

# Natural variation in viral susceptibility of *Caenorhabditis elegans*

Natural variation in viral susceptibility of *Caenorhabditis elegans*

Lisa van Sluijs

Lisa van Sluijs



# Propositions

1. Allelic functionality of antiviral pathways cannot be extrapolated from a single genetic background to others.  
(this thesis)
2. *Caenorhabditis elegans* males have fitness advantages in virus-infected populations.  
(this thesis)
3. Scientists waste disproportionate effort on mutual competition.
4. Digital conferences are necessary to improve inclusiveness in science.
5. The answer to the climate crisis needs to come from economists not technologists.
6. The production of houseplants is harmful to the environment.
7. Filter bubbles enhance (re-)pillarization of the society.

Propositions belonging to the thesis entitled

**Natural variation in viral susceptibility of *Caenorhabditis elegans***

Lisa van Sluijs  
Wageningen, 13 January 2021

# **Natural variation in viral susceptibility of *Caenorhabditis elegans***

Lisa van Sluijs

## **Thesis committee**

### *Promotor*

Prof. dr Jan E. Kammenga  
Professor at the Laboratory of Nematology  
Wageningen University & Research

### *Co-promotors*

Dr Gorben P. Pijlman  
Associate Professor, Laboratory of Virology  
Wageningen University & Research

Dr Mark G. Sterken  
Assistant Professor, Laboratory of Nematology  
Wageningen University & Research

### *Other members*

Prof. dr Bas J. Zwaan, Wageningen University & Research  
Prof. dr Astrid T. Groot, University of Amsterdam  
Prof. dr Ellen A. A. Nollen, University of Groningen  
Prof. dr Emily R. Troemel, University of California San Diego, USA

This research was conducted under the auspices of the Graduate School Production Ecology and Resource Conservation.



# Natural variation in viral susceptibility of *Caenorhabditis elegans*

Lisa van Sluijs

## **Thesis**

submitted in fulfilment of the requirements for the degree of doctor  
at Wageningen University  
by the authority of the Rector Magnificus,  
Prof. dr A.P.J. Mol,  
in the presence of the  
Thesis Committee appointed by the Academic Board  
to be defended in public  
on Wednesday 13 January 2021  
at 4 p.m. in the Aula.

Lisa van Sluijs

Natural variation in viral susceptibility of *Caenorhabditis elegans*,  
190 pages

PhD thesis, Wageningen University, Wageningen, the Netherlands (2021)  
With references, with summaries in English and Dutch

ISBN 978-94-6395-564-5

DOI <https://doi.org/10.18174/532060>

## Table of contents

Chapter 1	
General introduction	7
Chapter 2	
Why do individuals differ in viral susceptibility? A story told by model organisms	19
Chapter 3	
Balancing selection of the Intracellular Pathogen Response in natural <i>Caenorhabditis elegans</i> populations	33
Chapter 4	
Punctuated loci on chromosome IV determine natural variation in Orsay virus susceptibility of <i>Caenorhabditis elegans</i> strains Bristol N2 and Hawaiian CB4856	75
Chapter 5	
Virus infection modulates male sexual behavior in <i>Caenorhabditis elegans</i>	101
Chapter 6	
General discussion	139
Appendices	153
References	154
English summary	172
Nederlandse samenvatting	174
Acknowledgments	178
About the author	184
Publication list	185
PE&RC Training and Education Statement	186



