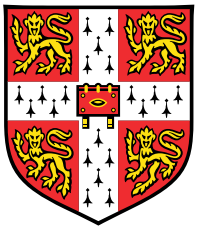
Double-stranded RNA uptake in *Caenorhabditis elegans*



Fabian Braukmann

Wolfson College

Department of Genetics

University of Cambridge

This dissertation is submitted for the degree of *Doctor of Philosophy*

September 2018

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text

It does not exceed the prescribed word limit for the relevant Degree Committee.

Fabian Braukmann

Double-stranded RNA uptake in *Caenorhabditis elegans*

Fabian Braukmann

Abstract

dsRNA uptake from the environment is a common phenomenon in many invertebrates and can be harnessed for acquiring pest control. In the nematode *Caenorhabditis elegans*, a transmembrane protein named *systemic RNAi deficient-2* (*sid-2)* is required for dsRNA uptake from the environment through the intestine. Furthermore, the tyrosine-protein kinase *sid-3* promotes import dsRNA. However, the molecular mechanism underlying environmental RNA uptake and its natural phenotypic roles are largely unknown. Here I characterized the intracellular properties of SID-2 and its role during the nematode development. I determined SID-2 localization at the intestinal apical membrane and the trans-Golgi-network (TGN). This localization is dependent on *sid-3* and its potential substrate *viro-2*, hence demonstrating a molecular interplay between the different SID proteins. The spatial organisation of SID-2 is conserved across species, suggesting a role for the TGN in environmental RNA uptake. By combing small RNA and transcriptome profiling with developmental studies I identified a novel role for systemic RNAi in worm morphology. A detailed phenotypical analysis ruled out a nutritional role for dsRNA uptake, implying that environmental RNA uptake affects morphology through gene regulation. These results uncover the first phenotypic role for the Sid pathway and identifies the TGN as a central cellular compartment for environmental dsRNA uptake.

Dedication

For my family.

Acknowledgements

I would like to thank my PhD supervisor Eric Miska for giving me the opportunity to work in his research group. He has provided tremendous support, guidance and advice throughout my time at Cambridge University.

I am indebted to David Jordan, Kin Man Suen and Melanie Tanguy whose help and supervision have been immensely important during my PhD.

I am thankful to Kay Harnish, Juanita Baker-Hay, Richard Buttler, Sylviane Moss and Charles Bradshaw whose support has been immensely important during my PhD.

I also want to thank all of the members of the Miska lab for making it such a great and collaborative place to work. In particular, I am grateful to Miranda Landgraf and Marc Ridyard, as well as to Daria Nikitin, Sophia Metz, Jon Donier, Alexandra Sapteschnig, Alper Akay, Jeremie Le Pen, Konrad Rudolf, Tomas di Domenico, Louise Veron, Eyal Maori, Sabrina Huber, Isabela Navarro, Omer Ziv, Asia Kosalka, Gregoire Vernaz, Nicolas Bologna, Alexandra Bezler, Giulia Furlan, Lise Frezal, help with experiments and for support in the most stressful times of my PhD.

I want to thank my friends Sarah, Robert, Eva, Philip, Mihoko, Keti.

And thank you Serena, for everything.

Table of Contents

1 Introduction 1

1.1 Application of environmental RNAi 1

1.1.1 Environmental RNAi in commercial application 1

1.1.2 Environmental RNAi in scientific application 2

1.2 Biological role of environmental RNAi 2

1.3 Molecular mechanism of environmental RNAi 3

1.1.3 Gene regulation by endogenous small interfering RNAs 3

1.1.4 Gene regulation by exogenous dsRNA 6

1.1.5 Uptake of environmental dsRNA 8

1.1.6 dsRNA uptake at the apical intestinal membrane 10

1.1.7 Transport of dsRNA within the animal 11

1.1.8 Genetics of dsRNA uptake into the germline 13

1.1.9 Genetic requirements for export and import of dsRNA 14

1.1.10 Endocytic import of dsRNA 14

1.1.11 Export of dsRNA 15

1.1.12 RNA mobility in plants 17

1.4 Evolution of environmental RNAi in *Caenorhabditis* 18

1.5 Aim of the thesis 20

2 Chapter – Identification of SID-2 interaction partners 21

2.1 Introduction 21

2.2 Results 22

2.2.1 A novel anti-SID-2 antibody 22

2.2.2 Identification of candidate SID-2 interaction partner via co-immunoprecipitation followed by mass spectrometry 25

2.2.3 The role of candidate SID-2 interaction partner in RNAi by feeding 28

2.2.4 Identification of candidate interaction partners via yeast two hybrid screening 32

2.3 Chapter summary 36

3 Chapter – SID-2 localisation 37

3.1 Introduction 37

3.2 Results 39

3.2.1 SID-2 localises at the apical membrane and in cytoplasmic punctuated foci in *C. elegans* 39

3.2.2 SID-2 localisation is conserved in *C. elegans* and *C. briggsae* 40

3.2.3 SID-2 associates with in the TGN-network and the recycling endosome 42

3.2.4 The tyrosine-protein kinase SID-3 and SID-3's potential substrate VIRO-2 affect SID-2 localisation 43

3.2.5 Change in SID-2 abundance upon viro-2 mutation cause no deficiency in environmental RNAi 46

3.2.6 dsRNA redistributes SID-2 subcellular in *C. elegans* and *C. briggsae*. 47

3.2.7 SID-2 is a dsRNA binding protein 49

3.3 Chapter summary 50

4 Chapter – Phenotypical consequences of sid mutations 51

4.1 Introduction 51

4.2 Results 52

4.2.1 Transcriptome analysis of *sid* mutants 52

4.2.2 small RNAome analysis in *sid* mutants 58

4.2.3 Phenotypic analysis of *sid* mutant animals 63

4.2.4 SID-2 functions non-nutritionally 67

4.3 Chapter summary 71

5 Discussion 72

5.1 Biochemical identification of novel factors of the dsRNA uptake pathway 72

5.1.1 Advantages and disadvantages of biochemical approaches 73

5.1.2 Technical Co-IP improvements 74

5.1.3 Yeast 2 Hybrid - Proteins in heterologous systems 75

5.1.4 The remaining potential SID-2 binding partner 75

5.1.5 Alternative approaches to identify novel *sid's* 76
5.1.6 Posttranslational modification of SID-2 76

5.2 Cell biology of SID-2 77

5.2.1 SID-2 role at the TGN 77
5.2.2 Subcellular localisation could be potentially regulated by *sid-3* and *viro-2* 78
5.2.3 SID-2 is a conserved dsRNA binding protein 79

5.3 Biological role of the *sid* pathway 80

5.3.1 Systemic RNAi for germline immortality? 81

6 Methods 83

6.1 *C. elegans* methods 83

6.1.1 Nematode culture 83
6.1.2 Strains 83
6.1.3 Genetic crosses 85
6.1.4 Bleaching and synchronization of *C. elegans* 85
6.1.5 RNAi experiments 85
6.1.6 dsRNA feeding assay 86

6.2 Molecular biology methods 86

6.2.1 Polymerase chain reaction 86
6.2.2 DNA electrophoresis 87
6.2.3 PCR product purification 87
6.2.4 DNA sequencing 87
6.2.5 Molecular cloning 87
6.2.6 Preparation of single worm lysate 88
6.2.7 *C. elegans* RNA extraction 88
6.2.8 quantitative real-time PCR 89
6.2.9 *In vitro* RNA synthesis 89
6.2.10 Small RNA sequencing 89
6.2.11 RNA sequencing 90

6.3 Microbiology methods 90

6.3.1 Bacteria strains used in this study 90
6.3.2 Bacterial transformation 91

6.3.3 Bacterial culture ... 91

6.4 Imaging ... 91

6.4.1 dsRNA soaking experiment ... 91

6.4.2 Isolation of *C. elegans* gut and immunhistochemistry ... 92

6.5 Biochemical analysis ... 93

6.5.1 Generation of *C. elegans* protein extract ... 93

6.5.2 SDS-PAGE ... 93

6.5.3 Western blot ... 94

6.5.4 Removal of antibodies from western blots ... 94

6.5.5 Generation and purification of custom antibodies ... 95

6.5.6 Immunoprecipitation ... 95

6.5.7 *In vitro* dsRNA binding assay ... 95

6.6 Bioinformatics ... 97

6.6.1 SID-2 foci analysis ... 97

6.6.2 Worm length measurement on wild bacteria and L1 ... 97

6.6.3 Small RNA and long RNA sequencing analysis ... 97

6.7 Worm length measurement ... 98

6.7.1 Single worm growth curves ... 98

6.7.2 Worm length at hatching ... 98

6.7.3 Length measurements in fixed Worms ... 99

7 References ... 100

8 Appendix ... 116

8.1 Quality control for deep sequencing files ... 116

8.2 Read trimming ... 117

8.2.1 Small RNA 3' Adapter trimming ... 117

8.2.2 Long RNA adapter trimming ... 118

8.2.3 Quality control for deep sequencing files after adapter trimming ... 119

8.3 Read alignment ... 120

8.3.1 Small RNA read alignment ... 120

8.3.2 Long RNA read alignment ... 122

8.4 Counting reads to genomic features 124
8.4.1 Counting small RNA reads to genomic features 124
8.4.2 Counting long RNA reads to genomic features 125
8.4.3 Counting small RNA reads to piRNA features 126
8.5 Identifying differentially expressed genes in RNA-Seq 127
8.5.1 Contents 127
8.5.2 Introduction 127
8.5.3 Load read count data 128
8.5.4 Filter method 128
8.5.5 Define samples and comparison 129
8.5.6 Normalizing Read Counts 129
8.5.7 Create table with statistics about each gene 130
8.5.8 Inferring Differential Expression with a Negative Bionomial Model 131
8.5.9 Create a table with significant genes 131
8.5.10 Create a MA plot with significant genes 132
8.6 Statistical testing for significant changes in the distribution of significantly expressed genes 134
8.6.1 Contents 134
8.6.2 Hypothesis 134
8.6.3 Load input files containing significant reads 134
8.6.4 Calculate the hypergeometical distribution 136
8.6.5 Calculate p-values for observed distributions 137
8.6.6 Plot gene distribution across the linkage groups 138
8.7 piRNA analysis in SID small RNA sequencing libraries 140
8.7.1 Contents 140
8.7.2 Introduction 140
8.7.3 Load read count data 140
8.7.4 Read count normalisation by library size 141
8.7.5 Calculate mean for samples 141
8.7.6 Errorbar plot 142
8.7.7 Calculate t-test 143
8.8 Quantify SID-2 foci 145
8.8.1 Contents 145

8.8.2 Introduction 145
8.8.3 Load images into data_structure 145
8.8.4 Identify non-intestinal areas in image using backgroundGUI 145
8.8.5 Identify intestinal areas in image using noisegroundGUI 146
8.8.6 Calcululates focimetrics for all samples 146
8.8.7 Detect foci in intestine and extract mean inensity and area of foci 146
8.8.8 Plot area and intensity on scatter plot to identify nonspecific dots using *sid-2* mutants 147

8.9 Significantly expressed small RNAs 149

8.9.1 *Sid-1 vs.* wild type small RNAs 149
8.9.2 *Sid-2 vs.* wild type small RNAs 155

8.10 Significantly expressed long RNAs 167

8.10.1 *Sid-1 vs.* wild type long RNAs 167
8.10.2 *Sid-2 vs.* wild type long RNAs 169

Table of Figures

Figure 1 Exogenous RNAi in the *C. elegans* germline.....6
Figure 2 Endogenous and exogenous small RNA types 8
Figure 3 Systemic RNAi screen controls and identification of *sid-1*.....9
Figure 4 Model for dsRNA transport in *C. elegans* 12
Figure 5 RNAi by feeding susceptibility is not predicted by presence of SID-2 19
Figure 6 Gene model of *C. elegans' sid-2* 23
Figure 7 SID-2 antibody specifically detects SID-2.....24
Figure 8 Silver staining and western blot of SID-2 immunoprecipitation.....26
Figure 9 SID-2 potential interacting partners 27
Figure 10 Ordinal scoring system for *rpb-2* RNAi by feeding.....29
Figure 11 Yeast 2 Hybrid SID-2 interaction candidate RhoGEF Y37A1B.17 and two RhoGTPases are not resistant to RNAi.....34
Figure 12 Yeast 2 Hybrid SID-2 interaction candidate RhoGEF Y37A1B.17 and two RhoGTPases do not increase the susceptibility to RNAi by feeding. 35
Figure 13 Novel subcellular localisation of SID-2 40
Figure 14 *C. elegans* and *C. briggsae* SID-2 protein alignment 41
Figure 15 SID-2 co-localises with recycling endosome and TGN marker proteins 43
Figure 16 Lack of Tyrosine-protein kinase SID-3 and SID-3's potential substrate VIRO-2 affect SID-2 abundance. 46
Figure 17 *Viro-2* is RNAi by feeding susceptible.....47
Figure 18 dsRNA causes redistribution of SID-2. 48
Figure 19 SID-2 binds dsRNA in vitro 49
Figure 20 Wild type and *sid-1* mutant animals have a similar transcriptome.53
Figure 21 Wild type and *sid-2* mutant animals have a similar transcriptome.54
Figure 22 Differentially expressed RNA function in developmental processes 55
Figure 23 *sdc-3* expression in RNAseq and qPCR.....57
Figure 24 Differentially expressed small RNAs in *sid-1* mutants 59

Figure 25 Differentially expressed small RNAs in *sid-2* mutants 60
Figure 26 Distribution of chromosomal location of significantly expressed small RNAs.....61
Figure 27 Reduced mean piRNA abundance in *sid* mutants. 62
Figure 28 Worm length increases in *sid-1* and *sid-2* mutant worms compated to wild type 64
Figure 29 L1 larvae are elongated in *sid* mutants. 65
Figure 30 *Sid* worm length is robust in the presence of wild bacteria.....66
Figure 31 SID-2 does not enhance nucleotide uptake for *pyr-1* dependent nutritional purposes.....68
Figure 32 *pyr-1* mutant can take up dsRNA.....70

Table of Tables

Table 1 Systemic RNAi screen strains.....10
Table 3 RNAi by feeding assay against the RNA polymerase II subunit B (*rpb-2*).....31
Table 4 Yeast two hybrid hits.....33
Table 4 List of *Caenorhabditis* strains.....83
Table 5 List of bacteria strains.....90
Table 6 Significantly expressed small RNAs between *sid-1* and wild type..149
Table 7 Significantly expressed small RNAs between *sid-2* and wild type..155
Table 8 Significantly expressed long RNAs between *sid-1* and wild type....167
Table 9 Significantly expressed long RNAs between *sid-2* and wild type....169

Abbreviations

aa	Amino acid
°	Degree
APS	Ammonium persulphate
bp	base pairs
C	Celsius
C.	Caenorhabditis
CAD	Carbamoyl phosphate synthetase, aspartate transcarbamoylase, and dihydroorotase activitiy
D.	Drosophila
dsRNA	Double-stranded RNA
EMCV	Encephalomyocarditis virus
ENTH	Epsin N-terminal homology
ER	Endoplasmic reticulum
FDR	Fase discovery rate
FI	Flow through
GFP	Green fluorescent protein
hp	Hairpin
HSV-1	Herpes simplex virus 1
IgG	Immunoglobulin G
In	Input
IP	immunoprecipitation
L	Ladder
miRNA	MicroRNA
MVB	Multivesicular bodies
piRNA	Piwi-interacting RNA
pri-miRNA	Primary miRNA
qPCR	Quantitative polymerase chain reaction
RdRP	RNA polymerases
RhoGEF	Rho guanine nucleotide exchange factor
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
SNARE	Soluble NSF Attachment Protein Receptor
TGN	Trans-Golgi-network
UTR	Untranslated region
WB	Western blot
YFP	Yellow fluorescent protein

1 Introduction

In plants and animals, small RNAs are key factors in controlling gene expression (Kim et al., 2009). Small RNAs consist of 21 to 30 nucleotides and mediate their functions via Watson and Crick base pairing with complementary ribonucleic acid (RNA). They exist in different flavours performing essential functions in both development and in maintaining genome stability (Aravin et al., 2003; Bartel, 2018; Girard et al., 2006; Grishok, 2013; Ruby et al., 2006; Weick and Miska, 2014).

1.1 Application of environmental RNAi

1.1.1 Environmental RNAi in commercial application

Small RNA mediated gene regulation can be artificially induced in many organisms through double-stranded RNA (dsRNA) taken up from the environment (Dhadialla and Gill, 2014). This process is known as environmental RNA interference (RNAi) (Ivashuta et al., 2015). Some of the animals competent for dsRNA uptake are parasites of plants important for crop production such as maize (Baum et al., 2007). An emerging technique aiming to control such parasites uses genetically modified crops expressing dsRNA to manipulate the expression of essential genes of the plant parasite (Zhang et al., 2017). For one of the major corn parasites, the western corn rootworm, the efficiency of such an approach has been demonstrated (Baum et al., 2007; Bolognesi et al., 2012; Gray et al., 2009). This strategy has been incorporated into a genetically modified corn breed named MON87411, which will likely to be the first commercially available corn product using dsRNA expression for pest control in plants (Fishilevich et al., 2016). Furthermore, the European food safety authority classified MON87411 as safe with respect to potential effects on human and animal health and the environment (Naegeli et al., 2018). Overall, RNAi has practical application for plant pest control.

1.1.2 Environmental RNAi in scientific application

In addition, environmental RNAi has been proven to be a very efficient technique, for example it has enabled the investigation of gene function (Ashrafi et al., 2003; Fraser et al., 2000; Gönczy et al., 2000; Kamath et al., 2000; Maeda et al., 2001; Piano et al.; Watson et al., 2013). This is especially true for the nematode *C. elegans* whose endogenous gene expression can be altered by soaking worms in dsRNA containing solutions (Tabara et al., 1998), or by feeding *C. elegans* on *E. coli* bacteria expressing dsRNA (Timmons and Fire, 1998). The efficacy of the technique was further improved by using *E. coli* strains deficient in dsRNA degradation (Timmons et al., 2001). The construction of bacteria libraries expressing dsRNA against every gene in the *C. elegans* genome allowed for an efficient and standardised way to investigate the consequences of gene knockdown and represents a key element of genetic research in the *C. elegans* field. Overall, environmental RNAi is a flexible tool to understand many biological processes.

1.2 **Biological role of environmental RNAi**

Because environmental RNAi can interfere with endogenous gene function, it is obvious that competence in environmental RNAi can have detrimental effect on the organism and the natural biological role of environmental RNAi remains unclear. In nematodes, some speculate that the pathway allows the worm to sample the environment and adapt its phenotype accordingly. In fact, stunning lab experiments showed that artificially supplied dsRNA can mediate changes in gene regulation that affect the phenotype over several generations (Alcazar et al., 2008; Buckley et al., 2012; Vastenhouw et al., 2006). Alternatively, environmental RNAi could have a role in immunity. In flies and mice, viral immunity is suggested to require dsRNA uptake from lysed cell during infection (Nguyen et al., 2017; Saleh et al., 2009). Supporting this hypothesis, RNAi by feeding in *C. elegans* confers immunity to Orsay virus infection, the only known

natural *C. elegans* virus (Ashe et al., 2015). Furthermore, the RNAi pathway contributes to viral defends in mammals and nematodes (Félix et al., 2011; Li et al., 2013; Maillard et al., 2013). Nevertheless, the systemic RNAi pathway did not affect Orsay virus infection in *C. elegans* under the tested conditions. (Ashe et al., 2015). Therefore, the biological function of the environmental RNAi remains a mystery.

1.3 Molecular mechanism of environmental RNAi

The mechanistic understanding of environmental RNAi has progressed immensely with the advancement in technological use of RNAi. The mechanism can be divided into two independent processes. One is the uptake and transport of the dsRNA from the environment into the animal, the other the processing of the dsRNA in the cell to regulate gene expression.

Much is known about the molecular details of intracellular processing of dsRNA, since the artificial dsRNA hijacks a highly conserved pathway, which generate small RNAs for endogenous gene regulatory purposes. Here, I describe first the processing and principle of endogenous small RNAs which have gene regulatory function, since the mechanism for environmental RNAi relies on similar principels and even uses partially the same molecular pathway.

1.1.3 Gene regulation by endogenous small interfering RNAs

To introduce the principles in biogenesis and function of small RNAs, I will describe two small RNAs pathways in *C. elegans:* the miRNA and the piRNA pathways.

The first small interfering RNA (siRNA) with gene regulatory function was identified in the nematode *C. elegans* and belongs to the microRNA (miRNA)

class which is defined by a highly conserved biosynthesis pathway (Ambros et al., 2003; Lee et al., 1993). miRNAs are encoded in the genome and are transcribed as long primary miRNAs (pri-miRNA) by the RNA polymerase II like proteins (Cai et al., 2004; Hamilton et al., 1999; Lau et al., 2001; Lee and Ambros, 2001; Lee et al., 2002, 2004a). In two subsequent steps, endonucleases in the nucleus and in the cytoplasm cut out the miRNA from its long precursor (Bernstein et al., 2001; Lee et al., 2003; Zhang et al., 2004). Finally, the miRNA is loaded into a protein of the Argonaute class forming the core of the RNA-induced silencing complex (RISC) and regulates mRNA with complementary sequence post-transcriptionally (Kawamata and Tomari, 2010). Overall, a long RNA is chopped in a small RNA, which together with a protein mediate gene regulation.

Interestingly, many other small RNAs exist in *C. elegans* differing in their biogenesis pathway and in the way they regulated gene expression (Grishok, 2013). Despite this diversitty, all small RNAs bind with an Argonaute protein to mediated their gene regulatory function (Peters and Meister, 2007). For example in *C. elegans*, piRNAs are germline specific expressed small RNA required for germline integrity (Weick and Miska, 2014). These piRNAs are 21 nucleotides long and start at their 5' end with a uridine. Similar to miRNAs, they are encoded in the genome (Ruby et al., 2006). In contrast to miRNAs, piRNAs initiated the production of novel small RNAs originating from an RNA templated. piRNAs initiate the production of the novel small RNAs by recruiting RNA dependent RNA polymerases (RdRP) to its targeted RNA. These novel small RNAs are called 22G RNAs, because the vast majority are 22 nt long and start at their 5' end with a guanosine. RdRPs use the target RNA as a template to generate many novel complementary small RNAs. Importantly, this mechanism allows the amplification of the original trigger for robust gene silencing. Since piRNAs initiate the production of the novel small RNAs, the piRNAs are called primary siRNAs, whereas the others are called secondary siRNAs.

Another population of small RNAs are the 26G RNAs (Conine et al., 2010; Gent et al., 2010; Vasale et al., 2010). They share features of miRNAs, piRNAs and 22G RNAs. As their name suggest, they are 26 nt long and start with a guanosine. Similar to piRNAs, they are germline specific and initiate the production of secondary small RNAs. However, in contrast to miRNA and piRNAs, they are not transcribed by RNA polymerase II (Pol II), instead, like 22G RNAs, they are transcribed by an RdRP. They function in spermatogenesis at elevated temperature.

1.1.4 Gene regulation by exogenous dsRNA

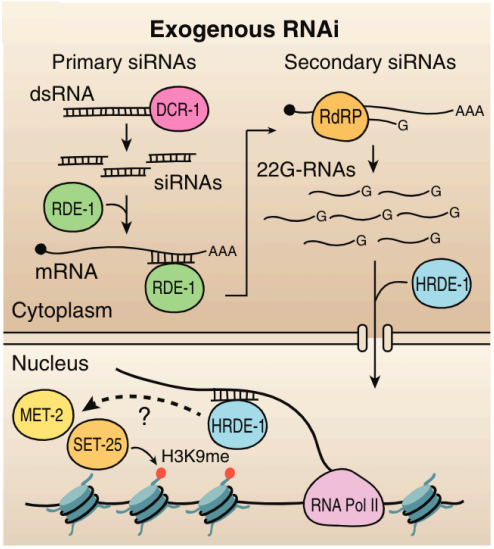


Figure 1 Exogenous RNAi in the *C. elegans* germline

In the cytoplasm, the RNAse III DCR-1 processes exogenous dsRNA into primary siRNAs, which are subsequently loaded into the argonaute RDE-1. The siRNA/RDE-1 complex binds mRNAs and triggers the production of secondary 22G RNAs by RNA dependent RNA polymerases. The secondary siRNAs bind the heritable RNAi Deficient-1 (HRDE-1) Argonaute and other Argonaute to trigger nuclear gene silencing via histone methylations mediated among others by the histone METhyltransferase-like MET-2 and SET (trithorax/polycomb) domain containing-25 proteins by a poorly understood mechanism (from Brown and

Environmental RNAi uses similar principles to mediate its gene regulatory functions once the exogenous dsRNA enters a cell. The exogenous dsRNA is processed by the cytoplasmic endonuclease DICER (Bernstein et al., 2001; Ketting et al., 2001; Knight and Bass, 2001). In addition, it functions in the biogenesis of endogenous miRNA (Grishok et al., 2001; Hutvágner et al., 2001; Ketting et al., 2001). *dicer* is a conserved gene in other organisms. Interestingly, in the *C. elegans* and in the *mus musculus* genome only one gene coding for *dicer* is present. In contrast, *Drosophila melanogaster* has two

paralogues*, dicer-1* and *dicer-2*. Here, the role for miRNA processing and exogenous dsRNA processing has been devided between the two proteins. DICER-1 efficiently processes miRNAs (Lee et al., 2004b), whereas DICER-2 is required for the processing of exogenous dsRNA (Lee et al., 2004b; Liu et al., 2003; Pham et al., 2004). Although, in *C. elegans* only one *dicer* exists, and it has the enzymatic domain required for processing the exogenous dsRNA, additional proteins are required to process the long exogenous dsRNA into small RNA duplexes. One of them is the dsRNA binding protein RDE-4 (Tabara et al., 2002; Thivierge et al., 2012). The resulting so-called primary small interfering RNAs (siRNAs) are loaded into the RDE-1 Argonaute protein (Tabara et al., 1999). Lastly, the signal is amplified by RdRPs similar to the piRNAs. In the soma, secondary small RNAs are generated by the RdRP *rrf-1* while in the germline, the RdRP *ego-1* synthesises secondary siRNAs (Aoki et al., 2007; Sijen et al., 2001; Smardon et al., 2000). In summary, once dsRNA enters the cell it is processed by enzymes of various endogenous pathways, however some proteins are specifically required for environmental RNAi processing.

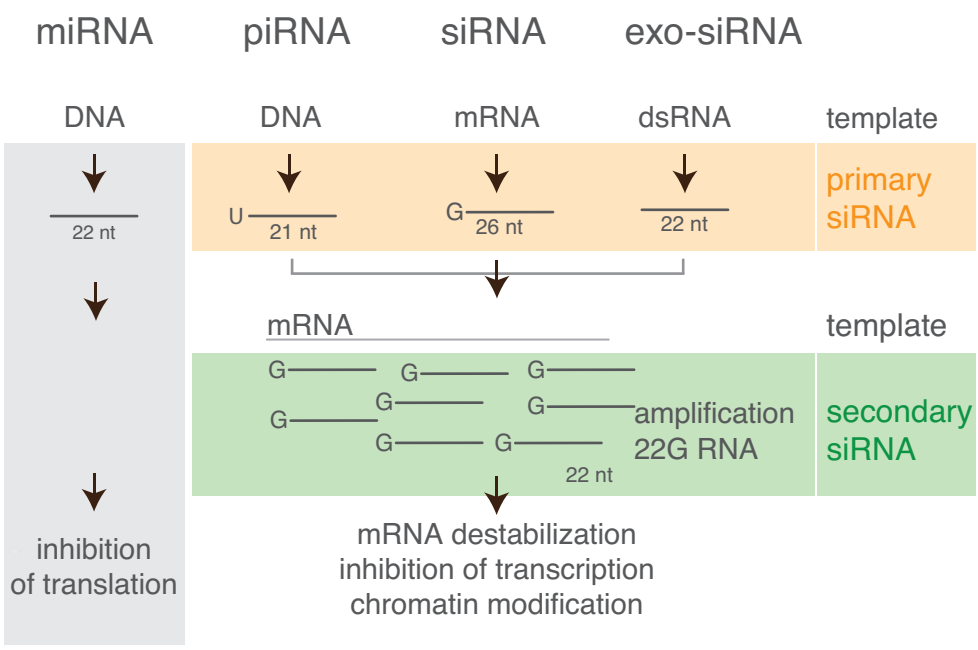


Figure 2 Endogenous and exogenous small RNA types

DNA, mRNA and dsRNA serve as template for primary siRNAs. miRNAs and piRNAs are transcribed from the genome, where as siRNA and exo-siRNAs originate from an RNA template. piRNAs, siRNAs and exo-siRNAs initiate the production of secondary siRNAs from their target mRNA. Primary and secondary siRNAs regulate gene expression.

1.1.5 Uptake of environmental dsRNA

Before dsRNA can be processed into biologically active small RNAs, it has to overcome two hurdles. First it has to enter the organism and second, it must be transported within the worm, since dsRNA is able to silence gene expression is almost every tissue. Genetic screens identified a class of genes, known as *systemic RNAi deficient (sid)* genes, which are required for the uptake of dsRNA from the environment and for its the transport in-between cells (Figure 3). However, these are not required for RNAi function itself (Hinas et al., 2012; Jose et al., 2012; Winston et al., 2002, 2007). The *sid* screen used RNAi by feeding and dsRNA expressed in the pharynx to knock down GFP in the body wall in combination with monitoring RNAi competence in the pharynx

to identify genes involved in RNA transport and uptake, but not required for RNAi within a cell.

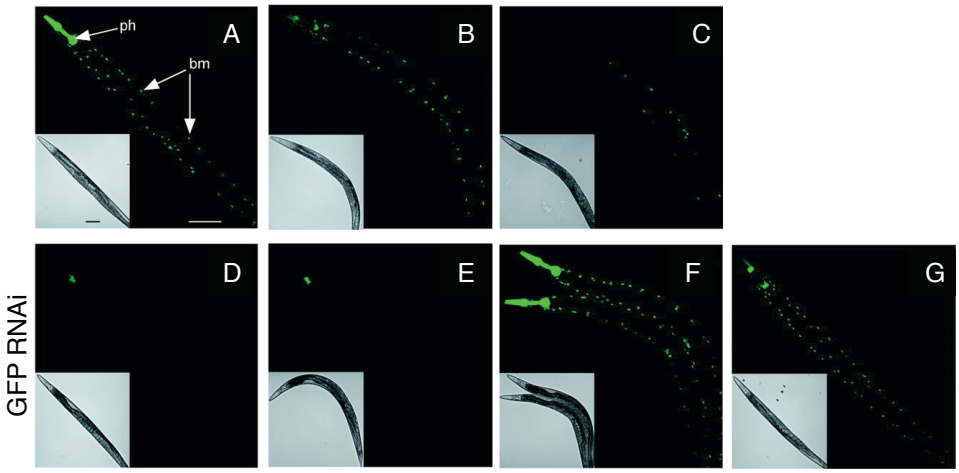


Figure 3 Systemic RNAi screen controls and identification of *sid-1*

In the systemic RNAi screen, a strain expressing GFP in the pharynx (ph) and in the bodywall (bw) (A) was used to show the basal GFP expression. To internally initiate RNAi, a GFP hairpin (hp) specifically expressed in the ph was introduced leading to a reduced expression reduced GFP in the ph and mild reduction in the body wall indicating systemic RNAi (B). To test the extent of knock-down in the bw, a strain additionally expressing a GFP hairpin in the bw was created showing strong reduction of GFP expression. To facilitate the screening process, strain B was exposed to GFP RNAi leading to efficient knock-down (E). The GFP RNAi efficiency was as efficient as in strain A not expressing GFP hp indicating that the GFP hp was not interfering with RNAi. To test the dependency of reduced GFP expression on the GFP RNAi by feeding and the hpGFP a mutation in the primary Argonaute (*rde-1)* was introduced in B. The systemic RNAi screen identified a strain carring a mutation in *sid-1* showing resistance to RNAi by feeding and pharyngeal induced silencing in the bw, but efficient silencing in the pharynx (G). A detailed description of the strains is given in Table 1. (Adapted from Winston et al. 2002)

Transgene	A	B	C	D	E	F	G
Pharynx	GFP	hpGFP GFP	hpGFP	GFP	hpGFP GFP	hpGFP GFP	hpGFP GFP
Bodywall	GFP	GFP	GFP	GFP	GFP	GFP	GFP
Genotype	wt	wt	wt	wt	wt	rde-1 (-/-)	sid-1 (-/-)
Treatment							
RNAi	no	no	no	GFP	GFP	GFP	GFP
GFP Expression							
Pharynx	on	reduced	n/a	off	off	on	on
Bodywall	on	reduced	reduced	off	off	on	off

Table 1 Systemic RNAi screen strains

Two additional screens were performed to identify genes requiered for dsRNA transport. In the *feeding defective (fed)* screen, first mutants were isolated, which were resistant to RNAi by feeding, then in a second step mutants deficient in transport were identified by testing for RNAi competence from direct injections into the germline (Timmons et al., 2003). The same strategy was used in the *RNAi spreading defective* (rsd) screen. It isolated mutants resistant to RNAi by feeding, but susceptible to RNAi injection (Tijsterman et al., 2004). All three screens isolated the gene ZK520.2, today known as *sid-2*. Two of the three screens, the *sid* screen and the *rsd* screen identified the gene C04F5.1 known as *sid-1.* Both genes were characterised extensively on the genetic and molecular level.

1.1.6 dsRNA uptake at the apical intestinal membrane

The single-pass trans-membrane protein SID-2 is required for the uptake of dsRNA from the environment into the gut (McEwan et al., 2012; Winston et al., 2007). Genetic mutations in *sid-2* cause resistance to RNAi by feeding (Winston et al., 2007). However, in these animals, the downstream RNAi

pathways are functional, as pharyngeal expressed RNA hairpins targeting GFP (hpGFP) are able to downregulate the GFP protein levels in the body wall (Winston et al., 2007). In addition, since the RNAi trigger is expressed in a tissue far from the RNAi target, this experiment suggests that dsRNA is transported from cell to cell without SID-2. Therefore SID-2 is not required for intercellular transport of RNA. In addition, GFP::SID-2 fusion proteins are expressed in the intestine and localise at the apical membrane (Winston et al., 2007), further support SID-2's role in dsRNA uptake from the environment. In experiments where worms were soaked in fluorescently labelled dsRNA, accumulation of fluorescent signal in the intestine was *sid-2* dependent (McEwan et al., 2012). Finally, expression of SID-2 in an heterologous system, such as *Drosophila* S2 cells, leads to the accumulation of intracellular cellular fluorescent signal if the labelled dsRNA is provided in the growth medium (McEwan et al., 2012). The accumulation of the fluorescent signal requires an environmental pH of 5 similar to the *C. elegans* intestinal pH 4-6, suggesting SID-2 requires an acidic pH to function (Chauhan et al., 2013). Additional soaking experimenters in S2 cells, this time using 32P-labeled dsRNA, indicate SID-2 is selective for dsRNA longer then 50 bp (McEwan et al., 2012). Finally, the uptake of dsRNA can be inhibited using drugs targeting the endocytosis pathway (McEwan et al., 2012). Overall, SID-2 enhances the ability of dsRNA uptake in *C. elegans* and in *Drosophilia* S2 cells.

1.1.7 Transport of dsRNA within the animal

Upon uptake in endocytic vesicles, dsRNA requires further proteins to enter the cytosol. One of the identified proteins required to make extracellular dsRNA accessible for DICER processing is the conserved putative dsRNA channel SID-1 (Winston et al., 2002). Similar to *sid-2* mutant animals, *sid-1* mutant animals are resistance to RNAi by feeding (Winston et al., 2002). However, SID-1 acts different from SID-2. In contrast to *sid-2* mutants, *sid-1* mutant animals expressing dsGFP in the pharynx are not able to silence GFP in the body wall. A series of genetic experiments indicate that SID-1 is important for

the import of dsRNA into the cell (Winston et al., 2002), both from the environment and from other tissues. First, SID-1 is not required for RNAi in general, since injection of dsRNA in the germline and expressing dsRNA in a cell show germline and cell specific silencing (Winston et al., 2002). Second, SID-1 is not required for the export of RNAi trigger, since expression of dsRNA in *sid-1* mutant tissue can silence transgene expression in *sid-1 (+)* tissues (Jose et al., 2009). Thirdly, RNAi by feeding can induce silencing in the body wall in *sid-1* mutant animals, if SID-1 is specifically expressed in the body wall (Jose et al., 2009). In addition, a series of dsRNA uptake experiments in *C. elegans* primary cells and *Drosophila* S2 cells further support the hypothesis that *sid-1* is important for dsRNA uptake (Shih and Hunter, 2011; Shih et al., 2009). Overall, these findings suggest that dsRNA can enter the animals via SID-2, however SID-1 is required to import the dsRNA into the cytosol and to be accessible to DICER and to initiate RNAi.

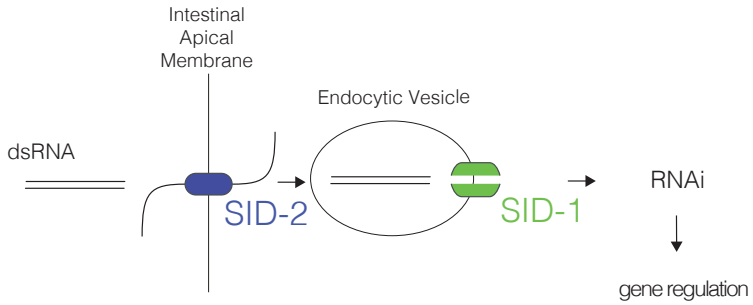


Figure 4 Model for dsRNA transport in *C. elegans*

The proposed dsRNA receptor SID-2 mediates the uptake dsRNA at the apical membrane. An endocytic vesicle is formed, in which the dsRNA enteres the cell. Next, the dsRNA channel SID-1 releases the dsRNA into the cytoplasm.

1.1.8 Genetics of dsRNA uptake into the germline

Further experiments describe how dsRNA can enter cells when present in the body cavity (pseudocoelom) of the worm. A series of experiments using fluorescently labelled dsRNA injection into the pseudocoelom indicated the importance of yolk in dsRNA uptake into the *C. elegans* germline (Marré et al., 2016; Wang and Hunter, 2017). In *C. elegans*, yolk is produced in the intestine and first exported into the pseudocoelom and subsequently imported into the germline through endocytosis dependent on the yolk receptor RME-2 (Grant and Donaldson, 2009; Kimble and Sharrock, 1983). Surprisingly, fluorescently labelled dsRNA injected into the pseudocoelom localises with yolk in the *C. elegans* germline, therefore it was speculated that dsRNA can enter the germline from the pseudocoelom via the yolk receptor RME-2 (Marré et al., 2016). This hypothesis was supported by experiments observing the localisation of fluorescently labelled dsRNA injected into the body cavity (pseudocoelom). Fluorescent signal was observed in wild-type germlines, but not in germlines of *rme-2* mutant animals (Marré et al., 2016), indicating that dsRNA uptake into the germline requires RME-2. Additional experiments examined the ability of the dsRNA taken up by RME-2 to initiate RNAi and found that RNAi is functional in *rme-2* mutant animals, suggesting that dsRNA taken up by rme-2 is not essential for RNAi, and therefore is an additional dsRNA uptake pathway (Wang and Hunter, 2017). This alternative pathway requires *sid-1*, since in *sid-1* mutant animals injected dsRNA does not cause any silencing in their offspring (Wang and Hunter, 2017; Winston et al., 2002). However, if *sid-1* function is restored in the offspring, silencing can be observed (Wang and Hunter, 2017). These experiments demonstrate that the RNAi silencing trigger can enter the embryo in a *sid-1* independent manner, presumably via *rme-2*. However, the RNAi silencing trigger requires SID-1 to be biologically active. Further experiments using *sid-1* and *rme-2* double mutants helped to dissect the roles of *sid-1* and *rme-2* for dsRNA uptake and RNAi. Similar to *sid-1* single mutant animals, *rme-2; sid-1* double mutants are deficient in RNAi in their offspring, indicating that the dsRNA is either not

accessible to or not present in the embryo. The *sid-1* and *rme-2* double mutant experiment showed that *rme-2* is required for dsRNA uptake, since restoring *sid-1* function in such animals did not restore RNAi function (Wang and Hunter, 2017). Together, these experiments suggest that dsRNA enters the cytoplasm of the germline via two routes in wild-type animals. One is an RME-2 and SID-1 dependent route, in which first dsRNA can first enter the germline via RME-2 mediated endocytosis, and then in second step the dsRNA leaves the endocytic vesicle to enter the cytoplasm via SID-1. Alternatively, dsRNA can enter the germline independent of RME-2 and can enter the cytoplasm via SID-1.

1.1.9 Genetic requirements for export and import of dsRNA

Other genes have been identified to play a role in dsRNA transport. Whereas *sid-1* and *sid-2* are absolutely required for uptake and transport of dsRNA, these other genes modulate its transport. Nevertheless, they help to understand how dsRNA transport works.

1.1.10 Endocytic import of dsRNA

Similar to *sid-1*, the genes *sid-3, rsd-3* and *sec-22* are involved in dsRNA import (Imae et al., 2016; Jose et al., 2012; Zhao et al., 2017). These proteins associate with different compartments of the endomembrane system. Together with work in *Drosophila* S2 cells, experiments in *C. elegans* strongly support a model of dsRNA endocytosis as ageneral mechanism of dsRNA uptake (Saleh et al., 2006). Furthermore, in the western corn rootworm, clathrin-dependent endocytosis affects RNAi efficiency providing additional evidence for the importance of the endocytosis pathway in dsRNA uptake (Pinheiro et al., 2018).

However, how the factors involved in dsRNA import relate to each other is not understood. Whereas *sid-3* and *rsd-3* positively regulate RNA transport, *sec-22* negatively regulates RNA transport. In addition, they all localise in different

subcellular regions, SID-3 localises at the apical membrane and or in the cytoplasm, RSD-3 at the trans-Golgi-network and SEC-22 at the late endosome (Imae et al., 2016; Jiang et al., 2017; Jose et al., 2012; Zhao et al., 2017). SEC-22 and RSD-3 both have homologues in the mammalian system acting in the endomembrane transport pathway. *Sec-22* encodes a SNARE (Soluble NSF Attachment Protein Receptor) protein, a family of proteins important for membrane fusion (Ungar and Hughson, 2003) and *rsd-3* encodes a protein having a ENTH (epsin N-terminal homology) domain, which is implicated in endo-membrane trafficking (Saint-Pol et al., 2004). Furthermore, *sid-3's* function in systemic RNAi and has an established role in viral entry (Jiang et al., 2017; Jose et al., 2012; Tanguy et al., 2017). The exact molecular mechanism of *sid-3* in the systemic RNAi pathway and viral entry remains unknown. *sid-3* encodes a conserved tyrosine kinase with a conserved CDC-42 binding domain potentially modulating cell polarity (Balklava et al., 2007), therefore it is possible that *sid-3* changes the cell identity or polarity to interfere with dsRNA uptake and viral entry. Overall, the identification of modulators of RNA uptake and transport strengthens the importance of the endo-membrane transport pathway in systemic RNAi and identifies individual genes for dsRNA import.

1.1.11 Export of dsRNA

Many tissues are capable of exporting dsRNA including the gut, pharynx, neurons and body wall muscle (Devanapally et al., 2015; Jose et al., 2009; Winston et al., 2002). The genes shown to act in RNA export are *mut-2* and *sid-5. mut-2* encodes for a putative nucleotidyltransferase important for transposon control and RNAi (Chen et al., 2005; Collins and Anderson, 1994). An additional role in dsRNA export was identified using genetic studies (Jose et al., 2011). The mechanism by which *mut-2* contributes to dsRNA export is unknown, however, because of its nucleotidyltransferase function, it is tempting to speculate that MUT-2 could physically mark dsRNA for export. Genetic experiments using tissue specific expression indicate MUT-2 acts downstream

of DICER but upstream of RDE-1, suggesting that MUT-2 is not required to transport long dsRNA but potentially for the mobility of short dsRNA.

