resistance phenotype of *clec-4(ok2050)* on BT18247, but no survival phenotype on PA14 or *S. marcescens*¹⁸. If not on BT, redundancy could be existent on PA14 or other pathogens. Therefore, exposing the knock-out mutant *clec-4* alone or in combination with *clec-41/clec-10* RNAi to other pathogen species (Gram-positive and -negative) as well as other strains of BT that produce toxins different from BT18247 could give valuable information about the specificity of *clec*-mediated immune responses and redundant dependencies. The former might be even more exciting as the transcriptional response of *clec-4* displays a rather broad up-regulation against many pathogens and stressors.

In vertebrates the recognition of fungus-derived components is covered particularly by C-type lectin receptors (e.g. Dectin-1 and -2 against *Candida albicans*⁶⁷), which seduces to assume a similar role for CTLD proteins in *C. elegans*. Given the fact that in the wild the worm might encounter bacterial as well as fungal microbes⁶⁸ testing *clec-4(ok2050)* survival on different fungi would lead host-pathogen interactions in *C. elegans* into a whole other direction.

In which tissue does *clec-4* act to regulate *C. elegans* immunity?

Analysis of the *pclec-4::GFP* reporter strain revealed that *clec-4* is not only expressed in the intestine but also the amphid neurons. For defining the tissue, in which *clec-4* functions to negatively regulate the immune response to BT infection, we plan to generate tissue-specific rescue and over-expression strains.

Does *clec-4* interact with known *C. elegans* immune signaling pathways?

It is not well understood how the expression of *clec-4* is regulated and if it acts up- or downstream of known immune signaling cascades. We thus would like to incorporate *clec-4* into existing immune pathways in *C. elegans*. It was shown previously that *clec-4* expression is under the control of the transcription factor ATFS-1, which functions in maintaining mitochondrial homeostasis during mitochondrial stress caused by PA14 infection²⁹, or of the GATA transcription factor ELT-2 upon BT infection (unpublished data from our group). Further, it is known for some other *clec* genes that they are regulated by the main *C. elegans* immune signaling pathways^{14,36,69}. Potential genetic interactions of *clec-4* with immune signaling pathway components will be tested by assessing survival of double-mutants, e.g. *daf-2(e1370);clec-4(ok2050)* or of *clec-4* mutant worms, which are additionally treated with RNAi targeting different pathway components. In addition a role of known immune signaling components, e.g. *tir-1* or *pmk-1* of the p38 MAPK pathway, in regulating *clec-4* expression will be examined with the help of the *pclec-4::GFP* reporter.

Which microbes does CLEC-4 bind to and does CLEC-4 have antimicrobial activity?

On the protein level, we could demonstrate that recombinant CLEC-4 seems to be able to bind to BT and yeast, but not to *E. coli* OP50 or PA14. The binding assay needs to be optimized and the results validated. Subsequently, different bacteria species could be tested in order to find binding specificity either on the species level or on a broader scale (e.g. Gram-negative vs. Gram-positive bacteria). As previously mentioned, CTLD proteins might also have antimicrobial activity. We will test if the recombinant CLEC-4 can act as antimicrobial protein by a variety of bacterial killing and inhibitory experiments *in vitro* (e.g. ^{62,70}). Considering the natural habitat of *C. elegans* in microbe-rich environments fungal pathogens should be added to the microbial spectrum for protein activity tests (e.g. ^{71,72}). Further, glucan arrays could assist to pinpoint the actual ligand of CLEC-4 and might also answer the question of specificity (e.g. ⁷³).

Acknowledgements

We thank Sabrina Butze for technical support; Andrei Papkou for help with the statistical analyses; Matthias Leippe and Jessica Gewiese-Rabsch for guidance with the protein work, the Evolutionary Ecology Genetics Department for valuable feedback; Philipp Dirksen for providing the yeast *Pichia* sp. and Dennis Kim for providing PA14. BP, WY, HS and KD are supported by grants from the German Science foundation (Grant DI 1687/1-1 to KD and SCHU 1415/9-2 to HS). WY is member of the IMPRS for Evolutionary Biology. KD is additionally supported by institutional funding from Kiel University.

References – Chapter III

1. Félix, M.-A. & Duveau, F. Population dynamics and habitat sharing of natural populations of *Caenorhabditis elegans* and *C. briggsae*. *BMC Biol.* **10**, 59 (2012).
2. Ashe, A. *et al.* A deletion polymorphism in the *Caenorhabditis elegans* RIG-I homolog disables viral RNA dicing and antiviral immunity. *Elife* **2**, e00994 (2013).
3. Zugasti, O. *et al.* Activation of a G protein-coupled receptor by its endogenous ligand triggers the innate immune response of *Caenorhabditis elegans*. *Nat. Immunol.* **15**, 833–838 (2014).
4. Kim, D. H. *et al.* A conserved p38 MAP kinase pathway in *Caenorhabditis elegans* innate immunity. *Science* **297**, 623–626 (2002).
5. Kurz, C. L. & Ewbank, J. J. *Caenorhabditis elegans*: an emerging genetic model for the study of innate immunity. *Nat. Rev. Genet.* **4**, 380–390 (2003).
6. Bogaerts, A., Beets, I., Schoofs, L. & Verleyen, P. Antimicrobial peptides in *Caenorhabditis elegans*. *Invertebr. Surviv. J.* **7**, 45–52 (2010).
7. Schulenburg, H., Hoepfner, M. P., Weiner, J. & Bornberg-Bauer, E. Specificity of the innate immune system and diversity of C-type lectin domain (CTLD) proteins in the nematode *Caenorhabditis elegans*. *Immunobiology* **213**, 237–250 (2008).
8. Drickamer, K. C-type lectin-like domains. *Curr. Opin. Struct. Biol.* **9**, 585–590 (1999).
9. Dambuza, I. M. & Brown, G. D. C-type lectins in immunity: recent developments. *Curr. Opin. Immunol.* **32**, 21–27 (2015).
10. Pees, B., Yang, W., Zárate-Potes, A., Schulenburg, H. & Dierking, K. High innate immune specificity through diversified C-type lectin-like domain proteins in invertebrates. *J. Innate Immun.* (2015). doi:10.1159/000441475
11. Boehnisch, C. *et al.* Protist-type lysozymes of the nematode *Caenorhabditis elegans* contribute to resistance against pathogenic *Bacillus thuringiensis*. *PLoS ONE* **6**, e24619 (2011).
12. Engelmann, I. *et al.* A comprehensive analysis of gene expression changes provoked by bacterial and fungal infection in *C. elegans*. *PLoS ONE* **6**, e19055 (2011).
13. Shapira, M. *et al.* A conserved role for a GATA transcription factor in regulating epithelial innate immune responses. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 14086–14091 (2006).
14. O’Rourke, D., Baban, D., Demidova, M., Mott, R. & Hodgkin, J. Genomic clusters, putative pathogen recognition molecules, and antimicrobial genes are induced by infection of *C. elegans* with *M. nematophilum*. *Genome Res.* **16**, 1005–1016 (2006).
15. Irazoqui, J. E. *et al.* Distinct pathogenesis and host responses during infection of *C. elegans* by *P. aeruginosa* and *S. aureus*. *PLoS Pathog.* **6**, e1000982 (2010).
16. Simonsen, K. T. *et al.* Quantitative proteomics identifies ferritin in the innate immune response of *C. elegans*. *Virulence* **2**, 120–130 (2011).
17. Sahu, S. N. *et al.* Genomic analysis of immune response against *Vibrio cholerae* hemolysin in *Caenorhabditis elegans*. *PLoS ONE* **7**, e38200 (2012).
18. Miltsch, S. M., Seeberger, P. H. & Lepenies, B. The C-type lectin-like domain containing proteins Clec-39 and Clec-49 are crucial for *Caenorhabditis elegans* immunity against *Serratia marcescens* infection. *Dev. Comp. Immunol.* **45**, 67–73 (2014).
19. Yang, W., Dierking, K. & Schulenburg, H. WormExp: a web-based application for a *Caenorhabditis elegans*-specific gene expression enrichment analysis. *Bioinformatics* **btv667** (2015).
20. Harris, T. W. *et al.* WormBase: A comprehensive resource for nematode research. *Nucleic Acids Res.* **38**, D463–D467 (2010).
21. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).

22. Hasshoff, M., Böhnisch, C., Tonn, D., Hasert, B. & Schulenburg, H. The role of *Caenorhabditis elegans* insulin-like signaling in the behavioral avoidance of pathogenic *Bacillus thuringiensis*. *FASEB J.* **21**, 1801–1812 (2007).
23. Leyns, F., Borgonie, G., Arnaut, G. & Waele, D. D. Nematicidal activity of *Bacillus thuringiensis* isolates. *Fundam. Appl. Nematol.* **18**, 211–218 (1995).
24. Schulte, R. D., Makus, C., Hasert, B., Michiels, N. K. & Schulenburg, H. Multiple reciprocal adaptations and rapid genetic change upon experimental coevolution of an animal host and its microbial parasite. *Proc. Natl. Acad. Sci.* **107**, 7359–7364 (2010).
25. Hobert, O. PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic *C. elegans*. *BioTechniques* **32**, 728–730 (2002).
26. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).
27. Yang, W. *et al.* Overlapping and unique signatures in the proteomic and transcriptomic responses of the nematode *Caenorhabditis elegans* toward pathogenic *Bacillus thuringiensis*. *Dev. Comp. Immunol.* **51**, 1–9 (2015).
28. Treitz, C., Cassidy, L., Höckendorf, A., Leippe, M. & Tholey, A. Quantitative proteome analysis of *Caenorhabditis elegans* upon exposure to nematicidal *Bacillus thuringiensis*. *J. Proteomics* **113**, 337–350 (2015).
29. Pellegrino, M. W. *et al.* Mitochondrial UPR-regulated innate immunity provides resistance to pathogen infection. *Nature* (2014). doi:10.1038/nature13818
30. de Hoon, M. J. L., Imoto, S., Nolan, J. & Miyano, S. Open source clustering software. *Bioinformatics* **20**, 1453–1454 (2004).
31. van der Linden, A. M. *et al.* Genome-wide analysis of light- and temperature-entrained circadian transcripts in *Caenorhabditis elegans*. *PLoS Biol.* **8**, e1000503 (2010).
32. Mueller, M. M. *et al.* DAF-16/FOXO and EGL-27/GATA promote developmental growth in response to persistent somatic DNA damage. *Nat. Cell Biol.* **16**, 1168–1179 (2014).
33. Greiss, S., Schumacher, B., Grandien, K., Rothblatt, J. & Gartner, A. Transcriptional profiling in *C. elegans* suggests DNA damage dependent apoptosis as an ancient function of the p53 family. *BMC Genomics* **9**, 334 (2008).
34. Miller, E. V., Grandi, L. N., Giannini, J. A., Robinson, J. D. & Powell, J. R. The conserved G-protein coupled receptor FSHR-1 regulates protective host responses to infection and oxidative stress. *PLoS ONE* **10**, e0137403 (2015).
35. Kerry, S., TeKippe, M., Gaddis, N. C. & Aballay, A. GATA transcription factor required for immunity to bacterial and fungal pathogens. *PLoS ONE* **1**, e77 (2006).
36. Alper, S., McBride, S. J., Lackford, B., Freedman, J. H. & Schwartz, D. A. Specificity and complexity of the *Caenorhabditis elegans* innate immune response. *Mol. Cell. Biol.* **27**, 5544–5553 (2007).
37. Hoekendorf, A. & Leippe, M. SPP-3, a saposin-like protein of *Caenorhabditis elegans*, displays antimicrobial and pore-forming activity and is located in the intestine and in one head neuron. *Dev. Comp. Immunol.* **38**, 181–186 (2012).
38. Hoekendorf, A., Stanisak, M. & Leippe, M. The saposin-like protein SPP-12 is an antimicrobial polypeptide in pharyngeal neurons of *Caenorhabditis elegans* and participates in defence against a natural bacterial pathogen. *Biochem. J.* (2012). doi:10.1042/BJ20112102
39. Bargmann, C. I. Chemosensation in *C. elegans*. *WormBook* (2006). doi:10.1895/wormbook.1.123.1
40. Melo, J. A. & Ruvkun, G. Inactivation of conserved *C. elegans* genes engages pathogen- and xenobiotic-associated defenses. *Cell* **149**, 452–466 (2012).
41. Schulenburg, H. & Ewbank, J. J. The genetics of pathogen avoidance in *Caenorhabditis elegans*. *Mol. Microbiol.* **66**, 563–570 (2007).