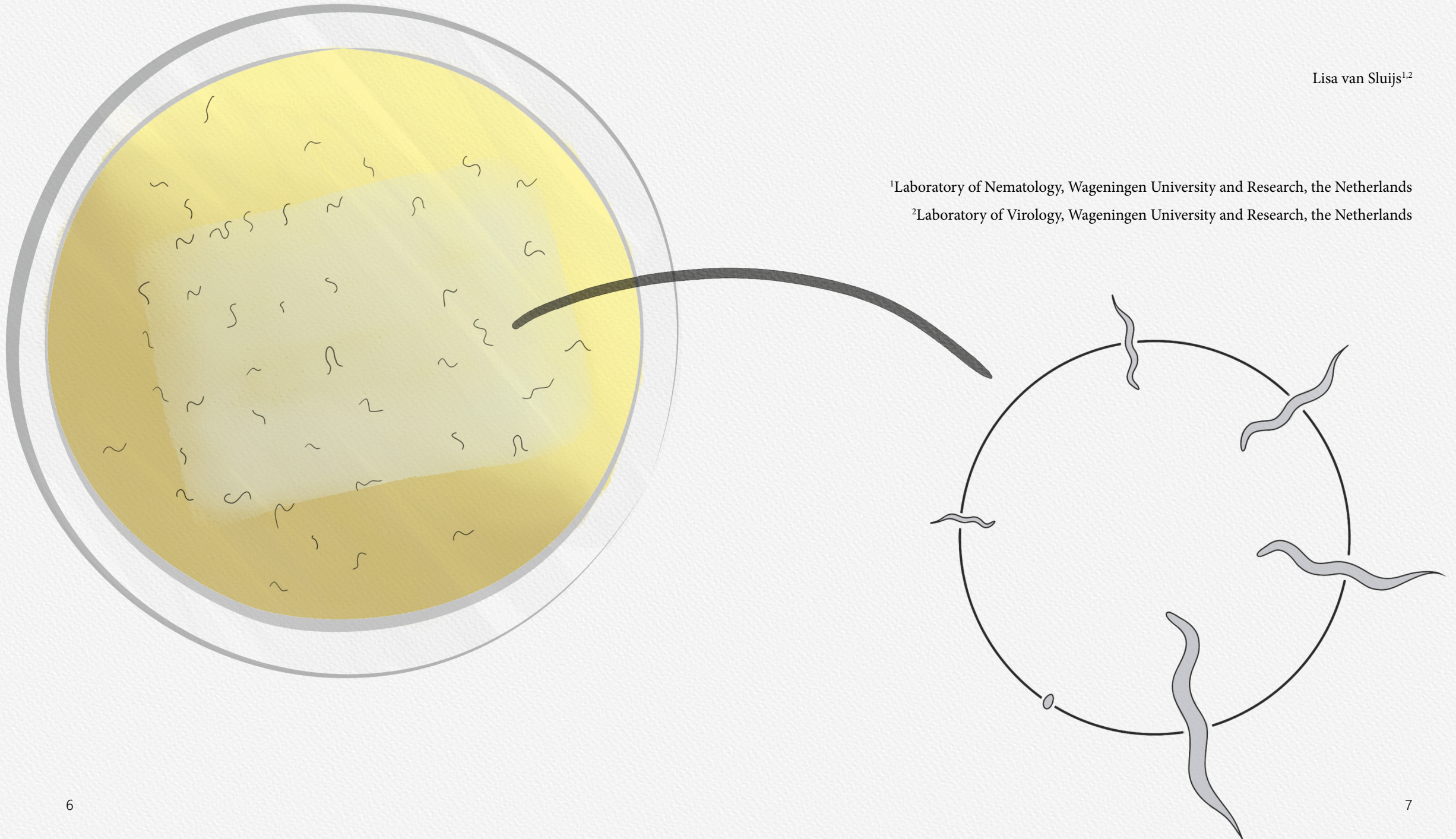
# Chapter 1

## General introduction

Lisa van Sluijs<sup>1,2</sup>

<sup>1</sup>Laboratory of Nematology, Wageningen University and Research, the Netherlands

<sup>2</sup>Laboratory of Virology, Wageningen University and Research, the Netherlands



## *An introduction to nematodes and *Caenorhabditis elegans**

Nematodes (roundworms), with their relatively simple body structures and apparent absence from our daily lives, may seem a fairly plain group of animals at first sight. In fact, many of us are not aware of their very existence. Still, if all animals in this world were to be counted one by one, eight out of ten animals would belong to the phylum Nematoda (van den Hoogen *et al.* 2019). These animals and their diverse range of lifestyles have for long fascinated biologists worldwide. Some nematodes are associated with humans, for example because they parasitize the human body or that of their cattle, pets and agricultural crops (Charlier *et al.* 2020; Jones *et al.* 2013; Pisarski 2019; Traversa 2012). Yet, most are free-living nematodes that are harmless to mankind. These nematodes feed on bacteria, fungi, protozoa (small, unicellular organisms) or other nematodes and contribute to (soil) food web stability (van den Hoogen *et al.* 2019). One of these free-living nematodes, the bacterivorous *Caenorhabditis elegans*, has played an exceptional role in science. This nematode species was selected as an experimental organism based on modesty: a simple body structure and its ease to be cultured in the lab (Brenner 1974). However, because of its success as a genetic model, this 1mm-long animal has not only contributed to understanding nematodes, but also many other life forms (Ankeny 2001; Corsi *et al.* 2015).

*C. elegans* makes a powerful genetic model, because of its androdieocious mating system which means both self-fertilizing hermaphrodites and males occur. Hermaphrodites predominate natural and laboratory populations and can be inbred until the genome is completely homozygous. Homozygous individuals can then be used to start a population containing genetically identical individuals. Males can be used to cross and combine genetic material. The small genome of the nematode is divided over six chromosomes spanning roughly 100 million base pairs (in comparison: the human genome contains 3 billion base pairs). Therefore, this nematode was selected as the first multicellular organism to become sequenced, originally mainly as a preparation before sequencing the human genome (The *C. elegans* Sequencing Consortium 1998; Travers 1999). By now, the *C. elegans* genome still is one of the best annotated genomes and genetic information is easily accessible via open sources such as WormBase (wormbase.org) (Lee *et al.* 2018; Stein *et al.* 2002; The *C. elegans* Sequencing Consortium 1998). This well-annotated genome further established the nematode as a model system for studies on genetics. As a result, *C. elegans* has been at the technological and scientific forefront of genetic and molecular research in multicellular organisms. Among the major achievements are the discovery of RNA interference (RNAi), molecular understanding of apoptosis, using fluorescent tags to directly visualize proteins, and discovery of lifespan altering pathways (Chalfie *et al.* 1994; Fire *et al.* 1998; Johnson & Friedman 1988; Kenyon *et al.* 1993; Sulston & Horvitz 1977; Sulston *et al.* 1983). Many of these studies were facilitated by the use of mutant animals: animals with an altered DNA sequence compared to their