In plants, elegant grafting experiment using genetically distinct source and sink tissue showed that 24 nt long small RNAs are mobile (Molnar et al., 2010). In *C. elegans,* the chemical properties of exported RNA have been inferred from genetic experiments. Tissue specific rescue experiments incorporating multiple copies of the dsRNA processing factors *rde-4* and *rde-1* indicate that long dsRNA and short dsRNA are mobile, however experiments using single-copy transgenes of *rde-4* indicate that short dsRNA is not mobile (Jose et al., 2011; Raman et al., 2017). It remains unclear how the cell organises export or if the two species of dsRNA are imported using different mechanism.

Similar to *mut-2, sid-5* is a factor involved in exporting RNA. SID-5 is an endosome-associated protein according to immunofluorescent staining and helps to transport dsRNA through the intestine in a *sid-1* independent manner (Hinas et al., 2012). In addition, *sid-5* enhances RNAi in the embryo of a *sid-1* deficient mother as restoring *sid-5* and *sid-1* in the embryo is required for efficient RNAi (Wang and Hunter, 2017). These results highlight the broad role of *sid-5* in the systemic RNAi pathway.

The function of these genes was tested for their role in RNAi by feeding, leaving the characterisation of their molecular role for RNA transport unclear. In addition, although many genes of the endocytosis pathway have been tested, most genes individually did not affect RNAi by feeding (Imae et al., 2016). The absence of effects from mutation in individual genes could be the result of redundancy in the pathway, and it makes the genetic dissection of dsRNA uptake more difficult. In order to elucidate this pathway further, it would be helpful to understand the endocytic route of dsRNA, in which vesicles it is taken up, when or if it leaves these vesicles and how the dsRNA is exported.

1.1.12 RNA mobility in plants

The molecular identity of mobile RNA is best understood in plants (Dunoyer et al., 2010; Melnyk et al., 2011; Zhang et al., 2019). Surprisingly, the molecular mechanism for mobile RNA differs depending on the distance of transport. For the short-distance transport from cell to cell the movement of the mobile RNA is proposed to occur through specific plasmodesmata channels (Voinnet et al., 1998). Experiments with tissue specific expression of the dsRNA processing enzyme DCL4 showed that it is required in the cell exporting the mobile RNA signal, suggesting that 21 nt short siRNA are mobile (Dunoyer et al., 2010). From this experiment, it was suggested that either short siRNA is bound by Argonaute proteins and the mobilised or that short siRNA duplexes are mobile. Experiments using either bombardment or microinjection of single stranded short siRNA or short siRNA duplexes had contradictory results regarding the ability of short dsRNA duplexes to induce silencing (Dunoyer et al., 2010; Yoo, 2004); therefore it is suggested that dicer products are a silencing signal, but additional analysis is required to understand which molecular identity the mobile RNA has.

The long-distance transport of RNA has been easier to disscect due to the spatial seperation. The silencing trigger is transported trough the phloem (Voinnet et al., 1998) and a combination of genetic grafting experiments and next generation sequencing identified 24 nt long siRNAs to be transported from the shoot to the sucrose sink (Dunoyer et al., 2010). An analysis of endogenous siRNA revealed that many small RNA pathways generate mobile RNAs but not all mobile RNAs from each pathway are mobile (Molnar et al., 2010).

Overall, the genetic studies in plants revealed two distinct pathways for short- and long-range RNA silencing. The shorter 21 nt siRNAs are required for the short-range RNA silencing via the cell-to-cell channels, whereas the longer 24 nt siRNAs are required for long-range RNA silencing via the phloem.

1.4 **Evolution of environmental RNAi in *Caenorhabditis***

Many animals are capable of dsRNA uptake and in *C. elegans* genes have been identified for this function. *Sid-2* is an essential gene for dsRNA uptake in *C. elegans* and conserved in the *Caenorhabditis* genus. However, many other *Caenorhabditis* species are not capable of RNAi by feeding (Nuez and Félix, 2012). Interestingly, introducing *C. elegans* SID-2 in other *Caenorhabditis* species is sufficient to make them susceptible to RNAi by feeding, suggesting that these worms generally have intact systemic RNAi pathways downstream of environmental dsRNA uptake (Nuez and Félix, 2012; Winston et al., 2007). The reciprocal experiment, introducing SID-2 from *C. briggsae,* a RNAi by feeding deficient *Caenorhabditis* species*,* in a *sid-2* mutant *C. elegans* does not restore RNAi by feeding competence (Winston et al., 2007). In addition, *Drosophila* tissue culture experiments measuring dsRNA uptake efficiency indicate that only *C. elegans* SID-2 increases dsRNA presents in these cells, whereas SID-2 from the RNAi by feeding deficient strain *C. briggsae* does not (McEwan et al., 2012). These experiments indicate fundamental differences in SID-2 function between the species.

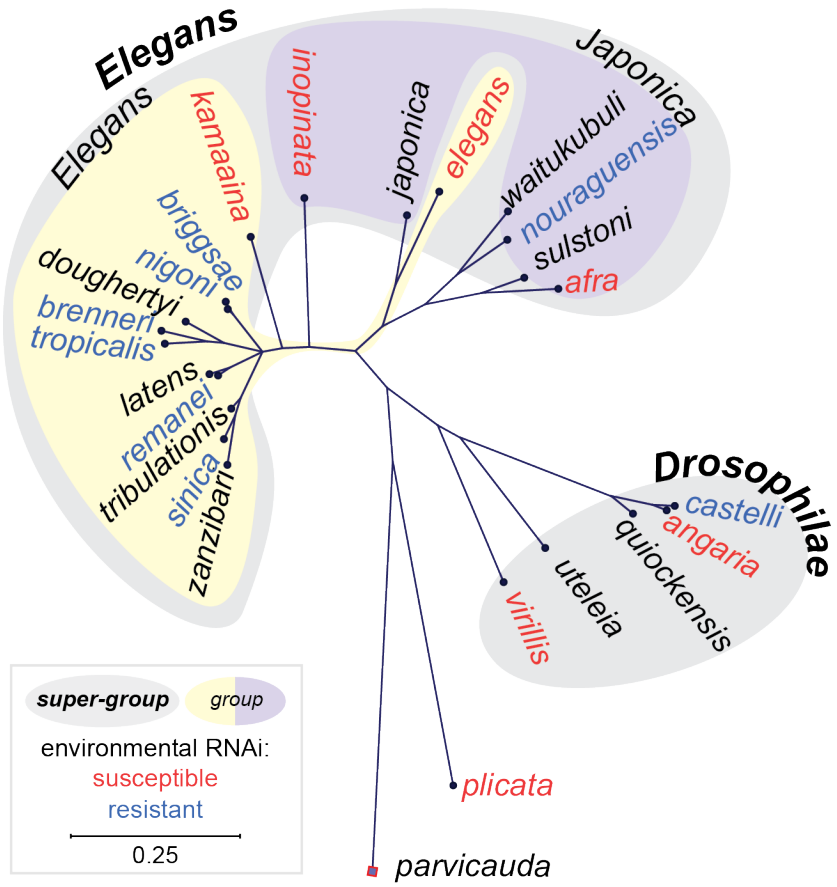


Figure 5 RNAi by feeding susceptibility is not predicted by presence of SID-2

Sid-2 gene tree reconstruction based on publicly available *sid-2* gene sequences. Sequeces were aligned with MUSCLE (v3.7). Ambiguous regions were removed with Gblocks. Gene tree was reconstructed using the maximum likelihood method implemented in the PhyML program (v3.0). Reliability for internal branch was assessed using the bootstrapping method and are shown at the nodes (100 bootstrap replicates). The genetic tree shows the evolution of *sid-2* and indicates the RNAi by feeding competence of the species by colour. Most RNAi by feeding negative species group together and RNAi by feeding positive species are found at various position in gene tree.

Mapping of the ability of species to perform environmental RNAi on the evolutionary tree based on ribosomal RNA shows that the gain and loss of the trait evolved convergently suggesting a high pressure in niche adaptation (Nuez and Félix, 2012). However, *C. elegans,* a worm competent in

environmental RNAi and *C. briggsae,* one deficient in environmental RNAi, have been found in the same natural environment (Barrière and Félix, 2014). Ongoing research investigating the ecology of *Caenorhabditis* might provide insight in the future.

1.5 **Aim of the thesis**

By and large, environmental RNAi is a successful laboratory technique with commercial applications in plant pest control. Furthermore, in *C. elegans*, several studies have helped to understand the molecular mechanism of environmental RNAi. A specific group of genes called the *sids* is required for dsRNA uptake and transport from the environment. However, little is known about how these proteins interact together at the subcellular level or about how RNA export is organised. Most importantly, although environmental RNAi is a wide spread phenomenon, not even in the simple model organism *C. elegans* the biological role of the pathway has been identified. In this thesis, I present and discuss my research regarding dsRNA uptake in *C. elegans*. I characterized the intracellular properties of SID-2 and its role during the nematode development. I determined SID-2 localization at the intestinal apical membrane and the trans-Golgi-network (TGN). This localization is dependent on *sid-3* and its potential substrate *viro-2*, hence demonstrating a molecular interplay between the different sid proteins. The spatial organisation of SID-2 is conserved across species, suggesting a role for the TGN in environmental RNA uptake. By combing small RNA and transcriptome profiling with developmental studies I identified a novel role for systemic RNAi in worm morphology. A detailed phenotypical analysis does not support a nutritional role for dsRNA uptake in pyrimidine synthesis mutant worms, implying that environmental RNA uptake affects morphology through gene regulation. These results uncover the first phenotypic role for the *sid* pathway and identifies the TGN as a central cellular compartment for environmental dsRNA uptake.

2 Chapter – Identification of SID-2 interaction partners

2.1 Introduction

The first member of the *sid* pathway was identified more than 15 years ago, but a full picture of dsRNA transport is still missing. A strong link between dsRNA uptake and the endocytic pathway has been established in general and individual members have been shown to localise at endomembrane compartments. It remains unclear how dsRNA transport is organised intracellularly. The identification of novel members of the *sid* pathway will elucidate the mechanism of dsRNA transport.

Several approaches have been undertaken to identify novel genes in the systemic RNAi pathway. Most successful was the random mutagenesis screen performed by the Hunter laboratory. They isolated *sid-1, sid-2, sid-3* and *sid-5,* all involved in dsRNA transport. This paved the way to elucidate the molecular mechanism of dsRNA transport in *C. elegans* (Hinas et al., 2012; Jose et al., 2012; Winston et al., 2002, 2007). In addition, the Mitani laboratory used a targeted genetic approach and examined mutants of the endocytosis pathway for a role in the *sid* pathway. While testing individual strains carrying mutation in endocytosis pathway genes, they found *rsd-3* caused a *sid* phenotype (Imae et al., 2016). Furthermore, a biochemical approach, a yeast 2 hybrid screen, identified SEC-22 as a binding partner of SID-5 (Zhao et al., 2017). Overall, genetic and biochemical approaches have proven invaluable to identify novel members of the *sid* pathway and have solidified the role of the endocytosis pathway in dsRNA uptake.

So far, the SID-2 protein is the most upstream component required for dsRNA uptake yet identified. However, the mechanism of dsRNA uptake via SID-2 on the molecular level remains largely unknown. SID-2 has been proposed to act

as an dsRNA receptor, however physical binding between dsRNA and SID-2 has not been shown. If SID-2 is a dsRNA receptor, how then does SID-2 mediate dsRNA transport? For example, are other proteins required to internalise the apical SID-2 with dsRNA? Do other proteins bind SID-2 to facilitate import? Is SID-2 endocytosed and if so, which endocytosis pathway is required?

Here, I describe my work, which aims to clearify the mechanism of dsRNA import *C. elegans* by identifying novel interaction partners of SID-2 using two biochemical approaches.

2.2 Results

2.2.1 A novel anti-SID-2 antibody

The gene *sid-2* codes for a 311 amino acid (aa) single pass transmembrane protein (McEwan et al., 2012; Winston et al., 2007). A signal peptide is located in the first 20 aa of the protein, which guide the ribosome to the rough endoplasmic reticulum (ER) for translation. The central transmembrane domain separates the 168 aa long extracellular N-terminus from the 100 aa long intracellular C-terminus Figure 6 (Winston et al., 2007). In summary, SID-2 is a classic transmembrane protein with one extracellular and intracellular domain.

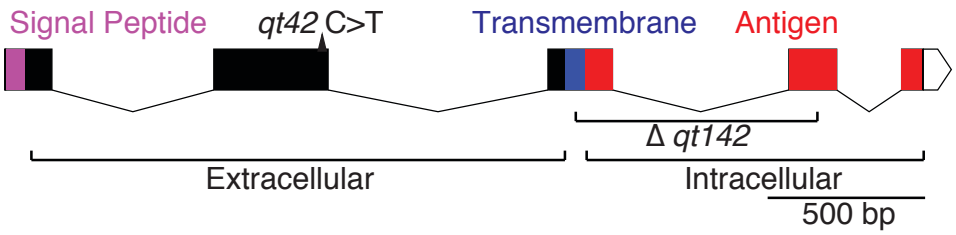


Figure 6 Gene model of *C. elegans' sid-2*

Sid-2 codes for a single pass transmembrane protein, with a N-terminal extracellular domain and an intracellular domain at the C-terminus. A SID-2 antibody was raised against the intracellular domain.
Rectangles indicate exons, rectangle with arrow head indicates the 3'UTR, connecting lines indicate introns, black colour indicates coding region for the extracellular domain, blue colour indicates the coding region for the transmembrane domain, red colour indicates the coding region of the intracellular domain and the SDIX antigen.

Genetic screens have identified several non functional *sid-2* alleles. Two alleles are used in the following experiments (gifts from Craig Hunter). One, the allele *sid-2(qt42)* introduces a C to T conversion causing a premature stop codon at the amino acid position 163. The other, *sid-2(qt142)* is a deletion of 794 bp causing a partial deletion of the transmembrane domain and the SID-2 C-terminus. Both *sid-2* alleles are promising tools for mutant analysis of SID-2.

Little is known about the molecular and cellular function of SID-2. Antibodies are important for many molecular biology techniques used to understand of a protein's function. However, the success of such techniques depends on the specificity of the antibodies. The evaluation of the specificity of novel SID-2 antibody is now possible with the availability of the *sid-2* mutant alleles described above. Therefore, I contracted the company SDIX to raise a SID-2 antibody. A poly-clonal SID-2 antibody was purified from rabbits after injection of a cDNA coding for the intracellular domain of SID-2. Then, the specificity of a polyclonal anti-SID-2 antibody was tested in immunoprecipitation (IP) and western blot (WB).

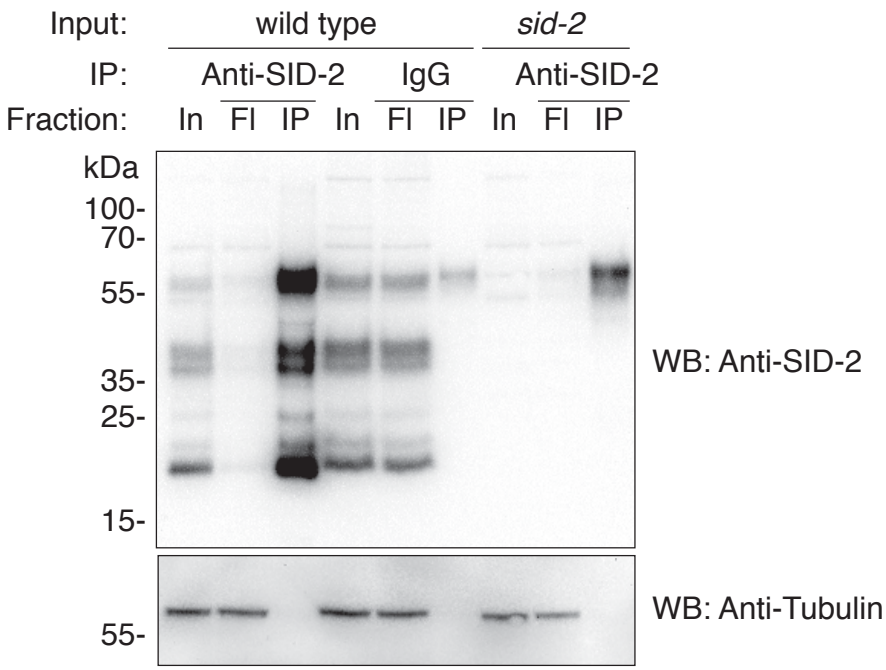


Figure 7 SID-2 antibody specifically detects SID-2

Specificity of SID-2 antibody was determined in a co immunoprecipitation assay. Top: Anti-SID-2 western blot of wild-type and *sid-2(qt142)* worm lysate using anti-SID-2 antibody and wild-type worm lysate using Immunoglobulin G before and after immunoprecipitation. Input (In), Flow through (Fl) and immunoprecipitated (IP). Bottom: Anti-Tubulin immunoblotting of the above membrane.

SID-2 has a predicted molecular weight of 32 kDaltons (kDa). The novel anti-SID-2 antibody generated signals at ~20 kDa, ~37 kDa, ~40 kDa and ~55 kDa in wild-type worm lysate (In), which were not present in *sid-2(qt142)* worm lysate (In). This suggests that the observed signal is specific to SID-2 and that serval modified versions and/or cleavage products of SID-2 with potential biological function exist in the worm (top panel; Figure 7). The observed difference was not due to uneven amounts of worm lysate, since anti-tubulin immunoblotting of the same membrane yielded similar amount of signal (bottom panel; Figure 7). Therefore, the custom anti-SID-2 antibody specifically detects SID-2 in western blot analysis.

Furthermore, incubating wild-type worm lysate with the anti-SID-2 antibody deplets it of SID-2 (wild-type Fl) and SID-2 binds to the anti-SID-2 antibody (wild-type IP*)*. Binding of SID-2 to the anti-SID-2 antibody is specific, since no binding nor depletion is observed in wild-type worm lysate incubated with IgG IP or Fl, respectively Figure 7.The observed SID-2 signal was not due to insufficient washing, since no tubulin signal was detected in wild-type worm lysate IP (Figure 7). Therefore, the anti-SID-2 antibody binds specifically to SID-2 in an immunoprecipitation assay.

2.2.2 Identification of candidate SID-2 interaction partner via co-immunoprecipitation followed by mass spectrometry

The anti-SID-2 antibody can be used to identify novel interaction partners of SID-2 by using it to perfrom a co-immunoprecipitation assay followed by mass spectrometry. Briefly, SID-2 interaction partners bound to SID-2 can be co-purified when SID-2 is immunoprecipitated using the anti-SID-2 antibody. Their identity is then revealed by mass spectrometry. The same immunoprecipitation experiment was performed in *sid-2(qt142)* worm-lysate to identify proteins that co-purify in the absence of SID-2 and are, therefore, not true interaction partners of SID-2. In other words, proteins which are detected more frequently in wild-type SID-2 immunoprecipitation than in *sid-2(qt142)* SID-2 immunoprecipitation are potential SID-2 interaction partners.

To estimate the amount of SID-2 protein enrichment in the immunoprecipitation, the immunoprecipitation was performed with wild-type and *sid-2(qt142)* mutant protein lysated and the quantity of protein was estimated in silver staining (Figure 8).

108

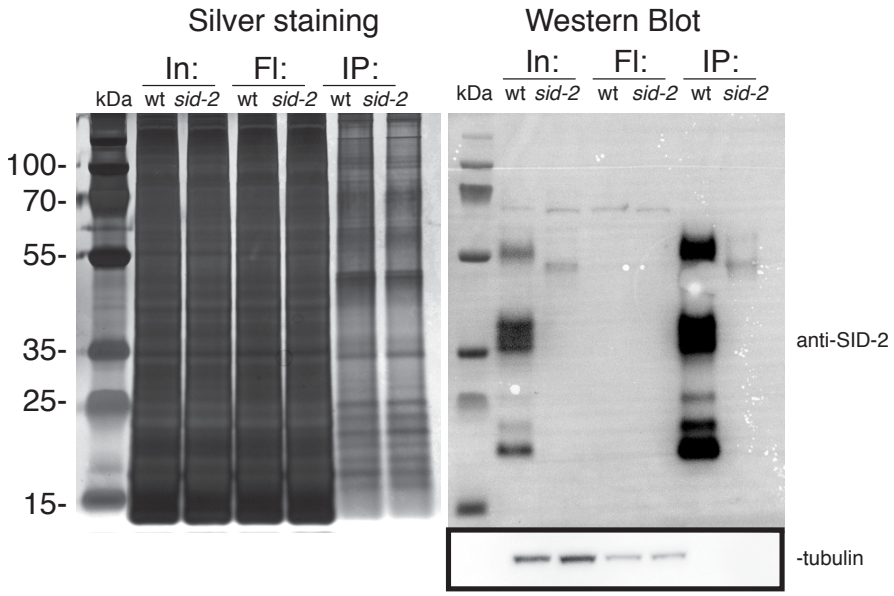


Figure 8 Silver staining and western blot of SID-2 immunoprecipitation

Silver staining of SID-2 immunoprecipitation shows similar amount of protein precipitation in wild type and *sid-2* mutant protein lystate. SID-2 specific bands were not detected. However, in the western blot anlaysis of the same material, SID-2 enrichment via SID-2 immunoprecipitation was detected. This results together suggest that SID-2 is specificlly enrichted in the SID-2 immunoprecipitation, however many additional proteins are precipitated.

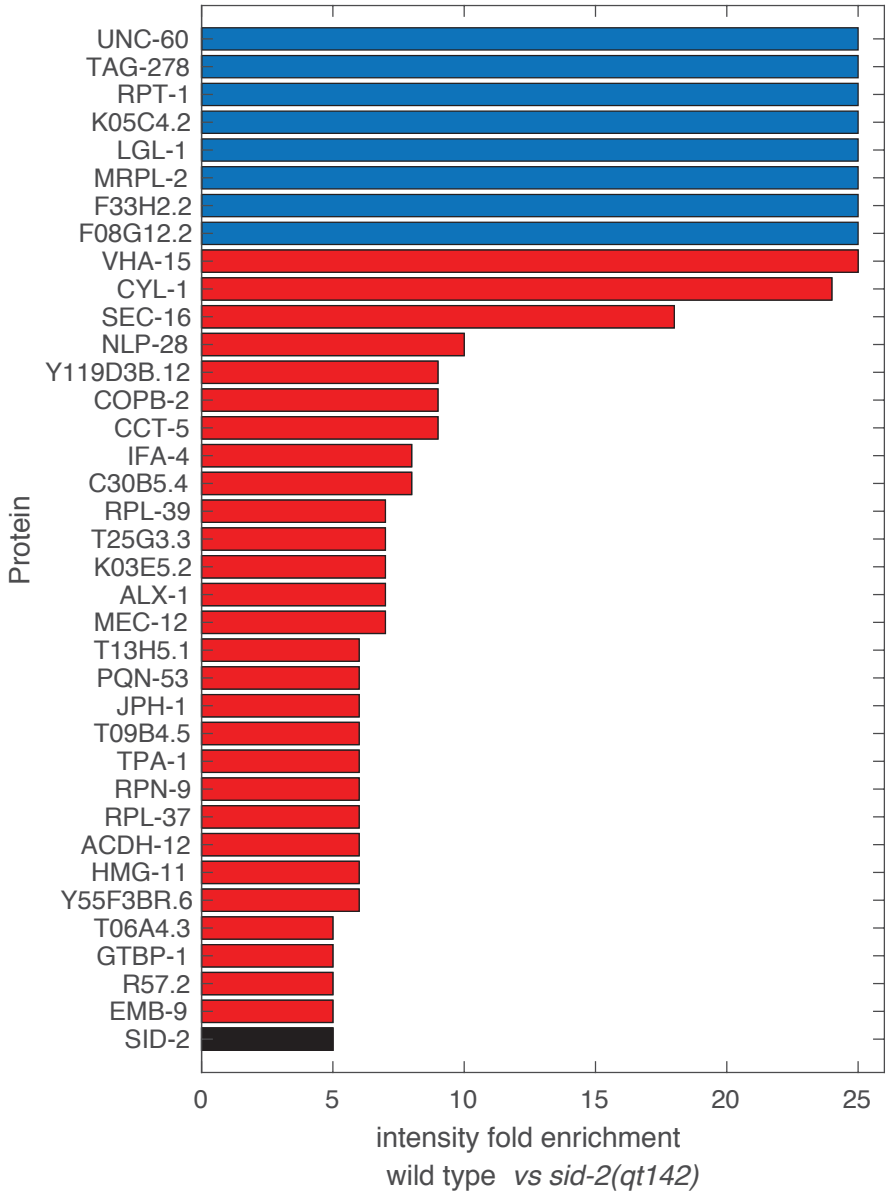


Figure 9 SID-2 potential interacting partners

SID-2 co-immunoprecipitation followed by mass spectromery identifies SID-2 interaction candidates. Only interaction candidates are displayed, which are detected in both wild type samples. The average of the protein mass spec signal internsity of wild type samples is divided by the signal intensity of *sid-2(qt142).* Proteins only detected in wild-type samples are coloured in blue and shown as 25-fold enrichment. Protein detected with a higher enrichment then SID-2 are coloured in red. SID-2's enrichment was 5 fold and is pictured in black.

SID-2 co-immunoprecipitation was performed on two biological replicates of wild-type worm lysate and one biological replicate of *sid-2(qt142)* worm lysate. Mass spectrometry analysis of the samples identified signal intensity for 906 different proteins. The ratio of the mean signal intensity of wild-type and *sid-2(qt142)* samples was calculated to identify candidates of SID-2 interaction. First, I observed a fivefold enrichment of SID-2 in wild-type IP compared to mutant control indicating enrichment of SID-2 was achieved in the immunoprecipitation conditions used. Furthermore, 36 proteins had a higher intensity enrichment than SID-2 in wild-type vs *sid-2(qt142)* and were detected in both wild-type samples. Therefore, these genes were classified as candidate SID-2 interaction partners Figure 9.

Of these 36 genes, the gene *vha-15* was particularly interesting. *Vha-15* encodes the subunit H of the cytoplasmic (V1) domain of vacuolar proton-translocating ATPase (V-ATPase). The orthologous gene in *Drosophila melanogaster,* VhaSFD, was previously identified to be required for dsRNA uptake (Saleh et al., 2006). The gene *vha-15* might not have been identified in the previous genetic screens which identified *sid-2* since it is likely essential for survival, as worms feed on *vha-15* RNAi bacteria are not viable (Saleh et al., 2006). Thus, *vha-15* is a potentially good starting point to start the validation process, however it is likely an essential gene and therefore hard to work on.

2.2.3 The role of candidate SID-2 interaction partner in RNAi by feeding

Interaction partners of SID-2 might affect the function of SID-2 and reduce the amount of dsRNA uptake. Therefore, the ability of animals with knock-downs in potential SID-2 interactors were assayed for dsRNA uptake using RNAi by feeding. Mutant alleles were available for 14 of the 36 candidates. Where possible, several alleles per mutant gene were tested in an RNAi assay. Worms were grown after L1 starvation on *rpb-2* (RNA polymerase II subunit B) dsRNA expressing bacteria and worm size was scored in comparison to wild-type or *sid-2(qt142)* worms Figure 10. RNAi of *rpb-2* slows development presumably

by reducing the function of the RNA polymerase II complex. In this experiment, strains carrying mutations in genes that aid dsRNA uptake via SID-2 should be less susceptible to the RNAi feeding of *rpb-2* than wild-type animals. The mutant alleles were obstained from the Caenorhabditis Genetics Center (CGC).

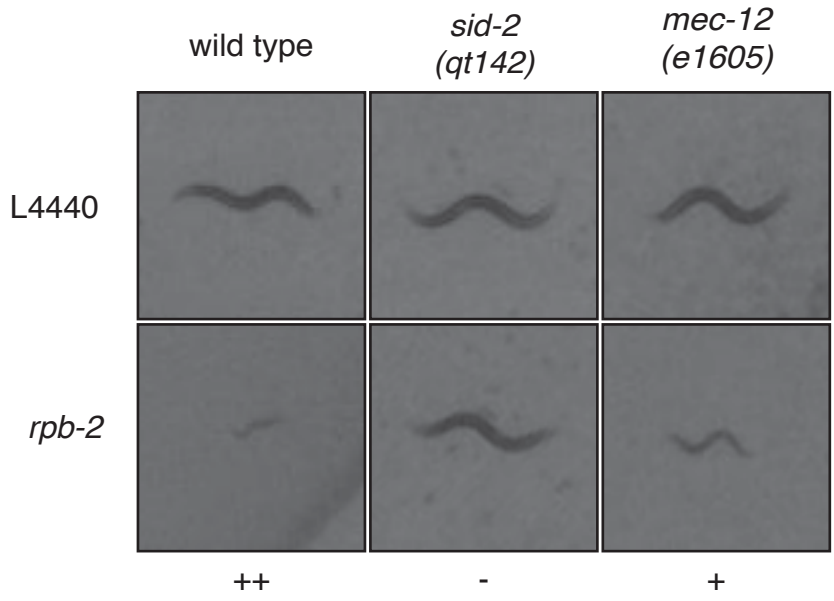


Figure 10 Ordinal scoring system for *rpb-2* RNAi by feeding.

Knock down of RNA polymerase II subunit B (*rpb-2*) interferes with development. Syncronised L1 larvae were fed for 72 hours with bacteria expressing either *rpb-2* dsRNA or the negative control dsRNA (L4440). Picture exemplify the effect of RNAi on wild type, *sid-2(qt142)* and *mec-12(e1605)* mutant animals. RNAi susceptibility is scored into three categories, (++) high, (+) intermediate susceptibility and (-) resistance based on worm size relative to the negative control. Wild type animals show high susceptibility (left), *mec-12(e1605)* animals show intermediate susceptibility (right), whereas *sid-2(qt142)* animals are resistant.

With the exception of the strain CB3284 *mec-12(e1605),* which showed resistance to *rpb-2* RNAi by feeding, all tested strains were as susceptible to *rpb-2* RNAi as wild type worms. However, unlike CB3284, the strain FX05083 *mec-12(tm5083)*, did not exhibit RNAi resistance Table 2 RNAi by feeding

assay against the RNA polymerase II subunit B (*rpb-2*). Hence, the RNAi resistance in CB3284 could be the result of a background mutation or could be specific to the allele *e1605.* Therefore, I concluded that all tested genes are not required for dsRNA uptake and do not reduce *sid-2* function *in vivo.*

Table 2 RNAi by feeding assay against the RNA polymerase II subunit B (*rpb-2*).

Strain	Allele	Gene	*rpb-2*	Replicates	Change
N2			++	5	wt
SX3072	*sid-2(qt142) III*	*sid-2*	-	5	794 bp deletion
VC552	*alx-1(gk275) III*	*alx-1*	++	1	525 bp deletion
VC841	*alx-1(gk338) III*	*alx-1*	++	1	836 bp deletion
VC900	*alx-1(gk412) III*	*alx-1*	++	1	965 bp deletion
PS3232	*cyl-1(sy433) V*	*cyl-1*	++	2	L to F substitution
VC1564	*cyl-1(ok1943) V*	*cyl-1*	++	2	Complex subsitution:1511 bp deletion insertion: AAAAAAA
MJ70	*emb-9(hc70) III*	*emb-9*	++	1	G to E substitution
JH3176	*gtbp-1(ax2029) IV*	*gtbp-1*	++	3	substitution Pending curation
JH3212	*gtbp-1(ax2068) IV*	*gtbp-1*	++	3	deletion Pending curation
JH3215	*gtbp-1(ax2073) IV*	*gtbp-1*	++	3	insertion/deletion Pending curation
RB1238	*hmg-11(ok1303) II*	*hmg-11*	++	1	271 bp deletion
RB1483	*ifa-4(ok1734) X*	*ifa-4*	++	2	711 bp deletion
VC1221	*ifa-4(ok1717) X*	*ifa-4*	++	2	894 bp deletion
VC2286	*jph-1(ok2823) I*	*jph-1*	++	2	637 bp deletion
KK1105	*lgl-1(tm2616) X*	*LGL-1*	++	1	insertion/deletion 213 bp deletion CACACCGAA (variant)
VC2471	*F56F10.4(ok3250) X*	*LGL-1*	++	3	604 bp deletion
FX05083	*mec-12(tm5083) III*	*mec-12*	++	2	233 bp deletion
CB3284	*mec-12(e1605) III*	*mec-12*	+	5	H to Y
VC3390	*rab-30(gk3322) III; pqn-53(gk3534) V; gkDf48 V*	*pqn-53*	++	1	deletion Pending curation
VC2592	*T13H5.1(ok3379) II*	*T13H5.1*	++	3	insertion/deletion 340 bp deletion TTGAAATTTGAGTT (variant)
VC1042	*tag-278(gk439) X*	*tag-278*	++	1	885 bp deletion
MJ563	*tpa-1(k530) IV*	*tpa-1*	++	3	Tc1 transposon insertion
MJ500	*tpa-1(k501) IV*	*tpa-1*	++	1	P to S
CB723	*unc-60(e723) V*	*unc-60*	++	1	Pending curation

154

2.2.4 Identification of candidate interaction partners via yeast two hybrid screening

Yeast two hybrid screens are an alternative technique to identify a proteins's interaction partners. Briefly, when a protein of interested (bait) and a candidate gene (prey) physically interact, a selection gene is activated. The activation of the reporter genes allows the selection of yeast cells in which the physical interaction occur. This is accomplished by fusing the bait to a transcription factor and the prey to a transcription factor activating domain. An interaction between the bait and the prey bring the transcription factor and its activating domain into close proximity; this triggers the expression of a reported gene with the transcription factor binding site. Sequencing of yeast clones activating the reporter gene allows the identification of potential direct binding partners.

The company Hybridgenics performed the yeast two hybrid screen using the intracellular domain of SID-2 as bait. A cDNA library of mix staged *C. elegans* RNA was used to generate the prey. The hits of the yeast two hybrid screen were classified in two categories, high and moderate confidence. Classification of a gene as a high confidence to hit required at least two independent identifications of that gene. Whereas classification as a moderate confidence hit required a single identification event. The screen yielded one high confidence hit, the Rho guanine nucleotide exchange factor (GEF) Y37A1B.17 and 8 moderate confidence hit Table 3. No additional validation has been performed for the individual screen hits.

Table 3 Yeast two hybrid hits

High confidence in the interaction	
Y37A1B.17	Rho guanyl-nucleotide exchange factor
Moderate confidence in the interaction	
C47D12.2	ortholog of human FLJ20071/FLJ90130
W03F8.4	ortholog of human TPRKB (TP53RK binding protein)
ASP-1	homolog of cathepsin D aspartic protease
CTC-1	mitochondrial protein cytochrome c oxidase subunit 1
NDUO-4	NADH-ubiquinone oxidoreductase chain 4
WAGO-4	PIWIL (Argonaute/PIWI) family argonaut
HIM-4	hemicentin belonging to the immunoglobulin superfamily
RPL-25.2	large ribosomal protein subunit L23a protein

RhoGEF activate Rho-family GTPases and regulate cell polarity and migration (Lundquist, 2006). Interestingly, the tyrosine kinase *sid-3* is required for *systemic RNAi* and has a predicted Rho-GTPase CDC-42 binding domain (Jose et al., 2012). In addition, other GEFs regulate vesicle transport, which is a mechanism important for dsRNA uptake (Novick, 2016). Therefore, I was interested in the role of Rho-family GTPases and the RhoGEF Y37A1B.17 in dsRNA uptake. I focused on two of the six Rho-family GTPases for which no or mild phenotypes have been reported, namely, *ced-10* and *crp-1*. CRP-1 localizes to the trans-Golgi network and to recycling endosomes and does not appear to affect actin cytoskeleton reorganization (Jenna et al., 2005). CED-10 localises to the early and recycling endosomes and is important for vesicle transport (Sun et al., 2012). I speculated that the RhoGEF and the RhoGTPases might affect the RNAi by feeding ability of worms negatively or positively. Therefore, I first tested for RNAi resistance by feeding the animals with bacteria expressing the *actin-5* (*act-5)* dsRNA which causes

195 developmental delay, and second, for RNAi sensitivity by feeding bacteria
196 expressing *dumpy-13 (dpy-13)* which reduces worm length.
197 *Act-5* RNAi was chosen to test RNAi by feeding resistance, because *sid-3*
198 animals were resistant to *act-5* RNAi by feeding (Jose et al., 2012). However,
199 all tested mutants of the RhoGEFs silenced the intestinal expression of *act-5*
200 to similar levels as wild-type and *eri-1(mg366) mutant* animals Figure 11. This
201 Indicates, that neither the tested RhoGEFs nor the Rho GTPases positively
202 regulate dsRNA uptake.
203

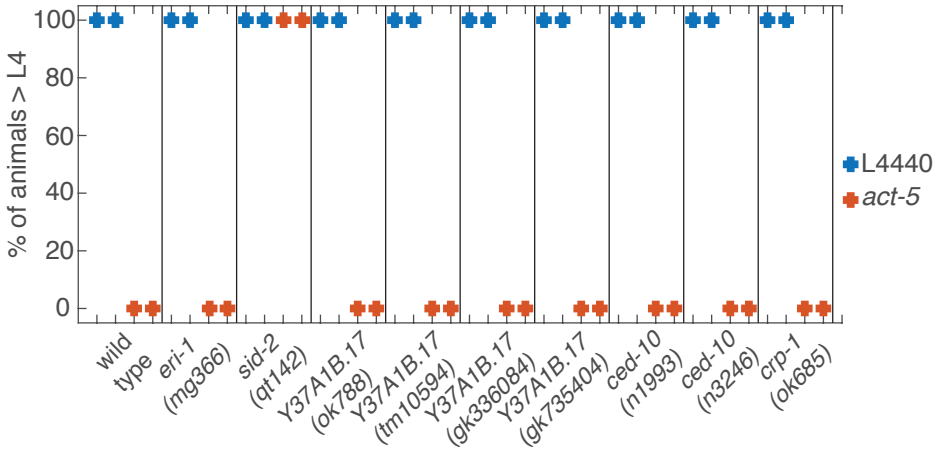


Figure 11 Yeast 2 Hybrid SID-2 interaction candidate RhoGEF Y37A1B.17 and two RhoGTPases are not resistant to RNAi

Mutants of SID-2 interaction candidate from the Yeast 2 Hybrid screen and candidate RhoGTPases were tested for RNAi resistance by feeding animals with *act-5* RNAi. *Act-5* RNAi by feeding causes developmental delay in RNAi competent anaimals. L4 animals were fed with bacteria expressing either *act-5* dsRNA or the negative control dsRNA (L4440). After 48 hours adults were removed from plates. After additional 48 hours, the percentage of animals developing to L4 larva stage and beyond are were counted. All animals fed with negative control bacteria grew at least to the L4 larva stage. None of the animals, except *sid-2(qt142)* mutant animals, which were all fed on *act-5* RNAi expressing bacteria, grew to the L4 larva stage. Data is shown for two biological replicates.

204
205

Secondly, the collagen gene *dpy-13* was chosen to test for increased sensitivity to RNAi by feeding. RNAi against *dpy-13* causes a mild reduction in worm length in wild-type animals, however in mutants with enhanced ability in RNAi by feeding, such as *enhancer of RNAi (eri-1)* mutant animals, greater reduction of worm length can be observed (Zhou et al., 2014). The lengths of F1 offspring were measured and wild-type worm length on *dpy-13* RNAi was only mildly reduced compared to negative control RNAi L4440 Figure 12.

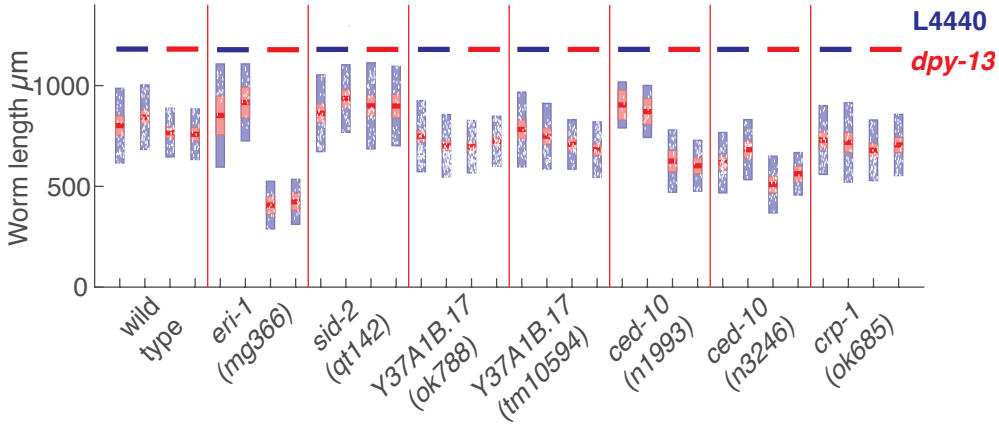


Figure 12 Yeast 2 Hybrid SID-2 interaction candidate RhoGEF Y37A1B.17 and two RhoGTPases do not increase the susceptibility to RNAi by feeding.

Mutants of SID-2 interaction candidate from the Yeast 2 Hybrid screen and candidate RhoGTPases were tested for RNAi enhancement by feeding animals with *dpy-13* RNAi. *Dpy-13* RNAi by feeding reduces worm length in RNAi by feeding competent animals. L4 animals were fed with bacteria expressing either *dpy-13* dsRNA or the negative control dsRNA (L4440). After 24 hours adults were removed from plates. After additional 48 hours, the worm length was measured. RNAi susceptibility was estimated based on worm size relative to the negative control. In wild type, *dpy-13* RNAi reduced worm length to a small extent. In mutants of *eri-1(mg366)*, a known enhancer of RNAi by feeding susceptibility, *dpy-13* RNAi strongly reduced worm length. In *sid-2(qt142)* mutants, *dpy-13* RNAi did not reduce worm length. In Yeast 2 Hybrid SID-2 interaction candidate RhoGEF Y37A1B.17, worm length was not reduced. In RhoGTPase *ced-10(n1993),* worm length was more reduced then in wild type. In RhoGTPase *ced-10(n3246),* worm length was more reduced then in wild type. In RhoGTPase *crp-1(ok685),* worm length was reduced to a simlilar degree as in wild type.

As expected, the worm length in *eri-1(mg366)* mutant animals was dramatically reduced and the worm length of *sid-2(qt142)* mutant animals was not affected by *dpy-13* RNAi. The RhoGEFs Y37A1B.17 and RhoGTPase *crp-1* did not increase the susceptibility to *dpy-13* RNAi. The reduction in worm length of *ced-10(n1993)* and *ced-10(n3246)* animals was stronger than in wild-type animals. However, the large difference in the variation of length reduction between the two alleles argues for background mutations causing the increased susceptibility. Therefore, I conclude that the tested RhoGEFs do not increase the susceptibility to RNAi by feeding.

2.3 Chapter summary

In this chapter, a novel anti-SID-2 antibody is described. Its specificity for SID-2 is established using western blot and immunoprecipitation assays. Furthermore, the results of a SID-2 co-immunoprecipitation assay using the novel anti-SID-2 antibody are laid out, which led to a list of potential SID-2 interacting partners. Attemps to validate these SID-2 interacting partners functionally via RNAi by feeding experiments demonstrates that these candidates do not positively regulate RNAi. Moreover, a yeast-two-hybrid screen identified the RhoGEF *Y37A1B.17* as potential interactor of the intracellular domain of SID-2. The functional validation of the RhoGEF and its potential RhoGTPases (*ced-10* and *crp-1*) indicates that these three genes are not enhancers of RNAi by feeding. However, the RhoGTPase *ced-10* is potentially a negative regulator of RNAi.