42. Griffiths, J. S. *et al.* Glycolipids as receptors for *Bacillus thuringiensis* crystal toxin. *Science* **307**, 922–925 (2005).
43. Sun, J., Singh, V., Kajino-Sakamoto, R. & Aballay, A. Neuronal GPCR controls innate immunity by regulating noncanonical unfolded protein response genes. *Science* **332**, 729–732 (2011).
44. Evans, E. A., Kawli, T. & Tan, M.-W. *Pseudomonas aeruginosa* suppresses host immunity by activating the DAF-2 insulin-like signaling pathway in *Caenorhabditis elegans*. *PLoS Pathog.* **4**, e1000175 (2008).
45. Neumann, K. *et al.* Clec12a is an inhibitory receptor for uric acid crystals that regulates inflammation in response to cell death. *Immunity* **40**, 389–399 (2014).
46. Thomas, J. H. Analysis of homologous gene clusters in *Caenorhabditis elegans* reveals striking regional cluster domains. *Genetics* **172**, 127–143 (2006).
47. Lynch, M. & Conery, J. S. The evolutionary fate and consequences of duplicate genes. *Science* **290**, 1151–1154 (2000).
48. Woollard, A. Gene duplications and genetic redundancy in *C. elegans*. *WormBook* (2005). doi:10.1895/wormbook.1.2.1
49. Fraser, A. Towards full employment: using RNAi to find roles for the redundant. *Oncogene* **23**, 8346–8352 (2004).
50. Kim, D. H. *et al.* Integration of *Caenorhabditis elegans* MAPK pathways mediating immunity and stress resistance by MEK-1 MAPK kinase and VHP-1 MAPK phosphatase. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 10990–10994 (2004).
51. *An Introduction to genetic analysis.* (W.H. Freeman, 1993).
52. Nelder, J. A. & Wedderburn, R. W. M. Generalized linear models. *J. R. Stat. Soc. Ser. Gen.* **135**, 370 (1972).
53. Dunn, O. J. Multiple comparisons among means. *J. Am. Stat. Assoc.* **56**, 52 (1961).
54. Takeuchi, T. *et al.* A C-type lectin of *Caenorhabditis elegans*: Its sugar-binding property revealed by glycoconjugate microarray analysis. *Biochem. Biophys. Res. Commun.* **377**, 303–306 (2008).
55. van Kooyk, Y. & Rabinovich, G. A. Protein-glycan interactions in the control of innate and adaptive immune responses. *Nat. Immunol.* **9**, 593–601 (2008).
56. Holmskov, U., Thiel, S. & Jensenius, J. C. Collectins and ficolins: Humoral lectins of innate immune defense. *Annu. Rev. Immunol.* **21**, 547–578 (2003).
57. Weis, W. I., Drickamer, K. & Hendrickson, W. A. Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature* **360**, 127–134 (1992).
58. Schwaeble, W., Dahl, M. R., Thiel, S., Stover, C. & Jensenius, J. C. The mannan-binding lectin-associated serine proteases (MASPs) and MASP19: Four components of the lectin pathway activation complex encoded by two genes. *Immunobiology* **205**, 455–466 (2002).
59. Drickamer, K. & Dodd, R. B. C-Type lectin-like domains in *Caenorhabditis elegans*: predictions from the complete genome sequence. *Glycobiology* **9**, 1357–1369 (1999).
60. Miki, T., Holst, O. & Hardt, W.-D. The bactericidal activity of the C-type lectin RegIII β against Gram-negative bacteria involves binding to lipid A. *J. Biol. Chem.* **287**, 34844–34855 (2012).
61. Vaishnava, S. *et al.* The antibacterial lectin RegIII γ promotes the spatial segregation of microbiota and host in the intestine. *Science* **334**, 255–258 (2011).
62. Yu, Y. *et al.* A short-form C-type lectin from *Amphioxus* acts as a direct microbial killing protein via interaction with peptidoglycan and glucan. *J. Immunol.* **179**, 8425–8434 (2007).
63. Kenyon, C., Chang, J., Gensch, E., Rudner, A. & Tabtlang, R. A *C. elegans* mutant that lives twice as long as wild type. *Nature* **366**, 461–464 (1993).
64. Lamitina, S. T. Transcriptional targets of DAF-16 insulin signaling pathway protect *C. elegans* from extreme hypertonic stress. *AJP Cell Physiol.* **288**, C467–C474 (2005).

65. Garsin, D. A. Long-lived *C. elegans daf-2* mutants are resistant to bacterial pathogens. *Science* **300**, 1921–1921 (2003).
66. Henderson, S. T. & Johnson, T. E. *daf-16* integrates developmental and environmental inputs to mediate aging in the nematode *Caenorhabditis elegans*. *Curr. Biol.* **11**, 1975–1980 (2001).
67. Hardison, S. E. & Brown, G. D. C-type lectin receptors orchestrate antifungal immunity. *Nat. Immunol.* **13**, 817–822 (2012).
68. Félix, M.-A. & Braendle, C. The natural history of *Caenorhabditis elegans*. *Curr. Biol.* **20**, R965–R969 (2010).
69. Troemel, E. R. *et al.* p38 MAPK regulates expression of immune response genes and contributes to longevity in *C. elegans*. *PLoS Genet.* **2**, e183 (2006).
70. Andrä, J., Herbst, R. & Leippe, M. Amoebapores, archaic effector peptides of protozoan origin, are discharged into phagosomes and kill bacteria by permeabilizing their membranes. *Dev. Comp. Immunol.* **27**, 291–304 (2003).
71. Park, H. G. *et al.* Dual function of a bee (*Apis cerana*) inhibitor cysteine knot peptide that acts as an antifungal peptide and insecticidal venom toxin. *Dev. Comp. Immunol.* **47**, 247–253 (2014).
72. Chikalovets, I. V. *et al.* A lectin with antifungal activity from the mussel *Crenomytilus grayanus*. *Fish Shellfish Immunol.* **42**, 503–507 (2015).
73. Arthur, C. M., Cummings, R. D. & Stowell, S. R. Using glycan microarrays to understand immunity. *Curr. Opin. Chem. Biol.* **18**, 55–61 (2014).
74. Bolker, B. M. *et al.* Generalized linear mixed models: A practical guide for ecology and evolution. *Trends Ecol. Evol.* **24**, 127–135 (2009).
75. Kaplan, E. L. & Meier, P. Nonparametric estimation from incomplete observations. *J. Am. Stat. Assoc.* **53**, 457 (1958).
76. Harrington, D. in *Encyclopedia of Biostatistics* (John Wiley & Sons, Ltd, 2005).

Supplementary Figures

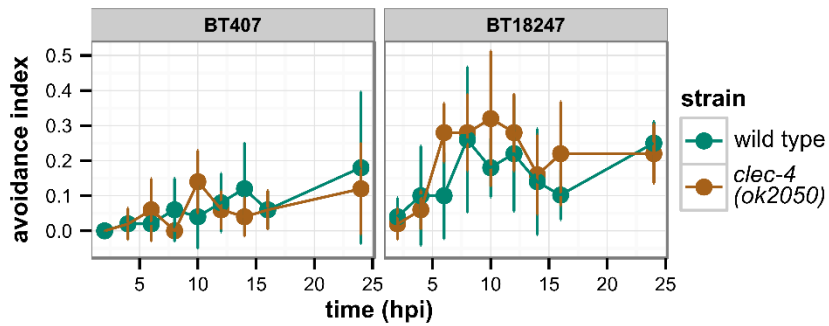


Figure S 1. Avoidance behavior of *clec-4(ok2050)* and wild type worms on BT. Worms were exposed to a mild concentration (1:400) of non-pathogenic BT407 and pathogenic BT18247. The avoidance index is defined as described in the “Materials and Methods” section. Means \pm SD of five replicates are shown. No significant differences between *clec-4(ok2050)* and wild type were determined by a generalized linear mixed model (GLMM)⁷⁴ per bacterial treatment, P values corrected for multiple comparisons with Bonferroni⁵³.

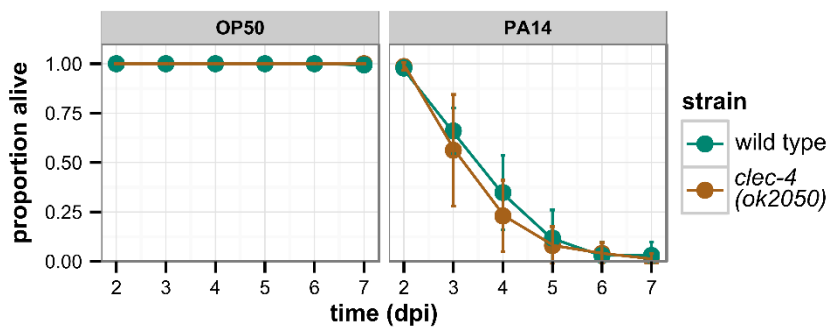


Figure S 2. Survival of *clec-4(ok2050)* and wild type worms on *E. coli* OP50 and PA14. Means \pm SD of five replicates are shown. dpi, days post infection. No significant differences between *clec-4(ok2050)* and wild type were determined by Kaplan-Meier⁷⁵ analysis and log-rank test⁷⁶, P values corrected for multiple comparisons with Bonferroni⁵³.

Additional Results

The following figures and tables present additional data and analyses which either did not meet our expectations, did not answer our questions of interest, or are not sufficient for a complete manuscript. This chapter should fulfill the purpose of giving detailed information for further studies concerning *clec*/CLEC genes and proteins.