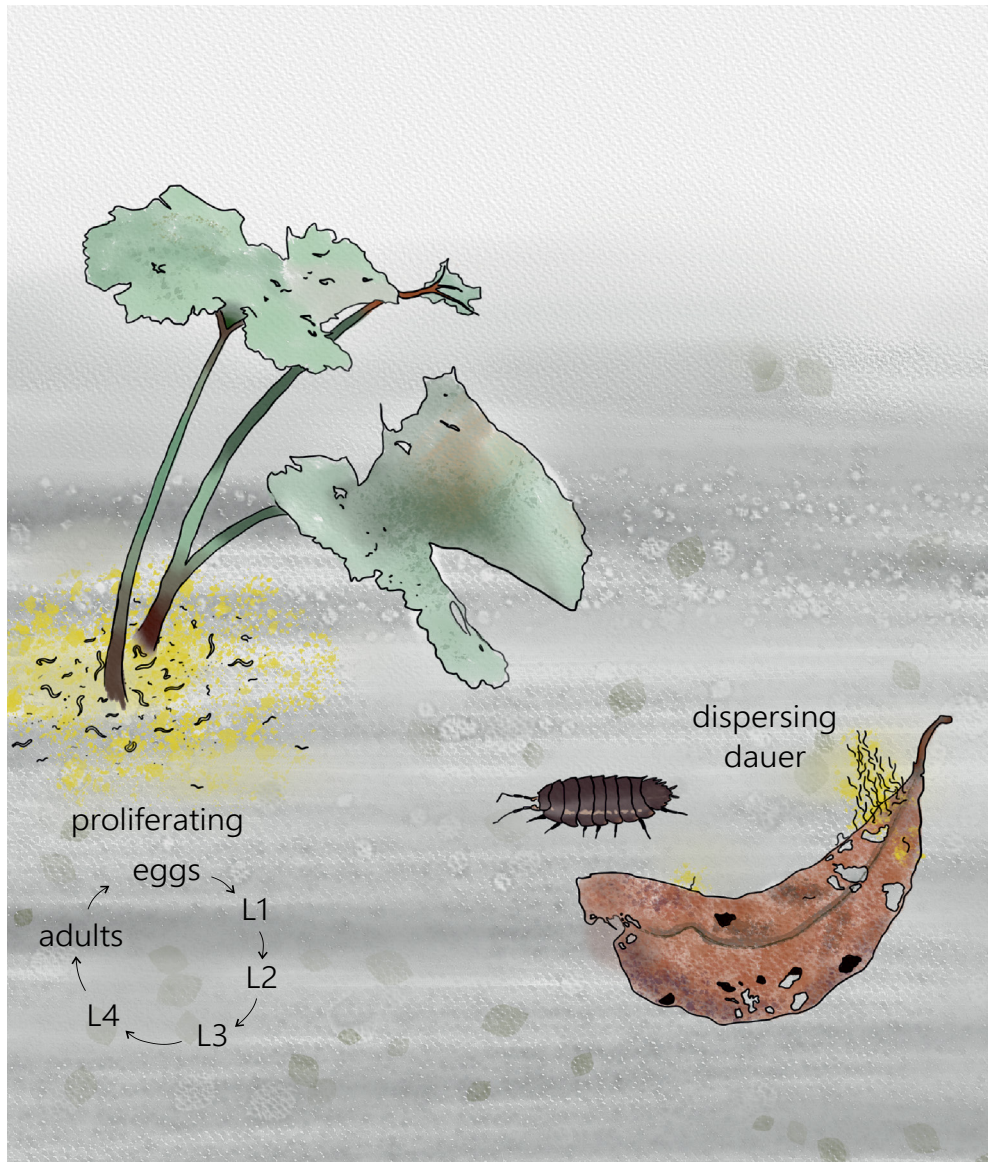
wild-type ancestor that are distributed among researchers via several resources (Caenorhabditis Genetics Center and (Mitani 2009; Thompson *et al.* 2013)). What started out with making random mutants using mutagens (Brenner 1974), is now replaced by precise genome editing techniques such as CRISPR/Cas9 (Chen *et al.* 2016; Farboud 2017). Genetic techniques will be used to (further) unravel functions of the roughly 19,000 genes in its genome.

One may think that *C. elegans* at this point must be among the most well-understood animals on earth. And in one way it is: every step in the nematode's development from egg to egg-laying adult (which takes roughly 3.5 days under optimal conditions) has been monitored and described (Sulston & Horvitz 1977; Sulston *et al.* 1983). Yet, surprisingly little is known about the natural life of the animal (Frézal & Félix 2015; Petersen *et al.* 2015). Since this ecological information has been lacking for years, several aspects of the nematode's biology have been understudied. This thesis focuses on one of these aspects: the interaction between *C. elegans* and the pathogens that infect this nematode in the wild.