3 Chapter – SID-2 localisation

3.1 Introduction

The endocytosis pathway imports macromolecules into the cell (Grant and Donaldson, 2009). A link between endocytosis and dsRNA uptake has been established by several studies. The first evidence was gathered in the *Drosophila* tissue culture system. Experiments using labelled dsRNA demonstrated that its uptake depends on the endocytosis pathway (Saleh et al., 2006). Furthermore, chemical inhibition of the endocytosis pathway in *C. elegans* reduced the efficiency of RNAi by feeding, suggesting the importance of endocytosis for dsRNA uptake in worms as well (Saleh et al., 2006). Subsequently, several studies showed that components of the systemic RNAi pathway co-localise with known components of the endocytosis pathway at the sub cellular level (Hinas et al., 2012; Imae et al., 2016; Jiang et al., 2017; Jose et al., 2012; Zhao et al., 2017). For example, a GFP-RSD-3 (RNAi spreading defective) fusion protein was shown to localise at the TGN (Imae et al., 2016). *Rsd-3* encodes for a homolog of epsinR which is implicated in clathrin mediated membrane trafficking at the trans-golgi-network (Hirst et al., 2003; Kalthoff et al., 2002). Furthermore, one of its protein domain is theENTH (epsin N-terminal homology) domain, which plays a role in endomembrane transport in other systems (Saint-Pol et al., 2004). Moreover, a suppressor of dsRNA import, SEC-22, is found, to localise to the late endosomes/multivesicular bodies (MVBs), and the single pass transmembrane protein SID-5 localises to the late endosome (Hinas et al., 2012; Zhao et al., 2017). The conserved tyrosine kinase *sid-3* is more loosely connected to vesicle transport. Due to its potential to activate the endomembrane trafficking small GTPase CDC-42, this protein is thought to enhance endosomal dsRNA import (Jose et al., 2012). However, because many components of the endocytosis pathway are either essential for viability or their functions can be compensated for by redundant proteins, it has been difficult to investigate their

involvement in dsRNA transport (Imae et al., 2016; Saleh et al., 2006). Thus, the exact route by which dsRNA enters the cell remains unknown.

Great progress in understanding how dsRNA is transported emerged from studies that focused on its uptake from the environment. First, RNAi by feeding was shown to require the single pass membrane protein SID-2. However, SID-2 was shown to be dispensable for RNAi transport and function once the dsRNA has entered the organism. Together, these two findings indicate that SID-2 is involved in the early steps of dsRNA import from the environment but is not required for dsRNA processing into small RNAs (Winston et al., 2007).

SID-2 is a conserved protein in the genus *Caenorhabditis,* but only a few *Caenorhabditis species* are susceptible to RNAi by feeding (Nuez and Félix, 2012; Winston et al., 2007). Surprisingly, deficiency of RNAi by feeding in the species *C. briggsae* and *C. remanei* could be rescued by the heterologous expression of *C. elegans' sid-2.* In a reciprocal experiment, the heterologous expression of *C. briggaes' sid-2* failed to restore RNAi by feeding in *sid-2* mutant *C. elegans.* This experiment indicated that, although *C. briggsae en*codes for a SID-2 homologue, it did not function in the uptake of dsRNA from the environment (McEwan et al., 2012; Nuez and Félix, 2012). Together, these results show a clear functional difference between the SID-2 proteins of the two species, however the molecular differences remain to be investigated.

A series of hybrid fusion protein expression experiments helped to characterise this functional difference by studying each domain of the protein. A fusion protein of the extracellular domain of *C. elegans'* SID-2 in combination with the *C. briggsae* intracellular domain was able to mediate RNAi by feeding, indicating an important role of the extracellular domain for dsRNA uptake (McEwan et al., 2012). In contrast, the *C. briggsae* extracellular domain fused to the intracellular domain of *C. elegans* failed to mediate dsRNA uptake (McEwan et al., 2012). Together, these experiments suggest that the

extracellular domain of *C. briggsae sid-2* functions differently from the *C. elegans* extracellular domain.

To further understand the molecular mechanism of the *C. elegans* SID-2 extracellular domain in dsRNA uptake, mutations in individual amino acids were introduced. Specifically, histidine residues were altered, since they have been shown to directly bind to the backbone of the dsRNA helix (Liu et al., 2008). Overexpression experiments indicated that one of the three histidines in the extracellular domain was required for dsRNA transport. However, mutations of all 3 histidines did not affect dsRNA transport (McEwan et al., 2012). Unfortunately, due to technical difficulties and contradicting results, it remains unknown if the extracellular domain of *C. elegans* SID-2 binds dsRNA. Overall, the fact that the SID-2 protein is conserved within the *Caenorhabditis* species while its function is not, raises the question of SID-2's natural function.

3.2 Results

3.2.1 SID-2 localises at the apical membrane and in cytoplasmic punctuated foci in *C. elegans*

Very little is known about the cell biology of dsRNA uptake in *C. elegans.* The localisation of SID-2 at the apical membrane of the intestinal cell has been reported on the basis of over-expression of SID-2::GFP fusion proteins in *C. elegans* and *D. melanogaster* cell lines (McEwan et al., 2012; Winston et al., 2007) (Figure 13). In order to test if the endogenous SID-2 protein also localised to the apical membrane, I decided to visualise it using the SID-2 antibody. I tested the specificity of the SID-2 antibody by performing immunofluorescence staining on dissected intestines of wild-type and *sid-2* mutant *C. elegans* (Figure 13). In *C. elegans* wild-type animals, staining was detected at the intestinal apical membrane and surprising in many cytoplasmic punctuated foci (Figure 13(I)). In contrast, the two functional null mutants, *sid-2(qt142)* and *sid-2(qt42)* did not yield any detectable staining (Figure 13(III + IV)), indicating that the detected signal in *C. elegans* wild-type animals is

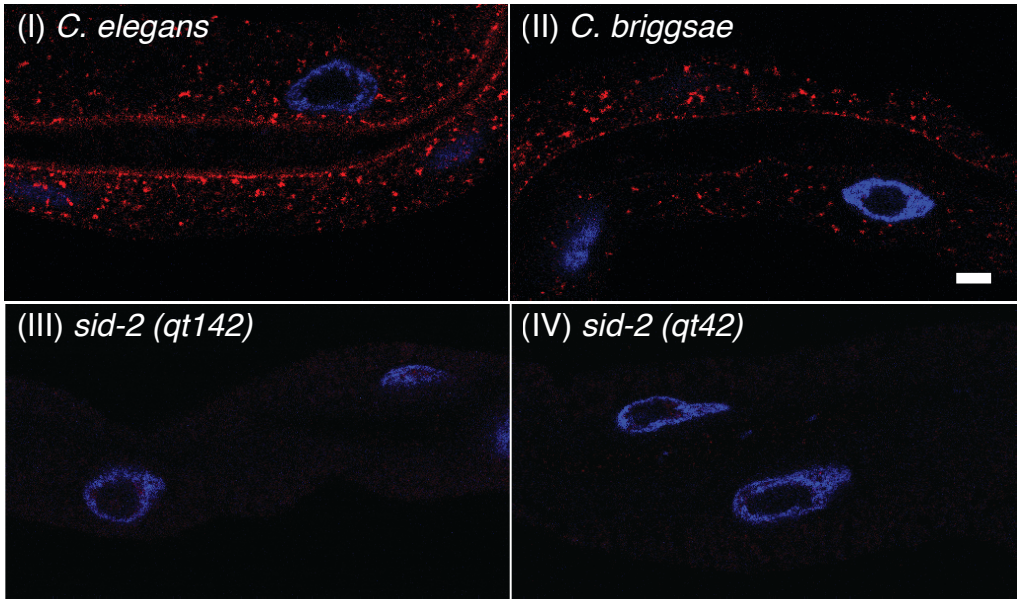


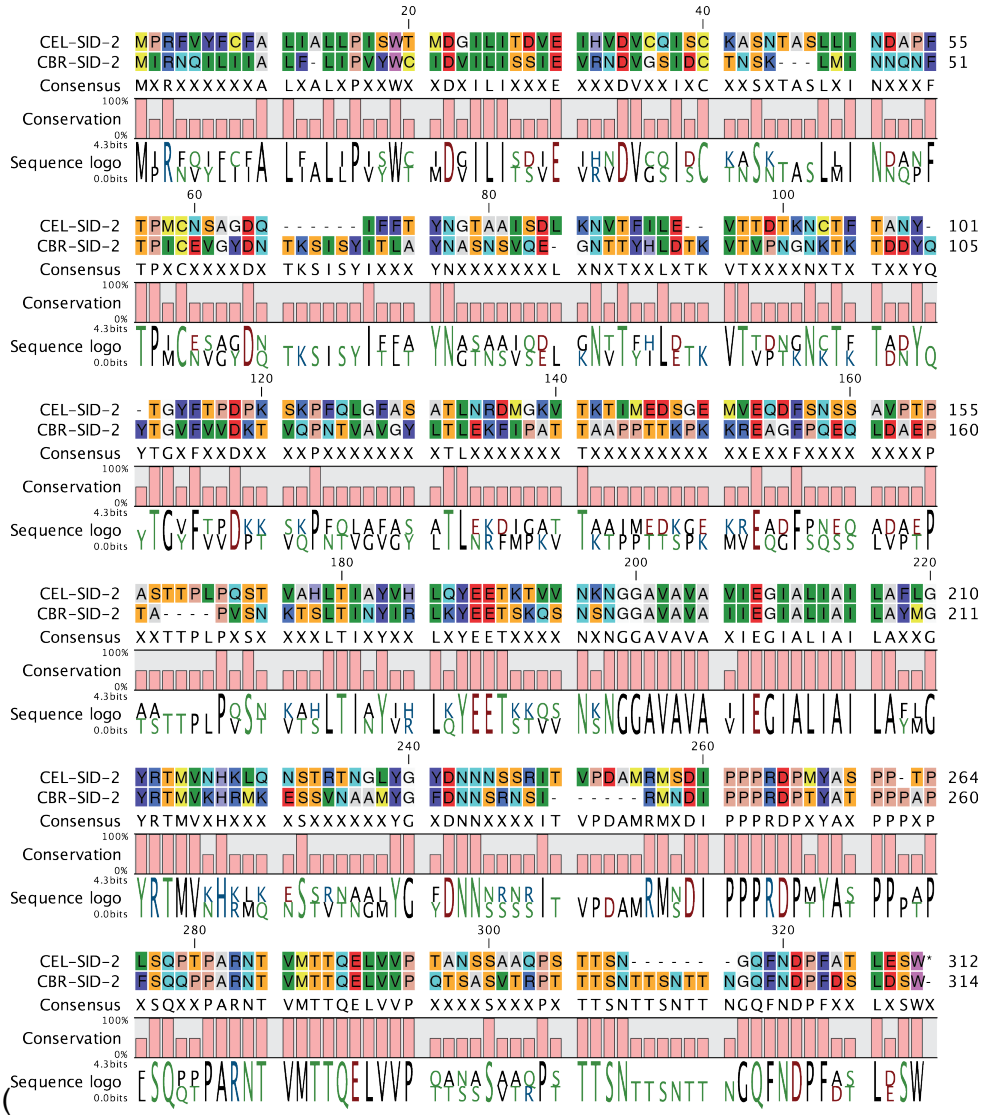
Figure 13 Novel subcellular localisation of SID-2

Confocal microscopy of SID-2 immunofluorescent staining. Adult intestines were dissected and fixed in formaldehyde and methanol. DNA was stained using DAPI (blue). SID-2 was immunostained using anti-SID-2 antibody. SID-2 antibody was visualised using an anti rabbit Alexa594 antibody (red). In (I) *C. elegans* wild type and (II) *C. briggsae* staining at the apical membrane and in punctuated foci was detected. In (III) *C. elegans sid-2(qt142)* and (IV) *C. elegans sid-2(qt42)* only DAPI staining was detected. Scale bar represents 5 µm.

specific signal for SID-2. This novel analysis shows that endogenous SID-2 localises at the apical membrane and, surprisingly, in punctuated foci in the cytoplasm as well.

3.2.2 SID-2 localisation is conserved in *C. elegans* and *C. briggsae*

Since SID-2 function is not conserved in *C. briggsae* but the protein structure



is (Figure 14), I speculated that the SID-2 antibody might be a useful tool to study the SID-2 in *C. briggsae,* as well. Therefore, I set out to test if the *C. elegans* SID-2 antibody is able to recognise the highly conserved *C. briggsae* SID-2. Similar to what I observed in *C. elegans*, I detected signal in cytoplasmic foci and at the apical membrane in *C. briggsae* intestinal cells (Figure 13 (II)). Hence, I assume that the *C. elegans* SID-2 antibody detects *C. briggsae* SID-

2 and conclude that SID-2 localisation is conserved in the genus *Caenorhabditis*.

Figure 14 *C. elegans* and *C. briggsae* SID-2 protein alignment

SID-2 protein sequence is highly conserved within the transmembrane domain and within the N-terminal region of SID-2 between *C. elegans* and *C. briggsae*.

3.2.3 SID-2 associates with in the TGN-network and the recycling endosome

Vesicle trafficking has been shown to be essential for dsRNA uptake into cells and for dsRNA transport from cell to cell. Therefore, I wondered whether SID-2 localises with known markers of cellular compartments of the vesicle trafficking pathway. I performed SID-2 and GFP co-immunofluorescent staining in dissected *C. elegans* intestines of worms expressing fluorescent protein fusions of cellular compartment markers. I found that SID-2 associates with SYN-16 at the trans-Golgi-Network (Figure 15 I) and near the medial Golgi marker mannosidase (Figure 15 II). SID-2 rarely overlapped with the autophagosome marker LGG-1 (Figure 15 III) and did not overlaps with the late endosome marker RAB-7 (Figure 15 IV). At the apical membrane, SID-2 colocalises with the recycling endosome marker RAB-11 (Figure 15 V). These results suggest that SID-2 associates with the TGN-network and the recycling endosome.

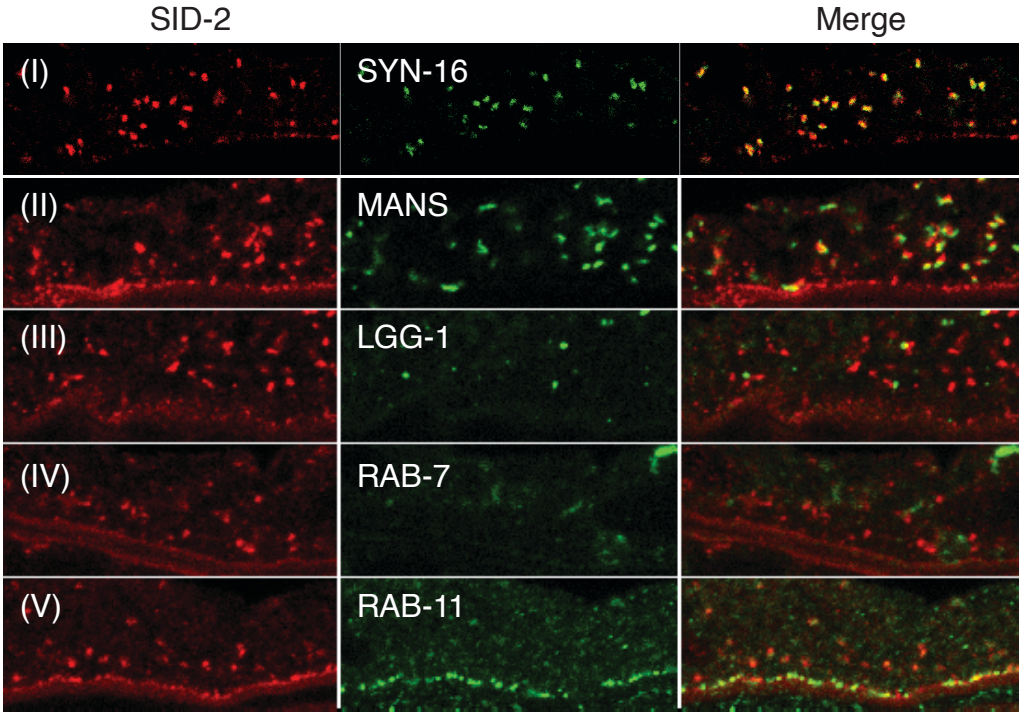


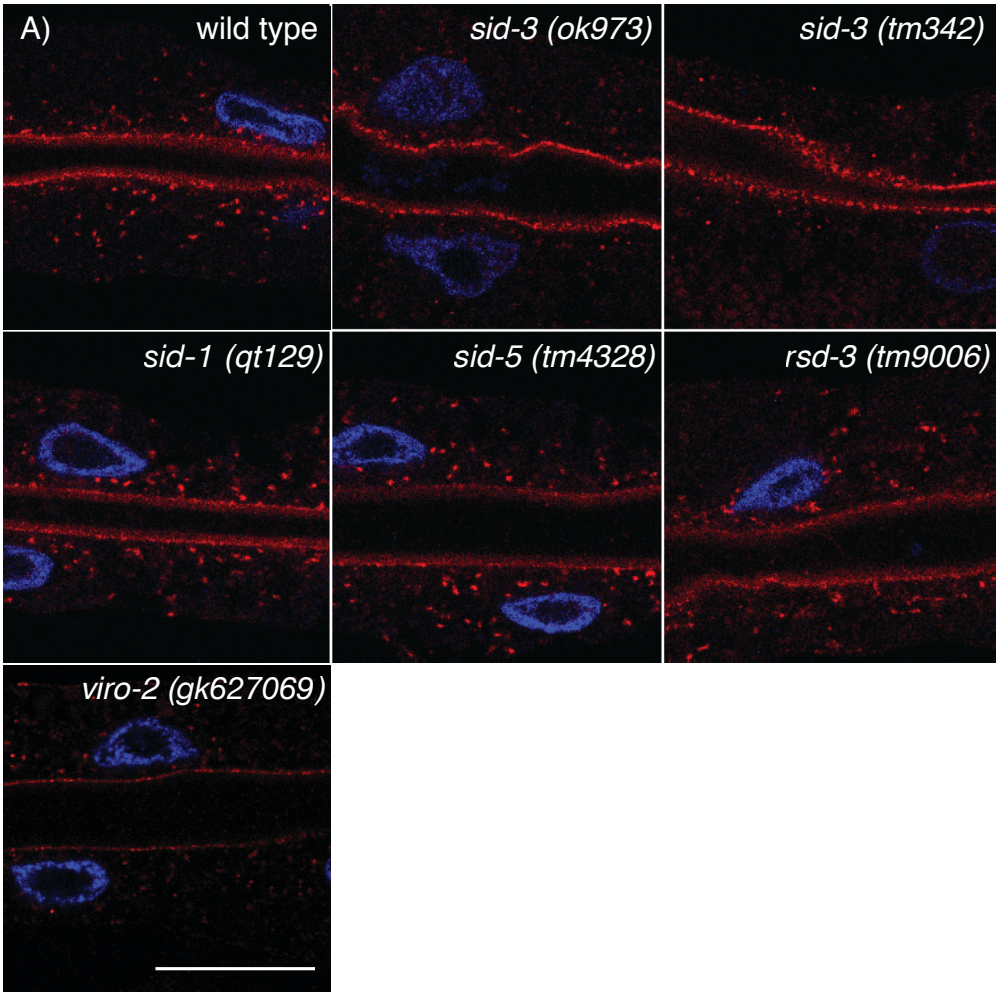
Figure 15 SID-2 co-localises with recycling endosome and TGN marker proteins

Confocal microscopy of SID-2 and GFP co-immunofluorescent staining. Adult intestines expressing fluorescent protein fusions to cellular compartment markers were dissected and fixed in formaldehyde and methanol. DNA was stained using DAPI (blue). SID-2 and GFP were immunostained using anti-SID-2 antibody or anti-GFP antibody, respectively. SID-2 antibody was visualised using an anti rabbit Alexa594 antibody (red). GFP antibody was visualised using an anti mouse GFP antibody. Left column shows SID-2 antibody staining in red. Middle column shows anti-GFP staining in green. (I) trans-golgi-network (YFP-SYN-16), (II) medial golgi (MANS-GFP), (III) autophagosomes (LGG-1::GFP), (IV) late endosomes (GFP-RAB-7) and (V) recycling endosomes (GFP::RAB-11).

3.2.4 The tyrosine-protein kinase SID-3 and SID-3's potential substrate VIRO-2 affect SID-2 localisation

Several genes have been shown to be required for systemic RNAi, however very little is known on how these genes mechanistically interact. The discovery of endogenous SID-2 localization prompted me to investigate the spatial

relationships between SID-2 and other components of the systemic RNAi pathway. Using quantitative SID-2 immunofluorescent staining in dissected *C. elegans* intestines of wild-type and mutant animals, I analysed SID-2 localization in *sid-3, sid-1, viro-2, sid-5* and *rsd-3* mutant animals. The cytoplasmic SID-2 abundance was significantly reduced in two independent mutants for the apical membrane tyrosine kinase *sid-3* (*sid-3(ok973)* and *sid-3(tm342)),* ($p < 0.01$ t-test) (Figure 16). This result suggests that *sid-3* regulates SID-2 localisation, perhaps by modulating the transport from the apical membrane to the TGN. To support the idea of a regulatory role for *sid-3*, I tested the SID-2 localisation in the potential substrate of SID-3 virus induced reporter off 2 (*viro-2*), which also associates with the apical membrane. Similarly, the cytoplasmic SID-2 abundance was significantly reduced in the *viro-2(gk627069)* mutant animals ($p < 0.01$ t-test) (Figure 16). Furthermore, I tested the ENTH (epsin N-terminal homology) domain protein RNAi Spreading Defective-3 (*rsd-3)*, which localises at the TGN, and found that its depletion/mutation did not affect the subcellular localisation of SID-2. Similarly, mutation of the endosome associated SID-5 and the dsRNA channel SID-1 did not impair SID-2 localization. In summary, SID-2 localisation is regulated by *sid-3* and *viro-2* but does not depend on the presence of several components of the systemic RNAi pathway (*sid-1, sid-5, rsd-3).*


A)
wild type
sid-3 (ok973)
sid-3 (tm342)
sid-1 (qt129)
sid-5 (tm4328)
rsd-3 (tm9006)
viro-2 (gk627069)

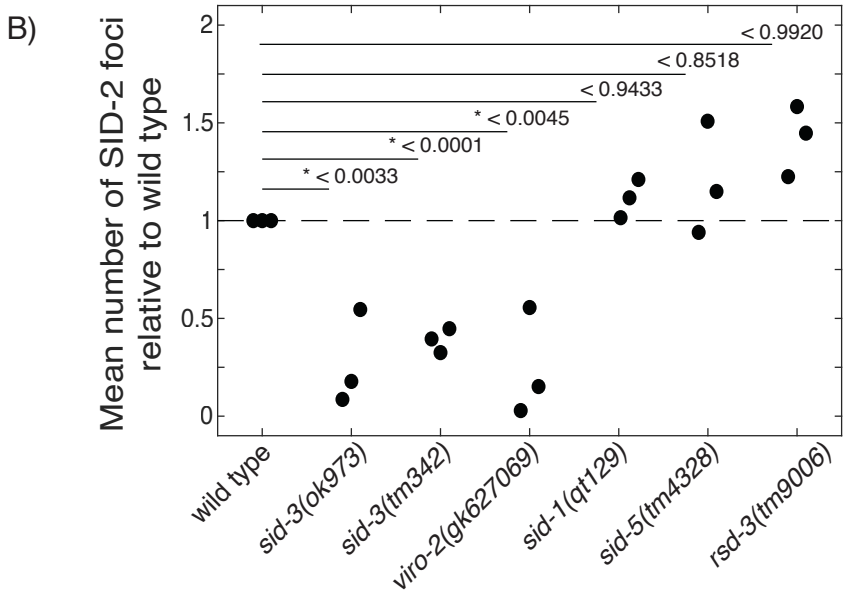

B)
Mean number of SID-2 foci relative to wild type
2
1.5
1
0.5
0
< 0.9920
< 0.8518
< 0.9433
* < 0.0045
* < 0.0001
* < 0.0033
wild type
sid-3(ok973)
sid-3(tm342)
viro-2(gk627069)
sid-1(qt129)
sid-5(tm4328)
rsd-3(tm9006)

Figure 16 Lack of Tyrosine-protein kinase SID-3 and SID-3's potential substrate VIRO-2 affect SID-2 abundance.

Confocal microscopy of SID-2 immunofluorescent staining. Adult intestines were dissected and fixed in formaldehyde and methanol. DNA was stained using DAPI (blue). SID-2 was immunostained using anti-SID-2 antibody. SID-2 antibody was visualised using an anti rabbit Alexa594 antibody (red). A) Representative images of wild-type or indicated mutant background is shown. Scale bar represents 20 µm. SID-2 localises at the apical membrane and at the TGN in wild type, *sid-1(qt129), sid-5(tm4328)* and *rsd-3(tm9006).* In *sid-3(ok973), sid-3(tm342)* and *viro-2(gk627069)* animals SID-2 localises at the apical membrane, but less SID-2 staining at the TGN is detected compared to wild type.
B) Quantification of relative number of detected SID-2 foci in mutant background compared to wild-type. Two-sample t-test, * indicates significant difference alpha <0.05 of mutant to wild-type after correction for multiple testing (Bonferroni correction)

3.2.5 Change in SID-2 abundance upon viro-2 mutation cause no deficiency in environmental RNAi

The genes *viro-2* and *sid-3* affect SID-2 subcellular localisation. Additionally, they both promote viral infection (Jiang et al., 2017; Tanguy et al., 2017). Furthermore, the human orthologues of SID-3 and VIRO-2 form a kinase-substrate pair (Yokoyama et al., 2005). Together, these findings suggest that VIRO-2 and SID-3 act in one pathway with one being upstream of the other. In order to understand if these two genes regulate SID-2 function, I asked whether *viro-2* mutations affects RNAi by feeding in the same way as it has been previous reported for *sid-3* mutations (Jose et al., 2012). Therefore, I performed an RNAi by feeding experiment targeting the mRNA *posterior segregation-1 (pos-1),* which is required for embryo development and scored the number of hatched eggs. While *sid-3* mutant animals showed partial RNAi by feeding resistance, *viro-2* mutant animals were susceptible to environmental RNAi like wild-type animals (Figure 17), indicating that, despite changing SID-2 localization, *viro-2* does not impair SID-2 function.

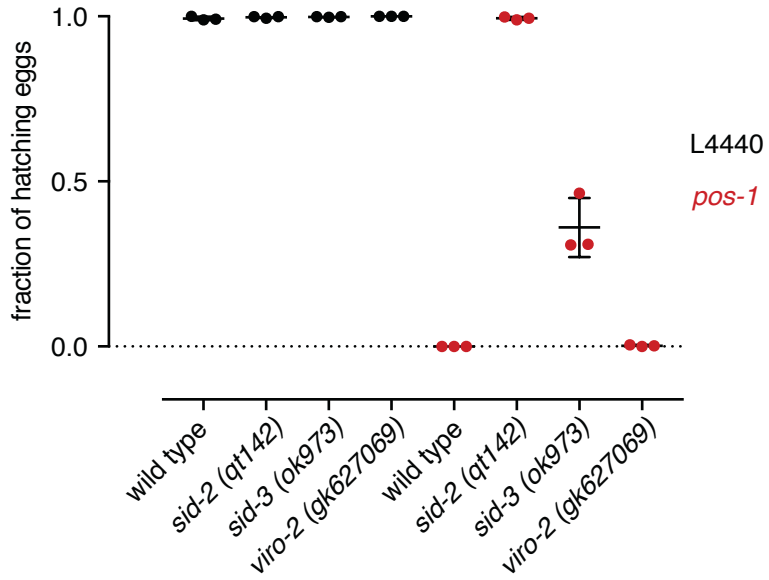


Figure 17 *Viro-2* is RNAi by feeding susceptible

Pos-1 RNAi by feeding experiments test if *viro-2* mutants are RNAi by feeding resistant. *Pos-1* RNAi by feeding is embryonic lethal. L1 animals were fed with bacteria expressing either *pos-1* dsRNA or the negative control dsRNA (L4440). After 72 hours adults were removed from plates. After additional 24 hours, the fraction of hatched eggs was counted. All eggs of animals fed with negative control bacteria hatched. No eggs of wild type animals and *viro-2(gk627069)* mutant animals fed on *pos-1* RNAi hatched. Around 30% of *sid-3(ok937)* mutant animals fed on *pos-1* RNAi hatched. All eggs of *sid-2(qt142)* mutant animals fed on *pos-1* RNAi hatched. Data is shown for two biological replicates.

3.2.6 dsRNA redistributes SID-2 subcellular in *C. elegans* and *C. briggsae*

Some plasma membrane receptors are recycled via the TGN after ligand binding (Johannes and Popoff, 2008). Since SID-2 is found both at the TGN and at the apical membrane, I asked if SID-2 localisation is dynamic. First

worms were starved for 1 hour to remove residual bacteria. Then worms were exposed for 3 hours to either NTPs or dsRNA or kept starved. Then I performed immunofluorescence staining of dissected intestines of *C. elegans* and *C. briggsae* adults. In *C. elegans,* SID-2 localises at the apical membrane and the TGN after starvation and exposure to NTPs, however the presence of dsRNA leads to changes of SID-2 subcellular localisation (Figure 18). Surprisingly, dsRNA redistributed SID-2 as well in *C. briggsae.* Therefore, I

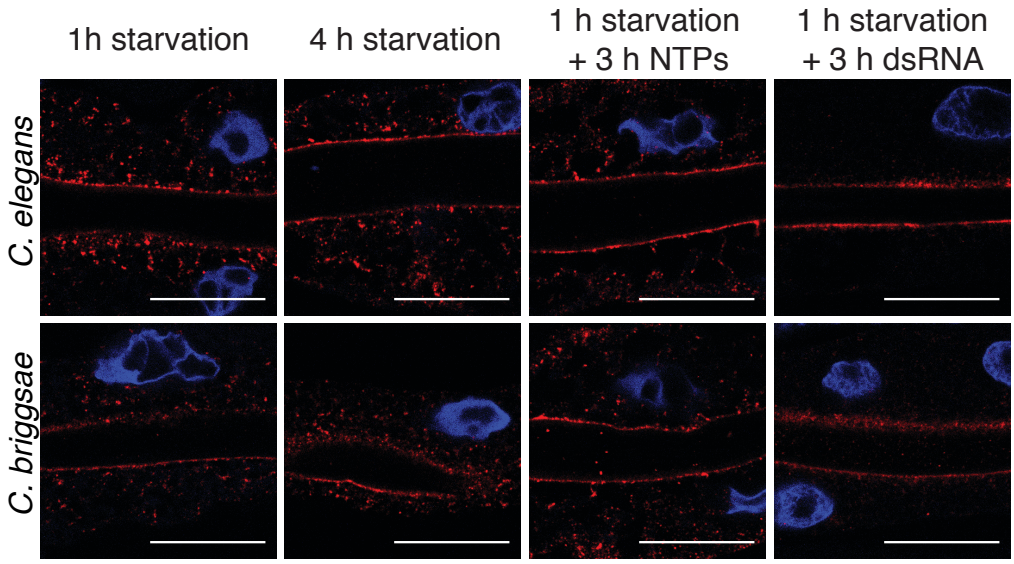


Figure 18 dsRNA causes redistribution of SID-2.

Confocal microscopy of SID-2 immunofluorescent staining. *C. elegans* and *C. briggsae* adult animals were starved for 1 hour, then either starved for an additional 3 hours, incubated for 3 hours with NTPs or dsRNA. Then intestines were dissected and fixed in formaldehyde and methanol. DNA was stained using DAPI (blue). SID-2 was immunostained using anti-SID-2 antibody. SID-2 antibody was visualised using an anti rabbit Alexa594 antibody (red). SID-2 localises at the apical membrane and at the TGN in *C. elegans* and *C. briggsae* after 1 hour of starvation, 4 hours of starvation, 1 hour of starvation and 3 hours incubation with NTPs. SID-2 localises at the apical membrane after 1 hour of starvation and 3 hours incubation with dsRNA. Representative images of indicated condition is shown chosen from 3 biological replicates. Scale bar represents 20 µm.

conclude that SID-2 is able to change localisation upon availability of dsRNA, both in *C. elegans* and *C. briggsae.*

3.2.7 SID-2 is a dsRNA binding protein

The redistribution of SID-2 by dsRNA might be due to direct contact with dsRNA. Therefore, I tested the dsRNA binding ability of the extracellular domain of SID-2 of both species *in vitro*. I incubated dsRNA with the extracellular domain of *C. elegans* and *C. briggsae* SID-2 fused to the Maltose binding protein (MBP) or with the MBP protein alone. DsRNA bound to the proteins was detected in significantly higher amounts on *C. elegans* and *C. briggsae* SID-2ex MBP fusion proteins compared to MBP proteins alone Figure 19. Therefore, I propose that the SID-2 extracellular domain has dsRNA binding abilities and speculate that SID-2 is a dsRNA receptor.

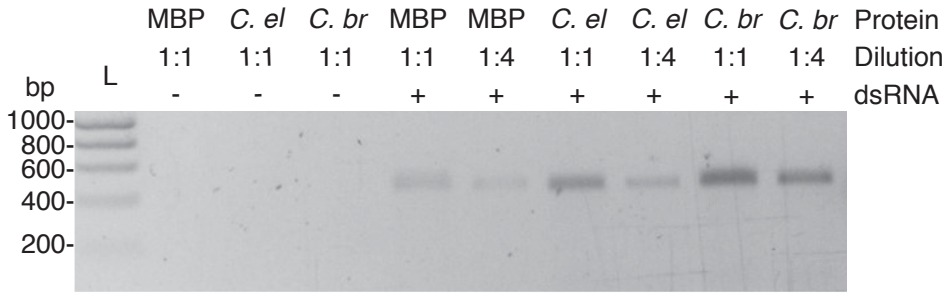


Figure 19 SID-2 binds dsRNA in vitro

Analytical agarose gel for in vitro bound dsRNA. Maltose binding protein (MBP)-*C. elegans* SID-2 extracellular(ex) fusion proteins, MBP-*C. briggsae* SID-2ex fusion proteins and MBP were expressed in *E. coli* bacteria. 10 µg or 2.5 µg purified recombinant protein was incubated with 2.5 µg of Luciferase dsRNA. Excess dsRNA was washed away. Then dsRNA was eluted and analysed on an agarose gel.
Each lanes show the nucleotide content of a samples (left to right): (L) dsDNA Ladder, 10 µg MBP not incubated with dsRNA, 10 µg MBP::C.el SID-2ex not incubated with dsRNA, 10 µg MBP::C.br SID-2ex not incubated with dsRNA, 10 µg MBP incubated with dsRNA, 2.5 µg MBP incubated with dsRNA, 10 µg MBP::C. el SID-2 incubated with dsRNA, 2.5 µg MBP::C. el SID-2 incubated with dsRNA, 10 µg MBP::C. br SID-2 incubated with dsRNA, 2.5 µg MBP::C. br SID-2 incubated with dsRNA. MPB binding protein fused to the extracellular domain of SID-2 bind more dsRNA then MBP alone.

3.3 Chapter summary

In this chapter, the specificity of a novel SID-2 antibody is shown via immunofluorescence staining. Moreover, this staining reveals the subcellular localisation of endogenous SID-2 in the intestine. SID-2 localises at the apical membrane and in punctuated foci in *C. elegans* and in *C. briggsae.* Co-Immunofluorescence staining using GFP fusion proteins of cell organelle markers uncovers SID-2 localisation at the TGN and at the recycling endosome. Subsequently, the co-localisation of SID-2 with genes of the endocytic pathway suggests that *sid-3* and *viro-2* regulate SID-2 directly or indirectly. However, the altered localisation does not abolish the ability of SID-2 to mediate RNAi by feeding. Furthermore, the subcellular localisation persists after starvation and in the presence of NTPs in environment. In contrast, dsRNA causes a redistribution of SID-2 in the intestine. Lastly, *in vitro* studies show that the *C. elegans* and *C. briggsae* SID-2 extracellular domain binds dsRNA and suggest that SID-2 is a dsRNA binding protein.

4 Chapter – Phenotypical consequences of *sid* mutations

4.1 Introduction

RNAi by feeding is a laboratory technique that has been very successful for gene function studies. In the model organism *C. elegans*, the gene family *sid* is required for the uptake and transport of dsRNA. Although the understanding of *C. elegans* natural biotic environments is growing, dsRNA uptake in such environments has not been identified (Schulenburg and Félix, 2017). Furthermore, the ability of dsRNA uptake and RNAi by feeding has been lost and gained several times in the Caenorhabditis genus (Nuez and Félix, 2012). This convergent evolution might indicate the need of RNAi by feeding in a highly specific ecological niche. However, the RNAi by feeding competent *C. elegans* and the incompetent *C. briggsae* strain have been found in the same environment (Barrière and Félix, 2005). This indicates that *Caenorhabditis* species with and without the ability of RNAi by feeding share a similar natural environment. Of course, the advantage of RNAi by feeding could be more specific to a certain phenotypic niches, therefore a greater understanding of the ecology in the *Caenorhabditis* genus would be desirable. Furthermore, an analysis of additional functions of the genes involved in dsRNA transported, herein after *sid* pathway, is required.

Extensive work was performed to characterise the molecular properties of *sid-1* and *sid-2* (Jose, 2015). Most of the work focus on its relationship to dsRNA on a molecular level. However, additional phenotypical data exist for *sidt2* the mammalian homologue of *sid-1.* First, tissue culture work suggest *sidt2* to be required for RNA degradation at the lysosome, a process the authors term RNautophagy (Aizawa et al., 2016). Second, additional tissue culture work indicates a role of SIDT2 in the transport of dsRNA from endocytic compartments into the cytoplasm. This SIDT2 dependent transport mechanism

is proposed to be required for proper immune response in encephalomyocarditis virus (EMCV) and herpes simplex virus 1 (HSV-1) infection in mice (Nguyen et al., 2017). Contrarily, no evidence for a role in viral immunity was found in *C. elegans* (Ashe et al., 2015). Similar to the proposed models of SIDT2 shuttling, a recent *C. elegans* study proposed a model in which *sid-1* transports endogenous dsRNA for germline integrity (Simon et al., 2018). The basis of this model are genetic studies of *sid-1* mutant animals in a genetic background causing germline sterility after several generations. However, molecular evidence supporting *sid-1* role in endogenous dsRNA transport is lacking and the endogenous role of *sid-1* remains unsolved.
Additionally, the natural function of *sid-2* remains more mysterious. One alternative hypothesis is that SID-2 takes up dsRNA for nutritional reasons, similar to nucleoside transporters in humans which take up nucleosides the product of digestive enzymes degrading polynucleotides (Carver and Allan Walker, 1995). However, these alternative hypothesises have yet to be tested. Here, the results of a phenotypical analysis on the molecular level and the organism level are described.

4.2 Results

4.2.1 Transcriptome analysis of *sid* mutants

The genes of the *sid* pathway transport experimentally supplied dsRNA to every cell to mediate gene regulation. However, the role of the *sid* pathway on gene regulation in the absence of dsRNA has not been addressed. Therefore, I performed transcriptome analysis in young adults of wild-type, two *sid*-1 mutants and two *sid-2* mutants. In wild-type and *sid-1* mutant animals, > 13000 transcripts were detected across the samples. Of these transcripts, 62 transcripts were significantly differentially expressed between wild-type and *sid*-1 mutant (FDR < 0.01) (Figure 20). The small percentage of differentially

expressed transcript suggests that *sid-1* is not a major regulator of transcription.

The same type of transcriptome analysis was performed to compare wild-type and *sid-2* mutant animals. Similarly, >13000 expressed transcripts were detected, of which 41 transcripts were significantly differentially expressed between wild type and *sid*-2 mutant (FDR < 0.01) Figure 21.

Again, the small percentage of differentially expressed genes indicate that *sid-2* is no major regulator of transcription.

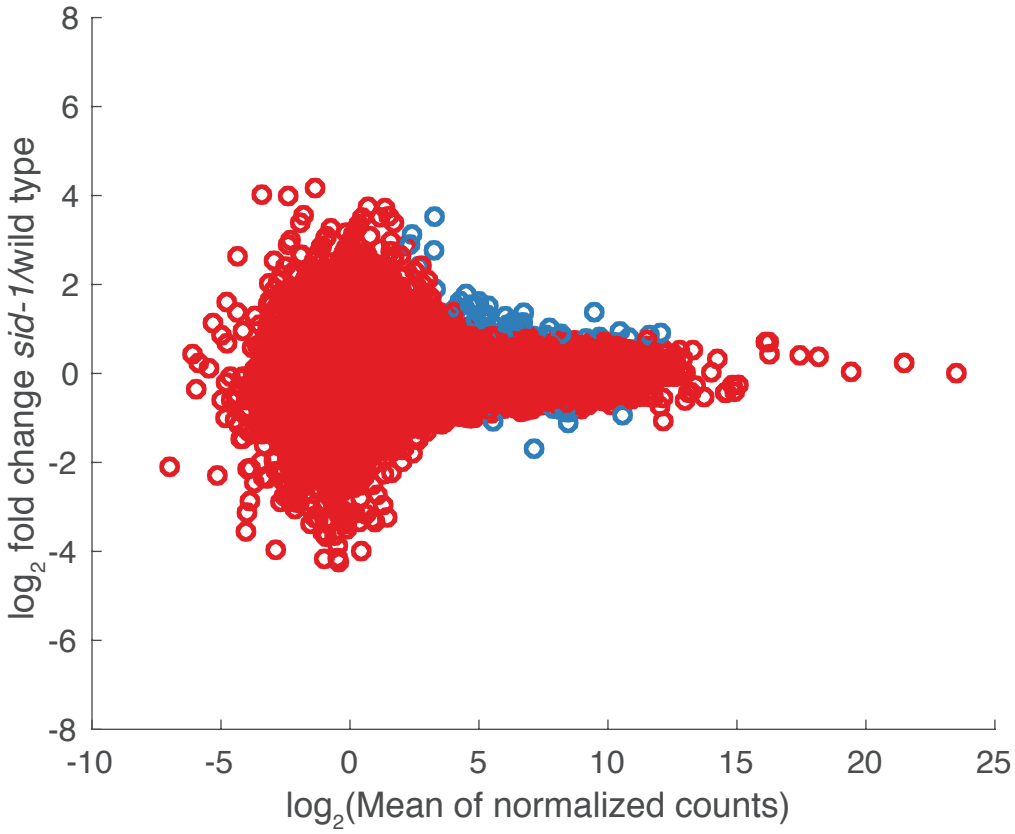


Figure 20 Wild type and *sid-1* mutant animals have a similar transcriptome

MA plot comparing the transcriptome of wild type and *sid-1* mutant young adult animals. Each blue circle represents a statistically significant differentially expressed transcripts (FDR <0.01). A red circle represents non-significant differentially expressed transcripts. The wild type data consist of three biological replicates. The *sid-1* mutant data consist six mutant samples. Three biological replicates for *sid-1(qt129)* and sid-*1(mj444).*

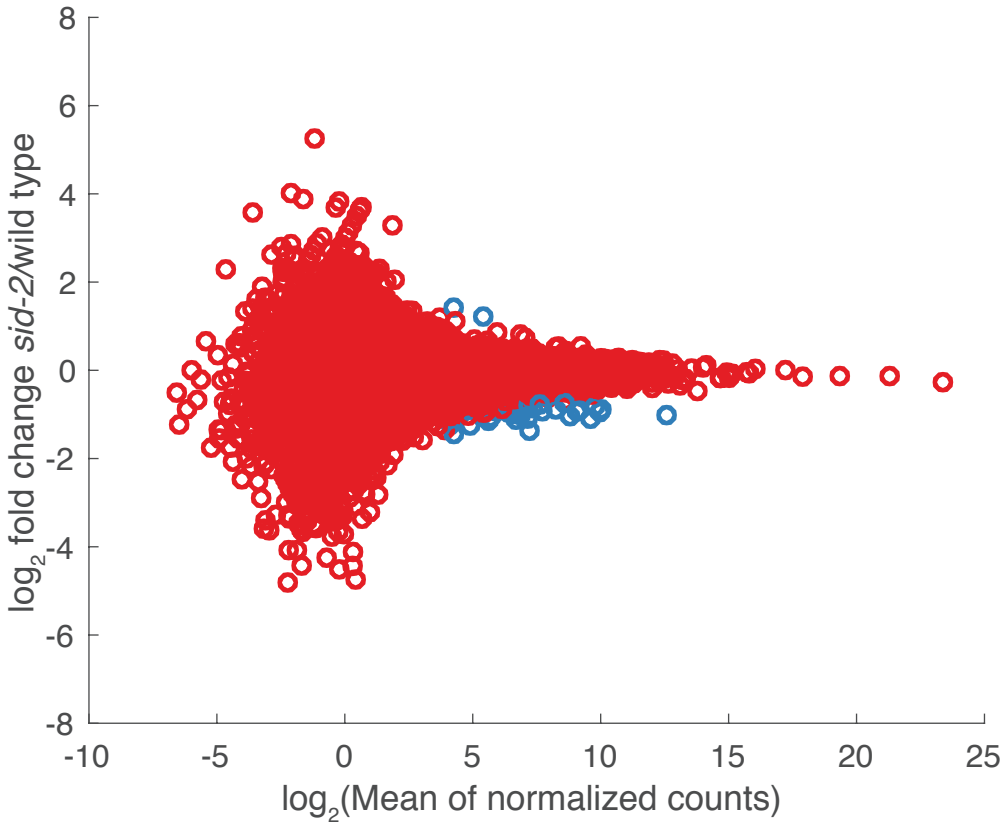


Figure 21 Wild type and *sid-2* mutant animals have a similar transcriptome

MA plot comparing the transcriptome of wild type and *sid-1* mutant young adult animals. Each blue circle represents a statistically significant differentially expressed transcripts (FDR <0.01). A red circle represents non-significant differentially expressed transcripts. The wild type data consist of three biological replicates. The *sid-2* mutant data consist of six mutant samples. Three biological replicates for *sid-2(qt142)* and *sid-2(mj465)*.