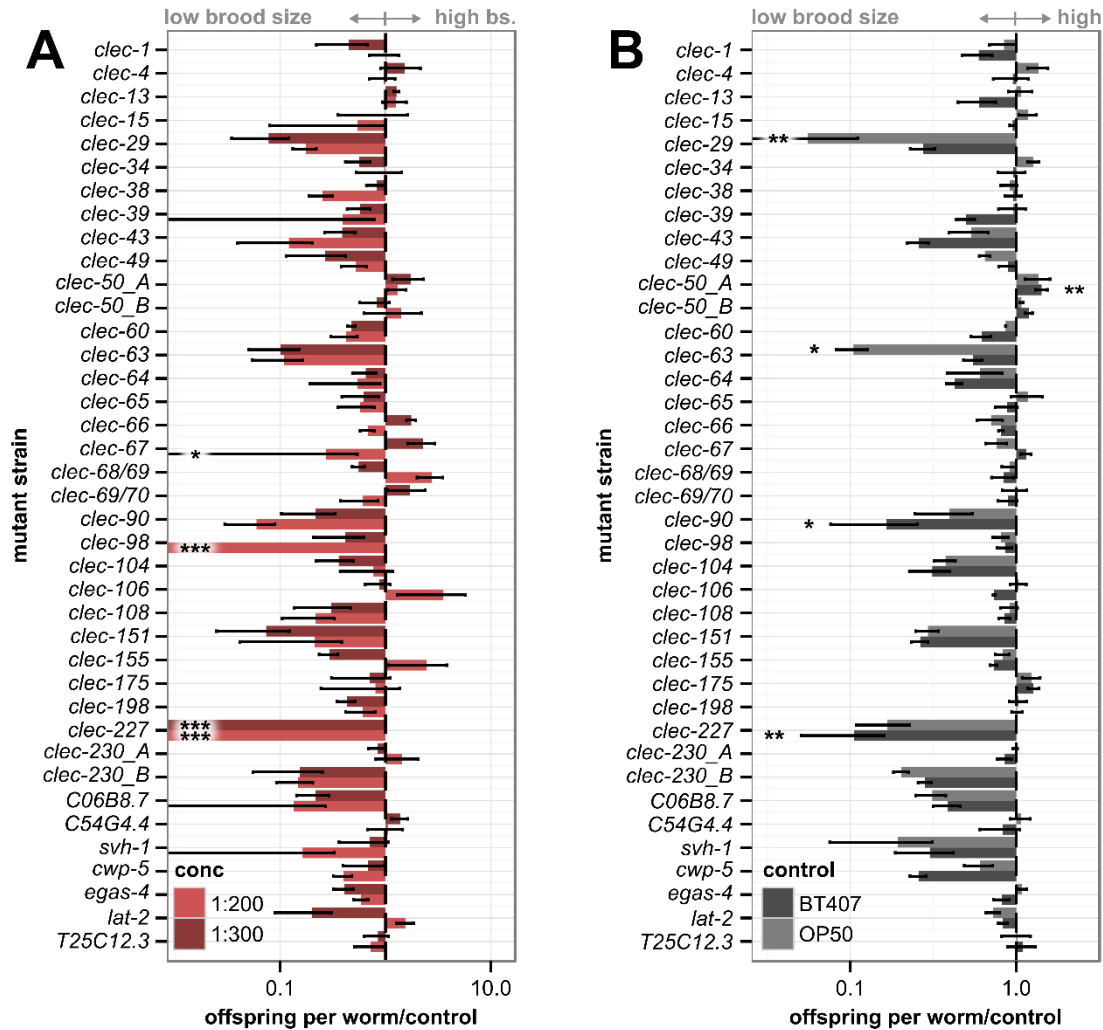


Figure 1. Brood size screen of *C. elegans clec* mutants upon exposure to BT and *E. coli* OP50. *clec* mutants were screened for brood size in parallel to survival (see Chapter II for the complete list of mutants used). Five synchronized L4 larvae per plate were exposed to (A) low concentrations (conc) of pathogenic BT18247 mixed with *E. coli* OP50 at OD 5 (two shades of red) and (B) non-pathogenic BT407 in concentration 1:200 as well as *E. coli* OP50 at OD 5. All individuals were washed off 72 hpi (hours post infection) with M9 supplemented with 0.1% Triton, frozen at -80 °C and the worm number was counted. The brood size was calculated as offspring per worm and normalized by the wild type control N2 ($x = 1$, vertical dashed line). Values < 1 indicate a lower brood size (less offspring per worm), values > 1 a higher brood size (more offspring per worm). Means of three replicates are shown; the error bars are the standard error (SE) of the mean. Significant differences were determined with a randomization test¹ per bacterial treatment, corrected for multiple comparisons with false discovery rate (FDR)² and denoted by asterisks ($*P < 0.025$, $**P < 0.005$, $***P < 0.0005$). Note that the x-axes are log₁₀ transformed. All significant *P* values can be found in Table 1.

Table 1. Randomization test¹ on brood size of *clec* mutants per bacterial treatment; *P* adjustment with FDR².

affected gene	genotype	strain	treatment	Figure	<i>P</i> value	significance level
<i>clec-227</i>	<i>clec-227(ok2618)</i> V	RB1984	BT18247, 1:200	1A	< 0.0005	***
<i>clec-98</i>	<i>ZK39.7(ok2085)</i> I	VC1600	BT18247, 1:200	1A	< 0.0005	***
<i>clec-227</i>	<i>clec-227(ok2618)</i> V	RB1984	BT18247, 1:300	1A	< 0.0005	***
<i>clec-67</i>	<i>F56D6.2(ok2770)</i> IV	RB2095	BT18247, 1:300	1A	0.0172	*
<i>clec-227</i>	<i>clec-227(ok2618)</i> V	RB1984	BT407	1B	0.0012	**
<i>clec-50</i>	<i>clec-50(ok2455)</i> V	RB1897	BT407	1B	0.0020	**
<i>clec-90</i>	<i>clec-90(tm3402)</i>		BT407	1B	0.0118	*
<i>clec-29</i>	<i>clec-29(ok3181)</i> V	RB2345	<i>E. coli</i> OP50	1B	0.0016	**
<i>clec-63</i>	<i>clec-63(ok2656)</i> II	RB2004	<i>E. coli</i> OP50	1B	0.0084	*

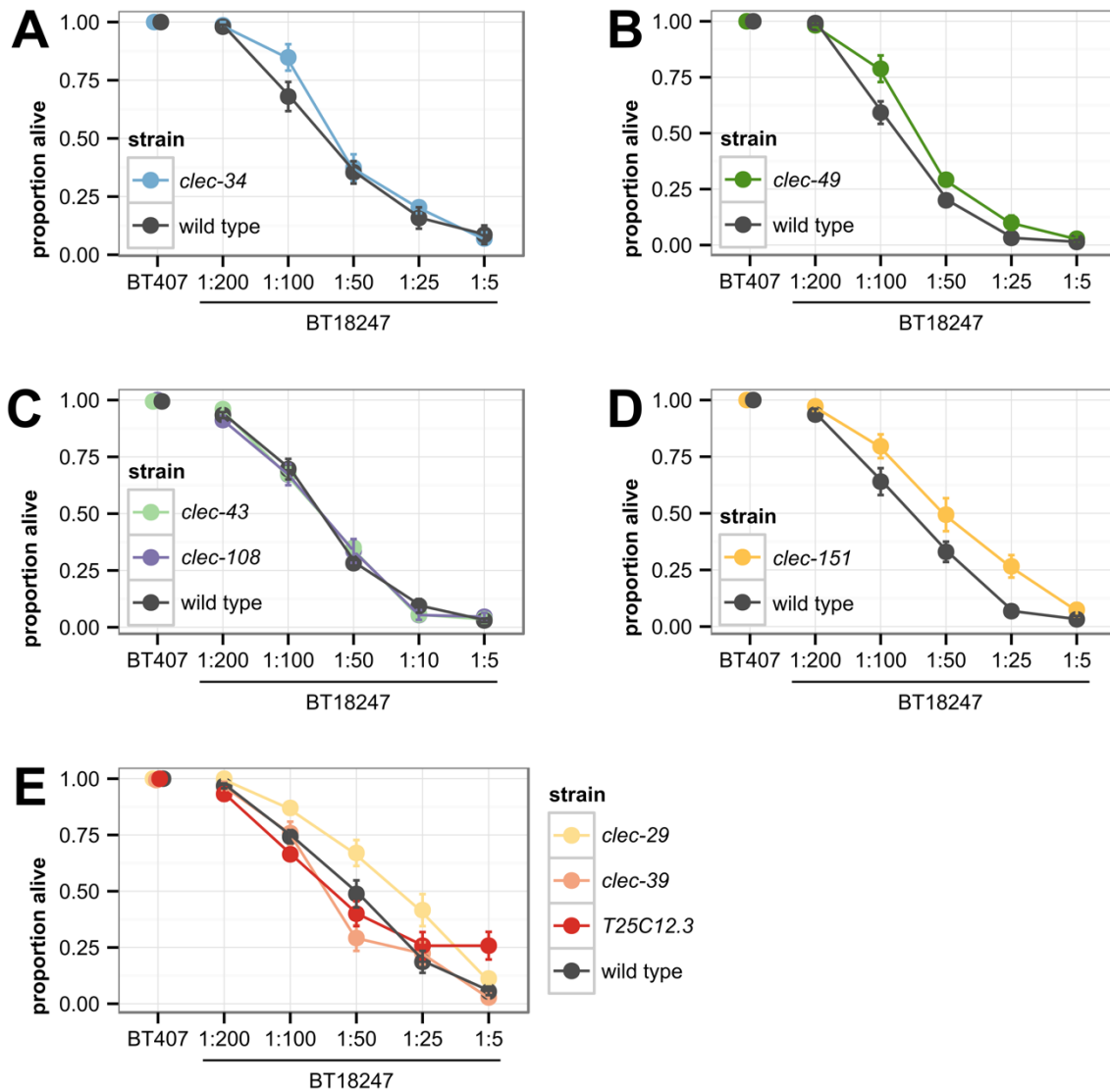


Figure 2. Survival of out-crossed *C. elegans clec* mutants upon exposure to BT18247. Based on the results of the *clec* mutant survival screen (Chapter II), we chose the most promising *clec* gene candidates for out-crossing and further analysis: The *clec-34* and *clec-151* mutants were resistant on both BT and PA14. *clec-49* and *clec-50* are in a cluster on chromosome V and seemed to mediate both resistance (*clec-49*) and susceptibility (*clec-50*) on BT and PA14. Due to technical reasons, we were not able to out-cross the *clec-50* mutant and thus could not analyze it further. The *clec-43* and *T25C12.3* mutants were highly susceptible on both BT and PA14. The *clec-108* mutant was highly susceptible on BT. The *clec-29* mutant was highly resistant on BT. As shown here after out-crossing 3x with wild type N2 we could confirm the resistance phenotypes of *clec-34*, *clec-49*, *clec-151*, and *clec-29* on BT, whereas we could not confirm the results for the other mutants. The assays were conducted with 30 synchronized L4 larvae per plate exposed to pathogenic BT18247 mixed with *E. coli* OP50 at OD 5 in serial dilutions in PBS (x-axes). Alive and dead animals were scored 24 hpi. Worms were considered to be dead when they did not respond to light touch. BT407 was used as non-pathogenic control in the highest BT concentration applied for BT18247 per assay. For a detailed description see Chapter II, “Materials and Methods”. Survival was defined as proportion of alive individuals (y-axes) and is displayed as means \pm SE of five replicates. Significant differences were determined with a generalized linear model (GLM)³, always compared against the wild type per survival assay, and corrected for multiple comparisons with Bonferroni⁴. All *P* values can be found in Table 2.

Table 2. GLM³ on survival of *clec* mutants per survival assay; *P* adjustment with Bonferroni⁴.

affected gene	genotype	strain	Figure	<i>P</i> value	significance level
<i>clec-34</i>	<i>clec-34(ok2120)</i> V	RB1701	2A	0.021	*
<i>clec-49</i>	<i>clec-49(ok2416)</i> V	RB1870	2B	< 0.001	***
<i>clec-43</i>	<i>clec-43(ok3188)</i> II	RB2350	2C	1.000	n. s. ^a
<i>clec-108</i>	<i>clec-108(tm1622)</i>		2C	1.000	n. s.
<i>clec-151</i>	<i>F10F2.7(ok2264)</i> III	RB1761	2D	< 0.001	***
<i>clec-29</i>	<i>clec-29(ok3181)</i> V	RB2345	2E	< 0.001	***
<i>clec-39</i>	<i>clec-39(ok2271)</i> V	RB1767	2E	0.345	n. s.
<i>T25C12.3</i>	<i>T25C12.3(tm4691)</i>		2E	0.679	n. s.
<i>clec-227^b</i>	<i>clec-227(ok2618)</i> V	RB1984	not shown	n. a. ^c	n. s.

^a n. s., not significant.

^b The phenotypic analyses with the *clec-227* mutant was performed by Marco Grasse; detailed information is to be found in his Master thesis.

^c n. a., not available.

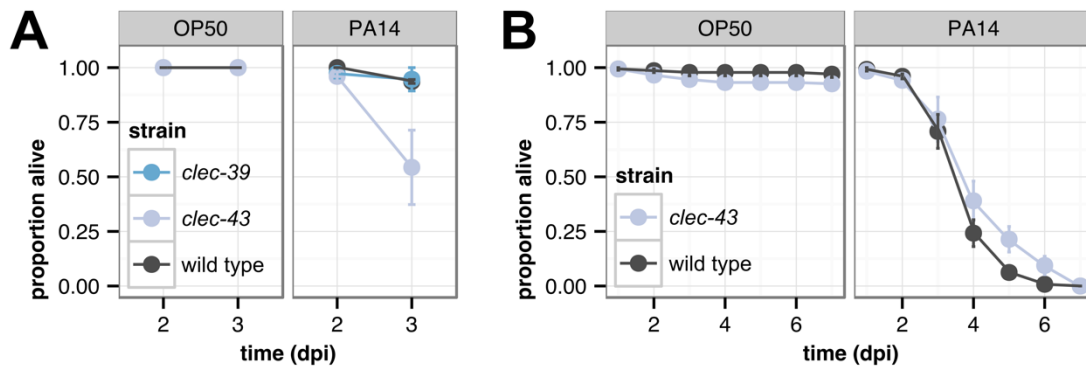


Figure 3. Survival assays of out-crossed *C. elegans clec* mutants upon exposure to PA14. Based on the results of the *clec* mutant survival screen (Chapter II), we chose the most promising *clec* gene candidates for out-crossing and further analysis: The *clec-43* mutant was highly susceptible on both BT and PA14 and the *clec-39* mutant was resistant on PA14. While we could not confirm the susceptibility phenotype of the *clec-43* mutant on BT (see Figure 2C), we could confirm its susceptibility on short-term exposure to PA14 (A) and obtained contradictory results on a long-term exposure being more resistant (B). The resistance phenotype of the *clec-39* mutant on PA14 was not confirmed. The assays were conducted at 25 °C with 30 synchronized L4 larvae (back-crossed 3x with wild type N2) per plate exposed to *E. coli* OP50 at OD 5 as non-pathogenic control or PA14 mixed with *E. coli* OP50 in dilution 1:3. Alive and dead animals were scored every day (dpi, days post infection). Worms were considered to be dead when they did not respond to light touch. For a detailed description see Chapter III, “Materials and Methods”. Survival was defined as proportion of alive individuals (y-axes) and is displayed as means \pm SE of five replicates. Significant differences were determined with Kaplan-Meier analysis⁵ and log-rank test⁶ per bacterial treatment and survival assay, and corrected for multiple comparisons with Bonferroni⁴. All *P* values can be found in Table 3.

