Zoology 119 (2016) 244–253

Contents lists available at ScienceDirect

Zoology

journal homepage: www.elsevier.com/locate/zool

GATA transcription factor as a likely key regulator of the *Caenorhabditis elegans* innate immune response against gut pathogens☆

Wentao Yang[a], Katja Dierking[a], Philip C. Rosenstiel[b], Hinrich Schulenburg[a,*]

[a] Department of Evolutionary Ecology and Genetics, Zoological Institute, University of Kiel, D-24098 Kiel, Germany
[b] Institute for Clinical Molecular Biology, University of Kiel, D-24098 Kiel, Germany

ARTICLE INFO

Article history:
Received 11 November 2015
Received in revised form 19 April 2016
Accepted 27 May 2016
Available online 27 May 2016

Keywords:
Caenorhabditis elegans
Innate immunity
GATA transcription factor
RNA-Seq
WormExp

ABSTRACT

Invertebrate defence against pathogens exclusively relies on components of the innate immune system. Comprehensive information has been collected over the last years on the molecular components of invertebrate immunity and the involved signalling processes, especially for the main invertebrate model species, the fruitfly *Drosophila melanogaster* and the nematode *Caenorhabditis elegans*. Yet, the exact regulation of general and specific defences is still not well understood. In the current study, we take advantage of a recently established database, WormExp, which combines all available gene expression studies for *C. elegans*, in order to explore commonalities and differences in the regulation of nematode immune defence against a large variety of pathogens versus food microbes. We identified significant overlaps in the transcriptional response towards microbes, especially pathogenic bacteria. We also found that the GATA motif is overrepresented in many microbe-induced gene sets and in targets of other previously identified regulators of worm immunity. Moreover, the activated targets of one of the known *C. elegans* GATA transcription factors, ELT-2, are significantly enriched in the gene sets, which are differentially regulated by gut-infecting pathogens. These findings strongly suggest that GATA transcription factors and particularly ELT-2 play a central role in regulating the *C. elegans* immune response against gut pathogens. More specific responses to distinct pathogens may be mediated by additional transcription factors, either acting alone or jointly with GATA transcription factors. Taken together, our analysis of the worm's transcriptional response to microbes provides a new perspective on the *C. elegans* immune system, which we propose to be coordinated by GATA transcription factor ELT-2 in the gut.

1. Introduction

Pathogens are ubiquitous. They comprise different species and genotypes that vary among locations and/or across time. They may even specifically adapt to host defences. In turn, host organisms are expected to adapt to these pathogen challenges, for example through the evolution of counter-adaptations to coevolving pathogen varieties and/or by tight regulation of immune responses, which allows individual hosts to react flexibly and quickly to an unpredictable pathogen threat. As a consequence, the immune system is expected to have a complex architecture that is fine-tuned to perceive a variety of pathogen-related signals in order to mount an appropriate and possibly highly specific defence response (Hughes and Nei, 1988; Sackton et al., 2007; Schulenburg et al., 2009). Comprehensive information is available on the complex organization of the immune system of higher vertebrates, which consists of innate and adaptive responses (Janeway et al., 2001). For invertebrates, the most detailed data sets are currently available for two model species, the fruitfly *Drosophila melanogaster* and the roundworm *Caenorhabditis elegans*, highlighting the involvement of several signalling cascades and various immune effectors (Buchon et al., 2014; Cohen and Troemel, 2015). Yet to date it is still unclear how exactly invertebrate immune responses are coordinated by general and possibly also by more specific regulators. To address this topic, we here focus on the nematode *C. elegans* and explore a recently established database, WormExp, which encompasses all gene expression studies for this organism (Yang et al., 2015b).

The nematode *C. elegans* has become a central model for dissecting the genetics of invertebrate immunity. It can be infected by various pathogens via several distinct infection routes (reviewed in Powell and Ausubel, 2008; Engelmann and Pujol, 2010; Marsh

☆ This article is part of a special issue entitled "Host-parasite coevolution - rapid reciprocal adaptation and its genetic basis".

* Corresponding author.
E-mail address: hschulenburg@zoologie.uni-kiel.de (H. Schulenburg).

http://dx.doi.org/10.1016/j.zool.2016.05.013

and May, 2012; Clark and Hodgkin, 2014), and a comprehensive toolbox is available for functional genetic analysis (Irazoqui et al., 2010b). The previous studies revealed that the nematode immune system is based on several signalling pathways conserved across invertebrates and vertebrates, including the p38 mitogen-activated protein kinase (MAPK), c-Jun N-terminal kinase (JNK) MAPK, extracellular signal-regulated kinase (ERK) MAPK, transforming growth factor-β (TGF-β), and also the insulin-like receptor (ILR) pathways (Engelmann and Pujol, 2010; Irazoqui et al., 2010a; Pukkila-Worley and Ausubel, 2012). Several transcription factors have been found to contribute to *C. elegans* immune defence, such as the GATA transcription factor ELT-2 (Shapira et al., 2006), the basic-region leucine zipper (bZIP) transcription factors ATF-7 (Shivers et al., 2010), ATFS-1 (Pellegrino et al., 2014), ZIP-2 (Estes et al., 2010), and SKN-1 (Papp et al., 2012), the basic helix-loop-helix (bHLH) transcription factor HLH-30 (Visvikis et al., 2014), the signal transducer and activator of transcription (STAT)-like transcription factor STA-2 (Dierking et al., 2011), and the activator protein 1 (AP-1) transcription factor dimer JUN-1/FOS-1 (Kao et al., 2011). Moreover, pathogen elimination involves certain antimicrobial peptides (reviewed in Dierking et al., 2016), including, for example, the caenacins and related peptides (Couillault et al., 2004), the caenopores (Mysliwy et al., 2010; Roeder et al., 2010), and additionally the generation of reactive oxygen species (ROS) (Chávez et al., 2009; Van Der Hoeven et al., 2011). While it remains unclear if and how pathogens are directly recognized by *C. elegans*, nematode defence can also be activated indirectly through a cellular surveillance system and/or damage signals, allowing the worms to respond to the cellular disturbance caused by an infection (Melo and Ruvkun, 2012; Zugasti et al., 2014; Ewbank and Pujol, 2016).

Although most studies on *C. elegans* immunity rely on functional genetic approaches, numerous transcriptomic and proteomic analyses were additionally performed to explore the set of genes which are activated or repressed upon pathogen exposure. These expression analyses considered a total of 21 pathogens and 6 non-pathogenic bacterial strains (see Section 3.1 and Table S1 in the supplementary online Appendix). Several of these studies involved more than one pathogen and showed an overlapping signature in the response to the various bacteria (Wong et al., 2007; Irazoqui et al., 2010a), indicating the presence of a common regulatory mechanism in the worm's immune system. Moreover, several expression analyses explored the downstream targets of immunity pathways, allowing us to assess the presence of shared or divergent signatures in the differentially expressed gene sets. Such expression analyses have been performed with mutants of the p38 MAPK, JNK MAPK, TGF-β and the ILR cascades (for example *pmk-1*(*km25*) for p38 MAPK (Bond et al., 2014), *jun-1*(*gk557*) for JNK (Uno et al., 2013), *dbl-1*(*ctIs40*) overexpression for TGF-β (Roberts et al., 2010) and *daf-16*(*mu86*) for ILR (Murphy et al., 2003)). Similar expression analysis was performed for mutants of GATA transcription factors, including a mutant for the GATA transcription factor gene *elt-2*, which is known to contribute to immunity against some gut-infecting pathogens such as *Pseudomonas aeruginosa, Salmonella typhimurium*, *Enterococcus faecalis, Cryptococcus neoformans* or *Burkholderia pseudomallei* (Kerry et al., 2006; Shapira et al., 2006; Lee et al., 2013). ELT-2 is the major transcriptional regulator in the *C. elegans* intestine, controlling the constitutive expression of most of the genes necessary for maintenance of intestinal function (McGhee et al., 2009). It was recently suggested to act as a master regulator for additionally activated immune defence responses in the intestine (Block and Shapira, 2015). The availability of numerous immunity-related gene expression studies provides a unique opportunity to explore the regulation of common and specific responses to pathogens.

The aim of the current article is to use the available gene expression studies to identify common and specific regulators of the *C. elegans* immune response to pathogens. To date, such an integrative analysis, which considers all available microbe- and immunity-related transcriptome data sets, has not yet been performed. It allows us to assess the following questions: (i) How similar or how different are the inducible responses to various microbes? (ii) How similar or how different are the responses to the gene sets controlled by distinct immunity regulators and pathways? These assessments may yield novel insights into the regulation of the worm's response to pathogens and also to non-pathogenic microbes. We specifically consider transcription factors and characterize the presence of transcription factor binding motifs in the promoter regions of differentially expressed genes. We take advantage of a recently established database, WormExp, in which we collated all of the more than 1800 available *C. elegans* gene expression data sets (Yang et al., 2015b). We used this database to extract gene sets that are differentially regulated upon exposure to pathogens and non-pathogenic bacteria. We further assessed gene expression data sets that comprise the downstream targets of transcription factors or of known immunity pathway components. The presence of common or specific regulators is analysed with the help of different statistical methods, including gene set enrichment (Subramanian et al., 2005) and de novo transcription factor-binding motif analyses (Shi et al., 2011).