### *The ecology of Caenorhabditis elegans*

The natural habitat of *C. elegans* is in rotting materials where many bacteria thrive (Cook *et al.* 2017; Félix & Duveau 2012; Frézal & Félix 2015). These habitats can be rotting fruits, plant stems or organic litter. Natural populations undergo so-called 'boom-and-burst' cycles, where exponential growth under favorable conditions is followed by a collapse of the thriving population after food runs out (Figure 1). Hermaphrodites usually prevail in natural populations with males being present in (frequently far) less than 1% of the cases. However, male frequencies differ per sample location and laboratory experiments indicate their presence can increase under stress caused by food depletion or pathogens (Andersen *et al.* 2012; Barrière & Félix 2005; Felix & Braendle 2010; Lynch *et al.* 2018; Masri *et al.* 2013; Morran *et al.* 2009; Richaud *et al.* 2018; Sivasundar & Hey 2005; Teotónio *et al.* 2006). From egg to adult, *C. elegans* passes through four larval stages (L1-L4). L1 nematodes will stop growing when little food is present and arrest in the L1 stage (Baugh 2013). If unfavorable conditions remain L2 nematodes can enter the dauer state: a stress-resistant life stage that can withstand long periods without food (in this state the nematodes even lack a functioning mouth to process any nutrition) (Diaz & Viney 2015). Dauers can travel to favorable conditions by themselves or can be carried by nematode vectors such as isopods, slugs, and snails (Lee *et al.* 2017; Schulenburg & Félix 2017). Vectors can be reached by the dauers when they stand on their tails and reach out towards passing animals. This behavior named nictation can even be observed in groups when several dauers together form a high column of swaying nematodes (a so-called 'dauer tower') (Figure 1) (Felix & Braendle 2010; Félix & Duveau 2012).





**Figure 1 Natural life cycle of *C. elegans*** – The natural lifecycle of *C. elegans* consists of cycles of dispersal to locations with new food sources. Typical habitats include rotting plant stems (such as from *Petasites hybridus*) where many bacteria can be found. Although *C. elegans* was originally described as a soil nematode, dry litter (such as dried out leaves) typically does not provide a suitable habitat. Nematodes will migrate to more favorable environments by jumping on or attaching themselves to larger animals (such as isopods). During this process dauers can stack together into large clumps called dauer-towers (such as on the right side of the dried leaf).

Once a favorable habitat is reached nematodes start feeding on bacteria. Contrary to the lab where *C. elegans* solely feeds on the bacterial model species *Escherichia coli*, in its natural habitat *C. elegans* is commonly found together with at least ten bacterial families (Samuel *et al.* 2016). Laboratory experiments that fed naturally associated bacteria to (wild) *C. elegans* showed that the microbiome in the *C. elegans* gut can differ significantly from the bacteria in its habitat (Berg *et al.* 2016; Dirksen *et al.* 2016).

Not all microbes encountered by *C. elegans* mean well for them (Figure 2). Several bacteria either colonize the intestine or the cuticula which leads to reduced offspring production (Félix & Duveau 2012; Hodgkin *et al.* 2000). Other bacteria can even poison nematodes via the toxins they produce (Marroquin *et al.* 2000; Tan *et al.* 1999). Fungal pathogens attach to the intestine or the cuticula and consume the nematode (Hodge *et al.* 1997; Jansson 1994). Also deadly are microsporidian pathogens which are commonly found in nematode populations worldwide (Zhang *et al.* 2016). Two of the most frequently found microsporidia, *Nematocida parisii* and *Nematocida ausubeli*, both infect the *C. elegans* intestine. Intestinal cells subsequently lose their food uptake capacity which in the end results in death 3-5 days after infection in the lab (Troemel *et al.* 2008; Zhang *et al.* 2016). A third microsporidian species, *Nematocida dispiodere*, infects and reproduces in the epidermis, neurons and muscles. Microsporidia are released after the nematode bursts (Luallen *et al.* 2016). Oomycetes are another group of deadly *C. elegans*-infecting pathogens that invade the animal after attaching to the cuticula and spread through the entire nematode's body forming pearl-like structures (Osman *et al.* 2018). Relatively mild symptoms are found when *C. elegans* is infected by the only virus known to infect this species to date: the Orsay virus (OrV) (Félix *et al.* 2011). This positive-stranded RNA virus causes progressive degradation of gut cells and reduces offspring production, but does not kill the nematode (Ashe *et al.* 2013; Félix *et al.* 2011). This virus was originally found in Orsay (France) and has two RNA segments: the first encodes the RNA dependent RNA polymerase (RdRP) and the second the structural genes (Jiang *et al.* 2014). The structural gene  $\delta$  forms long filaments and seems necessary for receptor binding and viral egress (Guo *et al.* 2014).



**Figure 2 Intestinal pathogens of *C. elegans*** – *C. elegans* feeds by pumping up bacteria (here shown in pink and yellow) via an organ called the pharynx. In this organ its food is crushed to pieces by the grinder after which it enters the digestive tract. Whilst *C. elegans* is feeding, intestinal pathogens like microsporidia and viruses are provided with the opportunity to enter the nematode's gut. Here, the Orsay virus enters intestinal cells and hijacks their cellular machinery to force the host to produce new viruses. The Orsay virus typically infects only a few of the twenty intestinal cells *C. elegans* has in total (here infected cells are illustrated in red). Still viral infection can damage these large cells severely. Infected cells will shed the new viruses back into the intestinal tract after they are ultimately released via the feces. There the virus will remain until another nematode accidentally ingests it. In reality an adult *C. elegans* hermaphrodite is 20,000 times larger than the Orsay virus (1 mm versus 500 Ångström respectively). A typical bacterium would measure around 5  $\mu\text{m}$ , thereby being 200 times smaller than the nematode.