Table 3. Kaplan-Meier analysis⁵ and log-rank test⁶ on survival of *clec* mutants per bacterial treatment; *P* adjustment with Bonferroni⁴.

affected gene	genotype	strain	treatment	Figure	<i>P</i> value	significance level
<i>clec-39</i>	<i>clec-39(ok2271)</i> V	RB1767	<i>E. coli</i> OP50	3A	n. a. ^a	n.s. ^b
<i>clec-43</i>	<i>clec-43(ok3188)</i> II	RB2350	<i>E. coli</i> OP50	3A	n. a.	n.s.
<i>clec-39</i>	<i>clec-39(ok2271)</i> V	RB1767	PA14	3A	n. a.	n.s.
<i>clec-43</i>	<i>clec-43(ok3188)</i> II	RB2350	PA14	3A	< 0.001	***
<i>clec-43</i>	<i>clec-43(ok3188)</i> II	RB2350	<i>E. coli</i> OP50	3B	0.094	n.s.
<i>clec-43</i>	<i>clec-43(ok3188)</i> II	RB2350	PA14	3B	0.010	**

^a n. a., not available.

^b n. s., not significant.

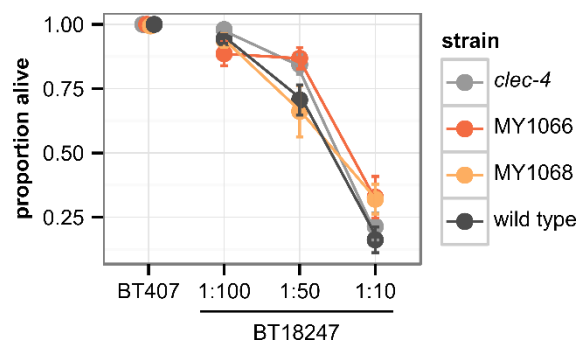


Figure 4. Survival of out-crossed *clec-4* mutants and two transgenic rescue strains upon exposure to BT18247.

Three transgenic rescue lines were generated, MY1066, MY1067, and MY1068, of which two were tested for their survival phenotype. For more information on the rescue lines please see Table 11. The assays were conducted with 30 synchronized L4 larvae per plate exposed to pathogenic BT18247 mixed with *E. coli* OP50 at OD 5 in serial dilutions in PBS (x-axes). Alive and dead animals were scored 24 hpi. Worms were considered to be dead when they did not respond to light touch. BT407 was used as non-pathogenic control in the highest BT concentration applied for BT18247 per assay. For a detailed description see Chapter II, “Materials and Methods”. Survival was defined as proportion of alive individuals (y-axes) and is displayed as means \pm SE of five replicates. Significant differences were determined with a GLM³ using Tukey contrasts⁷ as post-hoc test. All *P* values can be found in Table 4.

Table 4. GLM³ and Tukey contrasts⁷ on survival of *clec-4* and two transgenic rescue strains on BT.

strain 1 ^a	genotype	strain 2 ^a	genotype	<i>P</i> value	significance level
<i>clec-4</i>	<i>clec-4(ok2050)</i>	N2	wild type	0.013	*
<i>clec-4</i>	<i>clec-4(ok2050)</i>	MY1066 ^c	<i>clec-4(ok2050);yaEx59</i>	0.946	n. s. ^b
<i>clec-4</i>	<i>clec-4(ok2050)</i>	MY1068 ^c	<i>clec-4(ok2050);yaEx61</i>	0.469	n. s.
MY1066	<i>clec-4(ok2050);yaEx59</i>	MY1068	<i>clec-4(ok2050);yaEx61</i>	0.194	n. s.
N2	wild type	MY1066	<i>clec-4(ok2050);yaEx59</i>	0.002	**
N2	wild type	MY1068	<i>clec-4(ok2050);yaEx61</i>	0.396	n. s.

^a Comparison strain 1 against strain 2.

^b n. s., not significant.

^c More details about the transgenic strains are to be found in Table 11.

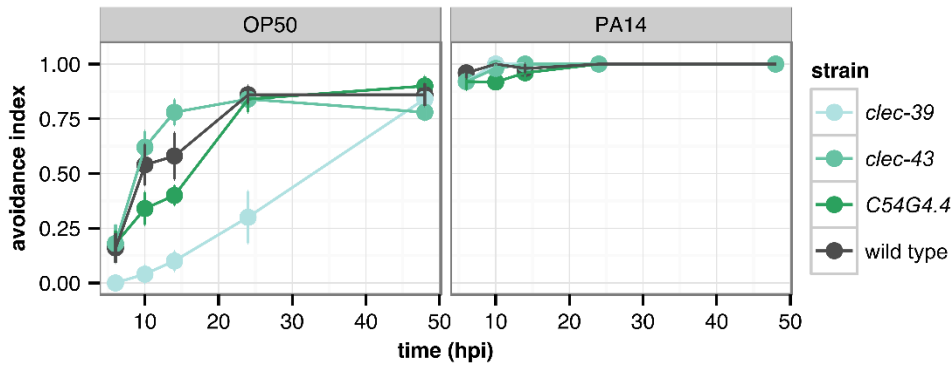


Figure 5. Avoidance assay of out-crossed *C. elegans* *clec* mutants on PA14. Based on the results of the *clec* mutant survival screen (Chapter II), we chose two *clec* genes which showed opposite survival phenotypes on PA14 – *clec-39* was more resistant, *clec-43* more susceptible – as well as *C54G4.4* with a profound pathogen avoidance behavior on BT (Chapter II). Before conduction of assay each worm strain was back-crossed 3x with wild type N2. Enriched 9 cm NGM (0.35% peptone) plates were inoculated with 30 μ l *E. coli* OP50 at OD 5 as non-pathogenic control or PA14 mixed with *E. coli* OP50 in dilution 1:3. On every plate an 80 μ l ring of *E. coli* OP50 at OD 5 was pipetted surrounding the spot in the middle at a 1.5 cm distance. The plates were incubated for 24 h at 37 °C and another 24 h at 25 °C at which the PA14 survival assay was conducted. Ten synchronized L4 larvae per plate were picked into the middle at 0 hpi. Worms residing on the bacterial spot were scored after several hours (x-axes). The avoidance index was defined as (initial total number of worms – worms on bacterial spot)/initial total number of worms. Means \pm SE of five replicates are shown. Significant differences were determined with a generalized linear mixed model (GLMM)⁸ always compared against the wild type per survival assay, and corrected for multiple comparisons with Bonferroni⁴. All *P* values can be found in Table 5.

Table 5. GLMM⁸ on avoidance behavior of *clec* mutants per bacterial treatment; *P* adjustment with Bonferroni⁴.

affected gene	genotype	strain	treatment	<i>P</i> value	significance level
<i>clec-39</i>	<i>clec-39(ok2271)</i> V	RB1767	<i>E. coli</i> OP50	<0.001	***
<i>clec-43</i>	<i>clec-43(ok3188)</i> II	RB2350	<i>E. coli</i> OP50	1.000	n.s. ^a
<i>C54G4.4</i>	<i>C54G4.4(ok2110)</i> I	RB1697	<i>E. coli</i> OP50	1.000	n.s.
<i>clec-39</i>	<i>clec-39(ok2271)</i> V	RB1767	PA14	1.000	n.s.
<i>clec-43</i>	<i>clec-43(ok3188)</i> II	RB2350	PA14	1.000	n.s.
<i>C54G4.4</i>	<i>C54G4.4(ok2110)</i> I	RB1697	PA14	1.000	n.s.

^a n. s., not significant.

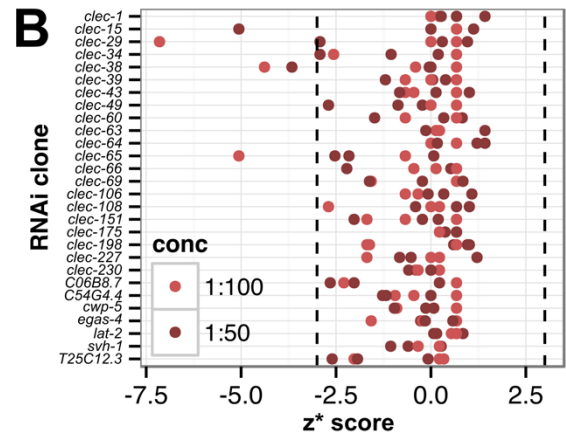
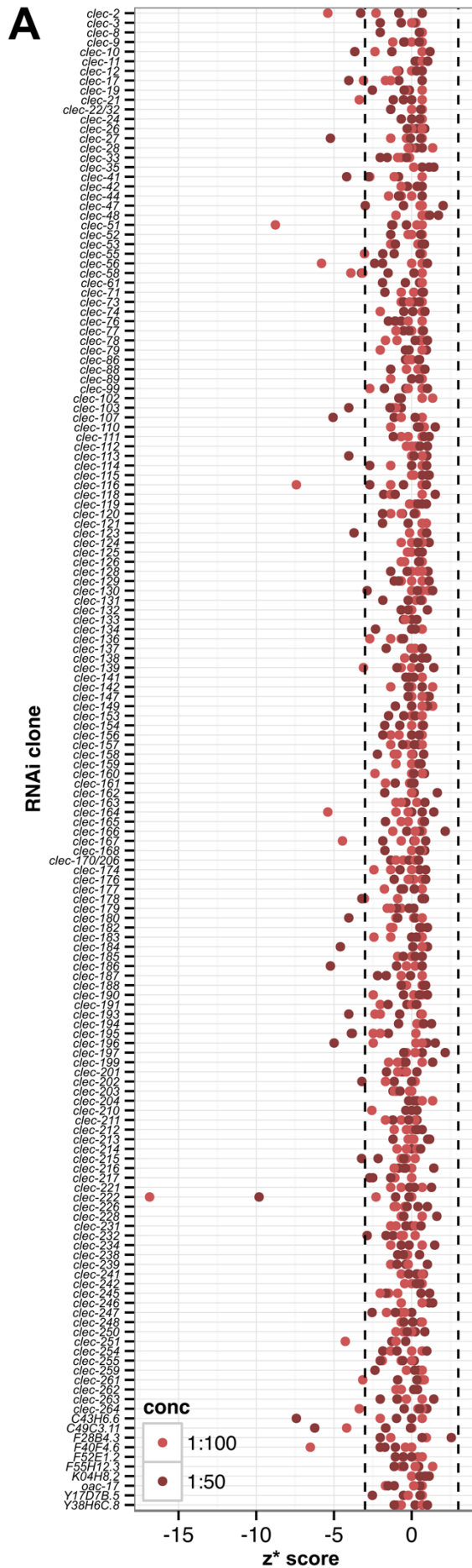


Figure 6. Systematic *clec* RNAi screen on BT. All RNAi clones used which were available from the Ahringer⁸ or Marc Vidal⁹ library (Table 6). L4440 vector inserts were sequenced (amplification with forward 5'-GTTTTCCAGTCACGACGTT-3' and reverse primer 5'-TGGATAACCGTATTACCGCC-3'⁹), the results are to be found in Table 6. Transgenic strain *fris7* was used, which harbors a red *pcol-12::RFP* construct next to green *pnlp-29::GFP*¹⁰. This particular strain was chosen as dead worms showed either blurry or no fluorescence in the epidermis while living worms showed distinct fluorescence. Starting from L1 larval stage worms were fed with the respective RNAi clone prior to infection according to standard protocols¹¹. Obvious phenotypic knock-downs were applied to ensure the success of RNAi (Table 7). 30-50 L4 larvae were pipetted onto PFM seeded with a mixture of BT18247 and *E. coli* OP50 at OD 5 in PBS. The two different shades of red represent the two different concentrations (conc). The worms were washed off the plates 24 hpi with M9 buffer and scored according to their fluorescence in M9 droplets. Each *clec* gene was tested in three completely independent replicates per BT concentration. For final hit selection (selection of candidate genes) the z^* score was employed¹²: $(Y_i - \bar{Y}_N)/MAD_N$ with Y_i = response variable (proportion of alive individuals), \bar{Y}_N = median of negative reference (median of all response variables per week), and MAD = median of absolute deviation. The dashed vertical lines represent the cut-off which is defined as median $\pm k \cdot MAD$ ($k = 3$, according to 3-sigma rule). Replicates outside the cut-off were considered as hits. Data of non-pathogenic control BT407 in dilution of 1:50 are not displayed as viability of worms was not impaired in any RNAi treatment (and the MAD would consequently be 0). (B) is a subset of the main data set (A) and only includes *clec* genes, for which knock-out mutants were tested beforehand (Chapter II).