2. Materials and methods

2.1. Data sets

All transcriptome and proteome data sets are available from WormExp (http://wormexp.zoologie.uni-kiel.de/wormexp/) (Yang et al., 2015b). We specifically focused on 30 studies in which the gene expression response to microbes was analysed (see Table S1, links to references in column I). We additionally considered gene sets from gene expression analysis of mutants of known immunity regulators and immunity pathway components, including, for example, mutants for the GATA transcription factors *elt-1*, *elt-2*, and *elt-3* (all based on microarrays). Information on confirmed transcription factor binding sites, inferred through ChIP-Seq (Furey, 2012), were obtained from the modencode database (http://www.modencode.org/) (Gerstein et al., 2010).

2.2. Analysis

Gene set enrichment analyses were performed with the help of Fisher exact tests, as implemented in WormExp. To identify possible transcription factors that mediate differential expression, we searched for enriched motifs in the promoter regions of a particular gene set with the program AMD (Shi et al., 2011). We focused on the core promoter region, which is widely assumed to range from 600 bp (base pair) upstream up to 250 bp downstream of the transcription start site (TSS) (Michaloski et al., 2006; Michaloski et al., 2011; Tabach et al., 2007; Trinklein et al., 2003) and which is also often most conserved across genes (Cheung et al., 2007) and thus suitable for comparative sequence analysis. We did not consider a larger region because the core promoter region should be most relevant for general regulatory elements and also because extraction of larger intergenic regions is not always straightforward in *C. elegans* due to the comparatively high gene density in the nematode genome (Hillier et al., 2005). The core promoter region was thus assessed for all of the genes per gene set, using *C. elegans* genome version WS235 from wormbase (http://www.wormbase.org) (Stein et al., 2001). ChIP-Seq data of transcription factor binding sites were directly used in motif analysis, based on genome version WS220. Motif logos are produced by Weblogo 3.0 (Crooks et al., 2004).

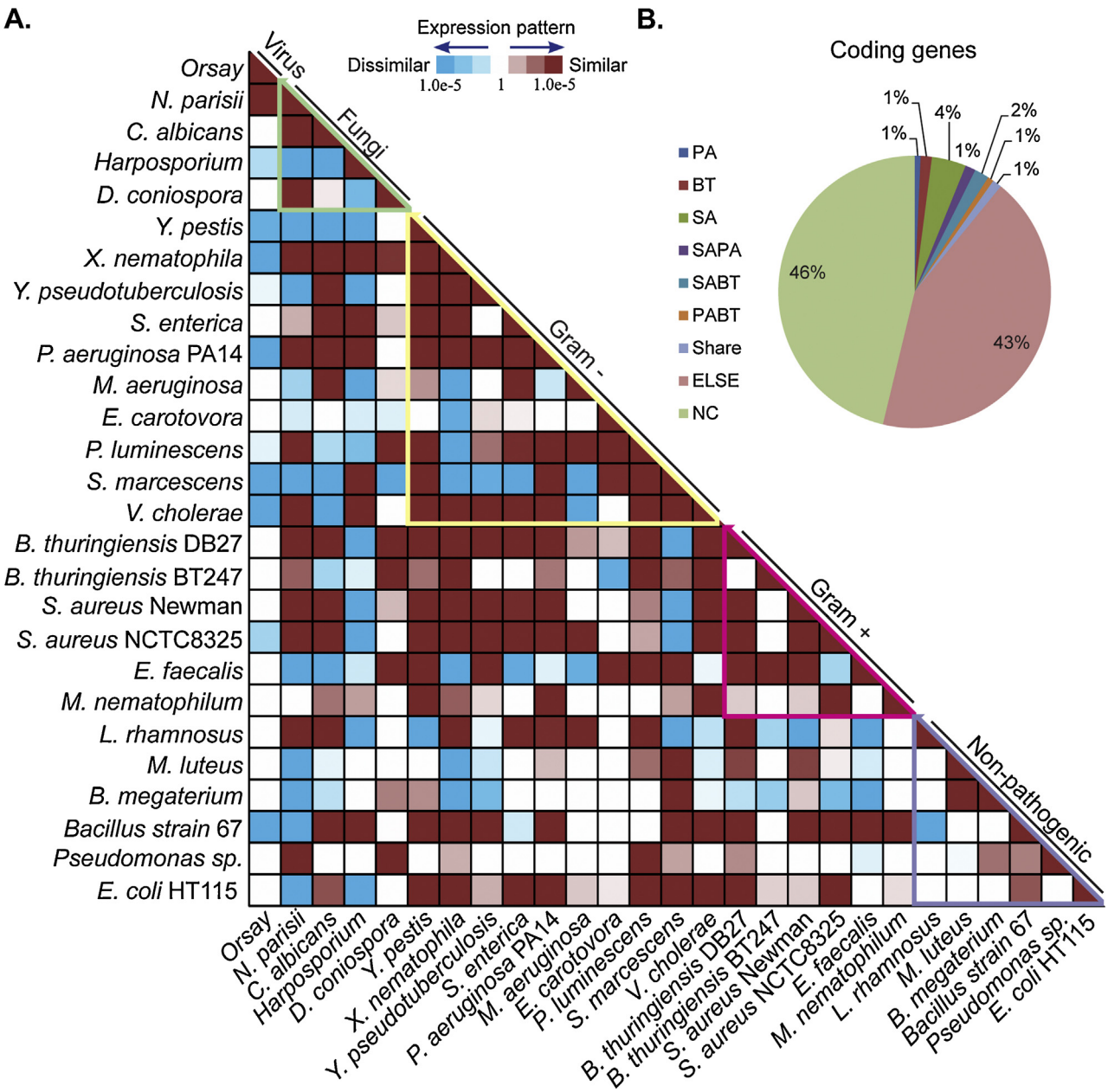

Fig. 1. Overlap among differentially expressed genes in response to pathogens and food bacteria. (A) Heat map for the significance of overlapping gene sets responding to virus, fungi, Gram-negative and Gram-positive bacteria, and also to non-pathogenic bacteria. The significance of overlapping gene sets is measured by a Fisher exact test with Bonferroni correction and indicated by colour intensity (see legend on top). Red indicates all overlaps for which genes show identical up- or downregulation, whereas blue represents all overlaps with opposite expression patterns (up in one case, down in the other). This illustration only shows the results for one data set per microbe strain in order to keep the overview manageable. For each pathogen, that data set has been included which produced the largest overlap with other data sets, in order to emphasise the presence of similar transcriptomic responses. Full results are shown in Fig. S1. (B) Percentage of all *C. elegans* coding genes influenced by pathogens. Number of coding genes is based on Wormbase version WS235. The colours highlight different pathogens or different combinations of pathogens, as indicated by the abbreviations in the legend on the left. For example, the orange colour is labelled PABT and indicates the overlapping gene set differentially regulated by both *Pseudomonas aeruginosa* and *Bacillus thuringiensis*. Abbreviations: NC, not changed by pathogens; PA, *P. aeruginosa*; BT, *B. thuringiensis*; SA, *Staphylococcus aureus*; Share, overlap among SA, PA and BT.