## Defense of *C. elegans* against the Orsay virus

When a pathogen infects *C. elegans*, the nematode will try to counteract infection. Infection activates an arsenal of anti-pathogenic defenses including molecular and behavioral responses. Antibacterial defense mechanisms have been thoroughly reviewed (Schulenburg & Félix 2017) and differ from the response to intracellular pathogens. The current knowledge of responses to the intracellular pathogen the Orsay virus has been reviewed recently (Félix & Wang 2019) and are the focus of this thesis (Figure 3).

### Intracellular Pathogen Response

Microsporidia, oomycetes, and the Orsay virus all trigger a transcriptional response named the Intracellular Pathogen Response (IPR). This response not only functions in counteracting pathogens, but is also activated under heat stress, exposure to toxic compounds or chemical blocking of the proteasome (Chen *et al.* 2017; Cui *et al.* 2007; Panek *et al.* 2020; Reddy *et al.* 2017, 2019). Activation by one of these signals upregulates the IPR genes. The IPR genes comprise 80 genes, many of which do not have a known function (Reddy *et al.* 2019). Most IPR genes with a known function are part of the ubiquitination system that targets proteins for degradation (Panek *et al.* 2020; Reddy *et al.* 2019). Degradation of excessive proteins, resulting from one of the sources of stress, may relieve cells from proteotoxic pressure.

Immune responses such as the IPR use cellular resources that cannot be invested in developmental processes. In case of the IPR, growth and immunity are balanced by the gene pair *pals-22* and *pals-25* (Reddy *et al.* 2019). *Pals*-genes are characterized by the shared ALS2CR12 protein signature (hence the name *pals*: protein with ALS signature) (Leyva-Díaz *et al.* 2017). Located adjacent to each other on chromosome III *pals-22* and *pals-25* control expression of the 80 IPR genes in an antagonistic manner (Reddy *et al.* 2019). *Pals-22* knock-out mutants develop slowly, but better withstand the stressors that activate the IPR than wild-type animals (Reddy *et al.* 2017). The *pals-22* mutant phenotype can be reversed by also knocking out *pals-25*. These double mutants regain wild-type expression of the IPR genes (Reddy *et al.* 2019). Besides *pals-22* and *pals-25* the *pals*-gene family contains 37 other members (Leyva-Díaz *et al.* 2017). Of these, 25 *pals*-genes are upregulated upon intracellular pathogen infection (Jiang *et al.* 2017; Leyva-Díaz *et al.* 2017). Although the transcriptional response to heat-stress resembles that of a microsporidian infection, both have a distinct biochemical effect (Panek *et al.* 2020).

Activation of the antiviral IPR depends on the presence of DRH-1, which is a homolog of the human RIG-I protein that recognizes viral RNA in infected cells (Ahmad & Hur 2015; Lässig & Hopfner 2017; Sowa *et al.* 2019). DRH-1 is required for IPR activation after viral infection, but not for IPR activation to counteract microsporidian infection or thermal stress. DRH-1 can activate the IPR when only viral RNA is present, suggesting that

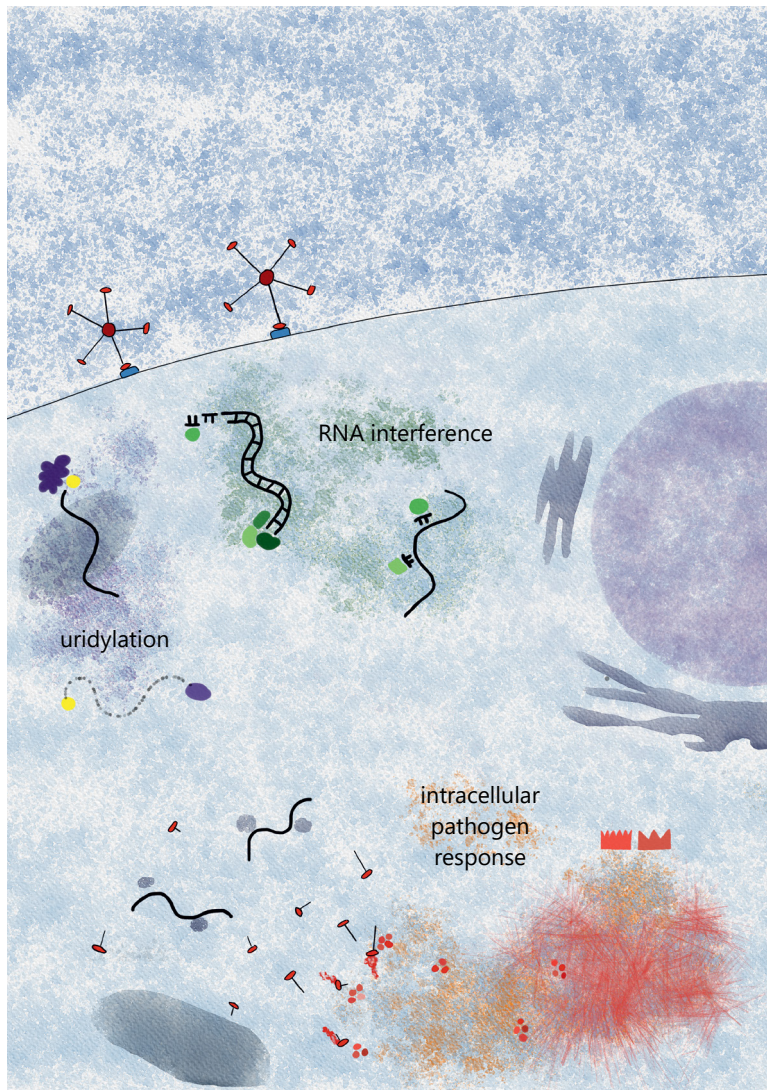
it functions as a pattern-recognition receptor to detect viral presence (Sowa *et al.* 2019). Crucially, DRH-1 is also necessary for activation of the antiviral RNA interference (RNAi) pathway (Ashe *et al.* 2013).