Table 6. RNAi clones of Ahringer library used in RNAi screen (Figure 6).

lab code	affected gene	Ahringer library ^a	remarks
327	<i>clec-1</i>	V-4K08	confirmed by sequencing
233	<i>clec-2</i>	II-2N06	confirmed by sequencing
189	<i>clec-3</i>	II-2L16	confirmed by sequencing
222	<i>clec-8</i>	II-8M14	confirmed by sequencing
170	<i>clec-9</i>	V-11M17	confirmed by sequencing
187	<i>clec-10</i>	II-1M19	confirmed by sequencing
234	<i>clec-11</i>	I-6E18	confirmed by sequencing
235	<i>clec-12</i>	I-6G02	confirmed by sequencing
311	<i>clec-15</i>	I-6K14	confirmed by sequencing
185	<i>clec-17</i>	I-6C22	confirmed by sequencing
188	<i>clec-19</i>	II-2G11	confirmed by sequencing
291	<i>clec-21</i>	n. a. ^b	separately obtained; confirmed by sequencing
238	<i>clec-22/clec-32</i>	V-11C10	confirmed by sequencing
106	<i>clec-24</i>	V-11C12	confirmed by sequencing
166	<i>clec-26</i>	V-11A20	confirmed by sequencing
239	<i>clec-27</i>	V-11A24	confirmed by sequencing
114	<i>clec-28</i>	V-11C14	confirmed by sequencing
333	<i>clec-29</i>	V-11O13	confirmed by sequencing
122	<i>clec-33</i>	V-11C18	confirmed by sequencing
332	<i>clec-34</i>	V-11O11	
240	<i>clec-35</i>	V-11C02	confirmed by sequencing
334	<i>clec-38</i>	V-11O15	confirmed by sequencing
331	<i>clec-39</i>	V-11O09	confirmed by sequencing
242	<i>clec-41</i>	V-8P17	confirmed by sequencing
201	<i>clec-42</i>	V-11P18	confirmed by sequencing
314	<i>clec-43</i>	II-1F11	confirmed by sequencing
258	<i>clec-44</i>	n. a.	obtained separately; confirmed by sequencing
182	<i>clec-47</i>	V-9N12	confirmed by sequencing
171	<i>clec-48</i>	V-11N18	confirmed by sequencing
335	<i>clec-49</i>	V-12F14	
211	<i>clec-51</i>	IV-4K03	confirmed by sequencing
212	<i>clec-52</i>	IV-4K07	confirmed by sequencing
146	<i>clec-53</i>	I-1L19	confirmed by sequencing
290	<i>clec-55</i>	n. a.	obtained separately; confirmed by sequencing
227	<i>clec-56</i>	V-9D13	confirmed by sequencing
151	<i>clec-58</i>	II-7G14	confirmed by sequencing
315	<i>clec-60</i>	II-7G20	confirmed by sequencing
301	<i>clec-61</i>	II-10I15	confirmed by sequencing
316	<i>clec-63</i>	II-8B17	
317	<i>clec-64</i>	II-8B19	confirmed by sequencing
318	<i>clec-65</i>	II-8B21	confirmed by sequencing
319	<i>clec-66</i>	II-8B23	
326	<i>clec-69</i>	IV-8J17	confirmed by sequencing
118	<i>clec-71</i>	IV-8H19	confirmed by sequencing

lab code	affected gene	Ahringer library ^a	remarks
243	<i>clec-73</i>	IV-8J13	confirmed by sequencing
213	<i>clec-74</i>	IV-8J09	confirmed by sequencing
113	<i>clec-76</i>	IV-8J05	confirmed by sequencing
134	<i>clec-77</i>	IV-8J15	confirmed by sequencing
244	<i>clec-78</i>	IV-2A15	confirmed by sequencing
250	<i>clec-79</i>	IV-2A19	confirmed by sequencing
125	<i>clec-86</i>	X-3B14	confirmed by sequencing
253	<i>clec-88</i>	II-4P15	confirmed by sequencing
183	<i>clec-89</i>	I-1B24	
293	<i>clec-99</i>	n. a.	obtained separately; confirmed by sequencing
141	<i>clec-102</i>	I-6I24	confirmed by sequencing
247	<i>clec-103</i>	I-6M06	confirmed by sequencing
310	<i>clec-106</i>	I-6F03	confirmed by sequencing
210	<i>clec-107</i>	I-6J21	confirmed by sequencing
313	<i>clec-108</i>	I-6L01	confirmed by sequencing
208	<i>clec-110</i>	I-6N19	confirmed by sequencing
300	<i>clec-111</i>	I-9F16	confirmed by sequencing
288	<i>clec-112^c</i>	n. a.	Marc Vidal Library ORF ID (WS9): Y71A12B.g; confirmed by sequencing
245	<i>clec-113</i>	I-7E05	confirmed by sequencing
209	<i>clec-114</i>	I-7K07	confirmed by sequencing
248	<i>clec-115</i>	I-7K03	confirmed by sequencing
254	<i>clec-116</i>	I-7K11	confirmed by sequencing
186	<i>clec-118</i>	II-1G02	confirmed by sequencing
246	<i>clec-119</i>	II-1D24	confirmed by sequencing
137	<i>clec-120</i>	II-2D02	confirmed by sequencing
136	<i>clec-121</i>	II-2N18	confirmed by sequencing
190	<i>clec-123</i>	II-2N12	
295	<i>clec-124</i>	n. a.	obtained separately; confirmed by sequencing
135	<i>clec-125</i>	II-3E02	confirmed by sequencing
251	<i>clec-126</i>	II-3C24	
191	<i>clec-128</i>	II-3E12	confirmed by sequencing
148	<i>clec-129</i>	II-3G06	confirmed by sequencing
147	<i>clec-130</i>	II-3G04	confirmed by sequencing
192	<i>clec-131</i>	II-3E14	confirmed by sequencing
255	<i>clec-132</i>	II-3G02	confirmed by sequencing
256	<i>clec-133</i>	II-3E24	confirmed by sequencing
257	<i>clec-134</i>	II-3E22	confirmed by sequencing
216	<i>clec-136</i>	II-3L18	confirmed by sequencing
259	<i>clec-137</i>	II-3L20	
229	<i>clec-138</i>	II-3L22	confirmed by sequencing
149	<i>clec-139</i>	II-3L24	confirmed by sequencing
297	<i>clec-141</i>	n. a.	obtained separately
150	<i>clec-142</i>	II-6G13	confirmed by sequencing
262	<i>clec-147</i>	IV-5G06	confirmed by sequencing
152	<i>clec-149</i>	III-1C18	confirmed by sequencing

lab code	affected gene	Ahringer library ^a	remarks
321	<i>clec-151</i>	III-2G16	
221	<i>clec-153</i>	III-2G18	confirmed by sequencing
299	<i>clec-154</i>	n. a.	obtained separately; confirmed by sequencing
154	<i>clec-156</i>	III-2B07	confirmed by sequencing
155	<i>clec-157</i>	III-2B09	
157	<i>clec-158</i>	III-3J07	
156	<i>clec-159</i>	III-3H19	confirmed by sequencing
219	<i>clec-160</i>	III-4K20	confirmed by sequencing
158	<i>clec-161</i>	III-5K13	confirmed by sequencing
117	<i>clec-162</i>	III-5M09	confirmed by sequencing
264	<i>clec-163</i>	III-5B18	
265	<i>clec-164</i>	IV-1C14	confirmed by sequencing
1	<i>clec-165</i>	IV-1M18	confirmed by sequencing
266	<i>clec-166</i>	IV-1M08	confirmed by sequencing
126	<i>clec-167</i>	IV-1M06	confirmed by sequencing
121	<i>clec-168</i>	IV-1M12	confirmed by sequencing
123	<i>clec-170, clec-206</i>	IV-1K24	confirmed by sequencing
104	<i>clec-174</i>	IV-8H23	confirmed by sequencing
323	<i>clec-175</i>	IV-2E04	confirmed by sequencing
129	<i>clec-176</i>	IV-2F23	confirmed by sequencing
130	<i>clec-177</i>	IV-2F21	confirmed by sequencing
307	<i>clec-178</i>	IV-2N16	confirmed by sequencing
268	<i>clec-179</i>	IV-3H01	confirmed by sequencing
2	<i>clec-180</i>	IV-3H02	confirmed by sequencing
231	<i>clec-182</i>	IV-4B01	confirmed by sequencing
205	<i>clec-183</i>	IV-4F16	confirmed by sequencing
269	<i>clec-184</i>	IV-6J13	confirmed by sequencing
270	<i>clec-185</i>	IV-6J17	
271	<i>clec-186</i>	IV-6H02	confirmed by sequencing
193	<i>clec-187</i>	IV-6F24	confirmed by sequencing
160	<i>clec-188</i>	IV-7N21	confirmed by sequencing
223	<i>clec-190</i>	IV-7N23	confirmed by sequencing
159	<i>clec-191</i>	IV-7J22	confirmed by sequencing
161	<i>clec-193</i>	IV-8C09	confirmed by sequencing
162	<i>clec-194</i>	IV-8C19	confirmed by sequencing
105	<i>clec-195</i>	IV-8G19	confirmed by sequencing
206	<i>clec-196</i>	IV-8A08	confirmed by sequencing
272	<i>clec-197</i>	IV-8E02	confirmed by sequencing
165	<i>clec-198</i>	IV-8E04	confirmed by sequencing
3	<i>clec-199</i>	IV-8E10	
273	<i>clec-201</i>	IV-8E14	confirmed by sequencing
274	<i>clec-202</i>	IV-8E16	confirmed by sequencing
4	<i>clec-203</i>	V-1H15	confirmed by sequencing
194	<i>clec-204</i>	V-1L02	confirmed by sequencing
176	<i>clec-210</i>	V-2F02	confirmed by sequencing

lab code	affected gene	Ahringer library ^a	remarks
178	<i>clcc-211</i>	V-5B19	confirmed by sequencing
276	<i>clcc-212</i>	V-5B21	confirmed by sequencing
278	<i>clcc-213</i>	V-5D11	obtained separately
179	<i>clcc-214</i>	V-5D07	confirmed by sequencing
303	<i>clcc-215</i>	V-14O18	confirmed by sequencing
200	<i>clcc-216</i>	V-5B08	confirmed by sequencing
196	<i>clcc-217</i>	V-5B12	confirmed by sequencing
277	<i>clcc-221</i>	V-7D13	confirmed by sequencing
220	<i>clcc-222</i>	V-7N23	confirmed by sequencing
109	<i>clcc-226</i>	V-9O03	confirmed by sequencing
328	<i>clcc-227</i>	V-9D09	confirmed by sequencing
304	<i>clcc-228</i>	V-16G24	confirmed by sequencing
329	<i>clcc-230</i>	V-10G19	confirmed by sequencing
131	<i>clcc-231</i>	V-10G17	confirmed by sequencing
133	<i>clcc-232</i>	V-10J05	confirmed by sequencing
280	<i>clcc-234</i>	V-11A13	confirmed by sequencing
112	<i>clcc-238</i>	V-11G20	confirmed by sequencing
281	<i>clcc-239</i>	V-11G18	confirmed by sequencing
305	<i>clcc-241</i>	V-16J02	confirmed by sequencing
107	<i>clcc-242</i>	V-11F15	confirmed by sequencing
292	<i>clcc-245</i>	n. a.	obtained separately
289	<i>clcc-246^c</i>	n. a.	Marc Vidal Library ORF ID (WS9): F16H6.2; confirmed by sequencing
116	<i>clcc-247</i>	V-12C13	confirmed by sequencing
119	<i>clcc-248</i>	V-12C23	confirmed by sequencing
202	<i>clcc-250</i>	V-12A14	confirmed by sequencing
203	<i>clcc-251</i>	V-12A10	
283	<i>clcc-254</i>	V-12A04	confirmed by sequencing
204	<i>clcc-255</i>	V-12O19	confirmed by sequencing
195	<i>clcc-259</i>	V-12F20	confirmed by sequencing
108	<i>clcc-261</i>	V-12P14	confirmed by sequencing
284	<i>clcc-262</i>	V-12P08	confirmed by sequencing
306	<i>clcc-263</i>	V-14B17	confirmed by sequencing
173	<i>clcc-264</i>	V-13A16	confirmed by sequencing
330	<i>C06B8.7</i>	V-10O19	confirmed by sequencing
296	<i>C43H6.6</i>	n. a.	confirmed by sequencing
230	<i>C49C3.11</i>	IV-8C24	confirmed by sequencing
9	<i>C54G4.4</i>	I-3P06	confirmed by sequencing
337	<i>cwp-5</i>	X-6M19	confirmed by sequencing
324	<i>egas-4</i>	IV-3D06	
124	<i>F28B4.3</i>	X-2K13	confirmed by sequencing
111	<i>F40F4.6</i>	X-2M01	confirmed by sequencing
180	<i>F52E1.2</i>	V-6C01	
10	<i>F55H12.3</i>	I-4M02	confirmed by sequencing
139	<i>K04H8.2</i>	I-7K13	confirmed by sequencing
5	<i>lat-2</i>	II-3F02	confirmed by sequencing

lab code	affected gene	Ahringer library ^a	remarks
236	<i>oac-17</i>	I-6N13	confirmed by sequencing
325	<i>svh-1</i>	IV-4K11	
336	<i>T25C12.3</i>	X-5B23	confirmed by sequencing
294	<i>Y17D7B.5</i>	n. a.	confirmed by sequencing
225	<i>Y38H6C.8</i>	V-13E13	confirmed by sequencing

^a Well plate position of RNAi clones taken from basic Ahringer library (unless stated otherwise), clones in bold letters are part of the supplementary set.