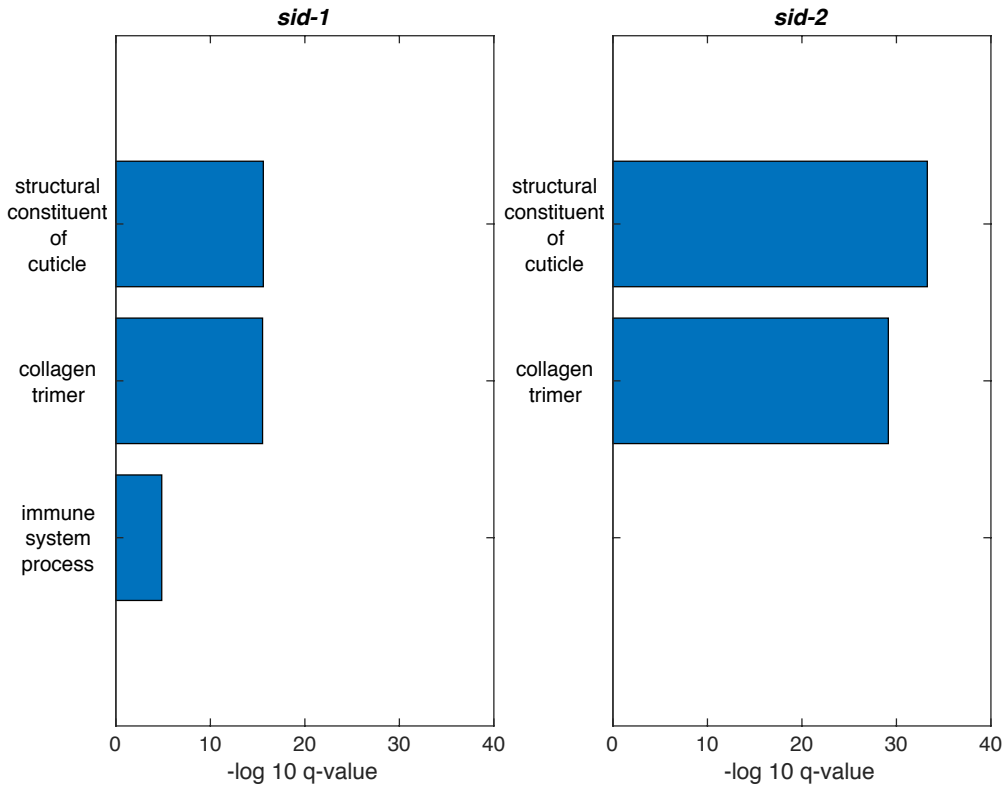


Figure 22 Differentially expressed RNA function in developmental processes

Gene ontology enrichment analysis. Wormbase enrichment analysis was performed on the significantly differentially expressed transcript of *sid-1* and *sid-2* mutant animals compared to wild-type. The enrichment anlaysis indicates the significantly differentially expressed transcript function in molting associated processes

The 62 and 41 significantly differentially expressed transcripts could be direct or indirect targets of *sid-1* and *sid-2,* respectively*,* and might help to understand the biological processes which are affected by *sid-1* and *sid-2.* A gene-ontology (GO-term) analysis allows to test such a hypothesis, by analysing if a certain types of features are overrepresented in subset of genes compared to its occurrence in the whole genome. The analysis was performed with the Wormbase integrated enrichment tool (Angeles-Albores et al., 2016). The analysis identified that upregulated transcripts in *sid-1* and *sid-2* mutant animals compared to wild type are significantly more often associated with genes functioning in collagen trimming or genes being structural constituents

of the cuticle (FDR < 0.05) (Figure 22). In addition, the *sid-1* mutant specific transcripts were found to be enriched for the immune system processes. This comprehensive global analysis suggests that the systemic RNAi pathway effects transcripts of the intestine and the epithelium.

Environmental dsRNA can regulate gene expression, for this function both *sid-1* and *sid-2* are required. Together these genes form the core of the *sid* pathway. Here, I hypothesize that the *sid* pathway could affect gene expression. If a gene is regulated by the *sid* pathway, the regulation should be revealed in *sid-1* and *sid-2* mutant animals. Furthermore, the change in gene expression should show the same tendency in both mutants. Therefore, *sid* regulated transcripts were identified by grouping *sid-1* and *sid-2* mutant animals and comparing them to wild-type animals. The *sex determination and Dosage Compensation defect 3* (*sdc-3)* transcript abundance was significantly reduced in these animals (FDR < 0.01). These analyses reveal *sdc-3* as a potential target of the *sid* pathway.

RNA sequencing allows the global analysis of transcripts. The great number of analysed genes allows comprehensive analysis of gene sets. However, the high number of statistical comparisons performed can lead of to the false positive detection of individual genes.

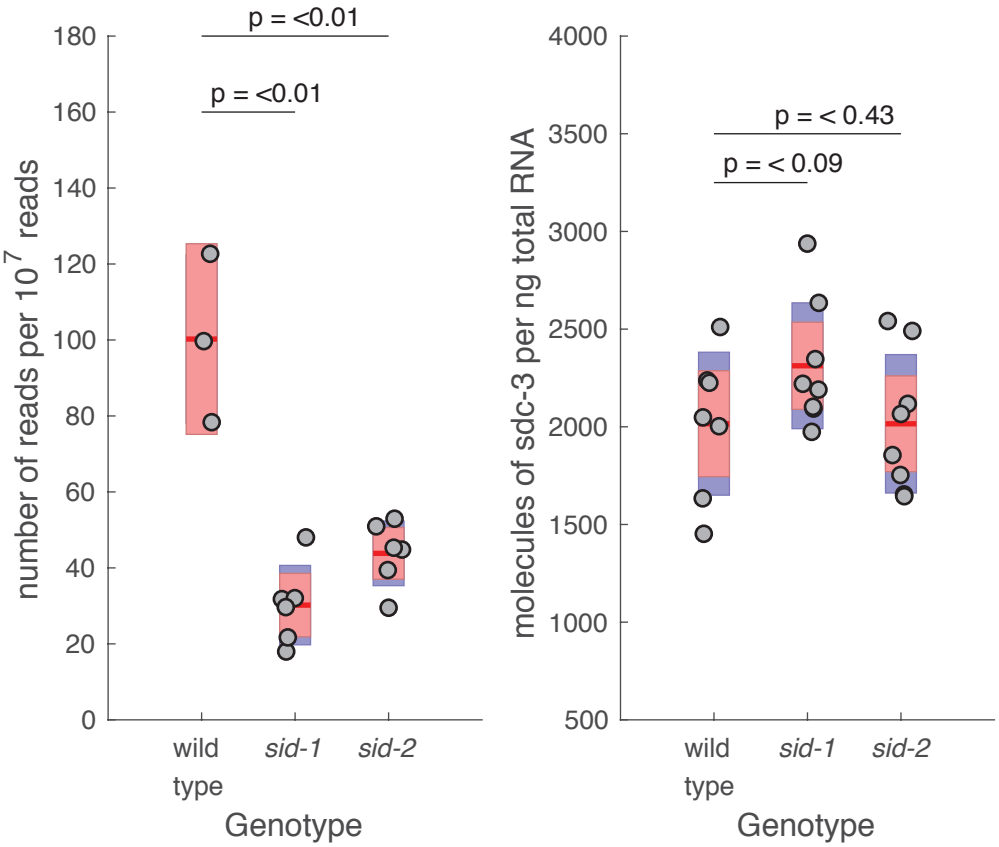


Figure 23 *sdc-3* expression in RNAseq and qPCR

Boxplot for *sdc-3* RNAseq result and *sdc-3* TaqMan qPCR results. Left, RNAseq normalised read counts for *sdc-3* for *wild-type, sid-1(qt129), sid-1(mj444), sid-2(qt142), sid-2(mj465).* Three biological replicates for each genotype are shown. Each dot represents one biological replicate, red bar represents the mean, red boxes the 95% confidence interval for the mean. Blue bar represents the standard deviation. Statistical analysis T-test, right tail.

Right, *sdc-3* molecules per ng of total RNA assesed by TaqMan qPCR in wild type*, sid-1(mj444)* and *sid-2(mj465).* Eight biological replicates for each genotype). Each dot represents one biological replicate, red bar represents the mean, red boxes the 95% confidence interval for the mean. Blue bar represents the standard deviation. Statistical analysis t-test.

A validation of the differential expression of the *sdc-3* gene was performed using TaqMan qPCR (Figure 23). Although, the RNA sequencing indicates high differential expression of the *sdc-3* transcript, the analysis via qPCR does not show a significant difference between wild-type and *sid* mutant animals in *sdc-3* expression levels. This suggests, that *sdc-3* is not under regulation of the *sid pathway.*

4.2.2 small RNAome analysis in *sid* mutants

Environmental dsRNA gene regulation is mediated by small RNAs, however, how changes in the *sid* pathway affects the endogenous small RNA steady state is unknown. Sequencing of the small RNAomes of wild-type and *sid* mutant animals was performed using 5' independent small RNA cloning (Weick et al., 2014). This technique allows the cloning of small RNAs independent of their 5' phosphorylation and allows the identifications of miRNAs, piRNAs and 22G. RNAs. First, differential expression analysis comparing *sid-1 mutant* with wild-type animals was performed. For the expression analysis, small RNAs were mapped to the genome. Small RNAs were then grouped on the gene level and the number of small RNAs mapping a gene was determined by counting all small RNAs mapping to the gene. The number of small RNAs for 167 genes was significantly different between wild-type and *sid-1* mutant animals (FDR < 0.01) Figure 24. The relatively few number of differentially abundant small RNAs indicate that *sid-1* is not essential for small RNA biogenesis or stability.

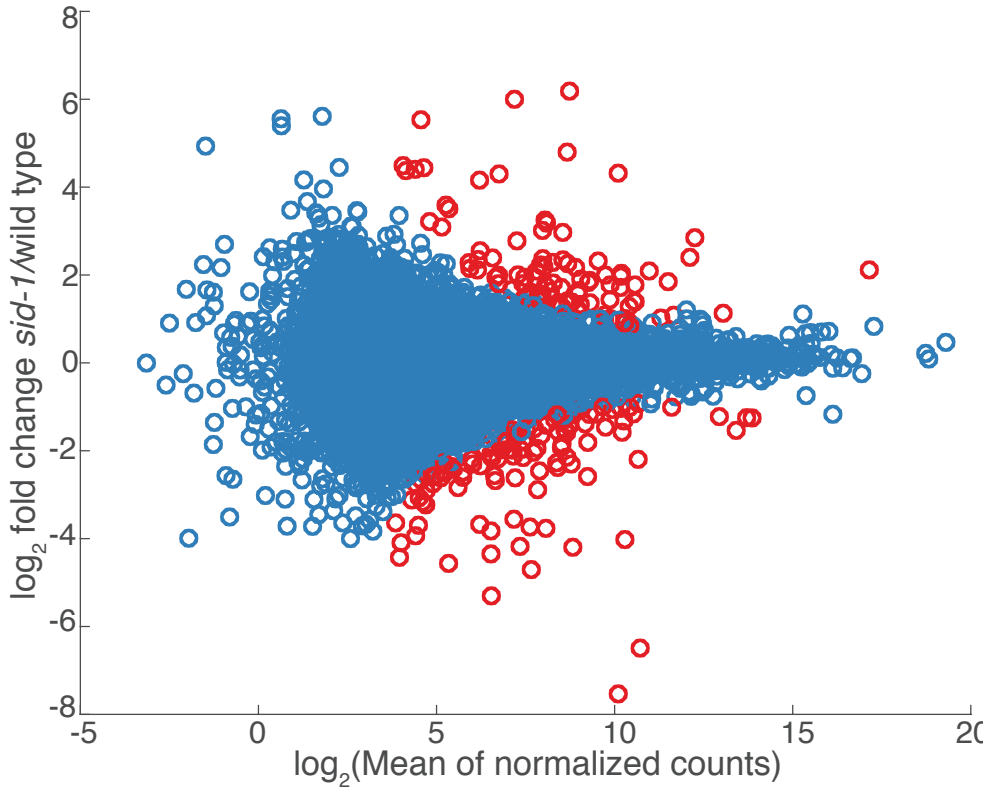


Figure 24 Differentially expressed small RNAs in *sid-1* mutants

MA plot comparing small RNA abundance in wild type and *sid-1* mutant animals. Each red circle represents a statistically significant differentially targeted transcripts (FDR <0.01). A blue circle represents non-significant differentially targeted transcript. The wild type data consist of three biological replicates. The *sid-1* mutant data consist of the pool of six mutant samples. Three biological replicates for *sid-1(qt129)* and *sid-1(mj444)*.

Second, differential expression analysis comparing *sid-2* mutant with wild-type animals was performed. 324 genes with significantly different levels of mapping small RNAs were detected (FDR < 0.01) (Figure 25). Similar to *sid-1* mutant animals, *sid-2* mutant animals are capable of generating small RNAs, suggesting that the *sid* pathway is not essential for small RNA biogenesis or stability.

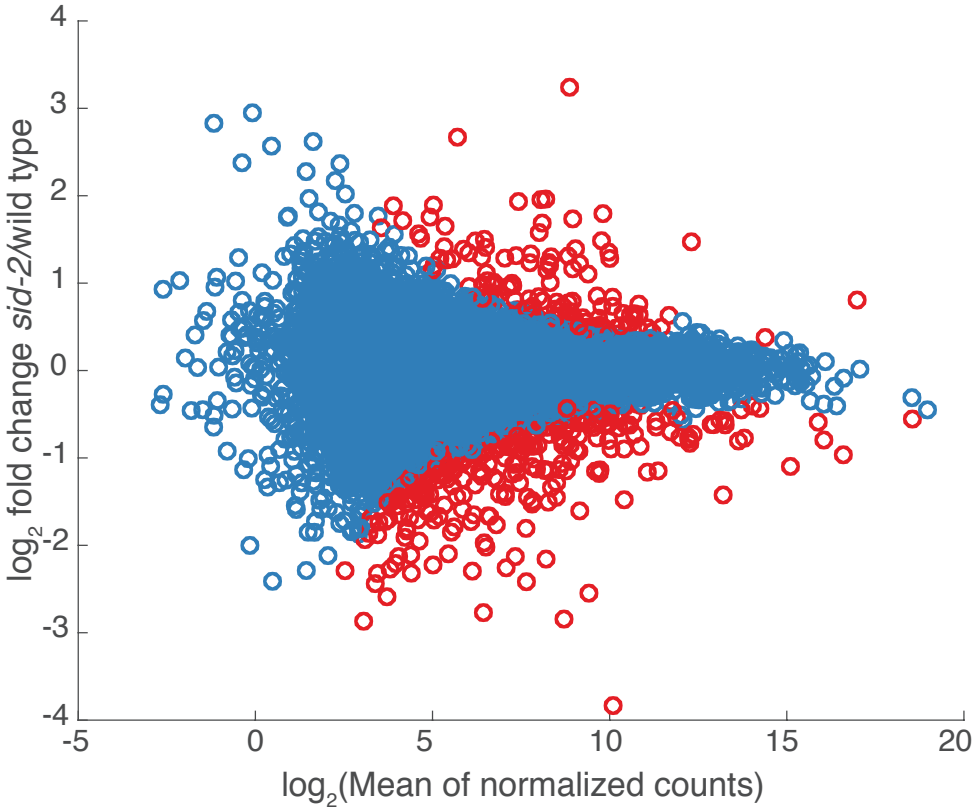


Figure 25 Differentially expressed small RNAs in *sid-2* mutants

MA plot comparing small RNA abundance in wild type and *sid-2* mutant animals. Each red circle represents a statistically significant differentially targeted transcript (FDR <0.01). A blue circle represents non-significant differentially targeted transcripts. The wild type data consist of three biological replicates. The *sid-2* mutant data consist of the pool of six mutant samples. Three biological replicates for *sid-2(qt142)* and *sid-2(mj465).*

Nevertheless, a number of small RNAs showed significantly different adundance and a global analysis could help to understand the role of the *sid* pathway genes in shaping the small RNAome. First, an identification of the chromosomal origin of the differentially expressed small RNAs was performed. In wild-type worms, smallRNAs mapped against 5050 genes and an even distribution among the linkage groups of these genes were found in wild-type samples. In contrast, the distribution of genes with significantly amounts of small RNAs is skewed in *sid-1* mutant animals. More than 40 percent of the

genes are located on Chromosome IV. In *sid-2* mutants, genes with differentially expressed amount of mapping smallRNAs are more frequently found on Chromosome IV (~50 percent) (Figure 26). Genes with changing amounts of mapping small RNAs are located significantly more often on the Chromosome IV in *sid-1* and *sid-2* mutant animals, (hypergeometric test (p < 10^-11) and (p < 10^-15)). Therefore, *sid* pathway affects abundance of Chromosome IV small RNAs.

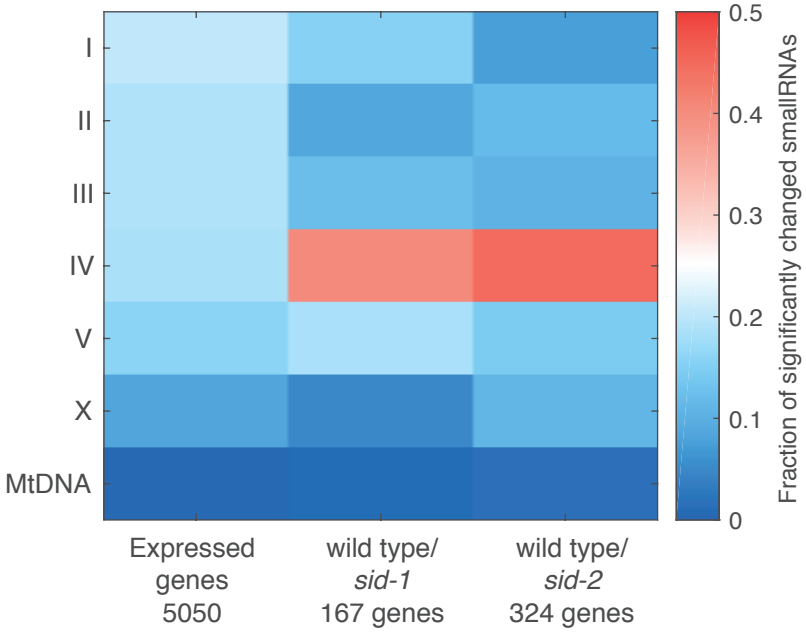


Figure 26 Distribution of chromosomal location of significantly expressed small RNAs.

Heatmap of significantly expressed small RNAs in sid-1 and sid-2 mutant animals. The fraction of expressed small RNA genes on each linkage group is plotted in the left column. The fraction of significantly differentially expressed small RNA genes for sid-1 mutant animals on each linkage group is plotted in the middle column. The fraction of significantly differentially expressed small RNA genes for sid-2 mutant animals on each linkage group is plotted in the right column. Differentially expressed small RNAs in sid-1 and sid-2 mutant animals are frequently located on Chromosome IV.

piRNAs are mostly transcribed from two clusters located on Chromosome IV (Ruby et al., 2006). Since a large fraction of significantly differentially expressed genes in *sid* mutant animals locate on Chromosome IV, piRNA abundance in *sid-1* and *sid-2* mutant animals were compared to wild type piRNA abundance. piRNA abundance was significantly less in all *sid* mutant animals when compared to wild-type animals (Figure 27).

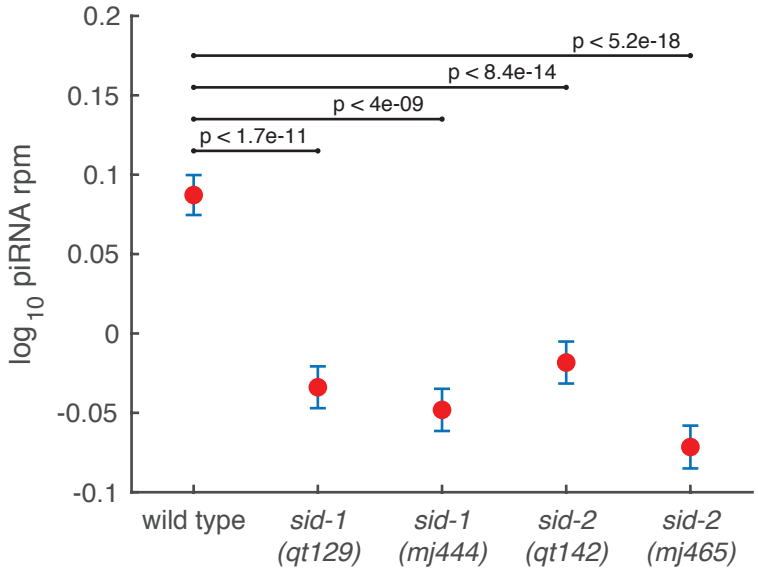


Figure 27 Reduced mean piRNA abundance in *sid* mutants.

Errorbar plot of mean normalised piRNA abundance in wild type, *sid-1(qt129), sid-1(mj444), sid-2(qt142)* and *sid-2(mj465)* mutant animals. Normalised abundance of each piRNA was determined in wild type, *sid-1* and *sid-2* mutant animals is calculated in reads per million from Illumina sequencing data mapped using STAR. The logarithm of the mean normalized piRNA abundance is plotted as the red dots with the blue bars as the standard error of the mean. 3 biological replicates are averaged for each genotype. Two sample right tailed T-tests were used for statistical analysis.

4.2.3 Phenotypic analysis of *sid* mutant animals

In the previous section, the effects of *sid-1* and *sid-2* mutations on the transcriptome and small RNAome have been described. Molecularly, differences in the transcription of genes related to the intestinal and epithelia systems and in piRNA abundance were observed suggesting developmental differences in *sid* mutant animals. This is because the epithelia system undergoes broad rearrangements during development due to cell fusions (Podbilewicz and White, 1994) and the size of the germline increases during development resulting in an increase of piRNA abundance (Batista et al., 2008; Hubbard and Greenstein, 2005). This molecular phenotype suggests that the *sid* pathway affects development in general. Therefore, the development of *sid* mutant animals was assessed with high temporal resolution using an automated video system. The worm length of *sid-1* and *sid-2* mutant worms was measured during the development from egg to adults. This analysis showed that within the first 12 hours of development, *sid-1 and sid-2* mutant animals are longer then wild type. In the period of 12 – 24 hours, the worm length was more similar. However, after 24 hours the worm length of *sid-1* and *sid-2* mutant animals is consistently longer than wild-type animals Figure 28. This differences in worm length suggest a regulatory role in development of the *sid* pathway. The worm length of *sid* mutant animals is greater than that of the wild-type animals at the beginning of the measurement. This measurement indicates that even at hatching, mutants in the *sid* pathway are longer than wild-type worms.

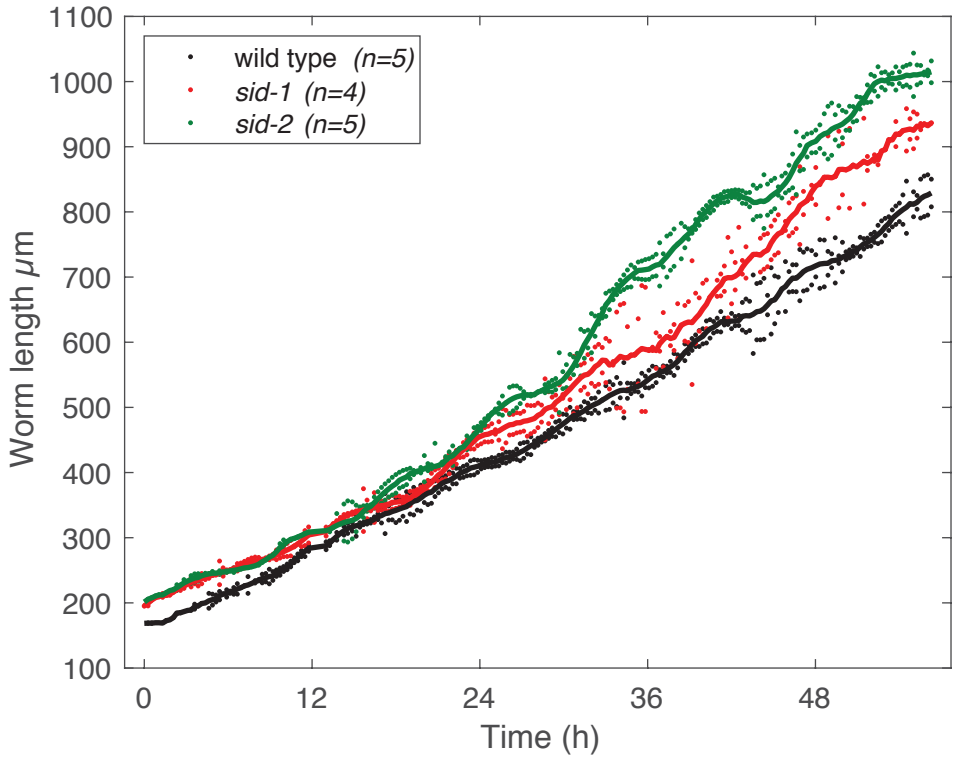


Figure 28 Worm length increases in *sid-1* and *sid-2* mutant worms compated to wild type

Worm length measurement through the development of wild type, *sid-1* and *sid-2* mutant animals. An automated imaging was used to measure worm length of individual worms throughout development. Individual eggs were placed on plates and worm length measurement was started at hatching. Worm length was measured in 15 minutes intervals until adulthood. Mean worm length for strains was plotted as line. Dot indicate the standard error of the mean. N indicates the number of observed worms.

To check if the length differences immediately at hatching were significant, worm length after hatching was measured with a higher resolution and magnification microscopy setup. This measurement identified that *sid-1* and *sid-2* mutant animals are significantly longer than wild-type animals already immediately at hatching (Figure 29), suggesting that the effects of *sid* mutation on worm development are acting via maternal contribution or within the egg.

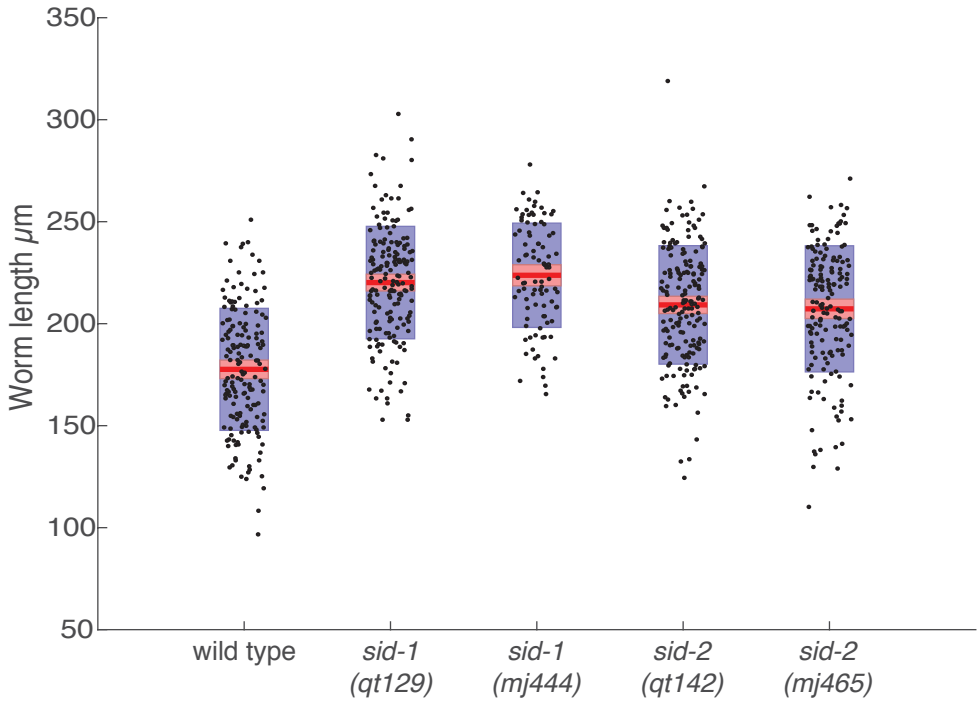


Figure 29 L1 larvae are elongated in *sid* mutants.

Worm length of *sid* mutant animals at hatching. Around 160 eggs were placed on a plate without food and worm length was measured at hatching. Black dots represent worm length of one hatched worm. Red line represents the mean. Red boxes the 95% confidence interval of the mean. Blue boxes the standard deviation.

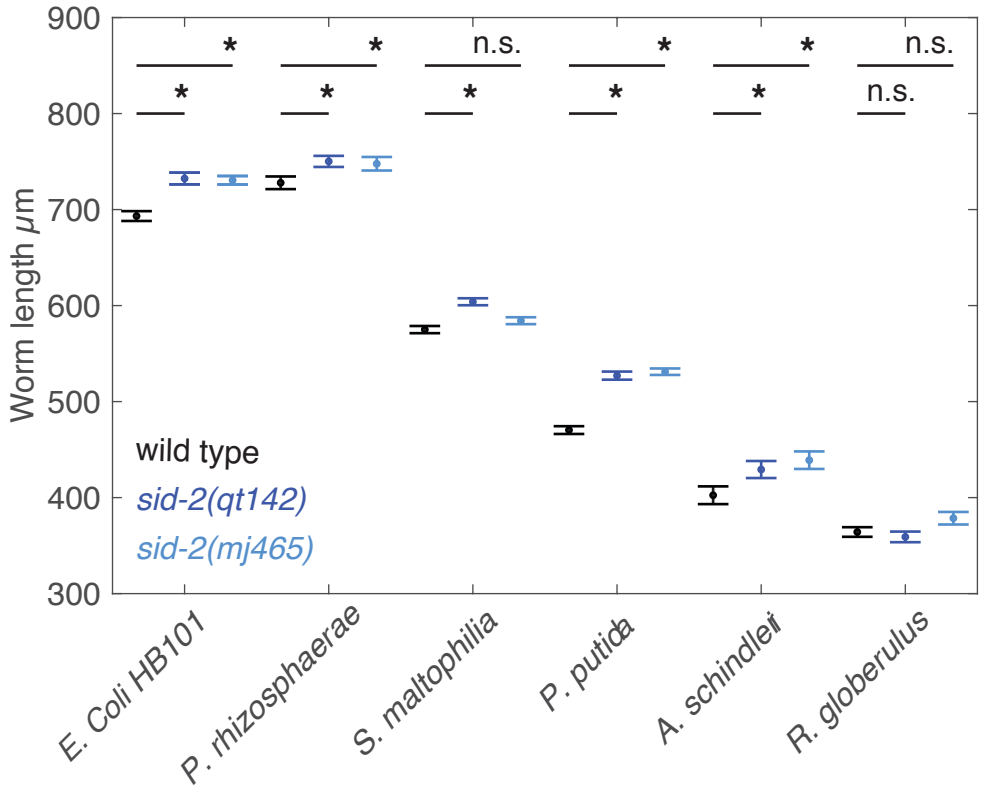


Figure 30 *Sid* worm length is robust in the presence of wild bacteria

Worm length measurement in the presence of wild bacteria. 250 synchronised L1 larvae were placed on a plate seeded with indicated bacteria. After 50 hours, worms were fixed and stained in Coomassie. Fixed worms were imaged on microscope slide and worm length measured. Dots represent mean worm length and bars the standard error of the mean. T-tests were used for statistical analyses. * represent statistically significant differences of the mean using Bonferroni correction ($q < 0.05$).

The difference in development is observed in worms fed *E. coli* HB101, however, these bacteria is not thought to be consumed by *C. elegans* in its natural habitat. Thus, the observed differences might be only robust within laboratory settings. To asses if the developmental differences causes by the *sid* mutation persist in other environments, wild-type and *sid-1* and *sid-2* mutant worms were fed from starved L1 stage with *bacteria* which were isolated in *C. elegans'* habitat (A kind gift from Marie-Anne Félix). These bacteria can be called wild bacteria. Although the bacteria had great influences on the development of the worm, the difference in worm length between *sid* mutant

and wild-type worms was maintained in most wild bacteria (Figure 30), arguing that the developmental differences are robust to environmental factors. This might suggest a more general role of *sid* in worm physiology.

4.2.4 SID-2 functions non-nutritionally

Previous studies described *sid-2* mutants only in nutrient rich environments potentially masking the role of SID-2 in nucleotide nutrition. Therefore, I tested if SID-2 enhances nucleotide uptake from the environment for nutritional purposes in a sensitised *C. elegans* strain lacking sufficient nucleotides in standard laboratory conditions. I made use of a specific allele of the gene *pyrimidine biosynthesis-1 (pyr-1)*. *Pyr-1* is orthologous to the human gene carbamyl phosphate synthetase I (*cps-1)* and codes for a single protein involved in pyrimidine biosynthesis with a predicted carbamoyl phosphate synthetase, aspartate transcarbamoylase, and dihydroorotase activitiy (CAD). The allele *pyr-1(cu8)* refers to single point mutation causing a Glutamine to Histidine amino acid substitution, which leads to embryonic development defects due to insufficient pyrimidine nucleotide production (Franks et al., 2006).

In this sensitised background, I tested the hypothesis that environmental dsRNA taken up by SID-2 contributes to nutrition. Therefore, I grew *C. elegans* wild type*, sid-2 mutant,* and two independent strains of *pyr-1, sid-2 double* mutant worms on *E. coli* bacteria, *E. coli* bacteria supplemented with Uracil, *E. coli* bacteria expressing a long dsRNA or *E. coli* bacteria expressing a short dsRNA. After 72 hours of growth, I assessed the defects in embryonic development by calculating the percentage of hatched eggs after 24 hours Figure 31.

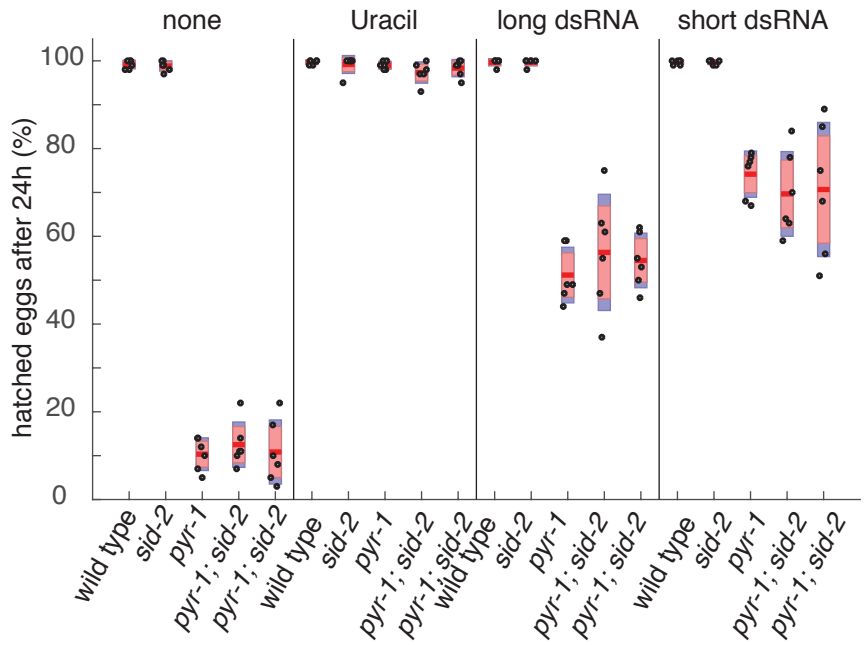


Figure 31 SID-2 does not enhance nucleotide uptake for *pyr-1* dependent nutritional purposes.

The effect of nucleobase supplement and dsRNA expression on hatching rate. Wild type, *sid-2(qt142), pyr-1(cu8),* or two different *pyr-1(cu8), sid-2(qt142)* double mutant strains were grown from L1 to adulthood in the present of various supplements. Egg hatching rate was determined. The basal egg hatching rate was determined without supplement (none), in the present of Uracil, or bacteria expressing short dsRNA or long dsRNA. Black dots represent hatching efficiency of eggs laid by adult worms growing in four different conditions. Three biological experiments with 2 technical replicates for each genotype was performed. Red line represents the mean. Red boxes the 95% confidence interval of the mean. Blue boxes the standard deviation.

When fed with *E. coli* bacteria, 100 % of the wild-type and *sid-2* embryos hatched, indicating *sid-2* does not affect embryonic development on its own. As previously demonstrated, *pyr-1* mutants showed a sever embryonic development defect with a hatching rate of around 10% as did *pyr-1; sid-2 double mutants.* This comparison indicates that, *sid-2* does not affect embryonic development in a *pyr-1* background in the absence of dsRNA.

Providing additional exogenous Uracil to the plates, was sufficient to completely rescue the embryonic development defect in all strains with a *pyr-1* mutation and was not toxic to wild type or *sid-2* mutants. Furthermore, it indicates that *pyr-1* mutant worms are capable of going into embryonic development perfectly in the presence of sufficient pyrimidines.

Next, bacteria expressing long dsRNA were fed to the worms. Again, the presence of dsRNA had no toxic effect on the embryonic development in wild type and *sid-2* mutants. In *pyr-1* mutants, the expression of dsRNA was able to improve the embryonic development to a hatching rate of 60%. Interestingly, the hatching rate in *pyr-1; sid-2 double* mutants was similar to the one in *pyr-1* single mutants. Similarly, no differences between *pyr-1* single mutants and *pyr-1; sid-2* double mutants were found when the worms were fed with bacteria expressing short dsRNA. These results indicate that the presence of dsRNA in the environment can improve the nucleotide nutrition. However, the nucleotides can be taken up independently or the dsRNA can be taken up in a *sid-2* independent manner, which is not accessible to the RNAi machinery. Alternatively, *pyr-1* mutants could be deficient in dsRNA, and therefore removing *sid-2* and providing dsRNA would not test if dsRNA contributes to the nucleotide pool.

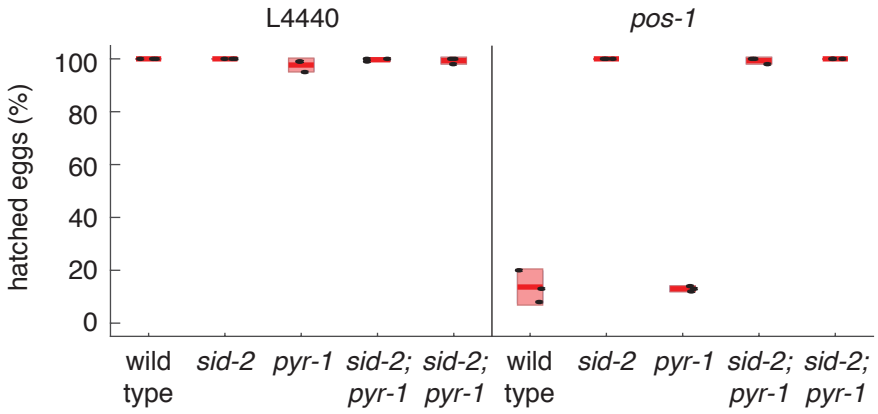


Figure 32 *pyr-1* mutant can take up dsRNA

Pos-1 RNAi by feeding assay in the presence of URACIL. The ability of dsRNA uptake of *pyr-1* mutant animals was accessed in an RNAi by feeding assay. Eggs of wild type, *sid-2, pyr-1,* or *pyr-1, sid-2* double mutant animals were examined after the parents were raised from L1 to adult stage on bacteria expressing negative control dsRNA or *pos-1* dsRNA. Black dots represent hatching efficacy of eggs from adult worms. Three biological replicates for each genotype was performed. Red line represents the mean. Red boxes the 95% confidence interval of the mean.

To test if *pyr-1* single mutants can take up dsRNA and *pyr-1; sid-2* double mutants are deficient in dsRNA uptake, I performed an RNAi by feeding assay against the *posterior segregation-1* (*pos-1)* mRNA a gene essential for embryo development in an environment supplied with exogenous Uracil Figure 32. When worms of any genotype were grown on negative control vector all embryonic hatched after 24 hours, indicating that in this condition embryoic development is unaffected. However, when wild-type or *pyr-1* single mutant worms were grown on bacteria expressing *pos-1* dsRNA only around 10% of the embryo hatched, suggesting that the development was affected and that wild-type and *pyr-1* single mutant worms are capable to take up dsRNA. In contrast, all strains with a *sid-2* mutation had a hatching rate of 100% after 24 hours indicating that they were unable to take up dsRNA. Therefore, I concluded that the *pyr-1* mutant can take up dsRNA and the *pyr-1; sid-2* double mutant is deficient in dsRNA uptake for RNAi. Together with the previous

experiment, I concluded that dsRNA uptake by *C. elegans* SID-2 is not for nutritional reasons.

4.3 Chapter summary

In this chapter, transcriptome profiling of *sid-1* and *sid-2* mutant animals uncovered significant differentially expression of transcripts associated with functions in the intestine and the epithelium. The genes *sdc-3* is expressed at a lower level in both mutants compared to wild type as indicated by RNA sequencing analysis, however qPCR analysis did not confirm this observation. Furthermore, small RNA sequencing indicates that only the abundance of a small fraction of small RNAs is affected by mutations in *sid-1* and *sid-2.* These small RNAs affected in *sid-1* and *sid-2* mutant animals originate from Chromosome IV. A comparison of piRNA abundance between *sid* mutants and wild-type worms shows significantly lower mean piRNA expression. Phenotypical analysis during the development indicate *sid-1* and *sid-2* are important for normal worm length throughout development. Specific measurements at hatching reveal *sid* mutant animals are larger at hatching than wild-type animals. The worm length difference is robust when animals were exposured to naturally occurring wild bacteria. Egg development indicates that *sid-2* does not contribute for dsRNA uptake for nutritional reasons.