3. Results

3.1. Overview of data sets used for analysis of the C. elegans response towards microbes

Our analysis focused on 30 previously published gene expression studies (Table S1). Of these, 23 considered the response to pathogens (virus, bacteria, and fungi) and 7 to non-pathogenic bacteria. From these studies, we extracted a total of 127 individual transcriptome data sets (in several cases, more than one set from the same study), for which the transcriptional response was examined at different time points after initial exposure (Table S1). These studies compared the response towards the test microbe with that towards a control bacterium, usually the standard laboratory food strain *Escherichia coli* OP50, or, in a few cases, a non-pathogenic strain of the pathogen species under study. Most of these expression data sets were directly taken from the original studies, while in a few cases the significantly differentially expressed gene sets were inferred by us from re-analysis of the raw data (Table S1). The number of significantly differentially expressed genes varied among studies from 12 to 4051 (Table S1). Several independent studies used the same pathogen strain (e.g., *Pseudomonas aeruginosa* PA14,

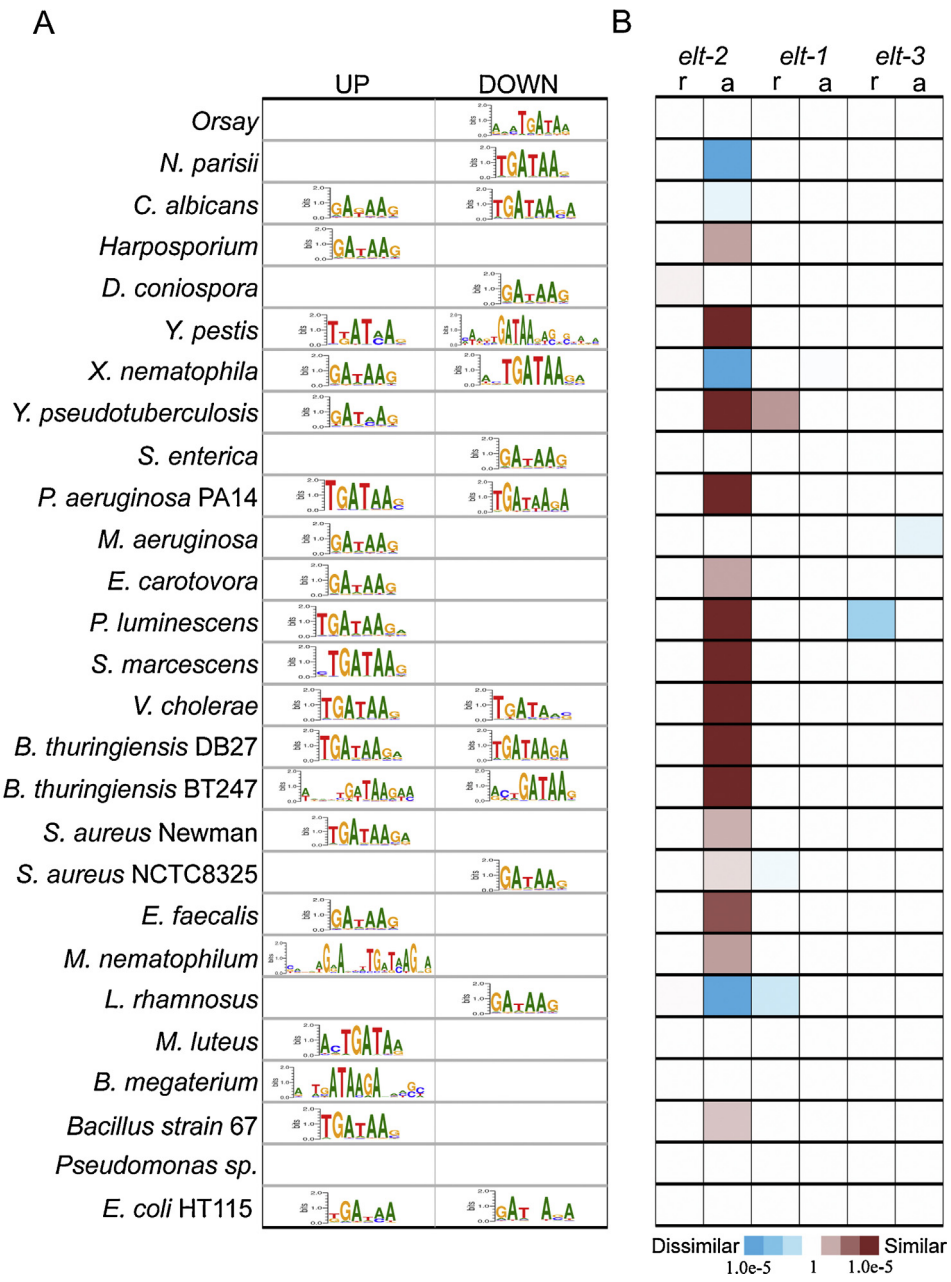

Fig. 2. Enriched GATA motifs and GATA transcription factor target genes for microbe-responsive gene sets. (A) The GATA motif is enriched for most microbe-related gene sets (either upregulated or downregulated). Microbes are presented in the same order as in Fig. 1. Further results are given in Table S1. (B) Only the gene set activated by the GATA transcription factor gene *elt-2* (but not *elt-1* or *elt-3*) produces a significant overlap with the microbe-responsive genes. Colour scale shows the significance of Bonferroni-corrected Fisher exact probabilities and the direction of enrichment, as in Fig. 1. Abbreviations: a, activated genes; r, repressed genes.

Bacillus thuringiensis BT247, and *Staphylococcus aureus* NCTC8325; Table S1), thus allowing us to evaluate consistency of inducible gene expression against specific bacteria. For our analysis, we excluded all data sets based on proteomic approaches, because these are not really comparable to transcriptomic data due to methodological biases and usually much lower coverage and thus a much smaller number of significantly differentially expressed genes (for example Yang et al., 2015a). We also excluded one transcriptomic study which assessed the response to *Bacillus subtilis* and which did not yield a single significantly differentially expressed gene under the here considered adjusted *p*-value (adjusted to take account of multiple testing using the false discovery rate, fdr). Moreover, for the analysis of transcriptome overlaps and motif inference we only show the results for one data set per microbe strain in order to keep figures as simple and thus as accessible as possible to the reader. For these cases, we chose the data set which produced the largest number of overlaps with other data sets in order to emphasize the presence of similar transcriptomic responses to microbes. Nevertheless, we provide the full results, based on all of the data sets, in Fig. S1 in the supplementary online Appendix.

3.2. *Pathogen exposure produces overlapping gene expression signatures in* C. elegans

We first assessed the percentage of coding genes in the *C. elegans* genome that are influenced by pathogen exposure (i.e., changed expression in response to at least one pathogen). Based on all data sets available, we found that 54% of all *C. elegans* coding genes are differentially expressed upon exposure to one of the considered pathogens (Fig. 1B). 11% are still differentially expressed in at least

one of the data sets if we only consider the three bacterial pathogens *P. aeruginosa* (indicated by PA), *B. thuringiensis* (BT), and *S. aureus* (SA), for which several independent data sets are available. Overall, a surprisingly large number of *C. elegans* genes are responsive to pathogenic microbes.

We next asked to what extent exposure to microbes causes similar or divergent expression responses in *C. elegans*. We found numerous significant overlaps among the considered pathogens and microbes (Fig. 1A; see full results in Fig. S1). In many comparisons, the overlap in gene expression goes in the same direction (i.e., either upregulated or downregulated by both of the two compared pathogens, as indicated by red colour in Fig. 1A). For example, the only two intracellular pathogens, the Orsay virus and the microsporidian *Nematocida parisii*, both seem to overlap in the up- and downregulated gene sets, confirming previous observations (Bakowski et al., 2014). A similar same-direction overlap is also found for those cases for which multiple independent data sets are available for the same pathogen species, which is the case for *B. thuringiensis*, *P. aeruginosa*, and also *S. aureus* (Fig. S2 in the supplementary online Appendix). These results strongly suggest that these pathogens produce a very robust transcriptome response that is independent of the platform and laboratory involved. In contrast, in other comparisons between different pathogens the changes in gene expression are in opposite directions (i.e., up in one pathogen but down in the other, indicated by blue colour in Fig. 1A). For example, the genes upregulated by the Orsay virus seem to be downregulated by *Yersinia pestis* and vice versa. Such opposite gene expression patterns may indicate specific responses to different pathogen types and/or infection routes.

In general, we found the largest number of same-direction overlaps among pathogenic bacteria (108 out of 210 comparisons among pathogens are significant at an adjusted p value < 0.05 and go into the same direction, Fig. 1A). The opposite pattern is particularly common for comparisons involving the Orsay virus (7 out of 20 comparisons between the virus and other pathogens). In fact, the virus only produces a same-direction overlap with the only other intracellular pathogen, the microsporidian *Nematocida parisii* (Fig. 1A). Other cases with a large number of opposite-direction overlaps include, for example, comparisons with the Gram-negative pathogen *Serratia marcescens* (11 out of 20 comparisons) and the Gram-positive pathogen *Enterococcus faecalis* (6 out of 20 comparisons).

Interestingly, certain non-pathogenic bacteria show significant same-direction overlaps with each other and also with pathogenic bacteria (for example, comparison between *E. coli* HT115 and *P. aeruginosa* PA14). As the non-pathogenic bacteria usually serve as food sources for *C. elegans*, this overlap may indicate that digestion and metabolism shape, at least partially, the response to pathogens. Alternatively, the food microbes which account for the observed overlap have at least some pathogenic effect, which was previously shown for *E. coli* OP50 (Gems and Riddle, 2000; Garigan et al., 2002).