^b n. a., not available.

^c Provided by the Ewbank lab (Olivier Zugasti).

Table 7. RNAi clones, which cause an obvious phenotype in wild type worms and were used as positive controls.

lab code	affected gene	Ahringer library ^a	phenotype	remarks
15	<i>plc-3</i>	II-7O15	sterile	confirmed by sequencing
16	<i>unc-22</i>	IV-6K06	twitching, paralyzed	confirmed by sequencing
355	<i>bli-3</i>	I-1A05	blistered	

^a Well plate position of RNAi clones taken from basic Ahringer library.

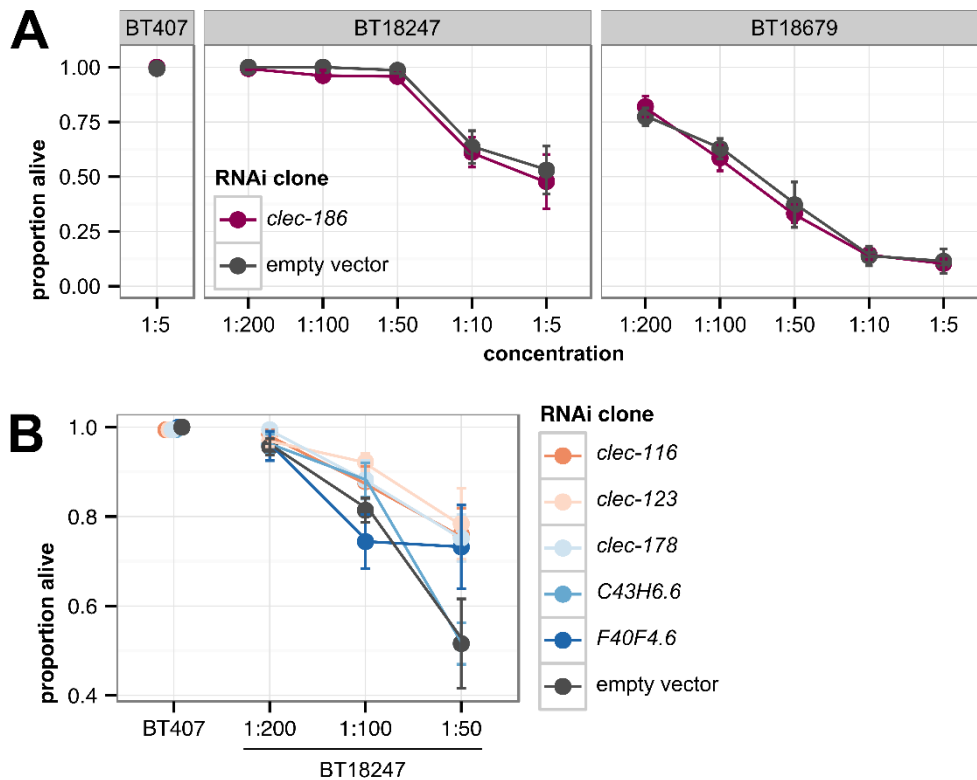


Figure 7. Survival assays of *clec* RNAi-treated *C. elegans* upon exposure to BT. Prior to infection wild type N2 worms were subjected to RNAi by feeding to knock-down specific *clec* genes (Table 6 and ¹¹). For each RNAi treatment 30 synchronized L4 larvae per plate were exposed to pathogenic BT18247 or BT18679 mixed with *E. coli* OP50 at OD 5 in serial dilutions in PBS (x-axes). Alive and dead animals were scored 24 hpi. Worms were considered to be dead when they did not respond to light touch. BT407 was used as non-pathogenic control in the highest BT concentration applied for BT18247 per assay. For a detailed description see Chapter III, “Materials and Methods”. Survival was defined as proportion of alive individuals (y-axes) and is displayed as means \pm SE of five replicates. Significant differences were determined with a GLM³, always compared against the wild type per bacterial treatment and survival assay, and corrected for multiple comparisons with Bonferroni⁴. All *P* values can be found in Table 8.(A) We were interested in further analyzing *clec-186* function, because its expression was found to be differentially expressed between the two pathogenic BT strains BT18247 and BT18679 in another study of our group (unpublished data). There was, however, no difference in survival between *clec-186* RNAi-treated and control worms, neither on BT18247, nor on BT18679. (B) Based on the results of the *clec* RNAi screen and additional characteristics collected from WormBase (Version WB250), five *clec* genes with at least one hit were chosen. The susceptibility phenotype of the *clec* RNAi screen (Figure 6) could not be confirmed in either of these genes, *clec-116*, *clec-123*, and *clec-178* even showed the opposite with being more resistant to BT.

Table 8. GLM³ on survival of *clec* knock-down strains per bacterial treatment and survival assay; *P* adjustment with Bonferroni⁴.

affected gene	RNAi clone	treatment	Figure	<i>P</i> value	significance level
<i>clec-186</i>	IV-6H02	BT18247	6A	0.105	n. s. ^a
<i>clec-186</i>	IV-6H02	BT18679	6A	0.574	n. s.
<i>clec-116</i>	I-7K11	BT18247	6B	<0.001	***
<i>clec-123</i>	II-2N12	BT18247	6B	<0.001	***
<i>clec-178</i>	IV-2N16	BT18247	6B	<0.001	***
<i>C43H6.6</i> ^b		BT18247	6B	1.000	n. s.
<i>F40F4.6</i>	X-2M01	BT18247	6B	0.312	n. s.

^a n. s., not significant.

^b Insert of original RNAi clone from Ahringer library could not be confirmed by sequencing; RNAi clone used here was separately received.

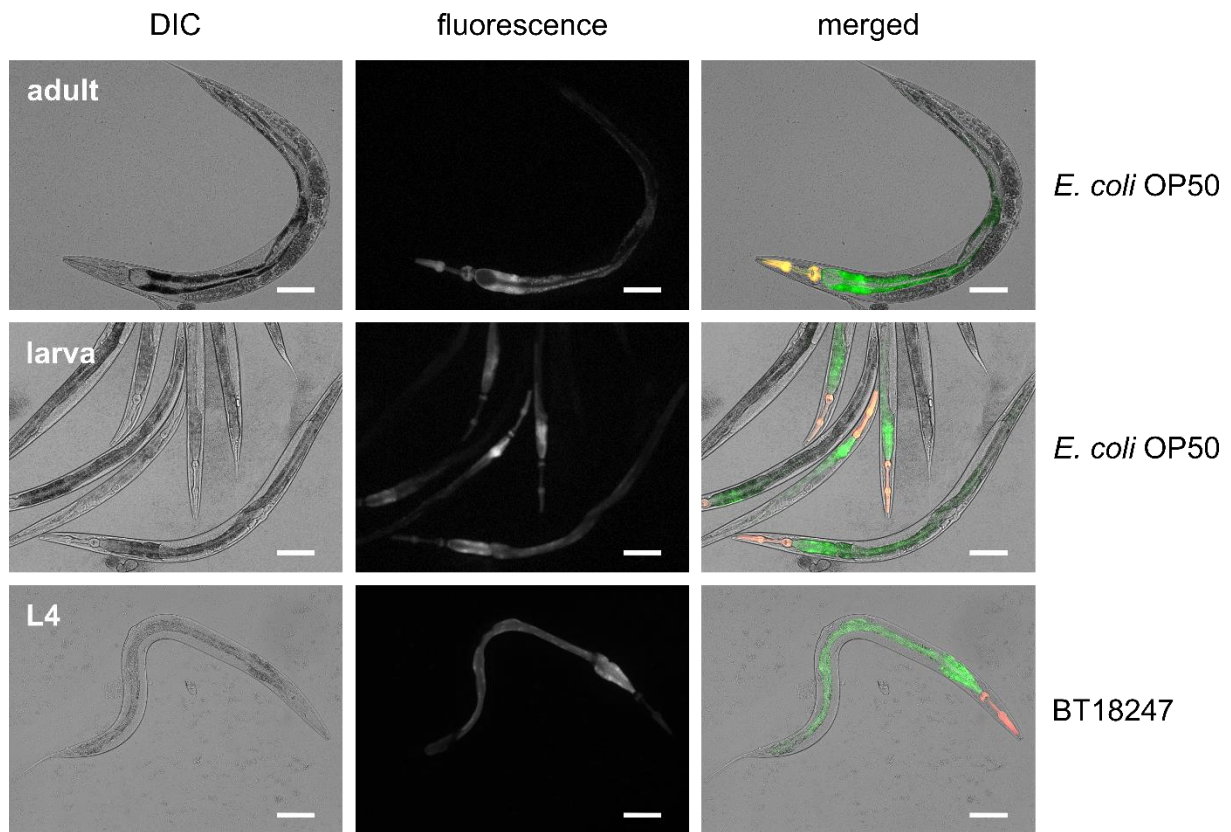


Figure 8. Expression of reporter strain *p_{clec-41}::GFP* (MY1077). Exemplary pictures of transgenic worms carrying the reporter construct *p_{clec-41}::GFP* and co-injection marker *p_{myo-2}::RFP* (expressed in pharynx). For more information about the generation of the transgenic strain see Table 9. Adult (upper panel) and larval stages (mid panel) exhibit constitutive expression of *clec-41* in cells of the intestine, both in the cytosol and the nuclei. No change of expression in worms infected with BT18247 (lower panel). The scale bars represent 100 μ m. Pictures were taken with Axio Observer Z.1 by Zeiss (Carl Zeiss AG, Jena, Germany). DIC, differential interference contrast; fluorescence, GFP filter; merged, both channels combined by using ImageJ 1.50d processing package Fiji (<http://fiji.sc/Fiji>).