5 Discussion

In this work I have described my studies on the dsRNA uptake in *C. elegans.* I described two biochemical approaches which identified potential novel *sid* factors and, for a subset of the identified factors, I provided evidence that they do not play a role in RNAi by feeding. In addition, using molecular biology techniques in whole animals I have characterised biological properties of endogenously expressed SID-2. Moreover, I show the phenotypical effects of dsRNA transport using genetic analysis coupled with high temporal resolution imaging and next-generation-sequencing techniques. In this section, I will interpret my findings in the context of currently available evidence to create novel understandings of dsRNA uptake. Furthermore, I take this opportunity to highlight relevant future approaches to help understanding of dsRNA uptake in *C. elegans*.

5.1 Biochemical identification of novel factors of the dsRNA uptake pathway

For the identification of novel factors required for dsRNA uptake, I use two complementary approaches, both based on the physical interaction with SID-2. One relies on the *in vitro* binding of naturally translated and modified SID-2 at endogenous levels*,* whereas the other captured interactors to artificially overexpressed SID-2 fragments in living yeast cells. Both technical approaches seemed promising in identifying SID-2 interacting partners, especially as some of the identified of proteins have been previously implicated in either vesicle transport (Y37A1B.17) or even in dsRNA uptake (VHA-15) (Saleh et al., 2006). However, both approaches yielded a non-overlapping set of potential

interaction partners of SID-2. All identified candidates can be found in Table 1 and Table 2.

5.1.1 Advantages and disadvantages of biochemical approaches

Established factors for dsRNA uptake were identified via a forward genetic approach which is strategy that uses gene perturbation to associate genes with a particular phenotype. Some setups allow a genome wide identification of involved genes. Specifically, random mutagenesis or genome wide RNAi libraries were very fruitful in *C. elegans* and *D. melanogaster* tissue culture (Saleh et al., 2006; Winston et al., 2002). A targeted approach tested a vesicle transport related gene for its function in systemic RNAi in *C. elegans* (Imae et al., 2016). Interestingly, work by *Saleh et al.* using an RNAi approach complemented the previous work, since the knock-down approach allowed the identification of otherwise lethal genes, which are technically challenging to identify in random mutagenesis or targeted mutant analysis (Kemphues, 2005). Unfortunately, dsRNA uptake studies in insect tissue culture has limited relevance for dsRNA uptake in *C. elegans.*

Despite, the ease and power of genetic studies in *C. elegans*, the identification of genes involved in dsRNA uptake for which mutations are lethal remains a technical challenge. A biochemical approach can prove powerful although the non-physiological conditions might lead to false interpretations (Falsone et al., 2008). Fortunately, the simplicity of the downstream validation allows the assessment of the candidate's role in dsRNA uptake in a reasonable time frame. In addition, a positive example for the success of a biochemical approach is the identification of *sec-22,* which was found using a SID-5 yeast two hybrid screen (Zhao et al., 2017).

5.1.2 Technical Co-IP improvements

Nevertheless, biochemical approaches performed during this thesis did not allow us to identify any novel factors for dsRNA uptake. In the case of the co-immunoprecipitation assay, a potential SID-2 interacting partner could be observed due to chance, technical bias or true interaction. In the future, several steps can be performed to improve this technical approach. First, an increased number of biological replicates would improve the statistical power of the experiment. Second, candidates identified due to technical bias can be eliminated by complementary alternative molecular biology techniques like immunofluorescent imaging. Third, protein-protein interactions can form only in the correct environmental conditions where hydrogen bonds, ionic interactions, Van der Waals forces and/or hydrophobic bonds can be established. Transmembrane protein-protein interactions are particularly sensitive to detergent, pH and salt concentration and optimisation of these conditions would increase the chance of identifying true binding partners (Avila et al., 2015; Pankow et al., 2016). Fourth, chemical crosslinking prior the immunoprecipitation might allow the detection of more transient binding partners (Vasilescu et al., 2004). Overall, the actions would improve the likelihood to detect novel SID-2 interactors required for dsRNA uptake.

Another challenge of the biochemical approach is the possibility of false negative classification of a candidate. A true interacting partner might not be detected in a follow up evaluation in an RNAi by feeding assay potentially caused by redundancy. Redundancy has been shown to occur, for example in several steps in the vesicle transport (Mayers et al., 2013; Mayor et al., 2014; Park et al., 2010). Furthermore, a recent analysis of more than 60 genes of the vesicle transport pathway did not identify any individual gene effecting RNAi by feeding (Imae et al., 2016). The work of Imae et al. suggests that dsRNA uptake mediated by SID-2 is part of a robust mechanism in which another factor can compensate the function in the absence the canonical protein. Together these possibilities make it very time-consuming to exhaustively solve the problem.

5.1.3 Yeast 2 Hybrid - Proteins in heterologous systems

Even though the Y2H approach provided some putative interactions candidates, further analysis did not allow us to identify a novel SID-2 interacting partner required for dsRNA uptake. A specific problem of the yeast 2 hybrid approach is the overexpression of individual factors. Additionally, in this experiment only the intracellular domain of SID-2 was expressed, therefore it did not provide the appropriate sub cellular context that may have been required for an interaction. Furthermore, expressing the *C. elegans* protein in the yeast system might alter the post-translational modifications of the protein which could be essential for certain interactions. The western blot analysis indicated that SID-2 occurs in different isoforms and some of them have substantially larger molecular weights, which might be due to larger protein modifications like glycosylation (Schachter, 2004). This protein modifications might not be recapitulated in yeast, since the number of yeast genes is much lower than in *C. elegans* and only ~ 20 % of the *C. elegans* genes have an ortholog in yeast (Chervitz et al., 1998).

5.1.4 The remaining potential SID-2 binding partner

Not all candidates have been tested for their negative and positive regulatory role in dsRNA uptake. So far, only candidates for which viable mutant strains already exist were tested. In the absence of an available mutant, an RNAi approach could be used to test the gene's function. In such an approach, first the candidate gene function is perturbed using RNAi by feeding, then the worms would be subjected to a second RNAi by feeding bacteria testing the ability of RNA uptake. A disadvantage of such an RNAi based assay is the incomplete depletion of the target might mask the candidate's role in dsRNA uptake. This problem can be avoided by generation of deletion strains using the *Crispr/Cas9* system (Paix et al., 2015), a strategy that could be considered for further experiments. More promising would be a detailed analysis of *vha-15* (Vacuolar H ATPase), since its *D. melanogaster* orthologue is required for

dsRNA uptake. This subunit is very likely essential for viability, since *vha-15* RNAi by feeding is lethal (Saleh et al., 2006). A single pilot *in vitro* binding assay was performed to gain validate the interaction between the two proteins. This experiment indicates that recombinant *C. elegans vha-15* binds to SID-2 *in vitro* (data not shown). Further experiments would be required to establish the binding between these two proteins. Subsequently, cellular studies in *C. elegans* could help to establish their localisation.

5.1.5 Alternative approaches to identify novel *sid's*

Novel approaches could help to identify other factors related to the uptake of dsRNA. The main factors of the *systemic* RNAi pathway, *sid-1* and *sid-2,* are superficially *wild-type.* The *sids* could cause a more severe phenotype if another gene would be non-functional as well. Genes which are only essential if the *sid* pathway is non-functional, can be identified in a synthetical lethality screen. Synthetic lethality is a method to isolate novel mutants, which cause lethality dependent on the absence of the gene of interest (Jorgensen and Mango, 2002). Identifying such genes can reveal pathways which interact with the *sids* and might help to understand the biological role of dsRNA uptake and transport.

5.1.6 Posttranslational modification of SID-2

The characterisation of the different molecular weight isoforms of SID-2 would be another line of very interesting research. One of the isoforms is a substantially larger molecule (Figure 7), and no *sid-2* alternative transcript can code for an isoform of such size, so the addition of post translational modifications to SID-2 is a likely hypothesis. Protein modifications play important roles in membrane protein function (Bonifacino and Traub, 2003; Hislop and von Zastrow, 2011; Low and Saltiel, 1988; Morgan and Burgoyne, 2004). Some of them, such as phosphorylation, ubiquitination and methylation only change the molecular weight by a few Daltons. However, glycosylation

can add highly complex polysaccharide branches to proteins (Spiro, 2002). Glycosylated proteins have critical functions in cellular trafficking, receptor activation and endocytosis (Ohtsubo and Marth, 2006). Most of the glycosylation events occur in the Golgi and especially place at the TGN in human cells (Rabouille et al., 1995). Furthermore, in mammals the toll-like receptors 3 binds dsRNA and is heavily glycosylated in the extracellular domain (Alexopoulou et al., 2001; Sun et al., 2006; Weber et al., 2004). The detection of 8 potential glycosylation sides by the NetOGlyc server in the extracellular domain of SID-2 provide a starting point in understanding how glycosylation affects dsRNA uptake (Steentoft et al., 2013). By and large, the identification of an endogenous modified protein with a clear function could be a good tool to further understand the mechanism of dsRNA uptake beyond genetics.

The mechanism by which SID-2 facilitates dsRNA uptake from the environment remains unclear. The identification of clear SID2 interactions partners through IP-MS and Yeast 2 Hybrid assay did not yield any functional hits, which could suggest that SID-2 is lacking strong and stable co-factors. Further experiments should be designed to detect more transient interactions using other approaches such as crosslinking. Here, I shown that SID-2s extracellular domain binds dsRNA *in vitro* and that *in vivo* dsRNA redistributes SID-2 subcellular localisation. Previously, work showed that SID-2 facilitates dsRNA uptake in an energy dependant mechanism and localises at the apical membrane (McEwan et al., 2012; Winston et al., 2007). Together the experiments strengthen the hypothesis that SID-2 is a dsRNA receptor.

5.2 Cell biology of SID-2

5.2.1 SID-2 role at the TGN

Very little is known about the fate of dsRNA which enters the cell via the intestinal apical membrane. Two models have been proposed, one suggests

the transport of dsRNA through the cell in a transcytosis process, while the second suggests a controlled release of dsRNA into the cytoplasm and subsequent release into the pseudocoelom (McEwan et al., 2012). Here, I identify that SID-2 is localised at the TGN and at the apical luminal membrane. Furthermore, in the presence of dsRNA, SID-2 redistributes to the apical membrane. It remains unclear what function SID-2 has at the TGN. SID-2 could only passes through the TGN after translation to be transported to the cell surface without any additional function except protein maturation, which is one of the main function of the TGN (Guo et al., 2014). In addition, SID-2 could bind cytoplasmic dsRNA at the TGN similar to the toll-like receptor 3 (TLR3) which recognises viral dsRNA products at the endosome (Alexopoulou et al., 2001). Otherwise, SID-2 might be recycled after endocytosis from the endosome pathway via retrograde transport system similar to the Wntless/MIG14 protein (Seaman, 2012). These hypotheses could be tested *in vivo* with labelled SID-2 and dsRNA with the aim to map the route of dsRNA entry into an organism.

5.2.2 Subcellular localisation could be potentially regulated by *sid-3* and *viro-2*

Alternatively, the role of the TGN could be more complex. The TGN could be a reservoir for storage and release of SID-2 in the presence of dsRNA. A similar mechanism has been reported for the mammalian dsRNA receptor TLR3. TLR3 localises at the rough endoplasmic reticulum and in the presence of dsRNA it localises to the endosome, where it binds dsRNA (Johnsen et al., 2006). If SID-2 is stored and released in the TGN, a mechanism to release SID-2 from the TGN is required. This mechanism could be regulated by the conserved tyrosine kinase SID-3 and its substrate VIRO-2. Both have been shown to inhibit virus uptake from the environment (Zhang et al., 2017). Their effect on SID-2 localisation on the TGN was surprising, since SID-3 localises at the apical membrane (Jiang et al., 2017) and in the cytoplasm (Jose et al., 2012). It is possible that they are part of a signalling cascade. In support, the mammalian *viro-2* ortholog *wasp* is required for TLR3 signalling (Sakuma et al.,

2015). It would be interesting, if *sid-3* and *viro-2* orthologous also affect the localisation of the human dsRNA binding protein TLR3.

Regulation of SID-2 localisation by SID-3 could be direct or indirect. For example, phosphorylation of SID-2 by SID-3 could affect SID-2 kinetics at the apical membrane altering SID-2 abundance at the TGN. Since, VIRO-2 is a potential substrate of SID-3, an indirect pathway is more likely. In addition, mutations in *viro-2,* in contrast to *sid-3,* does not affect the worm's ability to RNAi by feeding. This difference in phenotype suggest a signalling cascade involving additional proteins.

Furthermore, although *sid-3* mutant worms are more resistant to RNAi by feeding, they are more sensitive to RNAi by injections (Jose et al., 2012). This counter-intuitive finding supports a more complex function of *sid-3* in dsRNA uptake and RNAi efficiency. It is possible that, *viro-2*, similar to *sid-3,* affects dsRNA uptake, however at the same time enhances susceptibility to internalised dsRNA even more than *sid-3*. Therefore, a direct measurement of dsRNA uptake would help to untangle the role of VIRO-2 and SID-3 in both dsRNA uptake and in RNAi efficiency.

5.2.3 SID-2 is a conserved dsRNA binding protein

C. briggsae is a close relative of *C. elegans* and several studies have reported its deficiency in RNAi by feeding and RNAi by soaking, despite dsRNA transport and dsRNA mediated gene silencing being fully functional (McEwan et al., 2012; Nuez and Félix, 2012; Winston et al., 2007). Here I show, that *C. briggsae* SID-2 and *C. elegans* SID-2 both bind dsRNA *in vitro*. Furthermore, *in vivo C. elegans* as well as *C. briggsae* SID-2 changes its localisation in the intestinal cell in the presence of dsRNA suggesting a functional relationship between dsRNA and SID-2. Despite the homology between the SID-2 proteins of *C. elegans* and *C. briggsae*, these species have clear differences in their RNAi by feeding phenotypes. The similarities in the

behaviour of the *C. elegans* and *C. briggsae* SID-2 in the presence of dsRNA *in vitro* and *in vivo* suggest a functional difference downstream of RNA binding. Alternatively, it could be that *C. briggsae* is competent for RNAi by feeding, however the tested conditions were outside the range in which RNAi by feeding is possible, for example at different ion concentrations or different temperatures.

Overall, this work shows the importance of endomembrane transport for dsRNA uptake and highlights an important regulators of dsRNA receptor localisation.

5.3 Biological role of the *sid* pathway

Here, I establish a genetic link between the systemic RNAi pathway and worm morphology, specifically the worm length. Worm length has been previously shown to be affected not only by several endogenous pathways (Liang et al., 2007; Maduzia et al., 2002; Mörck and Pilon, 2006; Morita et al., 1999; Savage et al., 1996; Suzuki et al., 1999), but also by environmental factors such as the growth medium (Harada et al., 2016). Since the morphological changes are observed in *sid-1* mutants and biochemical data shows that *sid-1* functions in dsRNA uptake, it is possible that worm morphology is controlled by dsRNA which is transported within the animal. Such a function has not been described before neither in *C. elegans,* nor in other animals, in which exogenous dsRNA has only been involved in the host response to viral infection (Akira et al., 2006). In *C. elegans,* where many potential endogenous dsRNAs have been mapped in the genome, largely based on RNA sequencing data, no function for such dsRNAs has been identified (Whipple et al., 2015). Furthermore, I observed similar morphological changes in mutants of the protein SID-2, whose function seems to be the uptake of dsRNA from the environment, this suggests that exogenous dsRNA modulates germline development. The uptake of the exogenous dsRNA could be facilitated by SID-2 and subsequently the

internalised dsRNA could be imported by SID-1. Such a mechanism could also be the biological role of dsRNA uptake also in other animals.

5.3.1 Systemic RNAi for germline immortality?

The Ahmed laboratory proposed a more specific biological role for *sid-1.* They are interested in how the germline maintains its ability to regenerate indefinitely, termed germline immortality. Mutants which become sterile over the course of several generations have what is known as a mortal germline (*mrt)* phenotype (Ahmed and Hodgkin, 2000). The immortality of the germline is maintained by many pathways and traditionally factors controlling the telomere length were identified to cause a Mrt phenotype (Ahmed and Hodgkin, 2000). Subsequently, factors involved in the germline small RNA pathway were identify to be required for germline integrity including the piRNA Argonaute protein PRG-1 (Simon et al., 2014). Interestingly, mutations in the insulin receptor *daf-2* were able to reverse the effect (Simon et al., 2014). A function in germline stability has been shown for *sid-1* using genetic experiments. SID-1 is required to maintain germline immortality caused by the loss of piRNAs in the absence of *daf-2.* In other words, *sid-1*, *daf-2* and *prg-1* triple mutant worms show a mortal germline (Mrt*)* phenotype (Simon et al., 2018). This experiment indicates that SID-1 is required to reinstate germline immortality in a *prg-1* and *daf-2* double mutant suggesting that SID-1 promotes systemic RNAi in such animals. Here, I identify by small RNA sequencing a reduction in the number of piRNAs in *sid* mutant animals. Additional experiments are required to validate the effect of *sid-2* o

n piRNA abundance and if SID-2 affects piRNA abundance it would be interesting to test if *sid-2* can equally supress the *mrt* phenotype in *prg-1;daf-2* double mutants similar to SID-1.

The growth differences are observed in an environment of non-natural bacteria as well as many other wild bacteria suggesting that the trigger is a common molecule, rather than a molecule specific to a given bacteria species. Work in *D. melanogaster* demonstrates that nucleotides are sensed via the gustatory

receptor (Gr) subfamily 28 (Gr28) proteins and help flies to choose their food (Mishra et al., 2018). Such a mechanism to gather information could also be possible for *C. elegans*, however the dsRNA feeding experiments presented here already suggest that dsRNA uptake has no nutritional effect. This strengthen the hypothesis that the dsRNA uptake and transport are naturally designed for interaction with the environment via a RNAi based mechanism.

6 Methods

6.1 *C. elegans* methods

6.1.1 Nematode culture

C. elegans strains were grown essentially as described in (Brenner, 1974). Animals were kept at 20 °C unless otherwise indicated. The *E. coli* feeding strain used was HB101 (*Caenorhabditis* Genetics Center, University of Minnesota, Twin Cities, MN, USA). For general maintenance, animals were transferred from one nematode growth media (NGM) agar plate to another using a platinum wire under a Leica M50 dissecting microscope. Alternatively, chunks of one worm plate were cut out and deposited face down on a new plate. For freezing of strains, animals were left to starve on a 90 mm plate to generate L1 and L2/dauer larvae, then frozen in 2x Freezer buffer and M9 medium (0.3% KH_2PO_4, 0.6% Na_2HPO_4, 0.5% NaCl, 1 mM $MgSO_4$) and stored at -80 °C.

6.1.2 Strains

Table 4 List of *Caenorhabditis* strains

Strain	Genotype
AF16	
CB3284	*mec-12(e1605)III*
CB723	*unc-60(e723)V*
DA2123	*adIs2122[lgg-1::GFP + rol-6(su1006)]*
FX05083	*mec-12(tm5083)III*
FX30967	*(tm10594)IV*
GK932	*dkIs549[Pact-5-YFP-syn-16;Cb.unc-119(+)]*
GR1373	*eri-1(mg366)IV*
JH3176	*gtbp-1(ax2029)IV*
JH3212	*gtbp-1(ax2068)IV*
JH3215	*gtbp-1(ax2073)IV*

KK1105	*lgl-1(tm2616)X*
MJ500	*tpa-1(k501)IV*
MJ563	*tpa-1(k530)IV*
MJ70	*emb-9(hc70)III*
MT5013	*ced-10(n1993)IV*
MT9958	*ced-10(n3246)IV*
N2	
PS3232	*cyl-1(sy433)V*
RB1238	*hmg-11(ok1303)II*
RB1483	*ifa-4(ok1734)X*
RB855	*crp-1(ok685)V*
RB917	*Y37A1B.17(ok788)IV*
RT1315	*unc-119(ed3)III;pwIs503[vha-6p::mans::GFP + Cb-unc-119(+)]*
RT311	*unc-119(ed3)III;pwIs69[vha6p::GFP::rab-11 + unc-119(+)]*
RT327	*unc-119(ed3)III;pwIs72[vha-6p::GFP::rab-5+ unc-119(+)]*
RT476	*unc-119(ed3)III;pwIs170[vha-6p::GFP::rab- 7 + Cb unc-119(+)]*
SX3024	*sid-3(ok973)X*
SX3035	*sid-5(tm4328)X*
SX3046	*sid-1(qt129)V*
SX3072	*sid-2(qt142)III*
SX3112	*sid-2(qt42)III*
SX3116	*sid-1(mj444)V*
SX3237	*sid-2(mj465)III*
SX3329	*viro-2(gk627069)III*
VC1042	*tag-278(gk439)X*
VC1221	*ifa-4(ok1717)X*
VC1564	*cyl-1(ok1943)V*
VC20519	*Y37A1B.17(gk336084)IV*
VC2286	*jph-1(ok2823)I*
VC2471	*F56F10.4(ok3250)X*
VC2592	*T13H5.1(ok3379)II*
VC3390	*rab-30(gk3322)III;pqn-53(gk3534)V;gkDf48 V*
VC40635	*Y37A1B.17(gk735404)IV*
VC552	*alx-1(gk275)III*
VC841	*alx-1(gk338)III*
VC900	*alx-1(gk412)III*

6.1.3 Genetic crosses

For generation of male animals chromosome non-disjunction was induced by heat-shocking between 50 and 60 L4 larvae at 33 °C for 2 hours. Resulting male offspring were used to set up maintenance crosses with hermaphrodite L4 siblings. Crosses were performed with males outnumbering L4 larvae by at least 3-fold. Successful mating was assumed when roughly half the resulting progeny were males. After individualising F1 and subsequent F2 animals, individual F2s were genotyped for relevant alleles using standard PCR or pyrosequencing.

6.1.4 Bleaching and synchronization of *C. elegans*

C. elegans were grown on 90 mm agar plates to gravid adult stage, then washed several times in M9 medium, using an Eppendorf EP5810 centrifuge at 2000 rpm (800 g) to pellet animals. Bleaching to destroy adults and recover embryos was performed by adding 5–6 ml of bleaching solution (final concentration 177 mM NaOH, 177 mM NaOCl solution free chlorine 4–4.99%) to the worm pellet and vortexing vigorously for 5–7 minutes. To remove bleach, embryos were washed three times more. To generate a synchronised population of L1 larvae, embryos were left to hatch in 2 ml of M9 for 48 hours rotating on a spinning wheel. Starved L1 larvae were counted and plated as required.

6.1.5 RNAi experiments

Empty vector, *unc-22* (ZK617.1) and *pos-1* (F52E1.1) bacterial feeding clones were a kind gift from J. Ahringer's laboratory. Bacteria were grown in LBAmpicillin for 6 hours, then seeded onto 50 mm NGM agar plates containing

1 mM IPTG and 25 µg/ml Carbenicillin at a volume of 200 µl bacterial culture per plate and left to dry for 48 hours. Further details on RNA interference are described in (Kamath et al., 2003) and (Rual et al., 2004). For *pos-1* assays, 8 adults were dot bleached per strain onto a RNAi plates and left to lay eggs for 72 hours. Then the adult was removed, and non-hatched embryos were counted 24 hours later.

For *rpb-2* assays, synchronised L1 larva were bleached per strain and transferred onto a RNAi plates and worm size compared to wild type was assayed after 72 hours.

For *dpy-13* assays, 2-3 L4 per strain were transferred onto a RNAi plates, 24 hours later adults were removed from the plates, and another 48 hours later worm size was assayed.

For *act-5* assays, 2-3 L4 per strain were transferred onto a RNAi plates, 48 hours later adults were removed from the plates, and another 48 hours later number of animals at L4 stage and beyond was counted.

6.1.6 dsRNA feeding assay

For long dsRNA and short dsRNA feeding assays a *gfp* RNAi and L4440 empty vector control bacteria clone was used. Synchronised L1 larva were bleached per strain and transferred onto a RNAi plates and left to lay eggs for 72 hours. Then the adult was removed, and non-hatched embryos were counted 24 hours later.

6.2 Molecular biology methods

6.2.1 Polymerase chain reaction

For genotyping reactions PCR was performed using Taq DNA polymerase (NEB). Cycling conditions routinely used were: 94 °C 2 min (94 °C 30 s, 57 °C 30 s, 72 °C, 1min 45 s) x 35 cycles, 72 °C 10 min. For cloning PCR reactions, Q5 Taq Polymerase (NEB) was used according to manufacturer's instructions.

6.2.2 DNA electrophoresis

For visualisation of DNA on agarose gels, gels were prepared at a suitable concentration, usually 1% agarose in 1x TAE buffer (40 mM Tris acetate, 1 mM EDTA) and containing 0.1 mg/ml ethidium bromide. DNA samples were supplemented with 5x loading buffer (8.3 mM Tris-Cl pH 7.5, 50 mM EDTA pH 8, 0.125% Orange G dye, 50% glycerol) and appropriate molecular weight markers, either Hyperladder I, IV or V (Bioline), were loaded alongside DNA samples. After gel run, DNA was visualised on a UVItec Geldoc and camera UV illumination system.

6.2.3 PCR product purification

PCR products were purified using either QIAquick PCR purification kit (Qiagen) or, if excision from an agarose gel was required, using QIAquick gel extraction kit (Qiagen) according to manufacturer's instructions. Alternatively, for attB-PCR products to be used in Gateway cloning, PEG purification was performed according to the Gateway instruction manual (Life Technologies).

6.2.4 DNA sequencing

DNA samples were submitted for Sanger sequencing to Lark Technologies Inc (Beckman Coulter) in the required quantities. Sequencing results were analysed using CLC Main Workbench.

6.2.5 Molecular cloning

Plasmid constructs were generated performing Gibson Cloning (NEB) according to manufacturer's instructions.

6.2.6 Preparation of single worm lysate

Lysate of individual or small pools of animals were generated by transferring adult animals into 2x worm lysis buffer (KCl 100 mM, Tris pH8.2 20 mM, MgCl2 5 mM, IGEPAL 0.9 % (w,v), Tween 20 0.9 % (w,v), Gelatin 0.02 % (w/v)) supplemented with Proteinase K at a final concentration of 0.2 µg/µl. This was usually done by picking animals from plates into 10 µl lysis buffer placed inside the upturned lid of a PCR strip. Tubes were frozen at -80 °C for at least 10 min to crack open the nematodes. Tubes were then placed in a thermocylcer and incubated at 65 °C for 30 min, followed by 20 min at 95 °C. Lysates were used as template for PCR reactions.

6.2.7 *C. elegans* RNA extraction

C. elegans synchronised by L1 starvation arrest were grown on 90 mm plates to the required stage, then washed several times in M9 medium. Animals were snap-frozen in TRISure (Bioline) at a ratio of 1 volume of pellet to 10 volumes of TRISure and cracked open by five freeze-thawing cycles in liquid nitrogen. 0.2 volumes of chloroform were added to each sample and tubes were shaken for 15 sec, then left to settle for 3 min. To separate phases samples were centrifuged at 16,000g for 15 min at 4 °C and the upper aqueous phase was transferred to a new RNase-free tube. Isopropanol equivalent to 0.5 volumes of original TRISure volume was added and precipitation was facilitated by addition of 1 µ l glycogen (20 µ g/µ l). After mixing samples were incubated over night for approximately 16 hours at -20 °C. Precipitated RNA was pelleted by centrifugation at 16,000g for 15 min at 4 °C. The pellet was washed once with ice-cold 80% ethanol, dried for 3-5 min and resuspended in 30-50 µl DEPC water. RNA quantity was measured using a NanoDrop ND-1000 Spectrophotometer (Thermo Scientific).

6.2.8 quantitative real-time PCR

Sdc-3 qPCR was performed using TaqMan probe Ce02485564_g1 and the TaqMan® RNA-to-CT ™ 1-Step Kit. qPCR runs were performed using a STEP ONE PLUS Real Time PCR instrument (Applied Biosystems). Data was analysed using either the standard curve method (as detailed in the instrument manual). To calculate absolute molecules a standard curve using an *in vitro* transcribed sdc-3 standard was used.

6.2.9 *In vitro* RNA synthesis

As template the PCR product of a PCR reaction using the pGL3.Basic plasmid and the primer
M8775:
CTAATACGACTCACTATAGGGCGAGGATTCTAAAACGGATTACCA
and
M8776:
CTAATACGACTCACTATAGGGCGATCAGTGAGCCCATATCCTTG
was used. RNA *in vitro* transcription was performed using the MEGAscript™ T7 Transcription Kit (Ambion) according to manufactory protocol. RNA was purified using DNase treatment and Phenol/Chloroform purification followed by RNA precipitation using Isopropanol.

6.2.10 Small RNA sequencing

If required, 1 µg starting material was pre-treated with either 20U Tobacco Acid Pyrophosphatase (Epicentre) or 20U RNA 5′ Polyphosphatase (Epicentre) in a volume of 20 µl for 45 minutes at 37 °C. Enzymatically treated RNA was then purified using standard phenol chloroform extraction and sodium acetate/ethanol precipitation. The entire pre-treated sample was then used for TruSeq Small RNA library preparation according to manufacturer's instructions

(Illumina). In absence of pre-treatment, 1 µg total RNA was used as starting material for library preparation. PCR amplification of libraries was performed following the provided protocol with a total of 15 cycles. Quantitation and quality control of purified libraries were performed using a QuBit 2.0 Fluorometer (HsDNA kit) and a TapeStation 2200 Instrument

6.2.11 RNA sequencing

Libraries were prepared using the NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina® and the NEBNext® rRNA Depletion Kit (Human/Mouse/Rat) with 1 µg starting material.

6.3 Microbiology methods

6.3.1 Bacteria strains used in this study

For transformations of gibson constructs, α-selected silver efficiency competent cells (Bioline Reagents Ltd) were used. *C. elegans* RNAi by feeding constructs were expressed in HT115(DE3) cells.
Bacteria for growth measurement were a kind gift from Marie Anne Felix. Bacteria were grown for 48 hours in LB at 20 °C.

Table 5 List of bacteria strains

Strain	Species
JUb131-8	Rhodococcus globerulus
JUb271-22	Pseudomonas putida
JUb226-19	Stenotrophomonas maltophilia
JUb249-19	Pseudomonas rhizosphaerae
JUb259	Acinetobacter schindleri

6.3.2 Bacterial transformation

Appropriate competent bacterial cells were thawed on ice. 0.5–1 µl of plasmid DNA or 1 µl of a Gateway BP or 2 µl of a Gateway LR reaction were added to bacteria and reaction was left to thaw on ice for 30 min. Heat shock was performed at 42 °C for 30 sec, then tubes were immediately transferred to ice for 2 min. For recovery, 1 ml of pre-warmed SOC medium (Tryptone 2% (w/v), yeast extract 0.5 % (w/v), NaCl 8.6 mM, KCl 2.5 mM, MgSO4 20 mM, Glucose 20 mM) was added and samples were incubated at 37 °C, shaking for 1 hour. 100–200 µl bacterial cells were spread onto LB plates containing the appropriate antibiotic for selection.

6.3.3 Bacterial culture

For liquid culture, bacterial cells were incubated in autoclaved Luria-Bertani medium (LB medium) (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl, pH 7.5) at 37 °C in a shaking incubator (225 rpm). For selection, appropriate antibiotic, usually ampicillin or kanamycin, was added at a final concentration of 50–100 µg/ml. To grow bacterial clones on plates, LB-agar was prepared, and the antibiotic was added to mixture once it had cooled down to less than 50 °C.

6.4 Imaging

6.4.1 dsRNA soaking experiment

Adult *C. elegans* were washed 2 x in M9 and 1 x in soaking buffer (5X soaking buffer, 1.25 X M9 (Mg2+ free), 15 mM spermidine (SIGMA S2626), 0.25 % gelatine). Subsequently, incubated at RT on a rotator. Either 2.5 μM final luciferase dsRNA (~600 bp), 3 mM final NTPs or 1 x soaking buffer was added and worms incubator at RT for 3 hours. Worms were directly used for immunohistochemistry.

6.4.2 Isolation of *C. elegans* gut and immunhistochemistry

Microscope slides were coated with 50 µl of 0.3% polylysine per well and left for 20 min at RT. Excess liquid was removed and slides were baked for 10 min at 70 °C. For gut dissection, *C. elegans* were transferred into a drop of M9 medium, then transferred to a drop of 7 µl 2.5 mM tetramisol (in M9) in a coated well. Head end of animals were cut using a 21 gauge needle to release gonads. Formaledhyde for a final concentration of 1% was added and fixation was performed for 10 min at RT. A large coverslip (22x40 mm) was added perpendicular to the well applying some pressure, then slides were frozen on dry ice for at least 5 min. Cover slips were cracked off and slides were immediately transferred into ice-cold methanol and incubated at -30 °C for 5 minutes. Then samples were washed three times for 10 min in PBS, 0.2% Tween. For antibody incubation, a humid chamber was built from an empty tip box filled with moistened tissue, slides were place on top of the tip holder grill and 50 µl of antibody solution made up in PBS, 0.2% Tween was added to each well for incubation over night at 4 °C. Slides were washed twice in PBS, 0.2% Tween and 50 µl secondary antibody dilution in PBS, 0.2% Tween supplemented with DAPI at 100 ng/ml were added to each well before incubation at 37 °C in a humid chamber for 1 hour. After secondary antibody staining, slides were washed twice more, excess liquid was removed and samples were mounted with 7 µl Vectashield (Vector Laboratories, Inc.). Finished slides were stored at 4 °C in the dark. Primary antibodies used: chicken GFP ab13970 abcam 1:1000, rabbit SID-2 1:1000 (SDIX). Secondary antibodies: anti-rabbit IgG Alexa Fluor 594 1:1000 A11005 Molecular Probes and anti-chicken IgY Alexa Fluor 488 1:1000 703-545-155 Jackson ImmunoResearch Laboratories, Inc. Secondary antibody solution was supplemented with DAPI at 100 ng/ml.

6.5 Biochemical analysis

6.5.1 Generation of *C. elegans* protein extract

Synchronised populations of *C. elegans* were grown on 90 mm NGM agar plates to gravid adult stage. After several washes in M9 buffer and one final wash in cold 50 mM Tris pH 7.5 150 mm NaCl 0.5 mM EDTA, 0.5% NP40 + Complete Proteinase Inhibitor Cocktail (Roche) the majority of liquid was removed from the pellet and animals were snap-frozen in liquid nitrogen. To generate lysate, samples were homogenised using a Bioruptor Twin (Diagenode) and solution was cleared of debris by centrifugation at 16,000g, 4 °C for 20 minutes. Protein concentration was determined using Bradford protein assay reagent (Sigma), measuring absorbance at a wavelength of 595 nm on a UV/Visible spectrophotometer (Ultrospec 2100 pro, Amersham Biosciences) and calculating concentration of samples relative to a BSA standard curve.

6.5.2 SDS-PAGE

For tris-tricine sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), separating gels were prepared using 330 mM Tris-HCl pH 8.45, 0.1% SDS, 8% polyacrylamide mixture (Protogel30%: 0.8% w/v acrylamide:bisacrylamide, National Diagnostics), 0.1 % ammonium persulphate (APS) (Sigma Aldrich) and 0.1 % TEMED (Sigma). Isopropanol was used to cover separating gel during polymerization and removed by washing with ddH2O. Stacking gels were poured using 330 mM Tris-HCl pH 8.45, 0.1% SDS, 5% polyacrylamide mixture, 0.1% APS and 0.1% TEMED. 50–100 µg of sample protein extract were supplemented with 2x SDS sample buffer (100 mM Tris-HCl pH 6.8, 20% glycerol, 4% SDS, 20% β-Mercaptoethanol, 0.2% (w/v) bromophenol blue) and denatured at 95 °C for 5 min. Gel electrophoresis was performed in a Min Gel Tank Life Technology system in 1x Tris Tricine buffer (T1165-500ML Sigma Aldrich Co) at 150 V

samples were left to migrate until the protein ladder (PageRuler Plus pre-stained protein ladder 10–250 kDA, Fermentas) had reached the correct position.

6.5.3 Western blot

Transfer of SDS gels onto PDFM membrane (Hybond ECL, Amersham) was performed in transfer buffer (25 mM Tris-base, 190 mM glycine, 20% methanol) for 1 hour at 4 °C in a BioRad Mini Trans Blot apparatus at 250 mA. Membranes were blocked in 5% non-fat dry milk in TBS-T buffer for 60 min at RT. Primary antibody incubation was performed in a fresh batch of milk solution with antibody at appropriate dilutions over night at 4 °C and shaking. After three washes in TBST-T for 10 min, secondary antibody dilution in milk/TBS-T was added and membranes were incubated for 1-2 hours at RT. After 3 washes in TBS-T, bands were detected by using Immobilon Western Chemiluminescent HRP Substrate (Millipore) according to manufacturer's instructions, exposure of medical X-ray films (Super Rx, Fuji) to luminescent membrane and development of films on a Compact X4 automatic X-ray film processor (Xograph Imaging Systems Ltd). Primary antibodies used were: purified custom rabbit anti-SID-2 at 1:2000 (, purified monoclonal mouse anti α-tubulin clone DM1A (Sigma Aldrich) 1:10000. Secondary antibodies: ECL anti-mouse IgG HRP from sheep, ECL anti-rabbit IgG HRP from donkey (both GE Healthcare).

6.5.4 Removal of antibodies from western blots

Antibodies were stripped off blots in stripping solution Restore Plus Western Blot Stripping Buffer. After stripping, membranes were washed 3 times in TBS-T and blocked in 5% non-fat dry milk TBS-T.

6.5.5 Generation and purification of custom antibodies

Affinity purified Anti-SID-2 antibody were obtained from SDIX (Windham, United States) by injecting c-DNA coding for the SID-2 c-terminal Domain into rabbits.

SID-2 C-terminal Domain :

GYRTMVNHKLQNSTRTNGLYGYDNNNSSRITVPDAMRMSDIPPPRDPMYA
SPPTPLSQPTPARNTVMTTQELVVPTANSSAAQPSTTSNGQFNDPFATLES
W

Purified antibody stored in PBS pH 7.4 at -80 °C. For experiment the antibody was stored at 4 °C for weeks.

6.5.6 Immunoprecipitation

10 µL of SID-2 antibody was coupled to NHS-activated Magnetic Beads (Pierce) according to manufactory manual. Wormlysate was incubated with SID-2 antibody for 4 hours at +4°C. Finally, beads were wasged 3 times with 50 mM Tris pH 7.5 150 mm NaCl 0.5 mM EDTA, 0.5% NP40 + Complete Proteinase Inhibitor Cocktail (Roche). Proteins were eluted with 0.1 M Glycine pH 2.0.

6.5.7 *In vitro* dsRNA binding assay

Molecular cloning

Full length *sid-2* open reading frames were obtained from *C. elegans* and *C. briggsae* cDNA. The extracellular domains of *sid-2* were subcloned into pMAL-C5X vector using restriction enzymes Not1, Xmn1 (*C. elegans*) and EcoRI (*C. briggsae*). A thrombin cleavage site and strep-tag sequence were added to the C. elegans construct N-terminal and C-terminal to *C. elegans* SID-2, respectively.

Recombinant protein expression and purification

pMAL-C5X *sid-2* plasmids were transformed into BL21 (DE3) cells for expression. A single colony was inoculated into 25ml of LB supplemented with ampicillin to grow overnight. 5ml of the overnight growth was then transferred to 500ml of fresh LB in the presence of antibiotics the next day.Bacterial cells were grown until $O.D._{600}$ reached 0.8. 1mM IPTG was added to the cultures to induce protein expression for 3-4 hours and harvested.Bacterial cells were stored as pellets at -20^{o}C until required. Bacterial growth was performed at 37^{o}C throughout. Bacterial pellets were lysed by sonication on ice in buffer A supplemented with protease inhibitor cocktails (Sigma Aldrich cat. no: 11873580001). Lysed cells were cleared by centrifugation at 4^{o}C. Cleared lysates was then applied to amylose affinity matrix and washed with 10 column volumes with buffer A. The purity of proteins was determined by SDS PAGE to be >90% pure.

RNA pulldown asssay

2.5ug of RNA was incubated with MBP-tagged SID2 (*C. elegans* and *C. briggsae*) and MBP proteins in buffer B for overnight at 4^{o}C with gentle rotation. Unbound RNA was removed by 3 x 1ml buffer B washes. Protein:RNA complex was eluted with buffer C with agitation for 30min at room temperature. Elutions were analysed by 0.5% agarose gel stained with SYBR gold dye (Invitrogen cat. no: S11494). Intensity of RNA bands were quantified by Image J and normalised to amounts of protein eluted as assessed by Bio-rad protein assay (Bio-rad cat. no: 5000006).

Buffers

Buffer A: Phosphate buffered saline (PBS), 5mM β-mercaptoethanol (BME)

Buffer B: PBS, 5mM BME, supplemented with 450mM NaCl

Buffer C: PBS, 5mM BME, 20mM maltose

6.6 Bioinformatics

6.6.1 SID-2 foci analysis

Image analysis was performed in MATLAB. First, the gut tissue was identified in the picture using the the script “thresholding the minimum cross-entropy” by Fatma Gargouri. Next, SID-2 foci were detected using “Gray image thresholding using the Triangle Method” by Bernard Panneton. Since, the triangle method generate many foci of small size < 20 pixels, a size threshold was applied. Foci size was determined using MATLAB’s function “regionprops”. Finally, mean number of foci number above the threshold per gut was calculated.

6.6.2 Worm length measurement on wild bacteria and L1

Bacteria were grown for 48 hours at 20°C and 180 rpm. Bacteria were washed in M9 and seeded at OD10 on peptone free NGM plates. Synchronised ~250 L1 worms were seeded. After 46 hours, worms were washed of the plates in M9 and washed 3 x in M9. Then the worms were fixed using 1ml ice cold (~0°C) Methanol for 15 min, worms were stained in 1 ml of coomassie 50% methanol, 40% H2O, 10% glacial acid for 5 min. Staining was stopped using 1 ml destain 50% methanol, 40% H2O, 10% glacial acid. Worms were rehydrated in 1 ml PBST 0.1%. Worm picture were acquired using a self-build camera system. Length were extracted using MATLAB and size was calculated by skeletonising worms.

6.6.3 Small RNA and long RNA sequencing analysis

Small RNA and long RNA reads were aligned using STAR against the *C. elegans* genome WS235 (Dobin et al., 2013). Read counts per genetic element of the Wormbase genome annotation WS235 were calculated using feature counts (Liao et al., 2014). Reads were normalised using pseudo-

reference with geometric mean row by row (Anders and Huber, 2010). Statistical analysis was performed using Benjamini-Hochberg (BH) adjustment (Hochberg and Benjaminit, 1995)

6.7 Worm length measurement

6.7.1 Single worm growth curves

Populations of *C. elegans* were synchronized by coordinated egg laying. Single eggs were transferred to individual wells of a multi-well NGM-Gelrite plate inoculated with 2 μl of OD 10 *Escherichia coli* HB101 bacteria (~18 Million) and imaged periodically using a camera mounted to a computer controlled XY plotter (EleksMaker, Jiangsu, China) which moved the camera between different wells. Image processing was done in real-time using custom MATLAB scripts, storing both properties of objects identified as *C. elegans*, and sub images of regions around detected objects. Growth curves were aligned to egg-hatching time, which was manually determined for each animal.

6.7.2 Worm length at hatching

40 eggs were placed in each well of a multi-well NGM-Gelrite plate. Eggs were transferred being careful to minimize the amount of bacteria transferred to the multi-well plate, for reasons of image clarity. Images were captured every 10 minutes for a period of 24 hours. To determine animal length at hatching, a custom graphical user interface (GUI) was written which allowed the user to move through the images, and to click on the newly hatched larva. Measurements of worm length would be performed on larva inside of the clicked region and stored.

6.7.3 Length measurements in fixed Worms

Worms were washed of the plates using M9. Bacteria were removed with 3 washes in M9. Worms were fixed in 1 ml fixing solution (50% methanol and 10% glacial acetic acid) for 15 min at -20°C. Worms were stained ins taining solution (0.1% Coomassie Brilliant Blue R-250, 50% methanol and 10% glacial acetic acid) for 15 min. Worms were washed once in destaining solution (40% methanol and 10% glacial acetic acid) and rehydrated in M9.