### RNA interference

Antiviral RNAi reduces OrV infection by targeting viral RNA for degradation (Félix *et al.* 2011; Sterken *et al.* 2014). Sequencing of small RNAs in infected wild-type and mutant animals clarified which genes are involved during natural infection with the OrV (Ashe *et al.* 2013; Coffman *et al.* 2017). The antiviral RNAi process starts when OrV is sensed by DRH-1 (Sowa *et al.* 2019). This leads to processing of viral dsRNAs into 23-nt long double-stranded small interfering RNAs (primary siRNAs) with help of RDE-4 and DCR-1. Subsequently, RDE-1, RRF-1, and DCR-3 process and amplify the primary siRNAs into secondary siRNAs. Secondary siRNAs are single-stranded, 22-nt long, have a triphosphorylated 5'-end and a bias for having a guanine (G) as the first base pair. Finally secondary siRNAs are paired with viral RNA by Argonaute proteins which targets the viral RNA for degradation (Ashe *et al.* 2013; Coffman *et al.* 2017).

### Terminal uridylation of viral RNA

Terminal uridylation is the process by which a uridylyl residue (U) is added to the genetic sequence (Zigáčková & Vaňáčová 2018). Both viral RNA segments of OrV can be mono-uridylylated at the 3'-end by CDE-1, an uridylyltransferase (Le Pen *et al.* 2018). Since both OrV genome segments naturally end with an uridylyl residue an UU-tail forms which signals for RNA degradation. The exonuclease gene *xrn-2* and the exosome-component genes *exos-2* and *dis-3* help decay the viral RNA. Although CDE-1 can target siRNAs these small levels of uridylation appear not to change the antiviral RNAi functioning. Furthermore, IPR gene expression remains unaffected in CDE-1 mutants. Thus, antiviral terminal uridylation functions separately of the IPR and RNAi response.



**Figure 3 Antiviral defense in *C. elegans*** – The Orsay virus infects intestinal cells of *C. elegans* after which three antiviral mechanisms are activated. Two of these target the viral RNA: RNA interference (green) and uridylation (purple). Via uridylation the host cell places a tag on the viral RNA after which it is recognized and degraded. RNAi breaks the double-stranded intermediate of the replicating RNA in short pieces using protein complexes that detect viral RNA. The short pieces of viral RNA are loaded into host proteins to detect other viral RNAs. Further viral replication is blocked in this way. The third antiviral mechanism targets proteins. The mechanism of the Intracellular Pathogen Response (orange) is not entirely known, but two genes, *pals-22* and *pals-25* steer upregulation of 80 IPR genes. Part of the products encoded by the IPR genes form complexes which likely tag viral proteins for degradation.



## *Genetic variation in antiviral defense pathways*

Discovery and comprehension of antiviral pathways in *C. elegans* is mainly based on studies under standard laboratory conditions using hermaphrodites of the *C. elegans* reference strain N2 (Félix & Wang 2019; Huang & Kammenga 2020). Standardization of studies is highly important for the reproducibility and comparability of research but can limit understanding of real-world diversity. Diversity may result from environmental or genetic variation and can lead to perturbations in the molecular pathways studied (Kammenga 2017). In this thesis the emphasis lies on the effect of genetic and sex-specific variation on the antiviral defense mechanisms in *C. elegans*.

Immune response genes are for long known to be genetically highly variable, resulting from the ongoing conflict between co-adapting host and pathogen (Mitchell-Olds *et al.* 2007). Genetic variation in antiviral pathways is found in nature which is illustrated by the natural genetic variation present for the viral sensor *drh-1*. Several *C. elegans* strains collected worldwide contain deletions in the domain that recognizes viruses. This leads to higher viral susceptibility of these strains and malfunctioning of the RNAi pathway (and most likely also the IPR which was not yet described at that time (Sowa *et al.* 2019)) (Ashe *et al.* 2013).

Several resources to study the effect of genetic variation have been established over the last years. CeNDR provides a platform for collection, distribution and sequencing of wild *C. elegans* strains that are sampled worldwide (Cook *et al.* 2017). These wild strains can be used for Genome-Wide Association Studies (GWAS), a technique which is also frequently used to study human populations (Buniello *et al.* 2019). GWAS indicates genetic regions that can affect the trait of interest across multiple genetic backgrounds in the population (Ashe *et al.* 2013; Evans *et al.* 2017; Hahnel *et al.* 2018; Zdraljevic *et al.* 2019). Additionally, crosses between wild strains have been inbred to create Recombinant Inbred Lines (RILs) or Introgression Lines (ILs). Classically, two parents create a panel of strains that can be used to pinpoint genetic loci underlying phenotypic variation (Doroszuk *et al.* 2009; Li *et al.* 2006; Rockman *et al.* 2010). Multiparent inbred lines capture additional genetic variation. Currently, two sets of multiparent *C. elegans* inbred lines exist (Noble *et al.* 2017; Snoek *et al.* 2019).