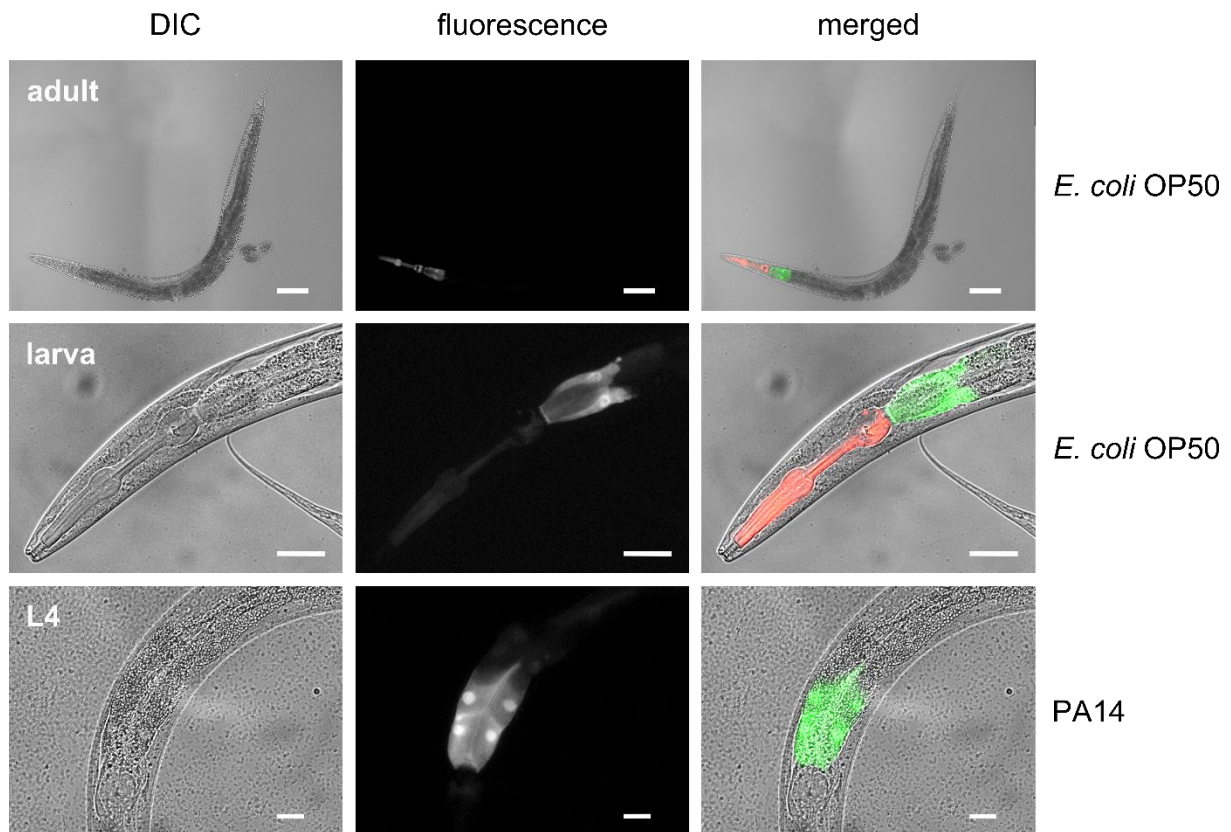


Figure 9. Expression of reporter strain *p_{clec-43}::GFP* (MY1040). Exemplary pictures of transgenic worms carrying reporter construct *p_{clec-43}::GFP* and co-injection marker *p_{myo-2}::RFP* (expressed in pharynx). For more information about the generation of the transgenic strain see Table 9. Adult (upper panel) and larval stages (mid panel) exhibit constitutive expression of *clec-43* exclusively in the first intestinal ring of the intestine int1, both in the cytosol and the nuclei. No change of expression in worms infected with PA14 (lower panel). The scale bars represent 100 μ m in the upper and 20 μ m in mid and lower panels. Pictures were taken with the confocal microscope LSM 700 (upper panel) and Axio Observer Z.1 by Zeiss (Carl Zeiss AG, Jena, Germany). DIC, differential interference contrast; fluorescence, GFP filter; merged, both channels combined by using ImageJ 1.50d processing package Fiji (<http://fiji.sc/Fiji>).

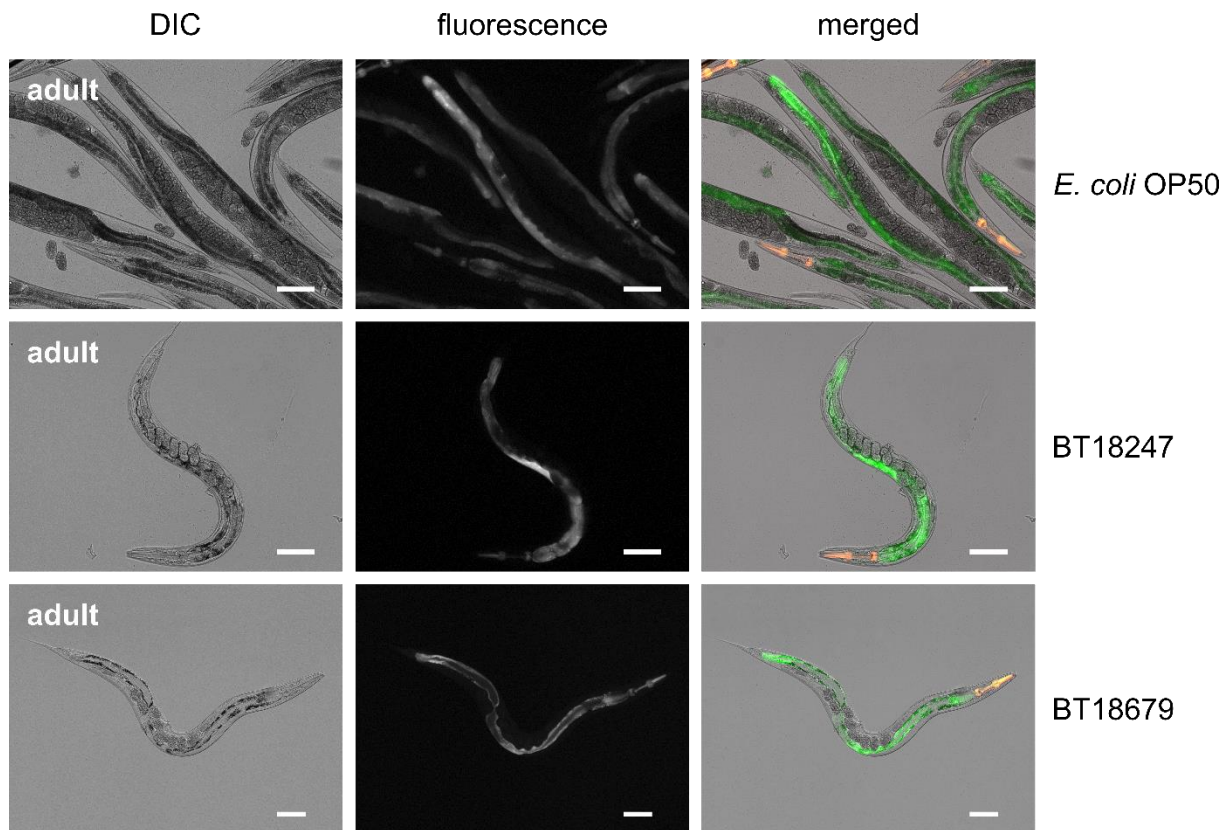


Figure 10. Expression of reporter strain *pclec-186::GFP* (MY1051). Exemplary pictures of transgenic worms carrying reporter construct *pclec-186::GFP* and co-injection marker *pmyo-2::RFP* (expressed in pharynx). For more information about the generation of the transgenic strain see Table 9. Adult and larval stages (not shown) exhibit constitutive expression of *clec-186* in cells of the intestine, both in the cytosol and the nuclei. No change of expression in worms infected with BT18247 (mid panel) nor BT18679 (lower panel). The scale bars represent 100 μ m. Pictures were taken with Axio Observer Z.1 by Zeiss (Carl Zeiss AG, Jena, Germany). DIC, differential interference contrast; fluorescence, GFP filter; merged, both channels combined by using ImageJ 1.50d processing package Fiji (<http://fiji.sc/Fiji>).

Table 9. Transgenic *clec* gene reporter strains and the primer sequences^a of the injected fusion construct¹³.

MY strain	genotype^b	primer A (5'→ 3')	primer A* (5'→ 3')	primer B (5'→ 3')^d
MY1038^c	wt; <i>yaEx33</i> (<i>pclec-8::GFP</i> ; <i>pmyo-2::RFP</i>)	accaatgccagcagtcacaatc (SHP582)	ctagcgttatggctccaagt (SHP583)	AGTCGACCTGCAGGCATGCAAGCTgattgaaatagtagtgc (SHP585)
MY1039	wt; <i>yaEx34</i> (<i>pclec-43::GFP</i> ; <i>pmyo-2::RFP</i>)			
MY1040	wt; <i>yaEx35</i> (<i>pclec-43::GFP</i> ; <i>pmyo-2::RFP</i>)	tctgtagttcgtgtctcagttagc (SHP569)	gcagtttattgaggctaccgatg (SHP568)	AGTCGACCTGCAGGCATGCAAGCTaatttcattttgcgattccttt (SHP570)
MY1041	wt; <i>yaEx36</i> (<i>pclec-43::GFP</i> ; <i>pmyo-2::RFP</i>)			
MY1042	wt; <i>yaEx37</i> (<i>pclec-227::GFP</i> ; <i>pmyo-2::RFP</i>)	acagcaccacacagtcgaatag (SHP573)	agacgaggaggagaagacgagtag (SHP572)	AGTCGACCTGCAGGCATGCAAGCTacagatcattacgcctgactgact (SHP574)
MY1043	wt; <i>yaEx38</i> (<i>pclec-227::GFP</i> ; <i>pmyo-2::RFP</i>)			
MY1044	wt; <i>yaEx39</i> (<i>clec-43::GFP</i> ; <i>pmyo-2::RFP</i>)			
MY1045	wt; <i>yaEx40</i> (<i>clec-43::GFP</i> ; <i>pmyo-2::RFP</i>)	tctgtagttcgtgtctcagttagc (SHP569)	gcagtttattgaggctaccgatg (SHP568)	AGTCGACCTGCAGGCATGCAAGCTagaaatattagttccataattgac (SHP571)
MY1046	wt; <i>yaEx41</i> (<i>clec-43::GFP</i> ; <i>pmyo-2::RFP</i>)			
MY1047	wt; <i>yaEx42</i> (<i>pC54G4.4::GFP</i> ; <i>pmyo-2::RFP</i>)	atccgccaacaactcttctatgct (SHP23)	attgccgtgcttaacagactcg (SHP24)	AGTCGACCTGCAGGCATGCAAGCTaagccgtgttttgcctgatcat (SHP25)
MY1048	wt; <i>yaEx43</i> (<i>pC54G4.4::GFP</i> ; <i>pmyo-2::RFP</i>)			
MY1051	wt; <i>yaEx44</i> (<i>pclec-186::GFP</i> ; <i>pmyo-2::RFP</i>)	attcgttcgggatggtcatg (SHP623)	ttgccctccgggataatac (SHP624)	AGTCGACCTGCAGGCATGCAAGCTcaaaaaagcagaaaataggagcat (SHP627)
MY1052	wt; <i>yaEx49</i> (<i>clec-4::GFP</i> ; <i>pmyo-2::RFP</i>)			
MY1054	wt; <i>yaEx50</i> (<i>clec-4::GFP</i> ; <i>pmyo-2::RFP</i>)	accagaagccgagaatccac (SHP415)	tttgaattccaccactgttcccac (SHP416)	AGTCGACCTGCAGGCATGCAAGCTatcgcgttctccactttctgtgt (SHP576)
MY1058	wt; <i>yaEx51</i> (<i>clec-4::GFP</i> ; <i>pmyo-2::RFP</i>)			
MY1053	wt; <i>yaEx45</i> (<i>clec-186::GFP</i> ; <i>pmyo-2::RFP</i>)	attcgttcgggatggtcatg (SHP623)	ttgccctccgggataatac (SHP624)	AGTCGACCTGCAGGCATGCAAGCTaaagtcacagaaacaactcc (SHP628)

MY strain	genotype	primer A (5'→ 3')	primer A* (5'→ 3')	primer B (5'→ 3')
MY1055	wt; <i>yaEx46</i> (<i>pC54G4.4::GFP</i> ; <i>pmyo-2::RFP</i>)			
MY1056	wt; <i>yaEx47</i> (<i>pC54G4.4::GFP</i> ; <i>pmyo-2::RFP</i>)	ttgaaatacagcagcggctcc (SHP618)	atccgccacaactcttctatgctc (SHP23)	AGTCGACCTGCAGGCATGCAAGCTaagccgtgttttctctgatcat (SHP25)
MY1057	wt; <i>yaEx48</i> (<i>pC54G4.4::GFP</i> ; <i>pmyo-2::RFP</i>)			
MY1062	wt; <i>yaEx55</i> (<i>clec-8::GFP</i> ; <i>pmyo-2::RFP</i>)			
MY1063	wt; <i>yaEx56</i> (<i>clec-8::GFP</i> ; <i>pmyo-2::RFP</i>)	accaatgccagcagtcctcaatc (SHP582)	ctagcgttatggctccaagt (SHP583)	AGTCGACCTGCAGGCATGCAAGCTgagcattgaacaaatttcca (SHP584)
MY1064	wt; <i>yaEx57</i> (<i>clec-8::GFP</i> ; <i>pmyo-2::RFP</i>)			
MY1069	wt; <i>yaEx62</i> (<i>pclec-4::GFP</i> ; <i>pmyo-2::RFP</i>)			
MY1070	wt; <i>yaEx63</i> (<i>pclec-4::GFP</i> ; <i>pmyo-2::RFP</i>)	cattcttcgccaaccgcaac (SHP729)	accagaagccgagaatccac (SHP415)	AGTCGACCTGCAGGCATGCAAGCTaagtaggacggaataaaattgcat (SHP417)
MY1071	wt; <i>yaEx64</i> (<i>pclec-4::GFP</i> ; <i>pmyo-2::RFP</i>)			
MY1075	wt; <i>yaEx68</i> (<i>pclec-41::GFP</i> ; <i>pmyo-2::RFP</i>)			
MY1076	wt; <i>yaEx69</i> (<i>pclec-41::GFP</i> ; <i>pmyo-2::RFP</i>)	ccacttgctcttccatgctg (SHP730)	gcttccatcaatctcgattcc (SHP731)	AGTCGACCTGCAGGCATGCAAGCTaagtaaagtagagtgaaccacat (SHP732)
MY1077	wt; <i>yaEx70</i> (<i>pclec-41::GFP</i> ; <i>pmyo-2::RFP</i>)			

^a Primers for *gfp* coding sequence plus 3'-UTR of *unc-54* amplified from Fire vector pPD95.75: primer C, 5'-agcttgcatgcctgcaggtcgact-3' (SHP408), primer D, 5'-aaggccctacggccgactagtagg-3' (SHP409), primer D*, 5'-ggaacagttatgtttgtatattggg-3' (SHP410).