3.3. *GATA motifs and GATA transcription factor targets are enriched in* C. elegans *pathogen response genes*

As an overlapping transcriptome signature may result from the action of the same regulator (i.e., a transcription factor), we next asked whether the core promoter regions of the different microbe-related gene sets are enriched for certain transcription factor binding motifs. Using a *de novo* motif analysis we demonstrate that the main motif enriched in the microbe-related gene sets is a GATA motif. Most pathogen-related gene sets are enriched for this motif, either in the up- or the downregulated gene sets and in several cases in both (Fig. 2A; full results are given in Table S1). GATA motifs are also enriched for non-pathogenic bacteria except for *Pseudomonas* sp. (Fig. 2A). Interestingly, the enriched

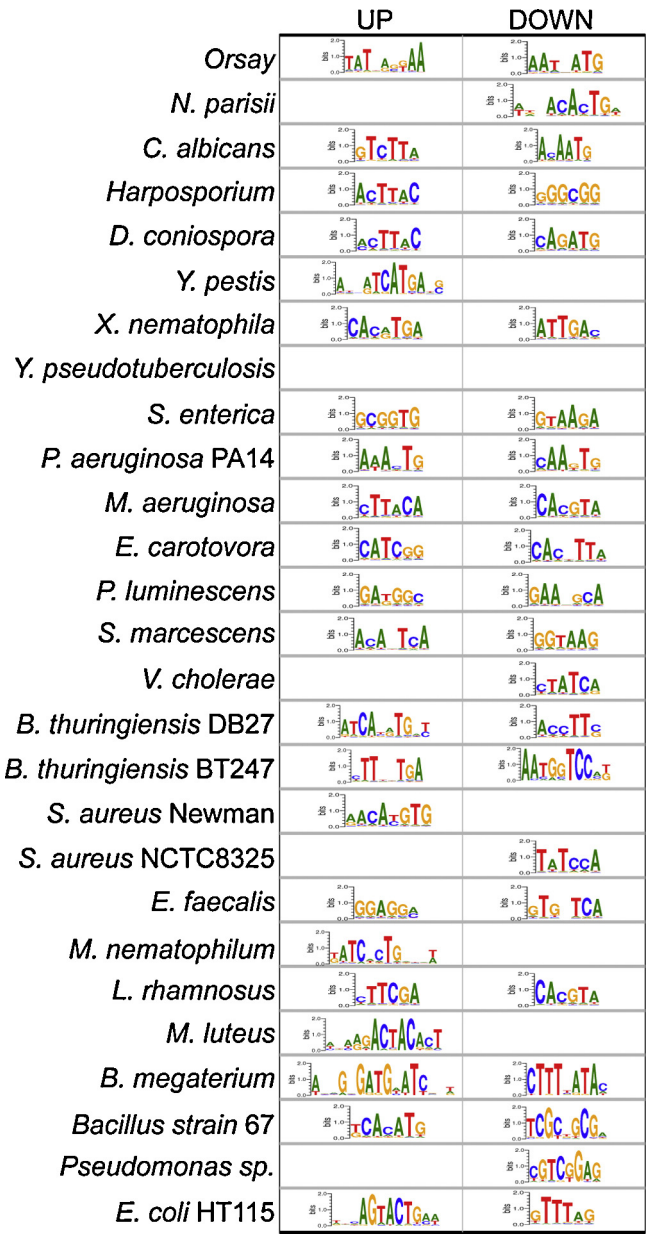

Fig. 3. Enriched non-GATA motifs for the microbe-responsive gene sets. One of the most enriched non-GATA motifs in the promoter regions of microbe-related gene sets is shown for the up- and downregulated genes. CACGTG is the Ebox motif enriched in the response to *X. nematophila*, *P. aeruginosa* PA14, *M. aeruginosa*, *S. aureus* Newman and *L. rhamnosus*. Full results are shown in Table S1.

GATA motifs can be divided into two distinct sub-families (Zhang et al., 2009): GATAAG and TGATAA. These two types are found for both pathogen- and non-pathogen-related gene sets and also for both up- and downregulated genes (Fig. 2A).

To further explore the particular importance of GATA transcription factor binding sites, we used the WormExp database to examine the overlap between the microbe-related gene sets and the sets of genes controlled in their expression by the three GATA transcription factor genes *elt-1*, *elt-2*, and *elt-3* (i.e., sets of differentially expressed genes in these mutants inferred from microarrays). While *elt-1* and *elt-3* regulate epidermal specification and differentiation, *elt-2* controls differentiation and function of the intestine (McGhee et al., 2009). An overrepresentation in the microbe-related data sets is mostly found for the *elt-2*-regulated genes and this enrichment is almost exclusively restricted to the

response to pathogenic taxa (Fig. 2B). Moreover, enrichment is also almost exclusively observed for the genes that are activated but not repressed by *elt-2* (Fig. 2B). The overlaps are either consistent (activated by *elt-2* and pathogens or vice versa) or opposite (activated by *elt-2* but repressed by pathogens or vice versa). Taken together, these results suggest that *elt-2* might serve as a central regulator in the *C. elegans* immune response to pathogens.

In addition to the GATA motif, we identified a few other motifs that are enriched in the microbe-related gene sets and possess palindromic or at least non-repetitive structures typical for transcription factor binding sites (Sorek et al., 2008) (Fig. 3, complete results in Table S1). One example is an Ebox motif (CACGTG) in the gene sets of *Xenorhabdus nematophila*, *P. aeruginosa* PA14, *Microcystis aeruginosa* and *Lactobacillus rhamnosus* (Fig. 3, Table S1). These additional motifs vary in the responses to different microbes, suggesting that they might mediate microbe-specific defences, possibly in interaction with a GATA transcription factor.

3.4. *GATA motifs are enriched in the targets of other immune-related transcription factors and signalling cascades*

We next hypothesised that the particular role of GATA-dependent transcription in *C. elegans* immunity is additionally reflected by an interaction between GATA transcription factors and other transcription factors previously implicated in the worm's immune system. Similarly, GATA transcription factors may also influence expression of the targets of the nematode's immunity-related signalling cascades.

To test the first hypothesis, we performed a *de novo* motif analysis on the transcription factor binding regions, which were previously identified through ChIP-Seq for the immune-related transcription factor genes *zip-2*, *skn-1*, *hlh-30*, *jun-1/fos-1*, and *daf-16*. Intriguingly, the GATA motif is enriched in the target binding regions of all of these transcription factors (Fig. 4A). In addition to the GATA motif, these binding regions are also enriched for other motifs, usually the motif predicted or validated for the respective transcription factor (Fig. 4A, right column). For instance, the *daf-16* ChIP-Seq data contains a significant signal for the consensus motif GTAAACA, which is indeed the confirmed binding site for the DAF-16 FOXO transcription factor (Tepper et al., 2013). Similarly, the *hlh-30* ChIP-Seq data is enriched for an Ebox motif, which also represents the previously reported binding sequence for this transcription factor (Visvikis et al., 2014). Such an Ebox motif is also overrepresented in the data for *zip-2*.

To further test the above hypotheses, we assessed the presence of the GATA motif in the promoter regions of the transcriptional targets of immunity regulators, including additional transcription factors such as *xbp-1* (Henis-Korenblit et al., 2010), *elt-2* (Block et al., 2015; Shapira et al., 2006), and *hif-1* (Bishop et al., 2004), or immune signalling pathway components such as *pmk-1* (Bond et al., 2014), *dbl-1* (Roberts et al., 2010), and *dkf-2* (Ren et al., 2009), or other types of regulators such as *npr-1* (Andersen et al., 2014). The targets were inferred from transcriptome data for these genes, as available from WormExp. A *de novo* motif analysis was then performed on the respective promoter regions of these targets. Promoter regions of *elt-2* activated genes showed enrichment for the GATA motif, confirming the accuracy of the motif analysis approach (Fig. 4B). The promoters of activated targets of the other considered genes (outside the circle) except *vhl-1* and *hif-1* (inside the circle) were similarly enriched for a GATA motif (Fig. 4B). Moreover, promoter regions of genes repressed by *npr-1* also contained the GATA motif, while we found the GATA enrichment for all the other considered genes always in the promoter region of activated targets.