Slides of fixed worms were imaged under backlit illumination in a high-resolution imaging system. Measurements were made using a custom GUI which detected all sufficiently dark objects and made measurements. These were used to classify each object as a worm or not, using a pre-trained support vector machine. The classified images were then presented to the user for verification. All assignments of objects as worms were thus checked manually by user to ensure accuracy.

7 References

Ahmed, S., and Hodgkin, J. (2000). MRT-2 checkpoint protein is required for germline immortality and telomere replication in C. elegans. Nature *403*, 159–164.

Aizawa, S., Fujiwara, Y., Contu, V.R., Hase, K., Takahashi, M., Kikuchi, H., Kabuta, C., Wada, K., and Kabuta, T. (2016). Lysosomal putative RNA transporter SIDT2 mediates direct uptake of RNA by lysosomes. Autophagy *12*, 565–578.

Akira, S., Uematsu, S., and Takeuchi, O. (2006). Pathogen Recognition and Innate Immunity. Cell *124*, 783–801.

Alcazar, R.M., Lin, R., and Fire, A.Z. (2008). Transmission Dynamics of Heritable Silencing Induced by Double-Stranded RNA in *Caenorhabditis elegans*. Genetics *180*, 1275–1288.

Alexopoulou, L., Holt, A.C., Medzhitov, R., and Flavell, R.A. (2001). Recognition of double-stranded RNA and activation of NF-κB by Toll-like receptor 3. Nature *413*, 732–738.

Ambros, V., Lee, R.C., Lavanway, A., Williams, P.T., and Jewell, D. (2003). MicroRNAs and Other Tiny Endogenous RNAs in C. elegans. Curr. Biol. *13*, 807–818.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Angeles-Albores, D., N. Lee, R.Y., Chan, J., and Sternberg, P.W. (2016). Tissue enrichment analysis for C. elegans genomics. BMC Bioinformatics *17*, 366.

Aoki, K., Moriguchi, H., Yoshioka, T., Okawa, K., and Tabara, H. (2007). In vitro analyses of the production and activity of secondary small interfering RNAs in C. elegans. EMBO J. *26*, 5007–5019.

Aravin, A.A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Gaasterland, T., Meyer, J., and Tuschl, T. (2003). The small RNA profile during Drosophila melanogaster development. Dev. Cell *5*, 337–350.

Ashe, A., Sarkies, P., Le Pen, J., Tanguy, M., and Miska, E.A. (2015). Antiviral RNA Interference against Orsay Virus Is neither Systemic nor Transgenerational in Caenorhabditis elegans. J. Virol. *89*, 12035–12046.

Ashrafi, K., Chang, F.Y., Watts, J.L., Fraser, A.G., Kamath, R.S., Ahringer, J., and Ruvkun, G. (2003). Genome-wide RNAi analysis of Caenorhabditis elegans fat regulatory genes. Nature *421*, 268–272.

Avila, J.R., Lee, J.S., and Torii, K.U. (2015). Co-Immunoprecipitation of Membrane-Bound Receptors. *13*, e0180.

Barrière, A., and Félix, M.-A. (2005). High local genetic diversity and low outcrossing rate in Caenorhabditis elegans natural populations. Curr. Biol. *15*, 1176–1184.

Barrière, A., and Félix, M.-A. (2014). Isolation of C. elegans and related nematodes. WormBook 1–19.

Bartel, D.P. (2018). Metazoan MicroRNAs. Cell *173*, 20–51.

Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S., et al. (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. Mol. Cell *31*, 67–78.

Baum, J.A., Bogaert, T., Clinton, W., Heck, G.R., Feldmann, P., Ilagan, O., Johnson, S., Plaetinck, G., Munyikwa, T., Pleau, M., et al. (2007). Control of coleopteran insect pests through RNA interference. Nat. Biotechnol. *25*, 1322–1326.

Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature *409*, 363–366.

Bolognesi, R., Ramaseshadri, P., Anderson, J., Bachman, P., Clinton, W., Flannagan, R., Ilagan, O., Lawrence, C., Levine, S., Moar, W., et al. (2012). Characterizing the Mechanism of Action of Double-Stranded RNA Activity against Western Corn Rootworm (Diabrotica virgifera virgifera LeConte). PLoS One *7*, e47534.

Bonifacino, J.S., and Traub, L.M. (2003). Signals for Sorting of

Transmembrane Proteins to Endosomes and Lysosomes. Annu. Rev. Biochem. *72*, 395–447.

Buckley, B.A., Burkhart, K.B., Gu, S.G., Spracklin, G., Kershner, A., Fritz, H., Kimble, J., Fire, A., and Kennedy, S. (2012). A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. Nature *489*, 447–451.

Cai, X., Hagedorn, C.H., and Cullen, B.R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA *10*, 1957–1966.

Carver, J.D., and Allan Walker, W. (1995). The role of nucleotides in human nutrition. J. Nutr. Biochem. *6*, 58–72.

Chauhan, V.M., Orsi, G., Brown, A., Pritchard, D.I., and Aylott, J.W. (2013). Mapping the Pharyngeal and Intestinal pH of *Caenorhabditis elegans* and Real-Time Luminal pH Oscillations Using Extended Dynamic Range pH-Sensitive Nanosensors. ACS Nano *7*, 5577–5587.

Chen, C.-C.G., Simard, M.J., Tabara, H., Brownell, D.R., McCollough, J.A., and Mello, C.C. (2005). A Member of the Polymerase β Nucleotidyltransferase Superfamily Is Required for RNA Interference in C. elegans. Curr. Biol. *15*, 378–383.

Chervitz, S.A., Aravind, L., Sherlock, G., Ball, C.A., Koonin, E. V, Dwight, S.S., Harris, M.A., Dolinski, K., Mohr, S., Smith, T., et al. (1998). Comparison of the complete protein sets of worm and yeast: orthology and divergence. Science *282*, 2022–2028.

Collins, J.J., and Anderson, P. (1994). The Tc5 family of transposable elements in Caenorhabditis elegans. Genetics *137*, 771–781.

Conine, C.C., Batista, P.J., Gu, W., Claycomb, J.M., Chaves, D.A., Shirayama, M., and Mello, C.C. (2010). Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in Caenorhabditis elegans. Proc. Natl. Acad. Sci. *107*, 3588–3593.

Devanapally, S., Ravikumar, S., and Jose, A.M. (2015). Double-stranded RNA made in C. elegans neurons can enter the germline and cause

transgenerational gene silencing. Proc. Natl. Acad. Sci. U. S. A. *112*, 2133–2138.

Dhadialla, T.S., and Gill, S.S. (2014). Insect Midgut and Insecticidal Proteins.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dunoyer, P., Schott, G., Himber, C., Meyer, D., Takeda, A., Carrington, J.C., and Voinnet, O. (2010). Small RNA Duplexes Function as Mobile Silencing Signals Between Plant Cells. Science (80-.). *328*, 912–916.

Falsone, S.F., Gesslbauer, B., and Kungl, A.J. (2008). Coimmunoprecipitation and Proteomic Analyses. (Humana Press), pp. 291–308.

Félix, M.-A., Ashe, A., Piffaretti, J., Wu, G., Nuez, I., Bélicard, T., Jiang, Y., Zhao, G., Franz, C.J., Goldstein, L.D., et al. (2011). Natural and Experimental Infection of Caenorhabditis Nematodes by Novel Viruses Related to Nodaviruses. PLoS Biol. *9*, e1000586.

Fishilevich, E., Vélez, A.M., Storer, N.P., Li, H., Bowling, A.J., Rangasamy, M., Worden, S.E., Narva, K.E., and Siegfried, B.D. (2016). RNAi as a management tool for the western corn rootworm, *Diabrotica virgifera virgifera*. Pest Manag. Sci. *72*, 1652–1663.

Franks, D.M., Izumikawa, T., Kitagawa, H., Sugahara, K., and Okkema, P.G. (2006). C. elegans pharyngeal morphogenesis requires both de novo synthesis of pyrimidines and synthesis of heparan sulfate proteoglycans. Dev. Biol. *296*, 409–420.

Fraser, A.G., Kamath, R.S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., and Ahringer, J. (2000). Functional genomic analysis of C. elegans chromosome I by systematic RNA interference. Nature *408*, 325–330.

Gent, J.I., Lamm, A.T., Pavelec, D.M., Maniar, J.M., Parameswaran, P., Tao, L., Kennedy, S., and Fire, A.Z. (2010). Distinct Phases of siRNA Synthesis in an Endogenous RNAi Pathway in C. elegans Soma. Mol. Cell *37*, 679–689.

Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature

442, 199.

Gönczy, P., Echeverri, C., Oegema, K., Coulson, A., Jones, S.J.M., Copley, R.R., Duperon, J., Oegema, J., Brehm, M., Cassin, E., et al. (2000). Functional genomic analysis of cell division in C. elegans using RNAi of genes on chromosome III. Nature *408*, 331–336.

Grant, B.D., and Donaldson, J.G. (2009). Pathways and mechanisms of endocytic recycling. Nat. Rev. Mol. Cell Biol. *10*, 597–608.

Gray, M.E., Sappington, T.W., Miller, N.J., Moeser, J., and Bohn, M.O. (2009). Adaptation and Invasiveness of Western Corn Rootworm: Intensifying Research on a Worsening Pest. Annu. Rev. Entomol. *54*, 303–321.

Grishok, A. (2013). Biology and Mechanisms of Short RNAs in Caenorhabditis elegans. In Advances in Genetics, pp. 1–69.

Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell *106*, 23–34.

Guo, Y., Sirkis, D.W., and Schekman, R. (2014). Protein Sorting at the *trans* -Golgi Network. Annu. Rev. Cell Dev. Biol. *30*, 169–206.

Hamilton, A.J., Baulcombe, D.C., Lendeckel, W., and Tuschl, T. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. Science *286*, 950–952.

Harada, S., Hashizume, T., Nemoto, K., Shao, Z., Higashitani, N., Etheridge, T., Szewczyk, N.J., Fukui, K., Higashibata, A., and Higashitani, A. (2016). Fluid dynamics alter Caenorhabditis elegans body length via TGF-β/DBL-1 neuromuscular signaling. Npj Microgravity *2*, 16006.

Hinas, A., Wright, A.J., and Hunter, C.P. (2012). SID-5 Is an Endosome-Associated Protein Required for Efficient Systemic RNAi in C. elegans. Curr. Biol. *22*, 1938–1943.

Hirst, J., Motley, A., Harasaki, K., Peak Chew, S.Y., and Robinson, M.S. (2003). EpsinR: an ENTH Domain-containing Protein that Interacts with AP-1. Mol. Biol. Cell *14*, 625–641.

Hislop, J.N., and von Zastrow, M. (2011). Role of Ubiquitination in Endocytic Trafficking of G-Protein-Coupled Receptors. Traffic *12*, 137–148.

Hochberg, Y., and Benjaminit, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B J. R. Stat. Soc. B *57*, 289–300.

Hubbard, E.J.A., and Greenstein, D. (2005). Introduction to the germ line. WormBook 1–4.

Hutvágner, G., McLachlan, J., Pasquinelli, A.E., Bálint, É., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme dicer in the maturation of the let-7 small temporal RNA. Science (80-.). *293*, 834–838.

Imae, R., Dejima, K., Kage-Nakadai, E., Arai, H., and Mitani, S. (2016). Endomembrane-associated RSD-3 is important for RNAi induced by extracellular silencing RNA in both somatic and germ cells of Caenorhabditis elegans. Sci. Rep. *6*, 28198.

Ivashuta, S., Zhang, Y., Wiggins, B.E., Ramaseshadri, P., Segers, G.C., Johnson, S., Meyer, S.E., Kerstetter, R.A., McNulty, B.C., Bolognesi, R., et al. (2015). Environmental RNAi in herbivorous insects. RNA *21*, 840–850.

Jenna, S., Caruso, M.-E., Emadali, A., Nguyên, D.T., Dominguez, M., Li, S., Roy, R., Reboul, J., Vidal, M., Tzimas, G.N., et al. (2005). Regulation of membrane trafficking by a novel Cdc42-related protein in Caenorhabditis elegans epithelial cells. Mol. Biol. Cell *16*, 1629–1639.

Jiang, H., Chen, K., Sandoval, L.E., Leung, C., and Wang, D. (2017). An Evolutionarily Conserved Pathway Essential for Orsay Virus Infection of Caenorhabditis elegans. MBio *8*.

Johannes, L., and Popoff, V. (2008). Tracing the Retrograde Route in Protein Trafficking. Cell *135*, 1175–1187.

Johnsen, I.B., Nguyen, T.T., Ringdal, M., Tryggestad, A.M., Bakke, O., Lien, E., Espevik, T., and Anthonsen, M.W. (2006). Toll-like receptor 3 associates with c-Src tyrosine kinase on endosomes to initiate antiviral signaling. EMBO J. *25*, 3335–3346.

Jorgensen, E.M., and Mango, S.E. (2002). The art and design of genetic screens: Caenorhabditis elegans. Nat. Rev. Genet. *3*, 356–369.

Jose, A.M. (2015). Movement of regulatory RNA between animal cells. *53*, 395–416.

Jose, A.M., Smith, J.J., and Hunter, C.P. (2009). Export of RNA silencing from C. elegans tissues does not require the RNA channel SID-1. Proc. Natl. Acad. Sci. *106*, 2283–2288.

Jose, A.M., Garcia, G.A., and Hunter, C.P. (2011). Two classes of silencing RNAs move between Caenorhabditis elegans tissues. Nat. Struct. Mol. Biol. *18*, 1184–1188.

Jose, A.M., Kim, Y.A., Leal-Ekman, S., and Hunter, C.P. (2012). Conserved tyrosine kinase promotes the import of silencing RNA into Caenorhabditis elegans cells. Proc. Natl. Acad. Sci. *109*, 14520–14525.

Kalthoff, C., Groos, S., Kohl, R., Mahrhold, S., and Ungewickell, E.J. (2002). Clint: A Novel Clathrin-binding ENTH-Domain Protein at the Golgi. Mol. Biol. Cell *13*, 4060–4073.

Kamath, R.S., Martinez-Campos, M., Zipperlen, P., Fraser, A.G., and Ahringer, J. (2000). Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in Caenorhabditis elegans. Genome Biol. *2*, research0002.1.

Kawamata, T., and Tomari, Y. (2010). Making RISC. Trends Biochem. Sci. *35*, 368–376.

Kemphues, K. (2005). Essential genes. WormBook 1–7.

Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev. *15*, 2654–2659.

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nat. Rev. Mol. Cell Biol. *10*, 126–139.

Kimble, J., and Sharrock, W.J. (1983). Tissue-specific synthesis of yolk proteins in Caenorhabditis elegans. Dev. Biol. *96*, 189–196.

Knight, S.W., and Bass, B.L. (2001). A role for the RNase III enzyme DCR-1 in

RNA interference and germ line development in Caenorhabditis elegans. Science *293*, 2269–2271.

Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. (2001). An Abundant Class of Tiny RNAs with Probable Regulatory Roles in Caenorhabditis elegans. Science (80-.). *294*, 858–862.

Lee, R.C., and Ambros, V. (2001). An Extensive Class of Small RNAs in Caenorhabditis elegans. Science (80-.). *294*, 862–864.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843–854.

Lee, Y., Jeon, K., Lee, J.-T., Kim, S., and Kim, V.N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. EMBO J. *21*, 4663–4670.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S., et al. (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415–419.

Lee, Y., Kim, M., Han, J., Yeom, K.-H., Lee, S., Baek, S.H., and Kim, V.N. (2004a). MicroRNA genes are transcribed by RNA polymerase II. EMBO J. *23*, 4051–4060.

Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., and Carthew, R.W. (2004b). Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell *117*, 69–81.

Li, Y., Lu, J., Han, Y., Fan, X., and Ding, S.-W. (2013). RNA Interference Functions as an Antiviral Immunity Mechanism in Mammals. Science (80-.). *342*, 231–234.

Liang, J., Yu, L., Yin, J., and Savage-Dunn, C. (2007). Transcriptional repressor and activator activities of SMA-9 contribute differentially to BMP-related signaling outputs. Dev. Biol. *305*, 714–725.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.

Liu, L., Botos, I., Wang, Y., Leonard, J.N., Shiloach, J., Segal, D.M., and Davies, D.R. (2008). Structural basis of toll-like receptor 3 signaling with double-stranded RNA. Science *320*, 379–381.

Liu, Q., Rand, T.A., Kalidas, S., Du, F., Kim, H.-E., Smith, D.P., and Wang, X. (2003). R2D2, a Bridge Between the Initiation and Effector Steps of the Drosophila RNAi Pathway. Science (80-.). *301*, 1921–1925.

Low, M.G., and Saltiel, A.R. (1988). Structural and functional roles of glycosyl-phosphatidylinositol in membranes. Science *239*, 268–275.

Lundquist, E. (2006). Small GTPases. WormBook.

Maduzia, L.L., Gumienny, T.L., Zimmerman, C.M., Wang, H., Shetgiri, P., Krishna, S., Roberts, A.F., and Padgett, R.W. (2002). lon-1 Regulates Caenorhabditis elegans Body Size Downstream of the dbl-1 TGFβ Signaling Pathway. Dev. Biol. *246*, 418–428.

Maeda, I., Kohara, Y., Yamamoto, M., and Sugimoto, A. (2001). Large-scale analysis of gene function in Caenorhabditis elegans by high-throughput RNAi. Curr. Biol. *11*, 171–176.

Maillard, P. V., Ciaudo, C., Marchais, A., Li, Y., Jay, F., Ding, S.W., and Voinnet, O. (2013). Antiviral RNA Interference in Mammalian Cells. Science (80-.). *342*, 235–238.

Marré, J., Traver, E.C., and Jose, A.M. (2016). Extracellular RNA is transported from one generation to the next in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. *113*, 12496–12501.

Mayers, J.R., Wang, L., Pramanik, J., Johnson, A., Sarkeshik, A., Wang, Y., Saengsawang, W., Yates, J.R., and Audhya, A. (2013). Regulation of ubiquitin-dependent cargo sorting by multiple endocytic adaptors at the plasma membrane. Proc. Natl. Acad. Sci. U. S. A. *110*, 11857–11862.

Mayor, S., Parton, R.G., and Donaldson, J.G. (2014). Clathrin-Independent Pathways of Endocytosis. Cold Spring Harb. Perspect. Biol. *6*, a016758–a016758.

McEwan, D.L., Weisman, A.S., and Hunter, C.P. (2012). Uptake of Extracellular Double-Stranded RNA by SID-2. Mol. Cell *47*, 746–754.

Melnyk, C.W., Molnar, A., and Baulcombe, D.C. (2011). Intercellular and systemic movement of RNA silencing signals. EMBO J. *30*, 3553–3563.

Mishra, D., Thorne, N., Miyamoto, C., Jagge, C., and Amrein, H. (2018). The taste of ribonucleosides: Novel macronutrients essential for larval growth are sensed by Drosophila gustatory receptor proteins. PLOS Biol. *16*, e2005570.

Molnar, A., Melnyk, C.W., Bassett, A., Hardcastle, T.J., Dunn, R., and Baulcombe, D.C. (2010). Small Silencing RNAs in Plants Are Mobile and Direct Epigenetic Modification in Recipient Cells. Science (80-.). *328*, 872–875.

Mörck, C., and Pilon, M. (2006). C. elegans feeding defective mutants have shorter body lengths and increased autophagy. BMC Dev. Biol. *6*, 39.

Morgan, A., and Burgoyne, R.D. (2004). Membrane Traffic: Controlling Membrane Fusion by Modifying NSF. Curr. Biol. *14*, R968–R970.

Morita, K., Chow, K.L., and Ueno, N. (1999). Regulation of body length and male tail ray pattern formation of Caenorhabditis elegans by a member of TGF-beta family. Development *126*, 1337–1347.

Naegeli, H., Birch, A.N., Casacuberta, J., De Schrijver, A., Gralak, M.A., Guerche, P., Jones, H., Manachini, B., Messéan, A., Nielsen, E.E., et al. (2018). Assessment of genetically modified maize MON 87411 for food and feed uses, import and processing, under Regulation (EC) No 1829/2003 (application EFSA-GMO-NL-2015-124). EFSA J. *16*.

Nguyen, T.A., Smith, B.R.C., Tate, M.D., Belz, G.T., Barrios, M.H., Elgass, K.D., Weisman, A.S., Baker, P.J., Preston, S.P., Whitehead, L., et al. (2017). SIDT2 Transports Extracellular dsRNA into the Cytoplasm for Innate Immune Recognition. Immunity *47*, 498-509.e6.

Novick, P. (2016). Regulation of membrane traffic by Rab GEF and GAP cascades. Small GTPases *7*, 252–256.

Nuez, I., and Félix, M.-A. (2012). Evolution of Susceptibility to Ingested Double-Stranded RNAs in Caenorhabditis Nematodes. PLoS One *7*, e29811.

Ohtsubo, K., and Marth, J.D. (2006). Glycosylation in Cellular Mechanisms of Health and Disease. Cell *126*, 855–867.

Paix, A., Folkmann, A., Rasoloson, D., and Seydoux, G. (2015). High

Efficiency, Homology-Directed Genome Editing in Caenorhabditis elegans Using CRISPR-Cas9 Ribonucleoprotein Complexes. Genetics *201*, 47–54.

Pankow, S., Bamberger, C., Calzolari, D., Bamberger, A., Yates, J.R., and IIIrd (2016). Deep interactome profiling of membrane proteins by co-interacting protein identification technology. Nat. Protoc. *11*, 2515–2528.

Park, J., Kim, Y., Lee, S., Park, J.J., Park, Z.Y., Sun, W., Kim, H., and Chang, S. (2010). SNX18 shares a redundant role with SNX9 and modulates endocytic trafficking at the plasma membrane. J. Cell Sci. *123*, 1742–1750.

Peters, L., and Meister, G. (2007). Argonaute Proteins: Mediators of RNA Silencing. Mol. Cell *26*, 611–623.

Pham, J.W., Pellino, J.L., Lee, Y.S., Carthew, R.W., and Sontheimer, E.J. (2004). A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in Drosophila. Cell *117*, 83–94.

Piano, F., Schetter, A.J., Mangone, M., Stein, L., and Kemphues, K.J. RNAi analysis of genes expressed in the ovary of Caenorhabditis elegans. Curr. Biol. *10*, 1619–1622.

Pinheiro, D.H., Vélez, A.M., Fishilevich, E., Wang, H., Carneiro, N.P., Valencia-Jiménez, A., Valicente, F.H., Narva, K.E., and Siegfried, B.D. (2018). Clathrin-dependent endocytosis is associated with RNAi response in the western corn rootworm, Diabrotica virgifera virgifera LeConte. PLoS One *13*, e0201849.

Podbilewicz, B., and White, J.G. (1994). Cell fusions in the developing epithelial of C. elegans. Dev. Biol. *161*, 408–424.

Rabouille, C., Hui, N., Hunte, F., Kieckbusch, R., Berger, E.G., Warren, G., and Nilsson, T. (1995). Mapping the distribution of Golgi enzymes involved in the construction of complex oligosaccharides. J. Cell Sci. *108 (Pt 4)*, 1617–1627.

Raman, P., Zaghab, S.M., Traver, E.C., and Jose, A.M. (2017). The double-stranded RNA binding protein RDE-4 can act cell autonomously during feeding RNAi in C. elegans. Nucleic Acids Res. *45*, 8463–8473.

Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. (2006). Large-Scale Sequencing Reveals 21U-RNAs and Additional MicroRNAs and Endogenous siRNAs in C. elegans. Cell *127*, 1193–

1207.

Saint-Pol, A., Yélamos, B., Amessou, M., Mills, I.G., Dugast, M., Tenza, D., Schu, P., Antony, C., McMahon, H.T., Lamaze, C., et al. (2004). Clathrin adaptor epsinR is required for retrograde sorting on early endosomal membranes. Dev. Cell *6*, 525–538.

Sakuma, C., Sato, M., Oshima, T., Takenouchi, T., Chiba, J., and Kitani, H. (2015). Anti-WASP intrabodies inhibit inflammatory responses induced by Toll-like receptors 3, 7, and 9, in macrophages. Biochem. Biophys. Res. Commun. *458*, 28–33.

Saleh, M.-C., van Rij, R.P., Hekele, A., Gillis, A., Foley, E., O'Farrell, P.H., and Andino, R. (2006). The endocytic pathway mediates cell entry of dsRNA to induce RNAi silencing. Nat. Cell Biol. *8*, 793–802.

Saleh, M.-C., Tassetto, M., van Rij, R.P., Goic, B., Gausson, V., Berry, B., Jacquier, C., Antoniewski, C., and Andino, R. (2009). Antiviral immunity in Drosophila requires systemic RNA interference spread. Nature *458*, 346–350.

Savage, C., Das, P., Finelli, A.L., Townsend, S.R., Sun, C.Y., Baird, S.E., and Padgett, R.W. (1996). Caenorhabditis elegans genes sma-2, sma-3, and sma-4 define a conserved family of transforming growth factor beta pathway components. Proc. Natl. Acad. Sci. U. S. A. *93*, 790–794.

Schachter, H. (2004). Protein glycosylation lessons from Caenorhabditis elegans. Curr. Opin. Struct. Biol. *14*, 607–616.

Schulenburg, H., and Félix, M.-A. (2017). The Natural Biotic Environment of *Caenorhabditis elegans*. Genetics *206*, 55–86.

Seaman, M.N.J. (2012). The retromer complex - endosomal protein recycling and beyond. J. Cell Sci. *125*, 4693–4702.

Shih, J.D., and Hunter, C.P. (2011). SID-1 is a dsRNA-selective dsRNA-gated channel. RNA *17*, 1057–1065.

Shih, J.D., Fitzgerald, M.C., Sutherlin, M., and Hunter, C.P. (2009). The SID-1 double-stranded RNA transporter is not selective for dsRNA length. RNA *15*, 384–390.

Sijen, T., Fleenor, J., Simmer, F., Thijssen, K.L., Parrish, S., Timmons, L.,

Plasterk, R.H., and Fire, A. (2001). On the role of RNA amplification in dsRNA-triggered gene silencing. Cell *107*, 465–476.

Simon, M., Sarkies, P., Ikegami, K., Doebley, A.-L., Goldstein, L.D., Mitchell, J., Sakaguchi, A., Miska, E.A., and Ahmed, S. (2014). Reduced insulin/IGF-1 signaling restores germ cell immortality to Caenorhabditis elegans Piwi mutants. Cell Rep. *7*, 762–773.

Simon, M., Spichal, M., Heestand, B., Frenk, S., Hedges, A., Godwin, M., Wellman, A., Sakaguchi, A., and Ahmed, S. (2018). DAF-16/Foxo suppresses the transgenerational sterility of prg-1 piRNA mutants via a systemic small RNA pathway. BioRxiv 326751.

Smardon, A., Spoerke, J.M., Stacey, S.C., Klein, M.E., Mackin, N., and Maine, E.M. (2000). EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. Curr. Biol. *10*, 169–178.

Spiro, R.G. (2002). Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. Glycobiology *12*, 43R-56R.

Steentoft, C., Vakhrushev, S.Y., Joshi, H.J., Kong, Y., Vester-Christensen, M.B., Schjoldager, K.T.-B.G., Lavrsen, K., Dabelsteen, S., Pedersen, N.B., Marcos-Silva, L., et al. (2013). Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J. *32*, 1478–1488.

Sun, J., Duffy, K.E., Ranjith-Kumar, C.T., Xiong, J., Lamb, R.J., Santos, J., Masarapu, H., Cunningham, M., Holzenburg, A., Sarisky, R.T., et al. (2006). Structural and functional analyses of the human Toll-like receptor 3. Role of glycosylation. J. Biol. Chem. *281*, 11144–11151.

Sun, L., Liu, O., Desai, J., Karbassi, F., Sylvain, M.-A., Shi, A., Zhou, Z., Rocheleau, C.E., and Grant, B.D. (2012). CED-10/Rac1 Regulates Endocytic Recycling through the RAB-5 GAP TBC-2. PLoS Genet. *8*, e1002785.

Suzuki, Y., Yandell, M.D., Roy, P.J., Krishna, S., Savage-Dunn, C., Ross, R.M., Padgett, R.W., and Wood, W.B. (1999). A BMP homolog acts as a dose-dependent regulator of body size and male tail patterning in Caenorhabditis

elegans. Development *126*, 241–250.

Tabara, H., Grishok, A., and Mello, C.C. (1998). RNAi in C. elegans: soaking in the genome sequence. Science *282*, 430–431.

Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell *99*, 123–132.

Tabara, H., Yigit, E., Siomi, H., and Mello, C.C. (2002). The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in C. elegans. Cell *109*, 861–871.

Tanguy, M., Véron, L., Stempor, P., Ahringer, J., Sarkies, P., and Miska, E.A. (2017). An Alternative STAT Signaling Pathway Acts in Viral Immunity in *Caenorhabditis elegans*. MBio *8*.

Thivierge, C., Makil, N., Flamand, M., Vasale, J.J., Mello, C.C., Wohlschlegel, J., Conte, D., and Duchaine, T.F. (2012). Tudor domain ERI-5 tethers an RNA-dependent RNA polymerase to DCR-1 to potentiate endo-RNAi. Nat. Struct. Mol. Biol. *19*, 90–97.

Tijsterman, M., May, R.C., Simmer, F., Okihara, K.L., and Plasterk, R.H.A. (2004). Genes Required for Systemic RNA Interference in Caenorhabditis elegans. Curr. Biol. *14*, 111–116.

Timmons, L., and Fire, A. (1998). Specific interference by ingested dsRNA. Nature *395*, 854–854.

Timmons, L., Court, D.L., and Fire, A. (2001). Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in Caenorhabditis elegans. Gene *263*, 103–112.

Timmons, L., Tabara, H., Mello, C.C., and Fire, A.Z. (2003). Inducible Systemic RNA Silencing in *Caenorhabditis elegans*. Mol. Biol. Cell *14*, 2972–2983.

Ungar, D., and Hughson, F.M. (2003). SNARE Protein Structure and Function. Annu. Rev. Cell Dev. Biol. *19*, 493–517.

Vasale, J.J., Gu, W., Thivierge, C., Batista, P.J., Claycomb, J.M., Youngman, E.M., Duchaine, T.F., Mello, C.C., and Conte, D. (2010). Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in

the ERGO-1/Argonaute pathway. Proc. Natl. Acad. Sci. U. S. A. *107*, 3582–3587.

Vasilescu, J., Guo, X., and Kast, J. (2004). Identification of protein-protein interactions usingin vivo cross-linking and mass spectrometry. Proteomics *4*, 3845–3854.

Vastenhouw, N.L., Brunschwig, K., Okihara, K.L., Müller, F., Tijsterman, M., and Plasterk, R.H.A. (2006). Long-term gene silencing by RNAi. Nature *442*, 882–882.

Voinnet, O., Vain, P., Angell, S., and Baulcombe, D.C. (1998). Systemic spread of sequence-specific transgene RNA degradation in plants is initiated by localized introduction of ectopic promoterless DNA. Cell *95*, 177–187.

Wang, E., and Hunter, C.P. (2017). SID-1 Functions in Multiple Roles To Support Parental RNAi in Caenorhabditis elegans. Genetics *207*, 547–557.

Watson, E., MacNeil, L.T., Arda, H.E., Zhu, L.J., and Walhout, A.J.M. (2013). Integration of Metabolic and Gene Regulatory Networks Modulates the C. elegans Dietary Response. Cell *153*, 253–266.

Weber, A.N.R., Morse, M.A., and Gay, N.J. (2004). Four *N* -linked Glycosylation Sites in Human Toll-like Receptor 2 Cooperate to Direct Efficient Biosynthesis and Secretion. J. Biol. Chem. *279*, 34589–34594.

Weick, E.-M., and Miska, E.A. (2014). piRNAs: from biogenesis to function. Development *141*, 3458–3471.

Weick, E.-M., Sarkies, P., Silva, N., Chen, R.A., Moss, S.M.M., Cording, A.C., Ahringer, J., Martinez-Perez, E., and Miska, E.A. (2014). PRDE-1 is a nuclear factor essential for the biogenesis of Ruby motif-dependent piRNAs in C. elegans. Genes Dev. *28*, 783–796.

Whipple, J.M., Youssef, O.A., Aruscavage, P.J., Nix, D.A., Hong, C., Johnson, W.E., and Bass, B.L. (2015). Genome-wide profiling of the C. elegans dsRNAome. RNA *21*, 786–800.

Winston, W.M., Molodowitch, C., and Hunter, C.P. (2002). Systemic RNAi in C. elegans Requires the Putative Transmembrane Protein SID-1. Science (80-.). *295*, 2456–2459.

Winston, W.M., Sutherlin, M., Wright, A.J., Feinberg, E.H., and Hunter, C.P. (2007). Caenorhabditis elegans SID-2 is required for environmental RNA interference. Proc. Natl. Acad. Sci. *104*, 10565–10570.

Yokoyama, N., Lougheed, J., and Miller, W.T. (2005). Phosphorylation of WASP by the Cdc42-associated Kinase ACK1. J. Biol. Chem. *280*, 42219–42226.

Yoo, B.-C. (2004). A Systemic Small RNA Signaling System in Plants. PLANT CELL ONLINE *16*, 1979–2000.

Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004). Single Processing Center Models for Human Dicer and Bacterial RNase III. Cell *118*, 57–68.

Zhang, J., Khan, S.A., Heckel, D.G., and Bock, R. (2017). Next-Generation Insect-Resistant Plants: RNAi-Mediated Crop Protection. Trends Biotechnol. *35*, 871–882.

Zhang, X., Lai, T., Zhang, P., Zhang, X., Yuan, C., Jin, Z., Li, H., Yu, Z., Qin, C., Tör, M., et al. (2019). Mini review: Revisiting mobile RNA silencing in plants. Plant Sci. *278*, 113–117.

Zhao, Y., Holmgren, B.T., and Hinas, A. (2017). The conserved SNARE SEC-22 localizes to late endosomes and negatively regulates RNA interference in *Caenorhabditis elegans*. RNA *23*, 297–307.

Zhou, X., Xu, F., Mao, H., Ji, J., Yin, M., Feng, X., and Guang, S. (2014). Nuclear RNAi contributes to the silencing of off-target genes and repetitive sequences in Caenorhabditis elegans. Genetics *197*, 121–132.

8 Appendix

8.1 Quality control for deep sequencing files

Quality control of small RNA Illumina sequencing files using fastqc. Parallelised using the work load manager slurm.

```
#!/bin/bash
#SBATCH --array=1-15                # Number for a parallel task
#SBATCH -o /mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/
fastqcLongRNA_%A_%a.out             # File to which STDOUT will be
written
#SBATCH -e /mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/
fastqcLongRNA_%A_%a.err             # File to which STDERR will be
written
#SBATCH --mail-type=END             # Type of email notification-
BEGIN,END,FAIL,   ALL
#SBATCH --mail-user=fb416@cam.ac.uk # Email to which notifications
will be sent

#Load specific input file path into arrayfile variable
arrayfile=`ls
/mnt/home1/miska/fb416/projects/braukmannfabian/264FB/seq/trim/
trnatrim/*trnatrim.trim.fq | awk -v line=$SLURM_ARRAY_TASK_ID '{if (NR
==       line) print $0}'`

#Define the output path
arrayout="$(dirname $arrayfile)/fastqc/"
 mkdir -p $arrayout

#Create a quality report using FastQC
/mnt/home1/miska/fb416/projects/res/programs/FastQCv.0.11.4/fastqc
$arrayfile - o $arrayout
```

8.2 Read trimming

8.2.1 Small RNA 3' Adapter trimming

Small RNA read trimming using a bash script and the work load manager slurm

```
#!/bin/bash
#SBATCH -n 1                    # Number of cores
#SBATCH -N 1                    # Ensure that all cores are on one
machine
#SBATCH -t 0-01:00              # Runtime in D-HH:MM
#SBATCH --mem=8G                # Memory pool for all cores (see also -
-mem-per- cpu)
#SBATCH --array=1-23            # Number for a parallel task
#SBATCH -o
/mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/cut_%A_%a.
out                             # File to which STDOUT will be written
#SBATCH -e
/mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/cut_%A_%a.
err                             # File to which STDERR will be written
#SBATCH --mail-type=END         # Type of email notification-
BEGIN,END,FAIL,   ALL
#SBATCH --mail-user=fb416@cam.ac.uk # Email to which notifications
will be sent

#Load specific input file path into arrayfile variable
arrayfile=`ls
/mnt/home1/miska/fb416/projects/braukmannfabian/264FB/seq/*.fq |  awk
-v line=$SLURM_ARRAY_TASK_ID '{if (NR == line) print $0}'`

#Define the output path
arrayout="$(dirname $arrayfile)/trim/$(basename ${arrayfile%%.fq})"
 mkdir -p $(dirname $arrayout)

#Trim reads using cutadapt
 /mnt/home1/miska/fb416/anaconda3/bin/cutadapt -m 18 -M 36 -a
TGGAATTCTCGGGTGCCAAGG $arrayfile -o $arrayout.trim.fq
```

8.2.2 Long RNA adapter trimming

Long RNA read trimming using a bash script and the work load manager slurm

```
#!/bin/bash
#SBATCH -n 1                    # Number of cores
#SBATCH -N 1                    # Ensure that all cores are on one
machine
#SBATCH -t 0-01:00              # Runtime in D-HH:MM
#SBATCH --mem=8G                # Memory pool for all cores (see also -
-mem-per- cpu)
#SBATCH --array=1-15            # Number for a parallel task
#SBATCH -o /mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/
cutlongRNA_%A_%a.out            # File to which STDOUT will be written
#SBATCH -e /mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/
cutlongRNA_%A_%a.err            # File to which STDERR will be written
#SBATCH --mail-type=END         # Type of email notification-
BEGIN,END,FAIL,   ALL
#SBATCH --mail-user=fb416@cam.ac.uk # Email to which notifications
will be sent

#Load input file path into path_input variable
path_input='/mnt/home1/miska/fb416/projects/braukmannfabian/264FB/seq/
longRNA/*.fq'

#Load specific input file path into arrayfile variable
 arrayfile=`ls $path_input | awk -v line=$SLURM_ARRAY_TASK_ID '{if (NR
== line)  print $0}'`

#Define the output path
arrayout="$(dirname $arrayfile)/trim/$(basename ${arrayfile%%.fq})"
 mkdir -p $(dirname $arrayout)

#Trim reads using cutadapt
 /mnt/home1/miska/fb416/anaconda3/bin/cutadapt -m 90 -M 100 -g NNNNN -
a AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC $arrayfile -o $arrayout.trim.fq
```

8.2.3 Quality control for deep sequencing files after adapter trimming

Quality control of Illumina sequencing files using fastqc. Parallelised using the work load manager slurm.

```
#!/bin/bash
#SBATCH --array=1-15                # Number for a parallel task
#SBATCH -o /mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/
fastqcLongRNA_%A_%a.out             # File to which STDOUT will be
written
#SBATCH -e /mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/
fastqcLongRNA_%A_%a.err             # File to which STDERR will be
written
#SBATCH --mail-type=END             # Type of email notification-
BEGIN,END,FAIL,   ALL
#SBATCH --mail-user=fb416@cam.ac.uk # Email to which notifications
will be sent

#Load input file path into arrayfile variable
arrayfile=`ls
/mnt/home1/miska/fb416/projects/braukmannfabian/264FB/seq/trim/
trnatrim/*trnatrim.trim.fq | awk -v line=$SLURM_ARRAY_TASK_ID '{if (NR
==       line) print $0}'`

#Define the output path
arrayout="$(dirname $arrayfile)/fastqc/"
 mkdir -p $arrayout

#Create a quality report using FastQC
/mnt/home1/miska/fb416/projects/res/programs/FastQCv.0.11.4/fastqc
$arrayfile - o $arrayout
```

8.3 Read alignment

8.3.1 Small RNA read alignment

Small RNA read alignment using STAR against the *C. elegans* genome WS235. Parallelised using the work load manager slurm.

```
#!/bin/bash
#SBATCH -n 8                    # Number of cores
#SBATCH --mem=64G               # Memory pool for all cores (see also
--mem-    per-cpu)
#SBATCH --array=1-15            # Number for a parallel task
#SBATCH -o /mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/
star_%A_%a.out      # File to which STDOUT will be written
#SBATCH -e /mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/
star_%A_%a.err      # File to which STDERR will be written
#SBATCH --mail-type=END         # Type of email notification-
BEGIN,END,FAIL,   ALL
#SBATCH --mail-user=fb416@cam.ac.uk # Email to which notifications
will be sent

#Define the genome path
genomepath=/mnt/home1/miska/fb416/projects/res/genomes/caenorhabditis/
cel/STAR

#Load specific input file path into arrayfile variable
arrayfile=`ls
/mnt/home1/miska/fb416/projects/braukmannfabian/264FB/seq/trim/
trnatrim/*.trim.fq | awk -v line=$SLURM_ARRAY_TASK_ID '{if (NR ==
line) print   $0}'`

#Define the output path
arrayout="$(dirname $arrayfile)/STAR/$(basename
${arrayfile%%.fq}.starcel235.)"
 mkdir -p $(dirname $arrayout)

#Map small RNA reads using Star against the C, elegans genome

/mnt/home1/miska/fb416/projects/res/git/STAR/bin/Linux_x86_64_static/S
TAR
--runThreadN 8
--genomeDir $genomepath
--readFilesIn $arrayfile
--alignIntronMax  1
--outFilterMismatchNmax 0
--scoreDelOpen -10000
--scoreInsOpen -10000
--outFilterMultimapNmax 10000
--winAnchorMultimapNmax 50
--outMultimapperOrder Random
--alignEndsType EndToEnd
```

```
--outFileNamePrefix $arrayout
--outTmpDir
/mnt/home1/miska/fb416/projects/res/git/STAR/tmp$SLURM_ARRAY_TASK_ID -
-outSAMtype BAM SortedByCoordinate
```

8.3.2 Long RNA read alignment

Long RNA read alignment using STAR against the *C. elegans* genome WS235. Parallelised using the work load manager slurm.

```
#!/bin/bash
#SBATCH -n 8                    # Number of cores
#SBATCH -N 1                    # Ensure that all cores are on one
machine
#SBATCH -t 0-03:00              # Runtime in D-HH:MM
#SBATCH --mem=64G               # Memory pool for all cores (see also -
-mem-    per-cpu)
#SBATCH --array=1-15            # Number for a parallel task
#SBATCH -o /mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/
starLongRNA_%A_%a.out
# File to which STDOUT will be written
#SBATCH -e /mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/
starLongRNA_%A_%a.err           # File to which STDERR will be written
#SBATCH --mail-type=END         # Type of email notification-
BEGIN,END,FAIL,   ALL
#SBATCH --mail-user=fb416@cam.ac.uk # Email to which notifications
will be sent

#Define the genome path
genomepath=/mnt/home1/miska/fb416/projects/res/genomes/caenorhabditis/
cel/STAR

#Load input file path into path_input variable
path_input='/mnt/home1/miska/fb416/projects/braukmannfabian/264FB/seq/
longRNA/  trim/*.fq'

#Load specific input file path into arrayfile variable
arrayfile=`ls $path_input | awk -v line=$SLURM_ARRAY_TASK_ID '{if (NR
== line)  print $0}'`

#Define output filename
filename_out="$(basename ${arrayfile%%.fq})"

#Define the output path
arrayout="$(dirname $arrayfile)/STAR/$(basename
${arrayfile%%.fq}.starcel235.)"

#Map long RNA reads using Star against the C, elegans genome
/mnt/home1/miska/fb416/projects/res/git/STAR/bin/Linux_x86_64_static/S
TAR
--runThreadN 8
--genomeDir $genomepath
--readFilesIn$arrayfile
--alignSJoverhangMin 8
--alignSJDBoverhangMin 1
--outFilterType BySJout
--alignIntronMin 20
--alignIntronMax 10000
```

```
--outFilterMismatchNmax 1
--scoreDelOpen -10000
--scoreInsOpen -10000
--outFilterMultimapNmax 10000
--winAnchorMultimapNmax 50
--outMultimapperOrder Random
--alignEndsType EndToEnd
--outFileNamePrefix $arrayout
--outTmpDir /mnt/home1/miska/fb416/
projects/res/git/STAR/tmp$SLURM_ARRAY_TASK_ID
--outSAMtype BAM SortedByCoordinate
```

8.4 Counting reads to genomic features

8.4.1 Counting small RNA reads to genomic features

Small RNA read alignment summarised over genomic features WBcel235.78. Parallelised using the work load manager slurm.

```
#!/bin/bash
#SBATCH -n 8                    # Number of cores
#SBATCH -N 1                    # Ensure that all cores are on one
machine
#SBATCH -t 0-02:00              # Runtime in D-HH:MM
#SBATCH --mem=64G               # Memory pool for all cores (see also
--mem-    per-cpu)
#SBATCH -o /mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/
fea_exon_%A_%a.out      # File to which STDOUT will be written
#SBATCH -e /mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/
fea_exon_%A_%a.err      # File to which STDERR will be written
#SBATCH --mail-type=END         # Type of email notification-
BEGIN,END,FAIL,   ALL
#SBATCH --mail-user=fb416@cam.ac.uk # Email to which notifications
will be sent

#Count small RNA reads to exons

/mnt/home1/miska/fb416/projects/res/programs/subread-1.5.2-
source/bin/featureCounts
-T 8
-M
--fraction
-F GTF
-a
/mnt/home1/miska/fb416/projects/res/genomes/annotations/cel/Caenorhabd
itis_elegans.WBcel235.78.gtf
-t exon
-g transcript_id
-o
/mnt/home1/miska/fb416/projects/braukmannfabian/264FB/seq/trim/STAR/fe
a_exon/fea.exon.transcriptid.txt
/mnt/home1/miska/fb416/projects/braukmannfabian/264FB/seq/trim/STAR/*.
bam
```

8.4.2 Counting long RNA reads to genomic features

Long RNA read alignment summarised over genomic features WBcel235.78. Parallelised using the work load manager slurm.

```
#!/bin/bash
#SBATCH --nodes=1                   # Number of cores
#SBATCH --ntasks=1                  # Number of tasks
#SBATCH --cpus-per-task=4           # Number of cpus per task
#SBATCH --mem-per-cpu=8G            # memory per coy
#SBATCH -o /mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/
fea_genelongrna_sense_%A_%a.out     # File to which STDOUT will be
written
#SBATCH -e /mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/
fea_genelongrna_sense_%A_%a.err     # File to which STDERR will be
written
#SBATCH --mail-type=END             # Type of email notification-
BEGIN,END,FAIL,   ALL
#SBATCH --mail-user=fb416@cam.ac.uk # Email to which notifications
will be sent

#Count long RNA reads to gene annotations

/mnt/home1/miska/fb416/projects/res/programs/subread-1.6.2-Linux-
x86_64/bin/featureCounts
-T 8
-p
-M
--fraction
-F GTF
-s 2
-O
-B
-a
/mnt/home1/miska/fb416/projects/res/genomes/annotations/cel/Caenorhabd
itis_elegans.WBcel235.78.gtf
-t gene
-g gene_name
-o
/mnt/home1/miska/fb416/projects/braukmannfabian/264FB/seq/longRNA/trim
_5p_1N_3p_adapter/STAR/fea_gene/fea.gene.genename.sortedbyname.sense.o
verlap.txt
/mnt/home1/miska/fb416/projects/braukmannfabian/264FB/seq/longRNA/trim
_5p_1N_3p_adapter/STAR/*.starcel235.Aligned.sortedByName.*bam
```

8.4.3 Counting small RNA reads to piRNA features

Small RNA read alignment summarised over genomic features WBcel235.78. Parallelised using the work load manager slurm.

```
#!/bin/bash
#SBATCH -n 8                    # Number of cores
#SBATCH -N 1                    # Ensure that all cores are on one
machine
#SBATCH -t 0-02:00              # Runtime in D-HH:MM
#SBATCH --mem=64G               # Memory pool for all cores (see also
--mem-per-cpu)
#SBATCH -o /mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/
fea_piRNA_%A_%a.out      # File to which STDOUT will be written
#SBATCH -e /mnt/home1/miska/fb416/projects/braukmannfabian/264FB/log/
fea_piRNA_%A_%a.err      # File to which STDERR will be written
#SBATCH --mail-type=END         # Type of email notification-
BEGIN,END,FAIL,   ALL
#SBATCH --mail-user=fb416@cam.ac.uk # Email to which notifications
will be sent


#Count small RNA reads to piRNA annotations

/mnt/home1/miska/fb416/projects/res/programs/subread-1.5.2-source/bin/
featureCounts
-T 8
-M
--fraction
-F GTF
-a
/mnt/home1/miska/fb416/projects/res/genomes/annotations/cel/Caenorhabd
itis_elegans.WBcel235.78.21ur.gtf
-t exon
-g gene_name
-o
/mnt/home1/miska/fb416/projects/braukmannfabian/264FB/seq/fea_piRNA/fe
a_piRNA.transcriptid.txt
/mnt/home1/miska/fb416/projects/braukmannfabian/264FB/seq/trim/STAR/*.
bam
```

8.5 Identifying differentially expressed genes in RNA-Seq

Script to test RNA-Seq data and small RNA-Seq data for differentially expressed genes using a negative binomial model.