Sex differences can also determine both the innate and adaptive immune response to viral pathogens in many species including fruit flies, birds, lizards, mice, and humans (Ingersoll 2017; Klein & Flanagan 2016; Scully *et al.* 2020). Males are thought to generally invest less resources in immunity than females, making them more prone to infection. Yet, also females can be the more susceptible sex depending on the pathogen (Ingersoll 2017; Klein & Flanagan 2016). Infections in *C. elegans* males have not been performed often, but illustrate that also for this species sex-dependent differences to pathogens occur. Males resist fungal infection by *Cryptococcus neoformans* better than hermaphrodites, whereas hermaphrodites are more

resistant to bacterial infections by *Serratia marcescens* and *Bacillus thuringiensis* (Lynch *et al.* 2018; Masri *et al.* 2013). Even though working with *C. elegans* males is more labor-intensive compared to hermaphrodites, there are several tools that make exploring sex-related traits especially attractive for this species. Valuable insights can be obtained by permanent or temporary manipulation of the sex-determination pathway (Cutter *et al.* 2019; Kasimatis *et al.* 2018; Timmons *et al.* 2014). In addition, large sets of gene-expression and neuronal data are freely available for both sexes (Cook *et al.* 2019; Gerstein *et al.* 2010).

## Outline of this thesis

This thesis focuses on the role of natural genetic variation in the host *C. elegans* in its interaction with the Orsay virus. The data presented here broadens the understanding of antiviral immunity in this thoroughly studied organism and provides insight into natural circumstances that shape this nematode's genome.

This chapter ([Chapter 1](#)) presents an introduction to the research field of host-pathogen interactions in the model organism *C. elegans*. Besides *C. elegans* other model organisms have been used to investigate host-virus interaction in wild isolates. [Chapter 2](#) reviews literature on studying the effect of genetic variation on host-virus interactions in three main model organisms and summarizes genetic resources that are available. In [Chapter 3](#) the CeNDR database, a genetic resource containing information of over 300 wild *C. elegans* isolates, is used to investigate natural genetic variation in the IPR genes that transcriptionally respond to viral infection. A two-parental *C. elegans* Recombinant Inbred Line panel is used in [Chapter 4](#) to pinpoint which genetic variants determine susceptibility to the Orsay virus. Sex-specific differences in *C. elegans* remain a less explored source of phenotypic variation. [Chapter 5](#) describes sex-determined viral susceptibility and describes what consequences viral infection has on mating dynamics. Finally, [Chapter 6](#) discusses the scientific findings in this thesis and puts them in perspective for future research.



## Chapter 2

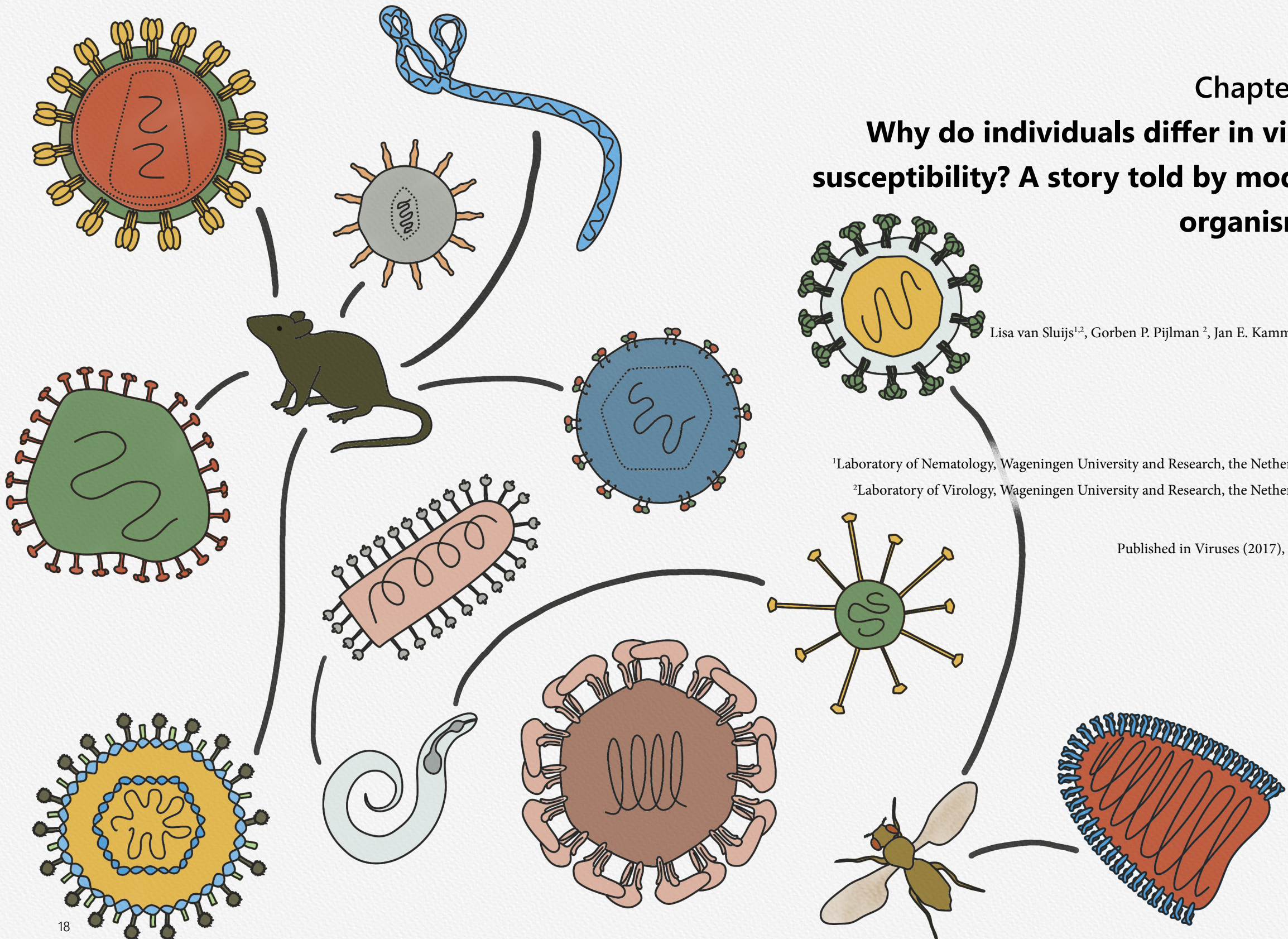
# Why do individuals differ in viral susceptibility? A story told by model organisms