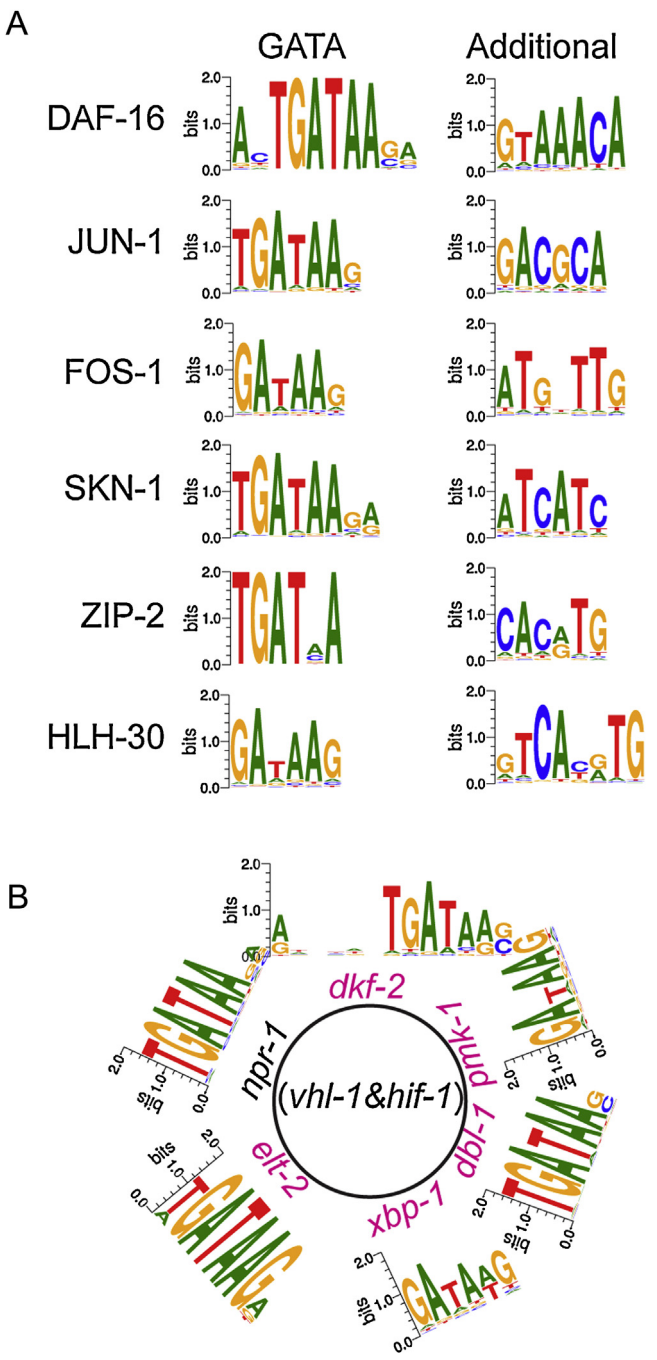

Fig. 4. GATA is enriched in binding sites in the targets of known regulators of the nematode immune response. (A) Enriched motifs in known targets of immunity-related transcription factors which were previously identified through ChIP-Seq. The enriched GATA motif is presented in the first column, while the enriched binding motifs for the considered transcription factors are shown in the second column. (B) Enriched motifs in the targets of various known immunity regulators, as inferred from previous transcriptomic studies with mutants or RNAi knockdown of the indicated genes. Pink regulator gene names (e.g., *elt-2*) indicate that only the activated target genes are enriched for the GATA motif. The *npr-1* regulator in black shows GATA enrichment in both activated and repressed genes. The targets of both *vhl-1* and *hif-1* (inside the circle) are not enriched for GATA.

3.5. *A set of 82 genes is regulated by both* elt-2 *and pathogens*

The above motif and gene set enrichment analyses identified a GATA transcription factor as the central regulator of the nematode's response to pathogens that likely accounts for the large overlap in the pathogen response genes. As *elt-2* is expressed exclu-

Table 1
Genes activated by *elt-2* and induced by at least five pathogens.*

Sequence name	Gene name	# Pathogens	Annotation
C50F7.5		12	No annotation
F55G11.5	*dod-22*	9	GO:0045087 innate immune response
Y46C8AL.2	*clec-174*	9	C type lectin domain- (CTLD-) containing proteins
F53A9.8		9	GO:0050830 response to Gram-positive bacterium
T28H10.3		9	Cysteine-type endopeptidase activity
ZC443.5	*ugt-18*	8	UDP glycosyltransferase
T01D3.6		7	GO:0005509 calcium ion binding
Y47H9C.1		7	No annotation
C09F12.1	*clc-1*	7	GO:0045087 innate immune response
Y39G10AR.6	*ugt-31*	6	UDP glycosyltransferase
ZK896.5		6	Epoxide hydrolase
C49C8.5		5	GO:0045087 innate immune response
F35C5.6	*clec-63*	5	C type lectin domain- (CTLD-) containing proteins
F35C5.9	*clec-66*	5	C type lectin domain- (CTLD-) containing proteins
F53B2.8		5	No annotation
F35E12.7	*dct-17*	5	GO:0045087 innate immune response
Y51A2D.13		5	Phospholipase
C32H11.12	*dod-24*	5	GO:0045087 innate immune response
C32H11.4		5	GO:0045087 innate immune response
H02F09.3		5	Serine/arginine repetitive matrix

* The full list of genes activated by *elt-2* and responsive to at least one pathogen is provided in Table S2 in the supplementary online Appendix.

sively in the intestine where it is the main regulator of constitutive transcription (McGhee et al., 2007), we finally asked what type of genes are both activated by *elt-2* and changed in inducible expression by at least one of the considered pathogens. The resulting list consists of 82 genes (Table S2 in the supplementary online Appendix), including 36 genes without any annotation and 15 putative immunity-related genes such as three lysozyme genes, three genes encoding caenopores/saposin domain-containing proteins, three galectin genes, five genes encoding C type lectin domain (CTLD)-containing proteins and one *nlp* gene (*nlp-40*). One of these 82 genes, namely *C50F7.5*, which currently has no functional annotation, is induced by 12 different pathogens (Table 1). 19 additional genes are induced by at least five different pathogens (Table 1), including several with a likely immune function such as the CTLD genes *clec-174*, *clec-63*, and *clec-66*.

4. Discussion

Microbes are central for our understanding of *C. elegans* biology, as the worm uses them as a food source and/or experiences them in its natural habitat as pathogens or commensals (Petersen et al., 2015). Therefore, we expect the response to microbes to be tightly regulated and this tight regulation should be manifested in the microbe-related gene expression data sets. Taking advantage of the recently established database WormExp with all available *C. elegans* expression data sets (Yang et al., 2015b), we demonstrate that the response to microbes and especially to gut-infecting pathogenic bacteria shows significant overlaps and is dominated by one of the *C. elegans* GATA transcription factors, ELT-2. We further demonstrate an enrichment of the GATA motif in the binding regions of previously identified immune-related transcription factors and also in the promoter regions of targets of other immune regulators and signalling cascades. ELT-2 is the central transcription factor controlling constitutive gene expression in the *C. elegans* intestine (McGhee et al., 2009), where infection with most bacterial pathogens and the corresponding immune responses are localised. Therefore, we propose that the ELT-2 GATA transcription factor is the primary regulator of the nematode's inducible intestinal immune defence, which is specifically activated in response to gut-infecting pathogens – in addition to the constitutively expressed genes involved in intestinal development and maintenance. The activated immune response is most likely regulated by ELT-2 in interaction with other transcription factors and/or other regulatory processes. In the following, we will discuss three main aspects of our results: (i) the similarity between the transcriptomic responses towards pathogens and food microbes; (ii) the central importance of the GATA transcription factor in regulating the response to gut pathogens; and (iii) possible mechanisms which could mediate a more specific response to different pathogens.

Although the strongest transcriptome overlap is found for the various responses against pathogens, these also show significant similarities to the expression changes induced by non-pathogenic bacteria and thus potential food organisms (Figs. 1 and S1). This may indicate that the worm's immune response incorporates components of the digestive machinery. Many of the studied pathogens cause an infection of the gut and thus the site of digesting food. Therefore, it may represent a highly economic strategy for the worm to use the same enzymes to process both types of microbes, including for example the members of the lysozyme family previously proposed to contribute to food digestion and immunity (Schulenburg and Boehnisch, 2008). This idea is supported by the observation that the transcriptional response to the non-pathogenic bacteria shows the smallest overlap and/or an opposite pattern to that found towards pathogens with a clearly distinct infection etiology, such as the intracellular Orsay virus, the intracellular microsporidian *N. parisii*, or the fungus *D. coniospora* which infects through the cuticle and not via the gut (Fig. 1) (Pujol et al., 2008; Marsh and May, 2012).

A non-exclusive alternative explanation may be that the available transcriptome studies with pathogens go beyond depicting only the nematode's immune response. In fact, most of these studies use the standard laboratory food *E. coli* OP50 as a control and often take any gene with expression differences to this bacterium to be part of the immune response. However, such expression differences may in these cases also be caused by exposing the worm to microbes with distinct characteristics (e.g., Gram-positive versus Gram-negative bacteria). Therefore, the observed expression differences could at least in part be determined by the worm's differential response to distinct microbial taxa, irrespective of their pathogenicity. A more precise characterisation of the immune system-specific expression response may in the future benefit from using several controls, including non-pathogenic varieties of the same pathogen taxon, as previously already used for *P. aeruginosa* (Troemel et al., 2006), *M. nematophilum* (O'Rourke et al., 2006), and *B. thuringiensis* (Boehnisch et al., 2011; Yang et al., 2015a).