8.5.1 Contents

- Introduction
- load read count data
- Filter method
- Normalizing Read Counts
- Create table with statistics about each gene
- Inferring Differential Expression with a Negative Bionomial Model
- Create a table with significant genes
- Create a MA plot with significant genes

8.5.2 Introduction

A typical differential expression analysis of RNA-Seq data consists of normalizing the raw counts and performing statistical tests to reject or accept the null hypothesis that two groups of samples show no significant difference in gene expression. This example shows how to inspect the basic statistics of raw count data, how to determine size factors for count normalization and how to infer the most differentially expressed genes using a negative binomial model.

```
function
diffexpression(inputfile,outputpath,filtering,cutoff,sample1id,sample2
id,order)
% order in which samples should be plotted %order = [1:5:15, 2:5:15,
4:5:15, 3:5:15, 5:5:15];
% inputfile, full path like '/Users/braukmann/Dropbox
(ericmiskalab)/Experiments/FB264/longRNAanalysis/cutadapt_5p_1N_3p_Ada
pter/fea_exon/fea.exon.transcriptid.xlsx'
% load read counts
% filter determines which genes are excluded from analysis, rRNA
srpRNA
% mtrRNA
% outputpath is the path where the results are writen to
'/Users/braukmann/Dropbox
(ericmiskalab)/Experiments/FB264/longRNAanalysis/cutadapt_5p_1N_3p_Ada
pter/tmp/fea_exon/'
% cutoff is the percentile like 85 percentile most abundant reads from
% which reads above are analysed
```

8.5.3 Load read count data

```
exonCountTable = readtable(inputfile);



%removes excessive Chr and Strand information
tmp = cellfun(@(x)
strsplit(x,';'),exonCountTable.Chr,'UniformOutput',0);
exonCountTable.Chr = cellfun(@(x) x(1),tmp);
tmp = cellfun(@(x)
strsplit(x,';'),exonCountTable.Strand,'UniformOutput',0);
exonCountTable.Strand = cellfun(@(x) x(1),tmp);
exonCountTable = exonCountTable(:,[1:6, order+6]);
```

8.5.4 Filter method

Filter out rRNA reads, signal recongnition particel, mtRNA if filtering == 1

```
if filtering == 1
    filter = {'R04F11.13','R04F11.12',...
        'F31C3.9','F31C3.8','F31C3.7','F31C3.11',...

'MTCE.7','MTCE.12','MTCE.23','MTCE.26','MTCE.31','MTCE.33','MTCE.35',.
..
        'R144.15','ZC155.8','B0285.12',...
        'ZK218.12', 'ZK218.16', 'ZK218.17', 'ZK218.18', 'ZK218.19',
'ZK218.20',...
        'Y102A5D.5', 'Y102A5D.6', 'Y102A5D.7', 'Y102A5D.8',
'Y102A5D.9', 'Y102A5D.10', 'Y102A5D.11', 'Y102A5D.12',...
        'T27C5.18', 'T09B4.23'};
end

if filtering == 0
    filter = [];
end

% generate index for lines that should be filtered out
index = zeros(length(exonCountTable.Geneid),length(filter));

for i = 1:length(filter)
    index(:,i) =
strncmp(filter{i},exonCountTable.Geneid,length(filter{i}));
end
```

```
index2 = find(sum(index,2));
% remove lines that should be filtered out
exonCountTable(index2,:) = [];

% identify and arrange data in order
% loads sample names
samples = exonCountTable(:,7:end).Properties.VariableNames;
```

8.5.5 Define samples and comparison

```
strsample1 = {sample1id};

strsample2 = {sample2id};



%generates a logical array with ones for sample group1
tmp = false(length(strsample1),length(samples));
for i = 1:length(strsample1)
    tmp(i,:) = strncmp(samples, strsample1(i) ,length(strsample1{i}));
end
Sample1 = logical(sum(tmp,1));

%generates a logical array with ones for sample group2
tmp = false(length(strsample2),length(samples));
for i = 1:length(strsample2)
    tmp(i,:) = strncmp(samples, strsample2(i) ,length(strsample2{i}));
end
Sample2 = logical(sum(tmp,1));
clear tmp

comparison_samples = logical(Sample1+Sample2);
```

8.5.6 Normalizing Read Counts

```
%estimate pseudo-reference with geometric mean row by row
counts = exonCountTable{:,7:end};

index_tmp = counts(:,comparison_samples) >=
prctile(counts(:,comparison_samples),cutoff,1);
index = sum(index_tmp,2) > 0;
```

```
counts_abovecutoff = counts(index,:);
pseudoRefSample = geomean(counts_abovecutoff,2);
nz = pseudoRefSample > 0;
sum(index)
sum(nz)
ratios =
bsxfun(@rdivide,counts_abovecutoff(nz,:),pseudoRefSample(nz));
sizeFactors = median(ratios,1);
normCounts = bsxfun(@rdivide,counts_abovecutoff(nz,:),sizeFactors);

% consider the mean
meanSample1 = mean(normCounts(:,Sample1),2);
meanSample2 = mean(normCounts(:,Sample2),2);

% compute the mean, fold change and the log2FC
meanBase = (meanSample2 + meanSample1) / 2;
foldChange = meanSample2 ./ meanSample1;
log2FC = log2(foldChange);
```

8.5.7 Create table with statistics about each gene

```
exonTabletmp = exonCountTable(index,:);

exonTabletmp = exonTabletmp(nz,:);

exonTable = table(meanBase,meanSample2,meanSample1,foldChange,log2FC,
exonTabletmp.Chr,...
    exonTabletmp.Start, exonTabletmp.End,
exonTabletmp.Strand,'VariableNames', {'meanBase',
strjoin(strsample2,'_'), strjoin(strsample1,'_'), 'foldChange',
'log2FC', 'Chr', 'Start', 'End', 'Strand'});
exonTable.Properties.RowNames = exonTabletmp.Geneid;

%add
for i = 1:length(comparison_samples)
    if comparison_samples(1,i) == 1
        column = length(exonTable.Properties.VariableNames);
        tmpname = samples(i);

eval(sprintf('exonTable.%s=counts_abovecutoff(nz,%d);',[tmpname{1}
'_counts'],i));
    end
end

for i = 1:length(comparison_samples)
    if comparison_samples(1,i) == 1
```

```
        column = length(exonTable.Properties.VariableNames);
        tmpname = samples(i);
        eval(sprintf('exonTable.%s=normCounts(:,%d);',[tmpname{1}
'_normcounts'],i));
    end
end
```

8.5.8 Inferring Differential Expression with a Negative Bionomial Model

```
tLocal =
nbintest(counts_abovecutoff(nz,Sample2),counts_abovecutoff(nz,Sample1)
,'VarianceLink','LocalRegression');

% Multiple Testing and Adjusted P-values
% compute the adjusted P-values (BH correction)
% what would be the fraction of false positives if all the genes with
adjusted P-values below a given threshold were considered significant?

padj = mafdr(tLocal.pValue,'BHFDR',true);

% add to the existing table
exonTable.pvalue = tLocal.pValue;
exonTable.padj = padj;
```

8.5.9 Create a table with significant genes

```
padj = 0.01;

sig = exonTable.padj < padj; %& (exonTable.log2FC < -1 |
exonTable.log2FC > 1);
exonTableSig = exonTable(sig,:);
exonTableSig = sortrows(exonTableSig,'padj');
numberSigGenes = size(exonTableSig,1);
writetable(exonTableSig,[outputpath,
strjoin(strsample1,'.'),'vs',strjoin(strsample2,'.'),'_sig_fea_WBcel23
5.78.xlsx'],'WriteRowNames',true);
writetable(exonTable,[outputpath,
strjoin(strsample1,'.'),'vs',strjoin(strsample2,'.'),'_exon_WBcel235.7
8.xlsx'],'WriteRowNames',true);
```

8.5.10 Create a MA plot with significant genes

```
close all
h = figure
index_significantvalues = exonTable.padj < padj;
scatter(log2(exonTable.meanBase),exonTable.log2FC,3,index_significantv
alues,'o', 'LineWidth', 6)
map = brewermap(2,'*Set1');
colormap(map)
ylabel(['log2 fold change ' strjoin(strsample2,'.') '/'
strjoin(strsample1,'.')]);
xlabel('log2(Mean of normalized counts)')
%set(gca, 'TickLabelInterpreter', 'none')
savefig(h,[outputpath,
strjoin(strsample1,'.'),'vs',strjoin(strsample2,'.')]);
saveas(h,[outputpath,
strjoin(strsample1,'.'),'vs',strjoin(strsample2,'.')],'epsc');
end
```

8.6 Statistical testing for significant changes in the distribution of significantly expressed genes

Script to test for differences in distribution using hypergeometric propability function

8.6.1 Contents

- Hypothesis
- Load input files containing significant reads
- Calculate the hypergeometical distribution
- Calculate p-values for observed distributions
- Plot gene distribution across the linkage groups

8.6.2 Hypothesis

```
%Hypothesis 0: gene from any chromosome have the same chance to be
differentially
%expressed
%Hypothesis 1: gene from any chromosome have not the same chance to be
differentially
%expressed

% use hypergeometric propability function to calculate the propablilty
of X
% in N dependent on M and K
% hygepdf = (X, M, K, N)
% X = number of significant genes on chromosome IV in N
% M = number of genes expressed in wt or sid2
% K = number of genes expressed on chr IV in wt or sid2
% N = number of significant genes in sid2 vs wt
```

8.6.3 Load input files containing significant reads

```
clear

% calculates N for comparisons made
% user defined
inputpath = '/Users/braukmann/Dropbox
(ericmiskalab)/Experiments/FB264/tmp20180322/tmp20180912/short/';
input_fileformat = '_sig_fea_WBcel235.78.xlsx';
outputpath = '/Users/braukmann/Dropbox
(ericmiskalab)/Experiments/FB264/tmp20180322/tmp20180912/short/';
```

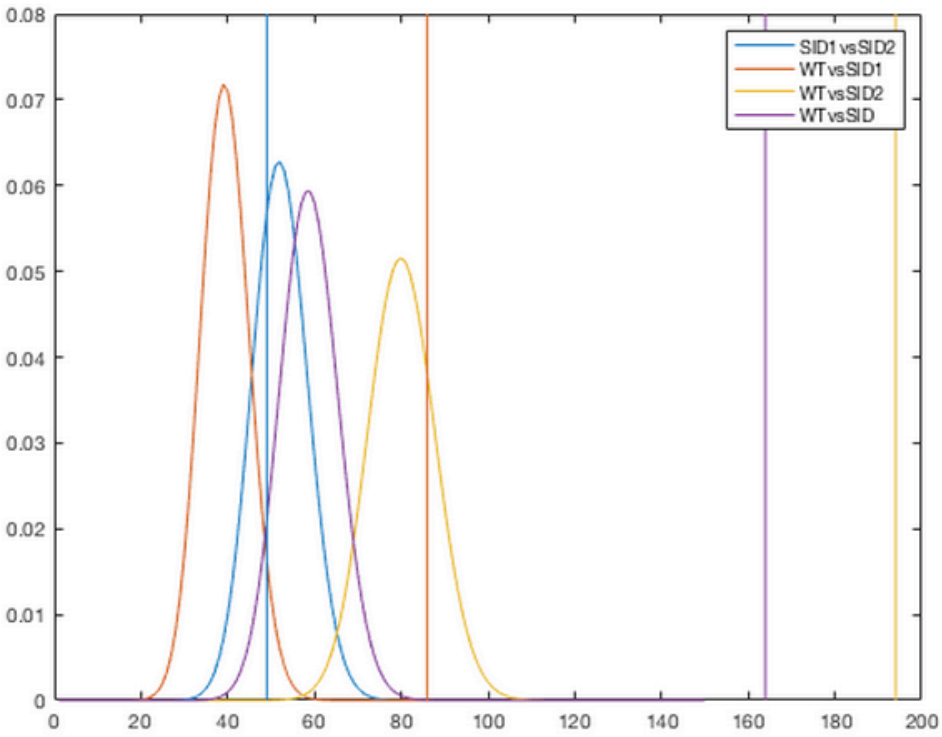
```
%get all files to process and load datatables
inputfiles = dir(inputpath);
inputfiles = inputfiles(cellfun(@(x)
endsWith(x,input_fileformat,'IgnoreCase',true),{inputfiles.name}));

for i = 1:length(inputfiles)
    inputfiles(i).sigTable = readtable([inputfiles(i).folder, '/',
inputfiles(i).name]);
end
% generate understandable name
for i = 1:length(inputfiles)
    tmp = split(inputfiles(i).name,'_');
    inputfiles(i).comparison = tmp(1,1);
end

% Count number of significant genes per chromosome

LinkageGroup = {'I', 'II', 'III', 'IV', 'V', 'X', 'MtDNA'};
SigLinkCounts = nan(length(LinkageGroup),length(inputfiles));

for i = 1:length(inputfiles)
    for j = 1:length(LinkageGroup)
        tmp =
sum(ismember(inputfiles(i).sigTable.Chr,LinkageGroup(j)));
        SigLinkCounts(j,i) = tmp;
    end
end

for i = 1:size(SigLinkCounts,2)
    inputfiles(i).SigCounts = SigLinkCounts(:,i);
    inputfiles(i).N = sum(SigLinkCounts(:,i));
end
% Calculate  M = number of genes expressed in wt or sid2 and K =
number of genes expressed on chr IV in wt or sid2
%user defined
inputpathM = '/Users/braukmann/Dropbox
(ericmiskalab)/Experiments/FB264/smallRNAanalysis/plusKin/fea_exon/all
genes';
input_fileformatM = '.xlsx';
outputpathM = '/Users/braukmann/Dropbox
(ericmiskalab)/Experiments/FB264/smallRNAanalysis/plusKin/fea_exon/all
genes';
%
% Get all files to process and load datatables
inputfilesM = dir(inputpathM);
inputfilesM = inputfilesM(cellfun(@(x)
endsWith(x,input_fileformatM,'IgnoreCase',true),{inputfilesM.name}));
```

```
for i = 1:length(inputfilesM)
    inputfilesM(i).exonTable = readtable([inputfilesM(i).folder, '/',
inputfilesM(i).name]);
end

% Generate understandable name
for i = 1:length(inputfilesM)
    tmp = split(inputfilesM(i).name,'_');
    inputfilesM(i).comparison = tmp(1,1);
end

%
LinkageGroup = {'I', 'II', 'III', 'IV', 'V', 'X', 'MtDNA'};
ExpressedFeatureChr = nan(size(LinkageGroup));

for i = 1:size(inputfilesM,1)
    inputfiles(i).M = sum(inputfilesM(i).exonTable.meanBase > 10^-5);
    Chr_tmp = inputfilesM(i).exonTable.Chr;
    featureindex_tmp = inputfilesM(i).exonTable.meanBase > 10^-5;
    expressedFeatures = Chr_tmp(featureindex_tmp);
    for j = 1:length(LinkageGroup)
        Chrsum_tmp = sum(ismember(expressedFeatures,LinkageGroup(j)));
        ExpressedFeatureChr(j) = Chrsum_tmp;
    end
    inputfiles(i).ExpressedFeatureChr = ExpressedFeatureChr;
    inputfiles(i).K = ExpressedFeatureChr(1,4);
end
```

8.6.4 Calculate the hypergeometical distribution

```
for i = 1:4
    inputfiles(i).ydistr =
hygepdf(1:150,inputfiles(i).M,inputfiles(i).K,inputfiles(i).N);
end

close all
Chr4counts = nan(1,4);
figure(1)
for i = 1:4
    h(i) = plot(inputfiles(i).ydistr,'-');
    legend_tmp(i) = inputfiles(i).comparison;
    hold on
end
ax = gca;

for i = 1:4
```

```
    colors(i,:) = ax.Children(i).Color;
end
for i = 1:4
    Chr4counts(i) = inputfiles(i).SigCounts(4);
    line([Chr4counts(i) Chr4counts(i)], ax.YLim,'color',colors(5-i,:))
end
legend([h(1) h(2) h(3) h(4)],legend_tmp)
```



8.6.5 Calculate p-values for observed distributions

```
for i = 1:4
    inputfiles(i).pvalue = 1-
sum(hygepdf(1:inputfiles(i).SigCounts(4),inputfiles(i).M,inputfiles(i)
.K,inputfiles(i).N));
end
```

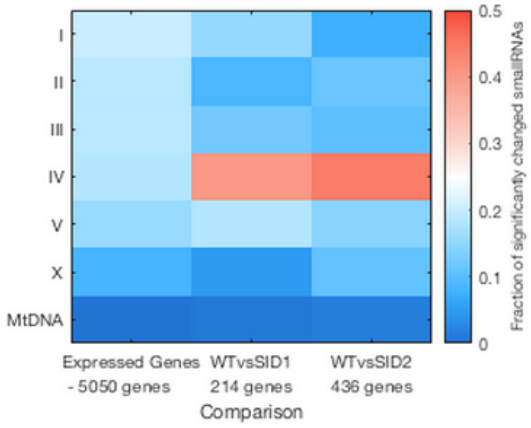
8.6.6 Plot gene distribution across the linkage groups

Obtain list of significanlty changed genes and their position on the linkage group

```
tmp = [];

for i = 1:4
    tmp(:,i,1) = inputfiles(i).SigCounts;
    tmp(:,i,2) = inputfiles(i).ExpressedFeatureChr';
end

% Calculate fraction of genes per linkage group
tmpratio = tmp./sum(tmp);
baseline = mean(tmpratio(:,:,2),2);
subselection = [1,3,4];
ratios = tmpratio(:,:,1);
ratios = cat(2,baseline,ratios);

% Plot linkage distributions
close all
fig = figure(2)
imagesc(ratios(:,subselection),[0 0.5])
%title('Relative number of significantly changed smallRNAs')
%map = jet(128);
map = diverging_map([0:0.001:1],[0,102,204]./255,[255,51,51]./255);
colormap(map)
colorbar('FontSize',16,'Direction','normal')
set(gca, 'TickLabelInterpreter', 'Tex')
cbar = colorbar;
cbar.Label.String = 'Fraction of significantly changed smallRNAs';
xlabel('Comparison')
xticks(1:size(ratios(:,subselection),2))
a = ['Expressed Genes\newline' [inputfiles.comparison]];
a = a(:,subselection);
b = [floor(mean(sum(tmp(:,:,2)))), sum(tmp(:,:,1))];
b = b(subselection);
xlabelinput = cell(size(a,1),size(a,2)+1);
for i = 1:length(a)
    xlabelinput(i) = {[a{i}, '\newline', num2str(b(i)), ' genes']};
end
xlabelinput(1) = {[a{1}, ' - ', num2str(b(1)), ' genes']};
xticklabels(xlabelinput)
%ylabel('Linkage Group')
yticks(1:length(LinkageGroup))
yticklabels(LinkageGroup)
set(gcf,'color','w');
set(gcf,'Units', 'centimeters','Position',[10,10,11,9])
```

```
set(gcf,'Units', 'centimeters','PaperPosition',[0,0,11,9])
set(gca, 'FontName', 'Helvetica')
set(gca,'FontSize',10)
set(gca,'TicklabelInterpreter','Tex')
print(fig,['/Users/braukmann/Dropbox (ericmiskalab)/'...

'Manuscripts/Thesis/Chapter3/sRNA/HyperGeometricSRNAHeatmap.eps'],'-
depsc');
```

```
fig =

  Figure (2) with properties:

      Number: 2
        Name: ''
       Color: [0.9400 0.9400 0.9400]
    Position: [680 678 560 420]
       Units: 'pixels'

  Use GET to show all properties
```



8.7 piRNA analysis in SID small RNA sequencing libraries

Script plot mean abundance of piRNA across the samples.

8.7.1 Contents

- Introduction
- Load read count data
- Read count normalisation by library size
- Calculate mean for samples
- errorbar plot
- Calcualte t-test

8.7.2 Introduction

Normalises the piRNA reads across the technical replicates and plots the mean abundance

8.7.3 Load read count data

```
clear

exonCountTable = readtable('/Users/braukmann/Dropbox
(ericmiskalab)/Experiments/FB264/smallRNAanalysis/plusKin/fea_piRNA/fe
a_piRNA.transcriptid.xlsx');
%removes excessive Chr and Strand information
tmp = cellfun(@(x)
strsplit(x,';'),exonCountTable.Chr,'UniformOutput',0);
exonCountTable.Chr = cellfun(@(x) x(1),tmp);
tmp = cellfun(@(x)
strsplit(x,';'),exonCountTable.Strand,'UniformOutput',0);
exonCountTable.Strand = cellfun(@(x) x(1),tmp);

%loads sample names
samples = exonCountTable(:,7:end).Properties.VariableNames;

%load total read counts
totalreadcounts = readtable('/Users/braukmann/Dropbox
(ericmiskalab)/Experiments/FB264/smallRNAanalysis/plusKin/star/star.su
mmary.txt');
totalreadcounts.Properties.RowNames = samples;
totalreadcounts.Properties.VariableNames = {'Sample_ID', 'Unique',
'Multi'};
% groups = {'WT', 'SID1', 'SID2', 'SX1316', 'PRG1'};
```

```
% groups =
{'WT_N2_01','WT_N2_02','WT_N2_03','SID2_SX3072_01','SID2_SX3072_02','S
ID2_SX3072_03'}
groups = {'WT_N2', 'SID1_SX3046', 'SID2_SX3072', 'SID1_SX3116',
'SID2_SX3237'};
order = [1 2 4 3 5];
xtlabel = {'\it{wild-} \newline\it{type}', '\it{sid-
1}\newline\it{(qt129)}', '\it{sid-2}\newline\it{(qt142)}', '\it{sid-
1}\newline\it{(mj444)}', '\it{sid-2}\newline\it{(mj465)}'};

%groups = {'WT', 'SID2_SX3072', 'SID2_SX3237'};
%order = [1 2 3];
%xtlabel = {'\it{wild-} \newline\it{type}', '\it{sid-
2}\newline\it{qt142}', '\it{sid-2}\newline\it{mj465}'};
for  i = 1:length(groups)
        tmp = strncmp(samples, groups{i}, length(groups{i}));
        group_index(i,:) = tmp;
end
```

8.7.4 Read count normalisation by library size

```
% extract count data from input file (exonCountTable), normalise reads
by Unique reads
counts = exonCountTable{:,7:end};
normCounts = counts./(totalreadcounts.Unique)';

%normCounts = counts./(sum(counts));
```

8.7.5 Calculate mean for samples

```
clear mean_count
for i = 1:length(groups)
    tmp = mean(normCounts(:,group_index(i,:)),2);
    mean_count(:,i) = tmp;
end

index1 = sum(mean_count,2) < 10^-6;
mean_count_abovecutoff = mean_count;
mean_count_abovecutoff(index1,:) = nan;
mean_count_plus_one = mean_count_abovecutoff*10^7+1;
```

8.7.6 Errorbar plot

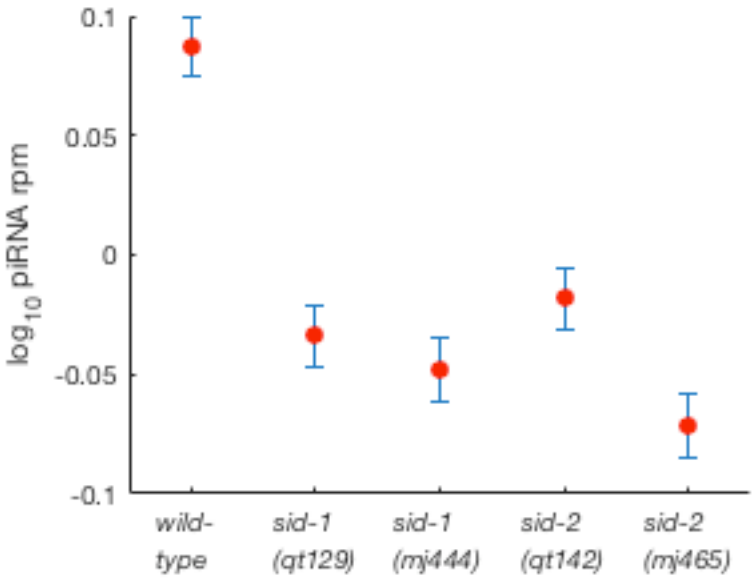
```
order = [1 2 4 3 5];

input = log10(mean_count_abovecutoff(:,order)*10^6);

input(isinf(input))=nan;

% calculate mean
mean_piRNAexpression = nanmean(input);

% calculate standard error of the mean
stderror_piRNAexpression = nanstd(input)./sqrt(sum(~isnan(input)));

% plot errorbar graph
fgr = figure(1)
errorbar(mean_piRNAexpression,stderror_piRNAexpression,'o','MarkerEdge
Color','red','MarkerFaceColor','red','LineWidth',1)
xlim([0.5 5.5])
xticks(1:5)
xticklabels(xtlabel(order))

ax = gca;
ax.LineWidth = 1;
ax.FontSize = 12;
ax.FontName = 'Helvetica';
ax.Box = 'off';
ax.YLabel.String = 'log_1_0 piRNA rpm';
ax.YLabel.FontSize = 12;
%ax.YLabel.Position = [0.275 3.4215 -1.0000];
ax.TickLabelInterpreter = 'Tex';
set(gcf,'color','w');
set(gcf,'Units', 'centimeters','Position',[10,10,11,8])
set(gcf,'Units', 'centimeters','PaperPosition',[0,0,11,8])
set(gca, 'FontName', 'Helvetica')
set(gca,'FontSize',10)
set(gca,'TicklabelInterpreter','Tex')
```

```
fgr =
```



```
  Figure (1) with properties:
```

```
      Number: 1

        Name: ''

       Color: [1 1 1]

    Position: [10 10 11 8]

       Units: 'centimeters'


  Use GET to show all properties
```

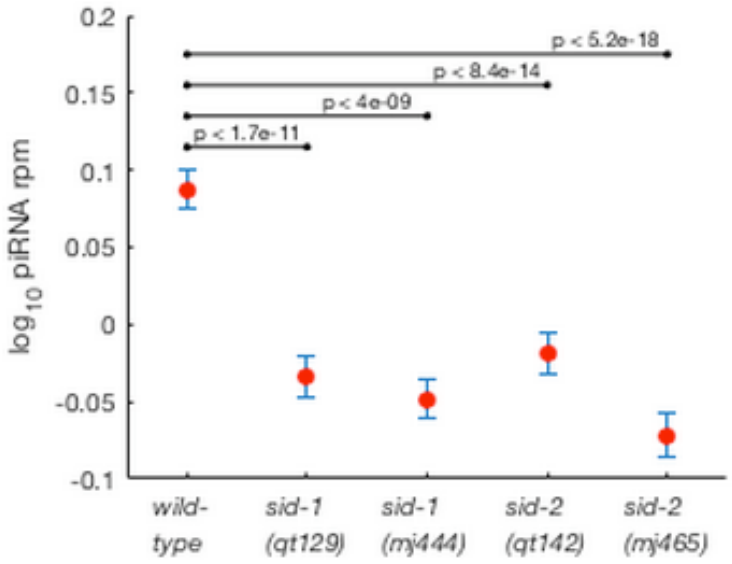


8.7.7 Calculate t-test

```
for i = 1:length(order)
    [htmp1, ptmp2] = ttest2(input(:,1),input(:,i),'Tail','right');
    test(i,1) = htmp1;
    test(i,2) = ptmp2;
end
```

```
test = test(order,:);
yposition = 0.075;
ysampleoffset = 0.02;
ytextoffset = 0.01
for i = 2:length(order)

line([1,1*i],[yposition+i*ysampleoffset,yposition+i*ysampleoffset],'co
lor','k','LineWidth',1,'Marker','.')
    text(1*i-0.5,yposition+ytextoffset+i*ysampleoffset,['p < '
num2str(test(i,2),2)],...
        'HorizontalAlignment','center','FontSize',8)
end

print(fgr,['/Users/braukmann/Dropbox (ericmiskalab)/'...
    'Manuscripts/Thesis/Chapter3/sRNA/piRNAexpressionerrorbar.eps'],'-
depsc');
```

```
ytextoffset =    0.0100
```



8.8 Quantify SID-2 foci

Script to detect SID-2 foci in immunofluorescent images

8.8.1 Contents

- Introduction
- Load images into data_structure
- Identify non intestinal areas in image using backgroundGUI
- Identify intestinal areas in image using noisegroundGUI
- Calcululates focimetrics for all samples
- Detect foci in intestine and extract MeanInensity of foci and Area of foci
- Plot area and intensity on scatter plot to identify nonspecific dots using sid-2 mutants

8.8.2 Introduction

SID-2 foci are detected using triangle thresholding (Zack GW, Rogers WE, Latt SA (1977), "Automatic measurement of sister chromatid exchange frequency", J. Histochem. Cytochem. 25 (7): 74153,)

8.8.3 Load images into data_structure

```
clear;

clc

input_path = '/Users/braukmann/Dropbox (ericmiskalab)/Experiments/FB212/imageanalysis/20170906_sid2insidmutants/cel2/';
file_ext = 'tif';
channel = 3;

data_structure = processFiles(input_path,data_structure);
```

8.8.4 Identify non-intestinal areas in image using backgroundGUI

```
backgroundGUI(data_structure);
```

8.8.5 Identify intestinal areas in image using noisegroundGUI

```
noisegroundGUI(data_structure)

%data must have field called image_name and exp_number

for i =  1:length(data_structure)
    data_structure = fociMap(data_structure,i);
end
```

8.8.6 Calcululates focimetrics for all samples

```
for i = 1:length(data_structure)
    data_structure = focimetrics(data_structure,i);
end
```

show images including foci for samples indicated in sample_index

```
figure

sample_index = [7,8,9,10,14,15,16,17];

for i = 1:length(sample_index)
    subplot(2,4,i)
    showfoci(sample_index(i),data_structure)
end
```

8.8.7 Detect foci in intestine and extract mean inensity and area of foci

```
foci_intensity = foci_summary(data_structure,'MeanIntensity');
foci_area = foci_summary(data_structure,'Area');

% apply cut off
intensity_cutoff = exp(4.9);
indx1 = repmat(intensity_cutoff,size(foci_intensity));
indx1 = foci_intensity > indx1;
foci_intensity(~indx1) = NaN;

area_cutoff = exp(1.5);
indx2 = repmat(area_cutoff,size(foci_intensity));
indx2 = foci_area > indx2;
```

```
foci_area(~indx2) = NaN;

indx3 = isnan(foci_area);
foci_intensity(indx3) = NaN;
indx4 = isnan(foci_intensity);
foci_area(indx4) = NaN;
```

8.8.8 Plot area and intensity on scatter plot to identify nonspecific dots using *sid-2* mutants

```
color = {'k', 'k', 'k', 'r', 'r', 'r', 'g', 'g', 'g', 'b', 'b', 'b',
'c', 'c', 'c', 'm', 'm', 'm', 'b', 'b', 'b', 'b',};
for i = [4:6 16:18]

scatter(log(foci_intensity(:,i)),log(foci_area(:,i)),'o',color{1,i})
    hold on
end

a = 0;
for i = [1,4,7,10,13,16,19,21]
    if i < 19
    a = a+1;
    tmp = foci_intensity(:,i:i+2);
    foci_intensity_grouped(:,a) = tmp(:);
    else
    a = a+1;
    tmp = foci_intensity(:,i:i+1);
    tmp(:,3) = nan(886,1);
    foci_intensity_grouped(:,a) = tmp(:);
    end
end

boxplot(foci_intensity_grouped,'Notch','on')
%boxplot(foci_intensity, 'Notch', 'on')
foci_intensity_mean = nanmean(foci_intensity);
foci_intensity_std = nanstd(foci_intensity);
strain = strain_position.Properties.VariableNames;

for h = 1:size(strain,2)
    strainbyexperiment = cell(size(exp_position));

    for i = 1:size(exp_position,2)
        strain_tmp = cell2mat(strain_position.(char([strain(h)])));
        experiment_tmp = cell2mat(exp_position{:,i});
        strainbyexperiment_tmp = intersect(strain_tmp,experiment_tmp);
        strainbyexperiment{i} = strainbyexperiment_tmp;
```

```
        clear strain_tmp experiment_tmp tmp strainbyexperiment_tmp
    end

    foci_mean = NaN(length(max(cellfun(@(x)
length(x),strainbyexperiment))),length(strainbyexperiment));
    for i = 1:length(strainbyexperiment)
        for j = 1:length(strainbyexperiment{i})
            foci_mean(j,i) =
mean([data_structure(strainbyexperiment{i}(j)).stats.MeanIntensity]);
        end
    end

    foci_mean(foci_mean == 0) = NaN;
    foci_mean_all{h} = foci_mean;
    clear foci_mean
end
```

8.9 Significantly expressed small RNAs

8.9.1 *Sid-1 vs.* wild type small RNAs

Table 6 Significantly expressed small RNAs between *sid-1* and wild type

Gene	*sid-1*	wild type	foldChange	padj
T09B4.12b	196520.59	94389.10	2.08	0
C25D7.14	149.55	2039.09	0.07	3.5285E-149
C25D7.2	150.43	2047.65	0.07	3.96E-149
Y69A2AR.31	321.07	3045.72	0.11	2.2985E-134
M01E11.4c	763.06	89.47	8.53	1.17385E-79
F36A2.3	1792.51	401.42	4.47	1.38698E-64
Y47D7A.16	497.28	2003.52	0.25	8.79752E-52
F37D6.6	683.40	129.63	5.27	6.49631E-48
Y57G11A.18	171.85	735.73	0.23	1.10186E-37
H14A12.5	259.40	32.43	8.00	3.06285E-37
C25D7.5	66.29	338.47	0.20	1.08991E-24
C28A5.4	114.62	422.71	0.27	1.59169E-21
F55A4.9	25.53	160.37	0.16	2.60513E-21
Y41E3.172	61.89	263.18	0.24	2.08627E-20
C39D10.10b	350.47	858.84	0.41	2.1101E-19
T16H12.2	1030.79	2205.85	0.47	5.02112E-19
K08D12.4	399.31	129.47	3.08	2.62627E-18
Y45G12C.16	549.70	196.43	2.80	5.08808E-18
C13B7.6	554.13	200.09	2.77	1.03739E-17
Y19D10A.12	408.67	135.68	3.01	1.2765E-17
C01B4.9	408.67	135.68	3.01	1.2765E-17
T01H8.2	176.33	39.73	4.44	4.16791E-17
Y57G11C.633	33.49	151.10	0.22	2.62477E-16
C25D7.12	84.88	309.81	0.27	3.2704E-16
T10B11.5	372.15	130.93	2.84	3.8618E-16
M03B6.6b	1563.65	771.79	2.03	1.51954E-15
C47B2.8	1026.59	459.70	2.23	7.42714E-15
C13B7.4	119.63	28.31	4.23	4.4419E-14
Y45G12C.9	119.63	28.31	4.23	4.4419E-14
F57A10.2	1537.97	776.78	1.98	9.92117E-14
B0261.8	7104.92	2646.62	2.68	1.97136E-13

Y64G10A.262	123.03	334.52	0.37	3.99558E-13
F42G8.5	65.11	223.49	0.29	4.27165E-13
T05G5.1a	2707.59	1311.46	2.06	4.43574E-13
Y64G10A.34	38.75	145.86	0.27	4.52642E-13
F11A6.2	1973.24	1066.89	1.85	2.44366E-12
K08A2.4	1142.12	571.59	2.00	6.3394E-12
T28F3.10a	3801.61	2002.94	1.90	6.37915E-12
R05A10.13	206.10	470.23	0.44	1.65391E-11
Y58A7A.2	650.98	306.59	2.12	3.45236E-11
C05D11.13	698.67	360.81	1.94	5.74857E-11
R07H5.9	563.46	257.65	2.19	6.72964E-11
C25D7.3	272.99	606.23	0.45	7.26026E-11
R10F2.1	1480.07	816.30	1.81	5.46318E-10
Y57G11C.259	32.41	115.87	0.28	6.06375E-10
C46A5.8	207.08	455.50	0.45	6.72483E-10
Y64G10A.41	142.20	333.38	0.43	8.21151E-10
Y55F3AM.21	1217.59	658.55	1.85	9.06097E-10
F49E11.2	517.85	229.76	2.25	1.0022E-09
Y53C10A.3	13.80	67.11	0.21	1.13608E-09
W08E3.2	351.71	154.58	2.28	1.452E-09
T23G11.1	829.04	432.60	1.92	1.98961E-09
T05G5.4	422.99	191.48	2.21	3.53718E-09
Y53G8B.3	422.75	791.61	0.53	4.57237E-09
T20H4.2	260.69	548.38	0.48	6.25567E-09
ZC416.1	803.22	425.93	1.89	7.7531E-09
F28D9.4.2	221.46	84.59	2.62	1.05143E-08
Y54E2A.4a	6174.14	2685.95	2.30	1.15816E-08
F33E2.9a	1854.54	1139.43	1.63	1.42641E-08
Y51H4A.99	868.16	1500.46	0.58	1.49769E-08
Y105C5A.590	85.03	210.53	0.40	4.3656E-08
F14F9.5	41.19	6.05	6.80	4.65322E-08
Y44E3A.1a	246.14	480.51	0.51	9.84957E-08
Y65B4BR.3	351.84	165.41	2.13	1.42026E-07
Y67D8A.3	331.28	152.96	2.17	1.73053E-07
Y73F8A.170	56.92	144.05	0.40	1.88962E-07
F22D6.10	647.13	366.14	1.77	2.65538E-07
F31E9.6	1887.63	1165.59	1.62	3.16832E-07
F25G6.11a	57.68	140.25	0.41	3.24685E-07
F09A5.5b	8067.95	13741.89	0.59	6.16994E-07