^b GFP construct injected at a concentration of approx. 10 ng/μl, RFP co-injection marker at a concentration of 25 ng/μl.

^c Injected by Eike Strathmann; transgenic strain not available anymore.

^d Capital letters, overlapping sequence to *gfp* vector.

Table 10. Transgenic strains crossed in *clec-43(ok3188)* mutant background.

MY strain	genotype	description
MY1049	<i>clec-43(ok3188)</i> II; <i>kcIs6 [ifb-2::CFP]</i> IV	<i>clec-43(ok3188)</i> crossed with <i>kcIs6 [ifb-2::CFP]</i> (provided by Olaf Bossinger); for more details see ¹⁴
MY1050	<i>clec-43(ok3188)</i> II; <i>dkIs92[pvha-6::GFP:rab-8]</i>	<i>clec-43(ok3188)</i> crossed with <i>dkIs92[pvha-6::GFP:rab-8]</i> (provided by Miyuki Sato); for more details see ¹⁵

Table 11. Transgenic *clec* gene rescue strains and the primer sequences of the injected amplicon or fusion construct¹³.

MY strain	genotype	injected conc ^a	primer A (5'→3')	primer A* (5'→3')	primer B (5'→3') ^d
MY1065	<i>clec-43(ok3188);yaEx58</i> (<i>clec-43</i> ; <i>pmyo-2::RFP</i>)	10 ng/μl	gcagtttattgaggctaccgatg (SHP568)		gttattggagagctgcca (SHP515)
MY1059 ^b	<i>C54G4.4(ok2110);yaEx52</i> (<i>C54G4.4</i> ; <i>pmyo-2::RFP</i>)				
MY1060 ^b	<i>C54G4.4(ok2110);yaEx53</i> (<i>C54G4.4</i> ; <i>pmyo-2::RFP</i>)	50 ng/μl	ttgaaatacagcagcgggcc (SHP618)		tcaccccatctttctgtgattg (SHP691)
MY1061 ^b	<i>C54G4.4(ok2110);yaEx54</i> (<i>C54G4.4</i> ; <i>pmyo-2::RFP</i>)				
MY1072	<i>C54G4.4(ok2110);yaEx65</i> (<i>C54G4.4</i> ; <i>pmyo-2::RFP</i>)				
MY1073	<i>C54G4.4(ok2110);yaEx66</i> (<i>C54G4.4</i> ; <i>pmyo-2::RFP</i>)	100 ng/μl	ttgaaatacagcagcgggcc (SHP618)		tcaccccatctttctgtgattg (SHP691)
MY1074	<i>C54G4.4(ok2110);yaEx67</i> (<i>C54G4.4</i> ; <i>pmyo-2::RFP</i>)				
MY1066 ^c	<i>clec-4(ok2050);yaEx59</i> (<i>clec-4::3'UTR_unc-54</i> ; <i>pmyo-2::RFP</i>)				
MY1067 ^c	<i>clec-4(ok2050);yaEx60</i> (<i>clec-4::3'UTR_unc-54</i> ; <i>pmyo-2::RFP</i>)	50 ng/μl	accagaagccgagaatccac (SHP415)	ttgaaattccaccactgttcccac (SHP416)	AGTCAGAGGCACGGGCGCGAGATG ttaatcggttctccactttctgt (SHP728)
MY1068 ^c	<i>clec-4(ok2050);yaEx61</i> (<i>clec-4::3'UTR_unc-54</i> ; <i>pmyo-2::RFP</i>)				

^a Approx. concentration of injected amplicon or fusion construct; RFP co-injection marker injected at a concentration of 25 ng/μl.

^b Injected by Anke Kloock.

^c Construct of *clec-4* gene fused to 3'-UTR of *unc-54* amplified from Fire vector pPD95.75: primer C, 5'-catctcgccccgtcctctgact-3' (SHP724), primer D, 5'-aaggccccctacggcggcactagtagg-3' (SHP409), primer D*, 5'-ggaaacagttatgttggatatattggg-3' (SHP410).

^d Capital letters, overlapping sequence to *unc-54* 3'-UTR.

References – Additional Results

1. Whitlock, M. C. & Schluter, D. *The analysis of biological data*. (Roberts and Company Publishers, 2009).
2. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**, 289–300 (1995).
3. Nelder, J. A. & Wedderburn, R. W. M. Generalized linear models. *J. R. Stat. Soc. Ser. Gen.* **135**, 370 (1972).
4. Dunn, O. J. Multiple comparisons among means. *J. Am. Stat. Assoc.* **56**, 52 (1961).
5. Kaplan, E. L. & Meier, P. Nonparametric estimation from incomplete observations. *J. Am. Stat. Assoc.* **53**, 457 (1958).
6. Harrington, D. in *Encyclopedia of Biostatistics* (John Wiley & Sons, Ltd, 2005).
7. Tukey, J. W. Comparing individual means in the analysis of variance. *Biometrics* **5**, 99 (1949).
8. Bolker, B. M. *et al.* Generalized linear mixed models: A practical guide for ecology and evolution. *Trends Ecol. Evol.* **24**, 127–135 (2009).
9. Rual, J.-F. *et al.* Toward improving *Caenorhabditis elegans* phenome mapping with an ORFeome-based RNAi library. *Genome Res.* **14**, 2162–2168 (2004).
10. Pujol, N. *et al.* Distinct innate immune responses to infection and wounding in the *C. elegans* epidermis. *Curr. Biol.* **18**, 481–489 (2008).
11. Ahringer, J. Reverse genetics. *WormBook* (2006). doi:10.1895/wormbook.1.47.1
12. Zhang, X. D. Illustration of SSMD, z score, SSMD*, z* score, and t statistic for hit selection in RNAi high-throughput screens. *J. Biomol. Screen.* **16**, 775–785 (2011).
13. Hobert, O. PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic *C. elegans*. *BioTechniques* **32**, 728–730 (2002).
14. Hüsken, K. *et al.* Maintenance of the intestinal tube in *Caenorhabditis elegans*: The role of the intermediate filament protein IFC-2. *Differentiation* **76**, 881–896 (2008).
15. Sato, T. *et al.* The Rab8 GTPase regulates apical protein localization in intestinal cells. *Nature* **448**, 366–369 (2007).

List of Abbreviations

ADF	amphid with dual ciliated sensory endings
AIY	amphid interneuron
AMP	antimicrobial peptide
ASABF	<i>Ascaris suum</i> antimicrobial factor
ATF	cAMP-dependent transcription factor
ATFS	activating transcription factor associated with stress
AWB	amphid wing "B"
AWC	amphid wing "C"
BRE	<i>Bacillus</i> -toxin resistant
BT	<i>Bacillus thuringiensis</i>
cAMP	cyclic adenosine monophosphate
Cas	CRISPR associated protein
CBB	Coomassie brilliant blue
CCP	complement control module
Ce-ABF	<i>C. elegans</i> antimicrobial factor
CGC	Caenorhabditis Genetics Center
CLEC	C-type lectin
CLR	C-type lectin receptor
CNC	caenacin
collectin	collagen-containing C-type lectin
CRD	carbohydrate-recognition domain
CRISPR	clustered regularly interspaced short palindromic repeats
Cry	crystal
CTLD	C-type lectin-like domain
CUB	C1r/C1s Uegf Bmp1
CW	conserved cysteine and tryptophan residues
DAF	abnormal dauer formation
DAMP	damage-associated molecular pattern
DCAR	dihydrocaffeic acid receptor
Dectin	dendritic cell-associated C-type lectin
DIC	differential interference contrast
DL	<i>Drosophila</i> lectin
DNGR	dendritic cell natural killer lectin group receptor
DRH	dicer related helicase
ELT	erythroid-like transcription factor family
ERK	extracellular signal-regulated kinase
FDR	false discovery rate
FoxO	Forkhead box protein class O
FREP	fibrinogen-related protein
FSHR	mammalian follicle stimulating hormone receptor
GFP	green fluorescent protein
GlcNAc	N-acetyl-glucosamine
GLH	germline RNA helicase
GLM	generalized linear model
GLMM	generalized linear mixed model
GPCR	G protein-coupled receptor
HECT	homologous to the E6-AP carboxyl terminus
HECW	HECT domain-containing E3 ubiquitin ligase
HLH	helix loop helix
HPLA	4-hydroxyphenyllactic acid

Ig	immunoglobulin
ILR	insulin-like receptor
ILYS	invertebrate lysozyme
INS	insulin-related protein
IRE	IRE kinase-related protein
ITAM	immunoreceptor tyrosine-based activation motif
ITIM	immunoreceptor tyrosine-based inhibitory motif
JNK	c-Jun N-terminal kinase
KGB	GLH binding kinase
LB	lysogeny broth
LPS	lipopolysaccharide
LYS	lysozyme
MAD	median of absolute deviation
MAMP	microbe-associated molecular pattern
MAPK	p38 mitogen-activated protein kinase
MASP	MBL-associated serine protease
MBL	mannose -binding lectin
MDL	myeloid DAP12-associated lectin
MEK-1	dual specificity MAPK kinase
MICL	myeloid inhibitory C-type lectin-like receptor
Mincle	macrophage-inducible C-type lectin
MR	mannose receptor
MyD88	myeloid differentiation primary response gene 88
NBRP	National BioResource Project
NF-κB	nuclear factor κ -light-chain-enhancer of activated B cells
NGM	nematode growth medium
NLP	neuropeptide-like peptide
NLR	NOD-like receptor
NOD	nucleotide-binding oligomerization domain
NPR	neuropeptide receptor
NSY	neuronal symmetry
OCTR	octopamine receptor
OD	optical density
OLL	outer labial lateral sensilla
PA	<i>Pseudomonas aeruginosa</i>
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PFM	peptone-free medium
PFT	pore-forming toxin
PMK	p38 MAPK
PRR	pattern recognition receptor
qRT-PCR	quantitative real-time PCR
Reg	regenerating gene family protein
RFP	red fluorescent protein
RIG	retinoic acid-inducible protein
RLR	retinoic acid-inducible gene-I-like receptor
RNA	ribonucleic acid
ROS	reactive oxygen species
SAPK	stress-activated protein kinase
SD	standard deviation
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SE	standard error
SEK	SAPK/ERK kinase

SPP	saposin-like protein
Syk	spleen tyrosine kinase
TIR	Toll-interleukin receptor
TLR	Toll-like receptor
TOL	TOLI (<i>Drosophila</i>) family
TPH	tryptophan hydroxylase
UPR	unfolded protein response
UPR^{mt}	mitochondrial UPR
UTR	untranslated region
VCBP	variable region-containing chitin-binding protein
VHP	VH1 dual specificity phosphatase family protein
VWA	von Willebrand factor A
XBP	x-box binding protein
YPD	yeast extract peptone dextrose
ZIP	bZIP transcription factor family