The large overlaps in the microbe-induced expression responses were associated with the presence of the GATA motif in the promoter regions of differentially expressed genes. Intriguingly, only the gene set representing activated targets of one of the analysed *C. elegans* GATA transcription factors, ELT-2, was significantly enriched for genes responsive to gut-infecting pathogens, but not for genes differentially regulated by intracellular pathogens or those infecting via the cuticle (e.g., the Orsay virus or the fungus *D. coniospora*; Fig. 2B), and if so, only in the opposite direction (the microspordian *N. parisii*; Fig. 2B). The same applies to the non-pathogenic microbes (Fig. 2B).

The ELT-2 GATA transcription factor is known to be primarily expressed in intestinal cells and to be the predominant transcription factor regulating differentiation and maintenance of the intestine (Fukushige et al., 1998; McGhee, 2013). ELT-2 is also known to regulate the inducible defence response to *P. aeruginosa* in the *C. elegans* intestinal epithelium by interacting with the two p38-activated transcription factors ATF-7 and SKN-1 (Shapira et al., 2006; Block et al., 2015). Moreover, worms in which *elt-2* is knocked down by RNAi are more susceptible to the bacterial pathogens *S. typhimurium* and *E. faecalis*, as well as the fungal pathogen *C. neoformans* (Kerry et al., 2006). In addition, ELT-2 was identified as a target for the manipulation of the host's immune defence by *Burkholderia pseudomallei* (Lee et al., 2013), emphasising its role in regulating pathogen-inducible responses. Furthermore, ELT-2 also controls non-infection stress responses in the intestine, such as the response to osmotic stress (Rohlfing et al., 2010), TOR-dependent hypoxia responses (Schieber and Chandel, 2014), and the response to high dietary zinc (Roh et al., 2015). ELT-2 also contributes to life span extension in calorically restricted *eat-2* mutants (Zhang et al., 2013) and survival of *rpn-10* mutants, which exhibit proteasome dysfunction (Keith et al., 2016). Based on these findings, Block and Shapira proposed a model in which ELT-2 functions as a master regulator of inducible responses in the intestine – in cooperation with different signal-activated transcription factors (Block and Shapira, 2015). Our data extends the previously available information on infection-induced responses by highlighting the involvement of ELT-2 in the activation of the response to a larger variety of pathogens, particularly those infecting the nematode gut (Fig. 3). As the transcriptional response to most pathogens is measured relative to control conditions with the standard food bacterium *E. coli* OP50, the role of ELT-2 in activating this defence response goes beyond its function in coordinating the constitutive expression of genes involved in development and maintenance of the intestine. Taken together, this strongly suggests that ELT-2 is the *C. elegans* GATA transcription factor that is centrally involved in the inducible response to gut-infecting pathogens. To date, we cannot yet exclude that another intestinal GATA transcription factor, such as ELT-7, for which transcriptional targets have not been characterised, additionally contributes to the pathogen response.

Such a central role of a GATA transcription factor (or several of these) is further supported by our finding of the GATA-binding motif in the targets of other regulators of the worm's immune response, for example various other transcription factors, such as ATF-7 (Shivers et al., 2010), ZIP-2 (Estes et al., 2010), SKN-1 (Papp et al., 2012), HLH-30 (Visvikis et al., 2014), JUN-1/FOS-1 (Kao et al., 2011), and DAF-16 (Garsin et al., 2003). The DAF-16 associated element (DAE) is a known GATA-like DNA motif found in promoters of DAF-16 target genes (Tepper et al., 2013). Zhang and colleagues showed that ELT-2 can bind to the DAE/GATA site and collaborates with DAF-16 to control tissue-dependent expression of multiple target genes, contributing to lifespan extension in the insulin-like receptor mutant *daf-2(e1370)* (Zhang et al., 2013). In addition, Block et al. highlighted that ELT-2 may cooperate with the SKN-1 and ATF-7 transcription factors to regulate p38 MAPK-dependent immune responses against the pathogen *P. aeruginosa* (Block et al., 2015).

The interaction of GATA transcription factors with other immune regulators may also contribute, at least in part, to more specific defence responses to different pathogens. Such more specific responses are indicated by those cases for which differential gene expression does not show any overlap or goes in the opposite direction (indicated by white and blue colours, respectively, in Figs. 1 and S1). Our finding of additional enriched binding motifs in the promoter regions of the pathogen-related gene sets (Fig. 3) suggests that the corresponding transcription factors could then mediate the more specific immune responses, either alone or in interaction with a GATA transcription factor. The latter is, for example, indicated in the cases with an enrichment of both the GATA motif and the Ebox motif (e.g., the genes activated by *X. nematophilum* or the genes repressed by *P. aeruginosa*; Figs. 2 A and Fig. 33), especially because the two transcription factors using the Ebox motif, ZIP-2 and HLH-30, both harbour the GATA motif in their validated binding regions (Fig. 4) (Estes et al., 2010; Visvikis et al., 2014).

5. Conclusion

Based on the analysis of transcriptomic responses to pathogenic and non-pathogenic microbes, we here propose that GATA transcription factors play a central role in regulating the *C. elegans* immune response. The ELT-2 GATA transcription factor seems to be particularly important in the defence against gut-infecting pathogens. GATA-dependent transcription may also mediate more specific immune responses against distinct pathogen taxa, possibly in interaction with other transcription factors. In future work it may be particularly interesting to assess how exactly ELT-2 interacts with other transcription factors and different immune signalling cascades to coordinate immune responses in the gut and to what extent other GATA transcription factors may additionally contribute to the response to pathogens and non-pathogenic microbes both inside and outside of the gut. Considering that pathogens may be detected through their effect on cellular homeostasis (Ewbank and Pujol, 2016) and considering that *elt-2* targets include genes important for the cytoprotective response in *C. elegans* (Shore et al., 2012), it may be particularly rewarding to find out how GATA-dependent transcription may relate to damage signals and the cellular surveillance system. Finally, our analysis provides a list of pathogen-responsive, ELT-2-regulated genes (Table 1) which could be important for mediating general and specific responses to pathogens. Their exact role, especially that of the gene *C50F7.5* without any current annotation, deserves particular attention using the available toolbox for functional genetic analysis in *C. elegans*.

Acknowledgements

We thank the members of the Schulenburg lab for advice and support. We are grateful for funding from the German Science Foundation within the Priority Program SPP 1399 on host–parasite coevolution (grants SCHU 1415/9 to H.S. and RO 2994/3 to P.C.R.). W.Y. was additionally supported by the International Max Planck Research School for Evolutionary Biology.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.zool.2016.05.013.

References

Andersen, E.C., Bloom, J.S., Gerke, J.P., Kruglyak, L., 2014. A variant in the neuropeptide receptor npr-1 is a major determinant of *Caenorhabditis elegans* growth and physiology. PLoS Genet. 10, e1004156.

Bakowski, M.A., Desjardins, C.A., Smelkinson, M.G., Dunbar, T.A., Lopez-Moyado, I.F., Rifkin, S.A., Cuomo, C.A., Troemel, E.R., 2014. Ubiquitin-mediated response to microsporidia and virus infection in *C. elegans*. PLoS Pathog. 10, e1004200.

Bishop, T., Lau, K.W., Epstein, A.C., Kim, S.K., Jiang, M., O Rourke, D., Pugh, C.W., Gleadle, J.M., Taylor, M.S., Hodgkin, J., 2004. Genetic analysis of pathways regulated by the von Hippel–Lindau tumor suppressor in *Caenorhabditis elegans*. PLoS Biol. 2, 1549–1560.

Block, D.H., Shapira, M., 2015. GATA transcription factors as tissue-specific master regulators for induced responses. Worm 4, e1118607.

Block, D.H., Twumasi-Boateng, K., Kang, H.S., Carlisle, J.A., Hanganu, A., Lai, T.Y.-J., Shapira, M., 2015. The developmental intestinal regulator ELT-2 controls p38-dependent immune responses in adult *C. elegans*. PLoS Genet. 11, e1005265.

Boehnisch, C., Wong, D., Habig, M., Isermann, K., Michiels, N.K., Roeder, T., May, R.C., Schulenburg, H., 2011. Protist-type lysozymes of the nematode *Caenorhabditis elegans* contribute to resistance against pathogenic *Bacillus thuringiensis*. PLoS One 6, e24619.

Bond, M.R., Ghosh, S.K., Wang, P., Hanover, J.A., 2014. Conserved nutrient sensor O-GlcNAc transferase is integral to *C. elegans* pathogen-specific immunity. PLoS One 9, e113231.

Buchon, N., Silverman, N., Cherry, S., 2014. Immunity in *Drosophila melanogaster* – from microbial recognition to whole-organism physiology. Nature Rev. Immunol 14, 796–810.