B0511.11	302.02	146.32	2.06	6.51977E-07
Y73F8A.304	645.18	1070.83	0.60	1.38946E-06
Y19D10A.9	41.18	8.83	4.66	1.42055E-06
F56A4.2	41.18	8.83	4.66	1.42055E-06
R13F6.9	635.80	368.89	1.72	1.42774E-06
Y105C5B.317	140.21	283.68	0.49	2.28807E-06
Y116A8C.25	948.72	1499.75	0.63	2.37256E-06
H14E04.3	254.30	482.04	0.53	2.37256E-06
F56A4.12	59.73	17.21	3.47	2.40092E-06
W10D5.1	1164.39	707.07	1.65	2.62973E-06
Y19D10A.5	34.72	7.54	4.60	3.94868E-06
C01B4.8	34.72	7.54	4.60	3.94868E-06
Y19D10A.11	62.69	18.58	3.37	4.08853E-06
K02B9.5b	1184.01	1770.26	0.67	5.74283E-06
T12E12.7b	346.10	188.07	1.84	6.6137E-06
Y41E3.12	291.51	148.90	1.96	8.23595E-06
F14D2.5	405.06	215.87	1.88	8.67909E-06
C56G7.3	1658.28	1062.86	1.56	8.67909E-06
F31C3.8	11680.47	18032.53	0.65	1.32483E-05
Y69A2AR.50	132.52	57.93	2.29	2.29566E-05
Y87G2A.2	111.04	234.25	0.47	2.30102E-05
C11D2.3	440.20	711.86	0.62	2.64155E-05
Y17G7B.8	152.49	299.31	0.51	2.64155E-05
F31C3.7	10518.69	16204.57	0.65	2.90769E-05
Y41E3.51	92.54	193.07	0.48	2.99767E-05
Y51H4A.197	8.68	33.95	0.26	2.99767E-05
Y41C4A.8	105.86	43.71	2.42	3.00992E-05
B0273.7	67.78	144.34	0.47	4.70627E-05
C18D4.6a	468.11	258.27	1.81	4.95618E-05
Y69A2AR.46	236.51	120.38	1.96	5.02361E-05
C27D8.3b	293.93	510.50	0.58	5.2995E-05
Y57G11A.53	289.01	489.36	0.59	5.57412E-05
Y65B4BL.7a	27.57	5.82	4.74	6.31858E-05
Y105C5A.667	9.78	35.20	0.28	6.92355E-05
B0513.4a	28.88	6.34	4.56	9.01154E-05
T24D8.8b	87.00	172.36	0.50	9.7407E-05
F08G2.12	644.05	406.25	1.59	9.84907E-05
Y105C5B.510	26.45	70.71	0.37	9.86856E-05
K07C5.13	784.09	488.18	1.61	0.00010329

K05C4.11	249.08	427.37	0.58	0.000131454
F21H11.3	5.52	25.59	0.22	0.000131454
Y45G12C.3	99.61	44.07	2.26	0.000178741
Y67A10A.48	46.27	104.23	0.44	0.000178741
Y51H4A.244	30.74	75.83	0.41	0.000178741
Y40H7A.130	75.06	149.08	0.50	0.000194001
Y41C4A.13	374.12	212.63	1.76	0.000197236
ZC178.2	110.61	219.14	0.50	0.000202489
Y48B6A.10	207.98	103.92	2.00	0.000204019
Y53G8AL.5b	175.84	312.17	0.56	0.000219741
Y105C5B.912	12.77	39.15	0.33	0.000258779
W01A11.6	60.48	124.05	0.49	0.000280626
C34B7.1	52.81	18.09	2.92	0.000302074
Y50D7A.5	550.66	341.10	1.61	0.000321981
C32D5.12	93.31	185.63	0.50	0.000325781
B0348.5	223.74	119.82	1.87	0.000347161
Y37A1B.276	6.25	25.83	0.24	0.00037649
ZC84.3	60.29	127.61	0.47	0.000389441
M199.36	326.63	527.42	0.62	0.000470421
Y40H7A.76	1143.24	1620.50	0.71	0.00050596
Y73B6BL.257	31.87	75.30	0.42	0.000506132
F13H8.8	318.15	524.83	0.61	0.000551916
F52B11.36	12.29	36.79	0.33	0.000587231
T23G11.11	294.43	170.70	1.72	0.000680053
Y56A3A.34a	2568.05	3643.95	0.70	0.000718517
Y105C5B.485	84.16	160.32	0.52	0.00073235
Y105C5A.774	40.20	89.25	0.45	0.00082745
F58E1.13	1367.25	1883.10	0.73	0.000831817
M04C9.7b	801.16	1156.86	0.69	0.000831817
T27E7.31	205.29	335.78	0.61	0.000842479
Y116A8C.146	567.44	850.13	0.67	0.000846095
ZK228.1	501.13	752.13	0.67	0.000989327
C47E8.11	620.06	411.65	1.51	0.001072159
Y45F10A.4	507.83	311.53	1.63	0.001072996
Y41G9A.7a	954.92	1367.17	0.70	0.001132877
C55C3.3	1487.65	1053.98	1.41	0.001172382
MTCE.9	6220.11	9501.03	0.65	0.001265188
B0286.1	314.21	186.81	1.68	0.001279005
H06O01.3a	95.92	46.11	2.08	0.001288537

C07G1.7	2958.97	2077.51	1.42	0.001431603
F42A9.5	134.57	68.66	1.96	0.001477882
F53A3.1	325.86	528.77	0.62	0.001577432
F55B11.57	13.95	38.30	0.36	0.001623978
F48A11.1	1005.42	689.33	1.46	0.001634016
Y73F8A.1047	40.25	86.14	0.47	0.001717226
Y37E11AL.1	3839.88	2636.97	1.46	0.001788712
T20H12.1	1461.85	1030.65	1.42	0.001832247
H16O14.3	89.69	166.78	0.54	0.001884803
W02D9.10	87.99	163.26	0.54	0.001884803
Y50E8A.3	42.12	13.82	3.05	0.001919663
Y9C9A.33	11.73	33.98	0.35	0.001978942
Y57G11C.207	55.39	108.76	0.51	0.002068272
F25H5.3c.1	142.52	71.37	2.00	0.002136862
F26D12.53	10.00	29.49	0.34	0.002141209
H12I19.36	16.62	43.10	0.39	0.00214175
Y8G1A.2	121.25	224.03	0.54	0.002142646
Y57G11C.651	6.37	22.51	0.28	0.002346494
T05C3.6a	1100.78	777.47	1.42	0.002499864
Y81G3A.1	295.29	175.33	1.68	0.002499864
Y67A10A.130	22.39	53.22	0.42	0.002602794
Y116A8C.363	20.06	49.66	0.40	0.002736862
T10D4.5	26.81	61.98	0.43	0.002880328
F42G4.6	307.49	184.96	1.66	0.002880328
T05C3.8	104.97	190.14	0.55	0.003105373
H05C05.4	160.74	269.86	0.60	0.003263761
C52D10.32	248.93	389.71	0.64	0.003498997
F54C9.4	420.11	261.49	1.61	0.003878478
Y57G11C.253	136.05	229.23	0.59	0.004061011
Y105C5A.319	111.30	194.58	0.57	0.004150809
F23B12.1	413.78	265.10	1.56	0.004327745
F56A4.4	82.92	39.38	2.11	0.004386976
C06E7.21	21.25	50.47	0.42	0.004386976
Y73F8A.340	18.38	44.73	0.41	0.004386976
F09A5.5a	10098.73	6828.29	1.48	0.004415579
F39G3.8	1613.14	1185.46	1.36	0.004542534
C11D2.7	351.58	528.14	0.67	0.004578337
F43G6.5	444.69	284.79	1.56	0.004662116
Y34F4.5a	83.01	37.63	2.21	0.005183889

ZC190.8	300.49	470.48	0.64	0.005375575
Y56A3A.30	139.14	74.52	1.87	0.005375575
Y10G11A.22	242.16	145.82	1.66	0.00570431
F56A12.3b	354.24	527.09	0.67	0.005838512
Y41G9A.7b	356.39	523.79	0.68	0.005912356
C44A5.1	241.59	140.93	1.71	0.005912356
Y105C5B.427	16.98	41.73	0.41	0.006189836
R12B2.2	676.55	957.63	0.71	0.006189836
MTCE.23	266.58	401.88	0.66	0.006374821
C11D2.83	21.30	50.03	0.43	0.006420585
C46A5.10a	419.21	606.05	0.69	0.006420585
C27H2.8	69.05	123.35	0.56	0.006420585
H12I19.12	378.36	556.63	0.68	0.006420585
C46G7.33	302.40	455.31	0.66	0.006867753
H08M01.12	26.69	60.52	0.44	0.006867753
C34E11.5b	66.29	117.56	0.56	0.006895646
Y57G11C.342	13.41	33.65	0.40	0.006949771
F10E9.2	1575.19	1176.52	1.34	0.007263755
T27E7.37	104.94	181.57	0.58	0.00726821
M199.37	11.95	31.15	0.38	0.007286417
F49C5.4	984.62	702.20	1.40	0.007286417
C52D10.57	176.64	99.11	1.78	0.007680788
Y44E3A.6b	23.32	51.22	0.46	0.007997072
T01A4.1a	298.64	183.65	1.63	0.008069242
D2096.10	280.49	418.75	0.67	0.008093956
T16H12.13	1439.06	1049.89	1.37	0.009202177
Y105C5A.784	102.99	174.74	0.59	0.009899123

8.9.2 *Sid-2 vs.* wild type small RNAs

Table 7 Significantly expressed small RNAs between *sid-2* and wild type

Gene	*sid-2*	wild type	foldChange	padj
T07D1.6b	309932.13	455404.11	0.68	0
C25D7.2	143.63	2047.65	0.07	4.2589E-193
C25D7.14	143.09	2039.09	0.07	1.8384E-192
M01E11.4c	845.96	89.47	9.46	4.74106E-90
F25H5.3e	199.41	1165.59	0.17	1.84578E-89
Y57G11A.18	102.48	735.73	0.14	1.20826E-69
F36A2.3	1393.30	401.42	3.47	1.13835E-38
H14E04.3	108.07	482.04	0.22	9.63651E-36
C25D7.5	63.41	338.47	0.19	1.21988E-34
C39D10.10b	282.17	858.84	0.33	5.78479E-32
C47B2.8	1287.20	459.70	2.80	1.26899E-30
F49E11.2	764.91	229.76	3.33	4.79458E-30
Y47D7A.16	718.97	2003.52	0.36	7.40036E-29
B0348.5	466.82	119.82	3.90	2.12247E-27
Y57G11C.633	22.16	151.10	0.15	3.27365E-27
R01H10.6	424.46	109.43	3.88	1.26969E-26
Y41E3.172	60.19	263.18	0.23	2.30007E-24
W06A11.4	1321.33	2958.03	0.45	2.84387E-23
C25D7.12	88.61	309.81	0.29	3.94968E-22
F42G8.5	46.76	223.49	0.21	3.94968E-22
F01F1.3	1447.01	566.18	2.56	1.14967E-21
K08D12.4	386.79	129.47	2.99	2.03949E-20
Y56A3A.34a	1644.13	3643.95	0.45	2.16104E-20
F09A5.5b	5122.56	13741.89	0.37	2.78087E-20
F57F5.1	760.65	289.54	2.63	6.43591E-20
M04C9.7b	511.43	1156.86	0.44	6.43591E-20
F37D6.6	417.35	129.63	3.22	6.43591E-20
R03H10.6	1453.15	597.29	2.43	2.30551E-19
F54C9.4	648.16	261.49	2.48	3.93004E-19
F25H5.3c.1	272.68	71.37	3.82	5.92187E-19
F13A7.11	529.89	1165.43	0.45	2.02098E-18
Y57G11C.259	23.61	115.87	0.20	1.86719E-17
C46A5.8	160.40	455.50	0.35	2.19295E-17

Y73F8A.304	482.21	1070.83	0.45	1.02136E-16
Y116F11B.9b	90.16	14.17	6.36	1.86967E-16
Y54E2A.4a	7453.76	2685.95	2.78	2.84076E-16
F53A3.1	193.54	528.77	0.37	4.0722E-16
F25G6.11a	35.76	140.25	0.25	5.53368E-16
Y64G10A.34	35.97	145.86	0.25	1.01666E-15
T24D8.8b	50.73	172.36	0.29	2.97949E-15
C06H5.7	795.43	340.39	2.34	3.20808E-15
Y64G10A.262	119.10	334.52	0.36	4.05325E-15
Y64G10A.41	115.62	333.38	0.35	5.04753E-15
Y58A7A.2	687.02	306.59	2.24	5.47522E-15
R05A10.13	186.12	470.23	0.40	7.18376E-15
C26C6.3	1315.51	2400.76	0.55	2.24805E-14
ZK930.10b	22424.77	47974.23	0.47	3.18118E-14
C28A5.6	109.90	300.27	0.37	3.58509E-14
ZC416.1	915.73	425.93	2.15	9.48771E-14
F59A7.7	211.13	514.46	0.41	1.21538E-13
Y41G9A.7a	734.24	1367.17	0.54	1.44711E-13
Y116F11B.11	463.16	187.99	2.46	2.33816E-13
C46A5.10a	273.64	606.05	0.45	3.78188E-13
Y40H7A.130	46.92	149.08	0.31	9.47795E-13
K02B9.5b	949.93	1770.26	0.54	1.2145E-12
Y73F8A.170	45.51	144.05	0.32	1.57122E-12
Y116F11B.24	437.04	185.58	2.35	2.91048E-12
Y41G9A.7b	238.78	523.79	0.46	5.14473E-12
Y105C5B.510	16.53	70.71	0.23	7.56053E-12
T05G5.4	440.81	191.48	2.30	1.59767E-11
E04F6.16	387.58	734.89	0.53	2.03857E-11
B0286.7b	731.24	1273.68	0.57	5.40458E-11
Y67A10A.48	31.45	104.23	0.30	7.51833E-11
B0286.1	413.46	186.81	2.21	1.00727E-10
Y67A10A.130	11.39	53.22	0.21	1.02685E-10
C36F7.1	313.28	124.29	2.52	1.25118E-10
F32D1.11	302.04	128.46	2.35	1.32772E-10
Y105C5A.319	71.64	194.58	0.37	1.55893E-10
MTCE.25	1253.29	2078.89	0.60	1.98342E-10
R08C7.14a	179.46	396.74	0.45	2.26848E-10
F36H1.8a	368.50	693.43	0.53	2.48802E-10
Y57G11C.207	34.64	108.76	0.32	3.35284E-10

Y41E3.51	72.92	193.07	0.38	3.96823E-10
ZC204.14	193.68	427.53	0.45	5.88901E-10
F59A7.9	179.53	393.21	0.46	7.62281E-10
ZK856.5	312.51	642.22	0.49	7.62281E-10
T20H4.2	262.49	548.38	0.48	8.40729E-10
F08G12.12b	64.74	172.64	0.37	8.47703E-10
Y57G11C.386	102.28	249.15	0.41	2.49928E-09
T27E7.31	153.76	335.78	0.46	2.74626E-09
MTCE.15	20.12	72.50	0.28	4.62098E-09
F57A10.2	1389.33	776.78	1.79	5.16917E-09
Y38H8A.5	234.29	94.19	2.49	5.70075E-09
K08A2.4	1032.52	571.59	1.81	6.04851E-09
F23D12.5	1096.34	660.08	1.66	8.19497E-09
Y105C5A.667	7.06	35.20	0.20	1.66363E-08
Y51H4A.244	21.94	75.83	0.29	2.69048E-08
Y39F10C.2	131.34	49.46	2.66	4.87879E-08
H06O01.3a	130.51	46.11	2.83	4.87879E-08
B0035.17b	2567.60	4174.89	0.62	4.95105E-08
Y41C4A.13	427.14	212.63	2.01	5.02095E-08
R02C2.1	1872.47	1202.81	1.56	6.09897E-08
F13D12.2	186.59	76.51	2.44	6.1057E-08
B0205.5	2279.23	1346.25	1.69	6.55621E-08
Y57G11C.253	102.10	229.23	0.45	1.13304E-07
Y51H4A.197	7.86	33.95	0.23	1.19542E-07
F20D12.5	588.54	1005.02	0.59	1.2213E-07
R10F2.1	1357.97	816.30	1.66	1.39643E-07
Y105C5B.485	67.12	160.32	0.42	1.48106E-07
Y105C5A.590	92.52	210.53	0.44	2.41377E-07
Y105C5B.912	10.12	39.15	0.26	2.59755E-07
C08F11.7	2997.39	5089.18	0.59	2.83705E-07
T10B5.4	142.63	305.56	0.47	3.14744E-07
Y73F8A.1047	28.67	86.14	0.33	3.28856E-07
F07G6.6	1191.58	1836.13	0.65	3.30543E-07
Y41E3.12	63.69	148.90	0.43	3.89008E-07
C34E11.6a	337.56	609.34	0.55	4.24311E-07
F56A12.3b	291.51	527.09	0.55	4.72217E-07
F54D10.3	2010.87	1281.31	1.57	4.76407E-07
F31C3.7	9267.33	16204.57	0.57	5.26456E-07
C34E11.5a	415.41	698.29	0.59	5.54475E-07

Y71H2AL.2	109.24	39.09	2.80	5.77986E-07
Y57G11C.651	3.74	22.51	0.17	7.25131E-07
Y71G12A.3a	576.69	337.24	1.71	7.29135E-07
H32C10.23	17.41	55.34	0.31	9.0789E-07
Y105C5B.84	42.30	104.85	0.40	9.0789E-07
T20H12.1	1621.56	1030.65	1.57	9.0789E-07
Y53G8AL.5b	154.90	312.17	0.50	9.80721E-07
F31C3.8	10549.45	18032.53	0.59	1.27118E-06
Y67D8A.4b	67138.23	130958.39	0.51	1.43029E-06
Y105C5A.349	31.67	85.14	0.37	1.50189E-06
Y43F8B.8	793.09	1212.64	0.65	1.90773E-06
Y73F8A.69	108.94	228.73	0.48	2.10549E-06
C34E11.5b	50.79	117.56	0.43	2.34125E-06
Y57G11A.57	20.65	60.02	0.34	3.06195E-06
Y51H4A.99	953.06	1500.46	0.64	3.1861E-06
R06A4.6	3973.39	2565.92	1.55	3.48861E-06
F13H8.8	299.01	524.83	0.57	4.31849E-06
C05D11.13	625.47	360.81	1.73	4.32889E-06
Y73F8A.380	1108.59	1708.29	0.65	4.94038E-06
F21H11.3	5.54	25.59	0.22	4.94038E-06
C06E7.21	15.42	50.47	0.31	6.09044E-06
ZK455.9b	100.73	212.00	0.48	6.18476E-06
Y57G11B.190	6.14	26.83	0.23	7.38146E-06
Y51H4A.408	14.87	46.11	0.32	7.39476E-06
Y57G11C.463	2.00	14.63	0.14	8.05482E-06
F55B11.17	8.59	29.96	0.29	8.34976E-06
C16H3.5a	777.51	1176.42	0.66	9.3072E-06
F36A4.16a	179.51	324.90	0.55	9.50413E-06
T27E9.6	459.51	267.24	1.72	9.60149E-06
Y57G11A.53	287.85	489.36	0.59	9.85756E-06
Y73B6BL.257	27.67	75.30	0.37	1.03255E-05
Y37E11AR.2	1981.04	3100.14	0.64	1.12383E-05
H14A12.5	84.85	32.43	2.62	1.15231E-05
Y105C5A.882	3.64	18.18	0.20	1.34634E-05
C16C8.21	2219.61	1425.93	1.56	1.4968E-05
W03G1.2	88.00	175.96	0.50	1.52486E-05
F13G11.19	8.48	30.26	0.28	1.52486E-05
Y50E8A.3	51.31	13.82	3.71	1.62055E-05
Y45F10D.10a	62.11	19.78	3.14	1.67814E-05

F52B11.36	11.25	36.79	0.31	1.73754E-05
Y105C5B.445	4.79	23.13	0.21	1.7823E-05
Y51H4A.910	137.24	250.87	0.55	1.9185E-05
T07F8.1	61.52	129.10	0.48	1.97615E-05
F13H10.4a	118.65	48.36	2.45	2.07549E-05
F26D12.53	7.85	29.49	0.27	2.15038E-05
T05G5.1a	2002.14	1311.46	1.53	2.24365E-05
T08D2.9	328.37	182.78	1.80	2.25695E-05
Y67A10A.29	128.70	241.37	0.53	2.33314E-05
H12I19.36	14.13	43.10	0.33	2.48877E-05
Y62E10A.4a	47.02	13.92	3.38	2.54414E-05
T27D12.5a	51.77	116.00	0.45	2.55257E-05
B0511.5	142.01	70.56	2.01	2.55257E-05
MTCE.12	862.40	1215.63	0.71	2.69259E-05
Y116A8B.80	145.13	262.16	0.55	3.21667E-05
Y43F8B.22	2030.45	3036.07	0.67	3.37695E-05
F44F1.7	995.24	657.09	1.51	3.65771E-05
T01B10.6a	49.98	110.58	0.45	4.18985E-05
Y17G7B.8	164.27	299.31	0.55	4.78434E-05
C25D7.3	375.72	606.23	0.62	5.17943E-05
C18A11.5b	512.00	307.48	1.67	5.3158E-05
Y41G9A.11b	2647.61	3815.36	0.69	5.5313E-05
W08F4.6	1796.48	1243.68	1.44	5.5313E-05
C04F12.5	125.45	61.53	2.04	5.5313E-05
Y105C5B.103	3.25	17.63	0.18	6.21843E-05
Y69A2AR.50	125.58	57.93	2.17	6.60352E-05
C44B7.11	203.75	104.00	1.96	6.66193E-05
Y37A1A.54	1329.02	1900.76	0.70	7.23486E-05
Y51H4A.34a	691.37	1024.78	0.67	7.56115E-05
C52D10.57	194.90	99.11	1.97	7.66304E-05
C16C8.20	201.68	100.99	2.00	8.23058E-05
Y105C5B.317	155.71	283.68	0.55	9.14647E-05
F55C9.5	766.16	1112.54	0.69	0.00010711
C45G9.7	582.77	914.95	0.64	0.000114749
R10F2.8	235.48	390.98	0.60	0.00012632
M199.37	9.69	31.15	0.31	0.000131518
B0454.9	1690.01	1190.30	1.42	0.000147126
Y7A5A.6	1662.20	1178.57	1.41	0.000149172
Y67A10A.41	139.23	242.96	0.57	0.000151099

T13H5.3	183.05	94.47	1.94	0.000151099
T05E11.9	984.02	1425.90	0.69	0.000157619
C35D6.19	58.54	21.93	2.67	0.000160175
Y48G9A.15a	7739.58	11909.78	0.65	0.000167223
ZK384.5b	3086.09	4485.31	0.69	0.000191698
Y105C5B.378	16.48	44.12	0.37	0.00019245
Y43F8B.29	333.76	524.31	0.64	0.000193434
Y57G11C.342	12.08	33.65	0.36	0.000207258
Y4C6A.54	596.04	394.02	1.51	0.000218269
T07D1.6a	34.68	82.03	0.42	0.000220443
K07A9.2	66.59	27.46	2.42	0.000220443
C32D5.12	101.36	185.63	0.55	0.000220443
Y43F8B.17	849.34	1196.18	0.71	0.000228387
F49C5.4	1005.16	702.20	1.43	0.000236268
F57G4.4	573.64	875.18	0.66	0.000244071
Y47D3A.20	93.23	36.66	2.54	0.00024464
Y102A5C.14	449.72	682.43	0.66	0.00025602
Y53G8B.3	527.31	791.61	0.67	0.000257235
F55A11.10	479.99	719.11	0.67	0.000258783
Y17D7C.3	16460.71	22231.93	0.74	0.000266759
Y105C5B.1120	646.97	956.78	0.68	0.000266759
MTCE.9	6186.17	9501.03	0.65	0.00030379
Y102A5C.35	942.78	1374.50	0.69	0.000306084
F21D9.6	566.85	378.29	1.50	0.000322753
F55H2.1a	463.93	289.98	1.60	0.000326188
Y44E3A.1a	730.42	480.51	1.52	0.000327093
R06A10.1	436.95	268.58	1.63	0.000327283
Y37A1B.83	15.15	40.59	0.37	0.000333066
E03A3.1	581.48	364.31	1.60	0.000333457
Y41E3.63	54.10	107.55	0.50	0.00035546
T23G11.11	284.29	170.70	1.67	0.000377402
Y116A8C.25	1035.27	1499.75	0.69	0.000399363
T27E7.34	106.36	195.49	0.54	0.000415038
Y57G11C.517	62.76	119.78	0.52	0.000434583
ZK1251.1	88.36	161.46	0.55	0.000438372
F48A11.1	998.20	689.33	1.45	0.0004705
Y67D8A.3	270.05	152.96	1.77	0.000497452
Y53C10A.3	26.90	67.11	0.40	0.000497452
F52G3.3	3287.94	2335.61	1.41	0.000504576

Y73F8A.930	29.82	66.26	0.45	0.000508525
F02H6.55	9.98	28.70	0.35	0.000525122
Y102A5C.38	351.18	201.97	1.74	0.000534331
F55B11.57	14.38	38.30	0.38	0.000534331
C44C1.1	579.72	393.48	1.47	0.000536238
K09A9.5	613.26	424.52	1.44	0.000536238
M199.36	340.98	527.42	0.65	0.000536238
F12E12.4	762.29	519.56	1.47	0.000536238
T21B4.13b	3534.67	6294.72	0.56	0.000565691
Y73F8A.372	78.85	147.48	0.53	0.000572622
F58A4.1	204.13	116.54	1.75	0.000576417
Y67A10A.35	14.44	39.67	0.36	0.00061522
Y105C5B.249	25.39	59.99	0.42	0.000636956
Y45F10D.32	14.01	36.87	0.38	0.000709514
C49C3.15	113.15	57.23	1.98	0.00070971
Y105C5B.1043	10.42	28.81	0.36	0.000715761
Y116A8C.105	4.68	17.22	0.27	0.000727071
Y55B1BR.3	2094.23	1498.72	1.40	0.000727071
Y116A8C.363	20.83	49.66	0.42	0.000750768
K10C2.9a	885.56	1203.69	0.74	0.000796897
H16O14.3	94.23	166.78	0.56	0.000849345
C10G6.7	13.09	34.76	0.38	0.000881612
B0286.7a	112.85	200.16	0.56	0.000892445
Y53G8AL.1	287.20	458.46	0.63	0.000917278
ZK353.2	136.72	69.47	1.97	0.000917758
T16H12.13	1498.03	1049.89	1.43	0.000992916
Y54F10BM.9	7833.95	11666.90	0.67	0.00100131
F08G2.12	616.31	406.25	1.52	0.00101812
F32A7.6	24.86	57.17	0.43	0.001031449
Y105C5B.178	143.73	234.46	0.61	0.001043365
T04D1.5a	7124.09	10564.01	0.67	0.001069018
B0273.7	79.74	144.34	0.55	0.001157629
Y105C5A.1195	15.37	39.59	0.39	0.001170845
H12I19.12	363.97	556.63	0.65	0.001204011
Y116A8C.436	37.67	77.96	0.48	0.001204011
Y41E3.129	8.09	25.39	0.32	0.001238366
C25H3.6c	36.44	12.33	2.96	0.001299587
T02G5.4	3543.37	6121.84	0.58	0.001326775
F57F5.5c	37.54	13.15	2.85	0.001361316

Y37A1A.61	5.92	19.57	0.30	0.001368906
Y48G9A.14b	6.40	20.80	0.31	0.001454181
C49F8.2	116.94	63.00	1.86	0.0014929
Y55F3AM.21	936.38	658.55	1.42	0.0014929
B0513.30	14.35	35.50	0.40	0.001575597
F34D6.11	181.12	292.30	0.62	0.001595569
F32A11.5	24257.40	18663.47	1.30	0.001595569
Y73F8A.1138	48.65	95.00	0.51	0.001611042
W08E3.2	251.87	154.58	1.63	0.001611042
M162.8	1030.37	1435.53	0.72	0.001615617
Y102A5C.8	713.22	1016.37	0.70	0.001631043
F52B11.24	22.81	52.26	0.44	0.00165374
Y105C5B.427	17.29	41.73	0.41	0.00165374
Y41E3.76	234.37	367.56	0.64	0.00167589
F55A11.12b	759.25	1060.01	0.72	0.001678098
C15H11.12	27.26	58.20	0.47	0.001694039
Y11D7A.10a	52.40	21.51	2.44	0.001741257
R09H3.3.2	777.33	533.81	1.46	0.001753382
K10C8.2	884.72	1204.17	0.73	0.001804537
C52E12.7b	3559.74	6048.33	0.59	0.001804537
F02H6.56	54.24	103.25	0.53	0.001882082
Y71H2B.8	59.63	24.79	2.41	0.001919141
K08E3.10	660.24	463.75	1.42	0.00194648
K02E2.6.1	374.60	553.37	0.68	0.001956373
F19C6.6b	32.13	66.19	0.49	0.001956373
C26C9.3	27.62	59.62	0.46	0.002090966
Y40H7A.28	730.30	1035.85	0.71	0.002113807
Y105C5B.388	22.64	51.09	0.44	0.002134408
Y45F10A.4	481.76	311.53	1.55	0.002183543
Y59H11AR.10	290.93	429.42	0.68	0.002256424
Y64G10A.25	8.21	24.03	0.34	0.002297272
C46G7.33	304.50	455.31	0.67	0.002372732
F31C3.9	49470.93	85846.13	0.58	0.002406764
MTCE.16	5888.58	9028.27	0.65	0.002406764
B0513.4a	23.40	6.34	3.69	0.002423853
Y37A1B.66	1.96	9.59	0.20	0.002423853
F20D1.11a	170.55	285.70	0.60	0.00246406
ZK520.2	27.39	8.34	3.28	0.00246406
F31E9.6	854.08	1165.59	0.73	0.002573471

Y41E3.417	36.93	75.35	0.49	0.002573471
Y105C5A.774	45.91	89.25	0.51	0.002637258
Y41E3.452	53.50	101.91	0.52	0.002708849
M88.6b	11363.39	15529.20	0.73	0.002817572
C40A11.8	201.58	319.06	0.63	0.002846588
Y54E10BR.7	1359.77	971.20	1.40	0.002867902
D1025.10	786.02	556.67	1.41	0.003001232
Y63D3A.9	423.22	279.81	1.51	0.003001232
F38C2.39	5.18	17.24	0.30	0.003037786
Y105C5A.318	44.70	85.19	0.52	0.00310313
Y37A1B.105	30.47	63.88	0.48	0.003159813
T05C3.8	107.86	190.14	0.57	0.00321042
R186.1	155.05	84.30	1.84	0.003293978
Y116A8C.127	45.20	87.44	0.52	0.00339662
DC2.5	599.43	425.72	1.41	0.003810819
C27H2.8	71.10	123.35	0.58	0.003810819
Y67A10A.309	17.78	40.97	0.43	0.003810819
B0198.1	208.43	128.64	1.62	0.003845493
M01F1.7	500.38	719.03	0.70	0.003845493
B0285.7	2515.15	1793.31	1.40	0.003897613
F29B9.9a	1051.51	773.53	1.36	0.003927731
F27C1.18	3947.59	6549.55	0.60	0.003927731
Y116A8C.70	33.72	67.43	0.50	0.003955733
Y73F8A.340	20.46	44.73	0.46	0.003963379
F59H5.6	14378.40	19056.25	0.75	0.003982159
T27E4.3	86.42	44.39	1.95	0.004089602
T27E4.9	86.42	44.39	1.95	0.004089602
F36H12.62	911.20	654.60	1.39	0.004163197
Y51H4A.374	39.76	79.51	0.50	0.004285732
R05A10.48	113.22	191.00	0.59	0.004392618
Y105C5B.118	16.64	38.14	0.44	0.004617987
Y73B6BL.203	104.15	173.45	0.60	0.004679569
C16H3.5b	1542.63	2016.31	0.77	0.004709654
C52E4.7	81.89	139.19	0.59	0.005105393
Y105C5A.1172	30.75	62.62	0.49	0.005105393
R02C2.4	1715.36	1307.57	1.31	0.005117253
F55B11.18	189.93	289.00	0.66	0.005117253
Y105C5A.784	104.21	174.74	0.60	0.005323009
Y105C5B.547	33.72	64.94	0.52	0.005323121

Y41E3.369	10.30	25.95	0.40	0.00535161
F21H11.2a	125.62	211.33	0.59	0.00535161
C04G2.14	222.65	329.11	0.68	0.005381911
Y37A1B.276	10.32	25.83	0.40	0.005381911
F57B9.8	654.53	891.40	0.73	0.005381911
H25K10.91	116.32	188.51	0.62	0.005393418
ZC84.3	72.82	127.61	0.57	0.005393418
W03G11.5b	306.95	447.82	0.69	0.005570783
C04G6.12	190.07	293.89	0.65	0.005662848
Y9C9A.33	14.39	33.98	0.42	0.005685901
Y46H3C.8	34.45	68.51	0.50	0.005772567
Y73F8A.32b	4.41	14.96	0.30	0.005780129
M03B6.6b	1080.20	771.79	1.40	0.005804614
R05A10.19	3.54	13.52	0.26	0.005804614
C15C7.5	296.54	192.75	1.54	0.00584721
Y60A3A.9	552.25	389.71	1.42	0.005861517
Y38H6C.20	261.58	168.30	1.55	0.005861517
Y57G11C.314	8.00	22.30	0.36	0.005861517
B0304.8	762.58	547.65	1.39	0.005861517
T09A5.13b	6.95	19.81	0.35	0.005861517
K11H12.12	26.56	55.21	0.48	0.005960281
H24K24.5	2698.21	1989.55	1.36	0.005980035
F18E9.1	46.03	20.01	2.30	0.005980035
Y105C5B.80	13.18	30.92	0.43	0.005980035
Y105C5A.208	75.28	130.57	0.58	0.006036928
Y40H7A.76	1227.53	1620.50	0.76	0.006048713
F14B6.1	1153.55	863.39	1.34	0.006161514
Y41E3.226	7.50	21.22	0.35	0.00617392
Y67A10A.114	16.60	36.77	0.45	0.006185425
Y57G11C.292	8.64	23.68	0.36	0.006194853
F31D5.1	597.36	427.30	1.40	0.006274963
C35D6.33	25.85	52.66	0.49	0.00628078
T09B4.12b	165070.19	94389.10	1.75	0.006385915
F55B11.50	7.82	21.26	0.37	0.006421716
F33E11.3	1020.37	747.32	1.37	0.006479462
F07G6.7	135.43	212.15	0.64	0.006506594
C32B5.14	546.22	393.26	1.39	0.006537277
C11H1.12b	25.95	53.07	0.49	0.006607995
C05D11.6a	21.98	47.13	0.47	0.006621919

Y57G11C.1113	4.39	14.93	0.29	0.006686022
Y64G10A.98	21.59	44.99	0.48	0.006703441
Y76A2B.3	2953.55	4052.67	0.73	0.006703441
Y105C5B.341	10.07	24.73	0.41	0.006845501
W01A11.6	70.15	124.05	0.57	0.006917516
Y105C5B.513	24.56	48.73	0.50	0.006935379
Y51H4A.543	9.02	24.54	0.37	0.006935379
H08M01.41	26.11	52.82	0.49	0.007121104
F35E2.5	220.20	324.27	0.68	0.007186392
T24A6.20	1376.90	1030.66	1.34	0.007217081
Y57G11B.210	20.93	42.43	0.49	0.007253723
T21F2.3	451.32	622.12	0.73	0.0073032
T11F9.17	159.30	248.86	0.64	0.007371595
B0273.33	13.20	31.84	0.41	0.007371595
Y7A9A.7	232.57	145.04	1.60	0.007476193
Y64G10A.270	7.31	20.55	0.36	0.00752935
K05F6.7	43.68	19.70	2.22	0.007541284
F58F9.16	3.97	14.45	0.27	0.00756061
R193.3	126.55	71.58	1.77	0.007613103
Y57G11C.708	4.55	14.66	0.31	0.007796447
C32B5.16	90.32	50.80	1.78	0.007799038
F40A3.5	1019.75	754.19	1.35	0.007992663
B0454.8	139.44	79.84	1.75	0.007992663
F45C12.11	324.25	218.58	1.48	0.008002834
R02E12.5	118.34	66.51	1.78	0.008047455
C53B7.4	201.98	128.10	1.58	0.008047455
F36H1.7a	1627.15	2144.49	0.76	0.008047455
Y57G11B.29	56.01	97.82	0.57	0.008065312
Y41G9A.9b	503.63	686.43	0.73	0.008488871
F26A1.12	108.90	61.43	1.77	0.008488871
Y105C5B.290	123.67	70.83	1.75	0.008752559
Y60A3A.14	387.66	264.65	1.46	0.008830598
C05B10.13	7.67	21.09	0.36	0.008958183
Y38C1AA.4a	17.75	5.72	3.10	0.009006031
F13H10.7b	36.18	68.90	0.53	0.009049565
C06A6.6b	125.33	195.43	0.64	0.009114681
Y8G1A.2	139.55	224.03	0.62	0.009130223
Y51H4A.464	12.15	28.62	0.42	0.009150557
F48E3.9	96.58	56.10	1.72	0.009260207

Y105C5B.713	15.92	35.37	0.45	0.009260207
MTCE.30	378.60	511.38	0.74	0.009260207
F20D12.12	120.30	68.78	1.75	0.009334282
R04A9.7	820.81	615.49	1.33	0.009432733
Y48G9A.15b	25.54	48.28	0.53	0.009600178
C06E7.56	8.35	21.91	0.38	0.009718322
K02B9.5a	338.94	482.19	0.70	0.009742634
F53G2.2	73.61	128.41	0.57	0.009742634
F19B6.4	82.67	137.30	0.60	0.009746451
F22F1.4a	48374.44	72847.06	0.66	0.009764717
Y53G8B.1	695.16	516.95	1.34	0.0098339
H12I19.88	80.18	131.10	0.61	0.009851119
Y81G3A.1	272.42	175.33	1.55	0.009879898

8.10 Significantly expressed long RNAs

8.10.1 *Sid-1 vs.* wild type long RNAs

Table 8 Significantly expressed long RNAs between *sid-1* and wild type

Gene	*sid-1*	WT	foldChange	padj
sdc-3	66.19	218.26	0.30	6.43098E-12
clec-209	1001.91	398.58	2.51	1.5241E-07
F56A4.2	1001.91	398.58	2.51	1.5241E-07
sax-2	261.57	567.96	0.46	1.5241E-07
col-162	152.61	60.92	2.51	2.16345E-07
col-14	85.81	30.62	2.80	2.62874E-07
col-73	90.70	38.35	2.37	6.99019E-07
sdz-24	331.29	161.36	2.05	7.85164E-07
CD4.18	276.25	143.13	1.93	1.617E-06
F21C10.11	229.01	112.28	2.04	1.34171E-05
F56A4.3	61.39	21.84	2.81	1.62857E-05
col-77	140.47	65.33	2.15	2.94902E-05
B0205.13	109.79	50.73	2.16	4.2508E-05
col-88	81.79	36.39	2.25	9.49403E-05
C18C4.17	1793.38	981.85	1.83	9.67862E-05
col-175	47.37	15.86	2.99	0.000100264
sup-6	2679.02	1496.90	1.79	0.000126393
R12E2.15	23.01	4.33	5.32	0.000126393
dpy-5	93.11	45.07	2.07	0.000126393
F58G1.7	2689.01	1504.47	1.79	0.000133442
C02E7.7	34.46	10.30	3.34	0.000231464
sup-39	1970.36	1135.49	1.74	0.000283518
C15F1.5	1986.07	1150.50	1.73	0.000366512
H27M09.8	695.18	425.74	1.63	0.000444736
H27M09.6	695.18	425.74	1.63	0.000444736
R12E2.7	49.31	18.38	2.68	0.000514691
F41F3.3	122.29	63.16	1.94	0.000517645
set-9	259.60	474.28	0.55	0.000519864
F35E12.5	267.85	154.55	1.73	0.000723048
grd-14	135.03	64.93	2.08	0.00098957
F08H9.10	2218.96	1331.01	1.67	0.00113556

cdr-2	173.26	99.73	1.74	0.001386807
dct-17	368.60	216.40	1.70	0.001395788
K11H12.4	241.11	136.82	1.76	0.001395788
tag-297	56.02	23.23	2.41	0.001538813
Y75B12B.12	415.77	266.75	1.56	0.001675956
K08D12.6	2947.76	6235.22	0.47	0.001675956
ceh-100	191.15	331.89	0.58	0.001838619
col-120	50.85	20.33	2.50	0.002068689
col-38	81.81	44.27	1.85	0.002068689
nduo-2	369.12	205.16	1.80	0.002074678
C06E8.5	37.24	13.01	2.86	0.002338098
swsn-1	185.34	322.32	0.58	0.002384285
C49C8.5	723.61	432.07	1.67	0.002414863
gst-10	308.92	185.32	1.67	0.004175428
col-104	50.86	22.68	2.24	0.005137927
T08G5.11	1012.80	602.49	1.68	0.00518613
str-238	1.95	10.73	0.18	0.005823054
Y73B3A.23	559.21	357.01	1.57	0.006518239
Y57G11C.35	251.56	156.51	1.61	0.006518239
asp-8	803.19	501.95	1.60	0.006673306
unc-22	344.20	694.38	0.50	0.006673306
col-60	53.30	23.88	2.23	0.006851714
cpg-1	1274.85	2449.73	0.52	0.006851714
mir-1831	59.27	29.37	2.02	0.007344948
Y66H1A.4	129.67	223.72	0.58	0.007702194
dod-17	387.75	246.22	1.57	0.007702194
C27H5.11	546.35	363.62	1.50	0.007702194
lin-41	372.75	651.30	0.57	0.008331938
Y47G6A.15	97.88	52.91	1.85	0.008782766
col-138	70.18	39.45	1.78	0.009398991

8.10.2 *Sid-2 vs.* wild type long RNAs

Table 9 Significantly expressed long RNAs between *sid-2* and wild type

Gene	*sid-2*	WT	foldChange	padj
sdc-3	83.82	218.26	0.38	3.31419E-12
T25C12.3	4031.96	8290.97	0.49	1.26979E-08
col-161	90.28	194.93	0.46	7.62238E-07
col-39	498.01	1075.77	0.46	7.62238E-07
col-146	293.17	609.61	0.48	2.77419E-06
col-7	399.42	762.32	0.52	2.77419E-06
col-62	370.00	715.43	0.52	4.24455E-06
dod-21	82.49	154.78	0.53	3.04099E-05
hmit-1.1	68.94	150.49	0.46	3.44008E-05
col-125	145.85	280.92	0.52	3.93213E-05
col-117	211.14	398.76	0.53	4.50015E-05
col-3	214.24	401.89	0.53	5.15225E-05
C34H4.1	77.93	156.41	0.50	5.6552E-05
col-13	685.58	1339.68	0.51	6.06036E-05
C32H11.9	80.65	145.17	0.56	8.99777E-05
col-168	77.76	158.62	0.49	0.000109399
col-12	771.29	1496.22	0.52	0.000128679
nob-1	73.73	144.74	0.51	0.000128679
col-167	129.67	244.55	0.53	0.000292368
T22B7.7	60.14	25.78	2.33	0.000410648
scl-22	117.71	213.68	0.55	0.000410648
col-166	211.26	376.29	0.56	0.001017576
col-147	735.63	1374.04	0.54	0.001142186
col-133	634.56	1134.32	0.56	0.001142186
sax-2	342.10	567.96	0.60	0.00141492
grl-21	30.52	68.40	0.45	0.001810991
C16E9.1	81.89	148.19	0.55	0.001942693
F21C10.11	62.14	112.28	0.55	0.002206185
hip-1	325.10	536.12	0.61	0.002206185
C54D10.10	10.24	28.46	0.36	0.002619036
Y47D7A.13	33.81	71.60	0.47	0.003206914
col-10	57.16	112.22	0.51	0.005040097
col-149	296.12	482.64	0.61	0.005040097

col-34	60.33	107.76	0.56	0.006766725
ram-2	106.31	186.84	0.57	0.007324168
C42D4.3	41.64	80.38	0.52	0.007426499
msra-1	296.38	169.98	1.74	0.007436119
col-130	17.59	42.58	0.41	0.008725207
F53A9.8	99.11	174.10	0.57	0.009564388
F17A9.5	196.50	306.26	0.64	0.009909316