Chávez, V., Mohri-Shiomi, A., Garsin, D.A., 2009. Ce-Duox1/BLI-3 generates reactive oxygen species as a protective innate immune mechanism in *Caenorhabditis elegans*. Infect. Immun 77, 4983–4989.

Cheung, L., Andersen, M., Gustavsson, C., Odeberg, J., Fernández-Pérez, L., Norstedt, G., Tollet-Egnell, P., 2007. Hormonal and nutritional regulation of alternative CD36 transcripts in rat liver – a role for growth hormone in alternative exon usage. BMC Mol. Biol. 8, 60.

Clark, L.C., Hodgkin, J., 2014. Commensals: probiotics and pathogens in the *Caenorhabditis elegans* model. Cell. Microbiol. 16, 27–38.

Cohen, L.B., Troemel, E.R., 2015. Microbial pathogenesis and host defense in the nematode *C. elegans*. Curr. Opin. Microbiol. 23, 94–101.

Couillault, C., Pujol, N., Reboul, J., Sabatier, L., Guichou, J.-F., Kohara, Y., Ewbank, J.J., 2004. TLR-independent control of innate immunity in *Caenorhabditis elegans* by the TIR domain adaptor protein TIR-1, an ortholog of human SARM. Nature Immunol. 5, 488–494.

Crooks, G.E., Hon, G., Chandonia, J.-M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190.

Dierking, K., Polanowska, J., Omi, S., Engelmann, I., Gut, M., Lembo, F., Ewbank, J.J., Pujol, N., 2011. Unusual regulation of a STAT protein by an SLC6 family transporter in *C. elegans* epidermal innate immunity. Cell Host Microbe 9, 425–435.

Dierking, K., Yang, W., Schulenburg, H., 2016. Antimicrobial effectors in the nematode *C. elegans* – an outgroup to the Arthropoda. Phil. Trans. R. Soc. B 371, 20150299.

Engelmann, I., Pujol, N., 2010. Innate immunity in *C. elegans*. In: Söderhäll, K. (Ed.), Invertebrate Immunity. Springer, New York, pp. 105–121.

Estes, K.A., Dunbar, T.L., Powell, J.R., Ausubel, F.M., Troemel, E.R., 2010. bZIP transcription factor zip-2 mediates an early response to *Pseudomonas aeruginosa* infection in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. U. S. A. 107, 2153–2158.

Ewbank, J.J., Pujol, N., 2016. Local and long-range activation of innate immunity by infection and damage in *C. elegans*. Curr. Opin. Immunol. 38, 1–7.

Fukushige, T., Hawkins, M.G., McGhee, J.D., 1998. The GATA-factor elt-2 is essential for formation of the *Caenorhabditis elegans* intestine. Dev. Biol. 198, 286–302.

Furey, T.S., 2012. ChIP-seq and beyond: new and improved methodologies to detect and characterize protein–DNA interactions. Nature Rev. Genet. 13, 840–852.

Garigan, D., Hsu, A.-L., Fraser, A.G., Kamath, R.S., Ahringer, J., Kenyon, C., 2002. Genetic analysis of tissue aging in *Caenorhabditis elegans*: a role for heat-shock factor and bacterial proliferation. Genetics 161, 1101–1112.

Garsin, D.A., Villanueva, J.M., Begun, J., Kim, D.H., Sifri, C.D., Calderwood, S.B., Ruvkun, G., Ausubel, F.M., 2003. Long-lived *C. elegans* daf-2 mutants are resistant to bacterial pathogens. Science 300, 1921–1921.

Gems, D., Riddle, D.L., 2000. Genetic, behavioral and environmental determinants of male longevity in *Caenorhabditis elegans*. Genetics 154, 1597–1610.

Gerstein, M.B., Lu, Z.J., Van Nostrand, E.L., Cheng, C., Arshinoff, B.I., Liu, T., Yip, K.Y., Robilotto, R., Rechtsteiner, A., Ikegami, K., 2010. Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. Science 330, 1775–1787.

Henis-Korenblit, S., Zhang, P., Hansen, M., McCormick, M., Lee, S.-J., Cary, M., Kenyon, C., 2010. Insulin/IGF-1 signaling mutants reprogram ER stress response regulators to promote longevity. Proc. Natl. Acad. Sci. U. S. A. 107, 9730–9735.

Hillier, L.W., Coulson, A., Murray, J.I., Bao, Z., Sulston, J.E., Waterston, R.H., 2005. Genomics in *C. elegans*: so many genes, such a little worm. Genome Res. 15, 1651–1660.

Hughes, A.L., Nei, M., 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature 335, 167–170.

Irazoqui, J.E., Troemel, E.R., Feinbaum, R.L., Luhachack, L.G., Cezairliyan, B.O., Ausubel, F.M., 2010a. Distinct pathogenesis and host responses during infection of *C. elegans* by *P. aeruginosa* and *S. aureus*. PLoS Pathog. 6, e1000982.

Irazoqui, J.E., Urbach, J.M., Ausubel, F.M., 2010b. Evolution of host innate defence: insights from *Caenorhabditis elegans* and primitive invertebrates. Nature Rev. Immunol. 10, 47–58.

Janeway, C.A., Travers, P., Walport, M., Shlomchik, M.J., 2001. Immunobiology: The Immune System in Health and Disease. Churchill Livingstone, London.

Kao, C.-Y., Los, F., Huffman, D.L., Wachi, S., Kloft, N., Husmann, M., Karabrahimi, V., Schwartz, J.-L., Bellier, A., Ha, C., 2011. Global functional analyses of cellular responses to pore-forming toxins. PLoS Pathog. 7, e1001314.

Keith, S.A., Maddux, S.K., Zhong, Y., Chinchankar, M.N., Ferguson, A.A., Ghazi, A., Fisher, A.L., 2016. Graded proteasome dysfunction in *Caenorhabditis elegans* activates an adaptive response involving the conserved SKN-1 and ELT-2 transcription factors and the autophagy-lysosome pathway. PLoS Genet. 12, e1005823.

Kerry, S., TeKippe, M., Gaddis, N.C., Aballay, A., 2006. GATA transcription factor required for immunity to bacterial and fungal pathogens. PLoS One 1, e77.

Lee, S.-H., Wong, R.-R., Chin, C.-Y., Lim, T.-Y., Eng, S.-A., Kong, C., Ijap, N.A., Lau, M.-S., Lim, M.-P., Gan, Y.-H., 2013. *Burkholderia pseudomallei* suppresses *Caenorhabditis elegans* immunity by specific degradation of a GATA transcription factor. Proc. Natl. Acad. Sci. U. S. A. 110, 15067–15072.

Marsh, E.K., May, R.C., 2012. *Caenorhabditis elegans*: a model organism for investigating immunity. Appl. Environm. Microbiol. 78, 2075–2081.

McGhee, J.D., 2013. The *Caenorhabditis elegans* intestine. Wiley Interdiscip. Rev. Dev. Biol. 2, 347–367.

McGhee, J.D., Sleumer, M.C., Bilenky, M., Wong, K., McKay, S.J., Goszczynski, B., Tian, H., Krich, N.D., Khattra, J., Holt, R.A., 2007. The ELT-2 GATA-factor and the global regulation of transcription in the *C. elegans* intestine. Dev. Biol. 302, 627–645.

McGhee, J.D., Fukushige, T., Krause, M.W., Minnema, S.E., Goszczynski, B., Gaudet, J., Kohara, Y., Bossinger, O., Zhao, Y., Khattra, J., 2009. ELT-2 is the predominant transcription factor controlling differentiation and function of the *C. elegans* intestine from embryo to adult. Dev. Biol. 327, 551–565.

Melo, J.A., Ruvkun, G., 2012. Inactivation of conserved *C. elegans* genes engages pathogen- and xenobiotic-associated defenses. Cell 149, 452–466.

Michaloski, J.S., Galante, P.A., Malnic, B., 2006. Identification of potential regulatory motifs in odorant receptor genes by analysis of promoter sequences. Genome Res. 16, 1091–1098.

Michaloski, J.S., Galante, P.A., Nagai, M.H., Armelin-Correa, L., Chien, M.-S., Matsunami, H., Malnic, B., 2011. Common promoter elements in odorant and vomeronasal receptor genes. PLoS One 6, e29065.

Murphy, C.T., McCarroll, S.A., Bargmann, C.I., Fraser, A., Kamath, R.S., Ahringer, J., Li, H., Kenyon, C., 2003. Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*. Nature 424, 277–283.

Mysliwy, J., Dingley, A.J., Stanisak, M., Jung, S., Lorenzen, I., Roeder, T., Leippe, M., Grötzinger, J., 2010. Caenopore-5: the three-dimensional structure of an antimicrobial protein from *Caenorhabditis elegans*. Dev. Comp. Immunol. 34, 323–330.

O'Rourke, D., Baban, D., Demidova, M., Mott, R., Hodgkin, J., 2006. Genomic clusters, putative pathogen recognition molecules, and antimicrobial genes are induced by infection of *C. elegans* with *M. nematophilum*. Genome Res. 16, 1005–1016.

Papp, D., Csermely, P., Soti, C., 2012. A role for SKN-1/Nrf in pathogen resistance and immunosenescence in *Caenorhabditis elegans*. PLoS Pathog. 8, e1002673.

Pellegrino, M.W., Nargund, A.M., Kirienko, N.V., Gillis, R., Fiorese, C.J., Haynes, C.M., 2014. Mitochondrial UPR-regulated innate immunity provides resistance to pathogen infection. Nature 516, 414–417.

Petersen, C., Dirksen, P., Schulenburg, H., 2015. Why we need more ecology for genetic models such as *C. elegans*. Trends Genet. 31, 120–127.

Powell, J.R., Ausubel, F.M., 2008. Models of *Caenorhabditis elegans* infection by bacterial and fungal pathogens. Methods Mol. Biol. 415, 403–427.

Pujol, N., Cypowyj, S., Ziegler, K., Millet, A., Astrain, A., Goncharov, A., Jin, Y., Chisholm, A.D., Ewbank, J.J., 2008. Distinct innate immune responses to infection and wounding in the *C. elegans* epidermis. Curr. Biol. 18, 481–489.

Pukkila-Worley, R., Ausubel, F.M., 2012. Immune defense mechanisms in the *Caenorhabditis elegans* intestinal epithelium. Curr. Opin. Immunol. 24, 3–9.

Ren, M., Feng, H., Fu, Y., Land, M., Rubin, C.S., 2009. Protein kinase D is an essential regulator of *C. elegans* innate immunity. Immunity 30, 521–532.

Roberts, A.F., Gumienny, T.L., Gleason, R.J., Wang, H., Padgett, R.W., 2010. Regulation of genes affecting body size and innate immunity by the DBL-1/BMP-like pathway in *Caenorhabditis elegans*. BMC Dev. Biol. 10, 61.

Roeder, T., Stanisak, M., Gelhaus, C., Bruchhaus, I., Grötzinger, J., Leippe, M., 2010. Caenopores are antimicrobial peptides in the nematode *Caenorhabditis elegans* instrumental in nutrition and immunity. Dev. Comp. Immunol. 34, 203–209.

Roh, H.C., Dimitrov, I., Deshmukh, K., Zhao, G., Warnhoff, K., Cabrera, D., Tsai, W., Kornfeld, K., 2015. A modular system of DNA enhancer elements mediates tissue-specific activation of transcription by high dietary zinc in *C. elegans*. Nucleic Acids Res. 43, 803–816.

Rohlfing, A.-K., Miteva, Y., Hannenhalli, S., Lamitina, T., 2010. Genetic and physiological activation of osmosensitive gene expression mimics transcriptional signatures of pathogen infection in *C. elegans*. PLoS One 5, e9010.

Sackton, T.B., Lazzaro, B.P., Schlenke, T.A., Evans, J.D., Hultmark, D., Clark, A.G., 2007. Dynamic evolution of the innate immune system in *Drosophila*. Nature Genet. 39, 1461–1468.

Schieber, M., Chandel, N.S., 2014. TOR signaling couples oxygen sensing to lifespan in *C. elegans*. Cell Rep. 9, 9–15.

Schulenburg, H., Boehnisch, C., 2008. Diversification and adaptive sequence evolution of *Caenorhabditis* lysozymes (Nematoda: Rhabditidae). BMC Evol. Biol. 8, 114.

Schulenburg, H., Kurtz, J., Moret, Y., Siva-Jothy, M.T., 2009. Introduction. Ecological immunology. Phil. Trans. R. Soc. B 364, 3–14.

Shapira, M., Hamlin, B.J., Rong, J., Chen, K., Ronen, M., Tan, M.-W., 2006. A conserved role for a GATA transcription factor in regulating epithelial innate immune responses. Proc. Natl. Acad. Sci. U. S. A. 103, 14086–14091.

Shi, J., Yang, W., Chen, M., Du, Y., Zhang, J., Wang, K., 2011. AMD, an automated motif discovery tool using stepwise refinement of gapped consensuses. PLoS One 6, e24576.

Shivers, R.P., Pagano, D.J., Kooistra, T., Richardson, C.E., Reddy, K.C., Whitney, J.K., Kamanzi, O., Matsumoto, K., Hisamoto, N., Kim, D.H., 2010. Phosphorylation of the conserved transcription factor ATF-7 by PMK-1 p38 MAPK regulates innate immunity in *Caenorhabditis elegans*. PLoS Genet. 6, e1000892.

Shore, D.E., Carr, C.E., Ruvkun, G., 2012. Induction of cytoprotective pathways is central to the extension of lifespan conferred by multiple longevity pathways. PLoS Genet. 8, e1002792.

Sorek, R., Kunin, V., Hugenholtz, P., 2008. CRISPR – a widespread system that provides acquired resistance against phages in bacteria and archaea. Nat. Rev. Microbiol. 6, 181–186.

Stein, L., Sternberg, P., Durbin, R., Thierry-Mieg, J., Spieth, J., 2001. WormBase: network access to the genome and biology of *Caenorhabditis elegans*. Nucleic Acids Res. 29, 82–86.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U. S. A. 102, 15545–15550.

Tabach, Y., Brosh, R., Buganim, Y., Reiner, A., Zuk, O., Yitzhaky, A., Koudritsky, M., Rotter, V., Domany, E., 2007. Wide-scale analysis of human functional transcription factor binding reveals a strong bias towards the transcription start site. PLoS One 2, e807.

Tepper, R.G., Ashraf, J., Kaletsky, R., Kleemann, G., Murphy, C.T., Bussemaker, H.J., 2013. PQM-1 complements DAF-16 as a key transcriptional regulator of DAF-2-mediated development and longevity. Cell 154, 676–690.

Trinklein, N.D., Aldred, S.J.F., Saldanha, A.J., Myers, R.M., 2003. Identification and functional analysis of human transcriptional promoters. Genome Res. 13, 308–312.

Troemel, E.R., Chu, S.W., Reinke, V., Lee, S.S., Ausubel, F.M., Kim, D.H., 2006. p38 MAPK regulates expression of immune response genes and contributes to longevity in *C. elegans*. PLoS Genet. 2, e183.

Uno, M., Honjoh, S., Matsuda, M., Hoshikawa, H., Kishimoto, S., Yamamoto, T., Ebisuya, M., Yamamoto, T., Matsumoto, K., Nishida, E., 2013. A fasting-responsive signaling pathway that extends life span in *C. elegans*. Cell Rep. 3, 79–91.

Van Der Hoeven, R., McCallum, K.C., Cruz, M.R., Garsin, D.A., 2011. Ce-Duox1/BLI-3 generated reactive oxygen species trigger protective SKN-1 activity via p38 MAPK signaling during infection in *C. elegans*. PLoS Pathog. 7, e1002453.

Visvikis, O., Ihuegbu, N., Labed, S.A., Luhachack, L.G., Alves, A.-M.F., Wollenberg, A.C., Stuart, L.M., Stormo, G.D., Irazoqui, J.E., 2014. Innate host defense requires TFEB-mediated transcription of cytoprotective and antimicrobial genes. Immunity 40, 896–909.

Wong, D., Bazopoulou, D., Pujol, N., Tavernarakis, N., Ewbank, J.J., 2007. Genome-wide investigation reveals pathogen-specific and shared signatures in the response of *Caenorhabditis elegans* to infection. Genome Biol. 8, R194.

Yang, W., Dierking, K., Esser, D., Tholey, A., Leippe, M., Rosenstiel, P., Schulenburg, H., 2015a. 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.

Yang, W., Dierking, K., Schulenburg, H., 2015b. WormExp: a web-based application for a *Caenorhabditis elegans*-specific gene expression enrichment analysis. Bioinformatics 15, 943–945.

Zhang, P., Judy, M., Lee, S.-J., Kenyon, C., 2013. Direct and indirect gene regulation by a life-extending FOXO protein in *C. elegans*: roles for GATA factors and lipid gene regulators. Cell Metab. 17, 85–100.

Zhang, Y., Wu, W., Cheng, Y., King, D.C., Harris, R.S., Taylor, J., Chiaromonte, F., Hardison, R.C., 2009. Primary sequence and epigenetic determinants of in vivo occupancy of genomic DNA by GATA1. Nucleic Acids Res. 37, 7024–7038.

Zugasti, O., Bose, N., Squiban, B., Belougne, J., Kurz, C.L., Schroeder, F.C., Pujol, N., Ewbank, J.J., 2014. Activation of a G protein-coupled receptor by its endogenous ligand triggers the innate immune response of *Caenorhabditis elegans*. Nature Immunol. 15, 833–838.