Lisa van Sluijs<sup>1,2</sup>, Gorben P. Pijlman<sup>2</sup>, Jan E. Kammenga<sup>1</sup>

<sup>1</sup>Laboratory of Nematology, Wageningen University and Research, the Netherlands

<sup>2</sup>Laboratory of Virology, Wageningen University and Research, the Netherlands

Published in *Viruses* (2017), 9, 284





## Abstract

Viral susceptibility and disease progression is determined by host genetic variation that underlies individual differences. Genetic polymorphisms that affect the phenotype upon infection have been well-studied for only few viruses as for example HIV-1 and Hepatitis C virus. However, even for well-studied viruses the genetic basis of individual susceptibility differences remains elusive. Investigating the effect of causal polymorphisms in humans is complicated, because genetic methods to detect rare or small-effect polymorphisms are limited and genetic manipulation is not possible in human populations. Model organisms have proven a powerful experimental platform to identify and characterize polymorphisms that underlie natural variation in viral susceptibility using quantitative genetic tools. We summarize and compare the genetic tools available in three main model organisms, *Mus musculus*, *Drosophila melanogaster* and *Caenorhabditis elegans*, and illustrate how these tools can be applied to detect polymorphisms that determine the viral susceptibility. Finally, we analyze how candidate polymorphisms from model organisms can be used to shed light on the underlying mechanism of individual variation. Insights in causal polymorphisms and mechanisms underlying individual differences in viral susceptibility in model organisms likely provide a better understanding in humans.

## Introduction

It is common knowledge that individual people differ in their susceptibilities to different viruses. But exactly why individuals differ in viral susceptibility is hardly known. Viral susceptibility is a complex phenotypic trait for which there is large variation among individuals regarding infection establishment and development of disease symptoms. The phenotype upon infection is determined by host genes, the environment and their interactions. Like for many other traits the genetic architecture is complex, which means that viral susceptibility is associated with multiple genes or loci. Whereas most genes and loci have a small effect on phenotypic traits, few of them have a large phenotypic effect (Bloom *et al.* 2013; Manolio *et al.* 2009; Park *et al.* 2011). Individual phenotypic differences are due to polymorphisms in genes or loci that affect the presence, function and interaction of host factors such as RNAs and proteins (Gasch *et al.* 2016; Li *et al.* 2016).

The detection of polymorphic variants in humans is often based on Genome Wide Association Studies (GWAS) which are currently the most widely used approaches to link genetic variation with viral infection. For instance GWAS detected genetic variants associated with variation in HIV-1, Hepatitis C, dengue and Influenza A virus infection (Al-Qahtani *et al.* 2013; Dang *et al.* 2014; van Manen *et al.* 2012; Wei *et al.* 2015; Zignego *et al.* 2014; Zúñiga *et al.* 2012). Causal polymorphisms discovered by GWAS are often polymorphic regions that have a large effect on the phenotype. The identification of multiple small-effect polymorphisms is far more challenging due to requirement of large and genetically highly diverse populations (McLaren *et al.* 2015). By definition GWAS correlates genotypic variation with phenotypic variation based on statistical association tests and as such, GWAS does not provide insight into the underlying molecular mechanisms (van der Sijde *et al.* 2014). Moreover, GWAS is a population level readout that is difficult to translate to the individual level. Other approaches to find polymorphisms that determine viral susceptibility in humans include specific patients and twin studies. Studies in specific patients typically focus on severe outcome of disease and can thereby identify large-effect polymorphisms (Dupuis *et al.* 2003; Zhang *et al.* 2007). Twin studies are a classical approach to compare the effect of genetics and environment and have identified multiple polymorphisms involved in infectious diseases (Chapman & Hill 2012). But twin studies also underline the importance of environment, especially as ambient environmental factors can trigger the adaptive immune system to develop further and become more efficient (Brodin *et al.* 2015). Both twin and special patient studies require that human subjects are investigated. These may be difficult to find, especially for rare or poorly studied viral infections.

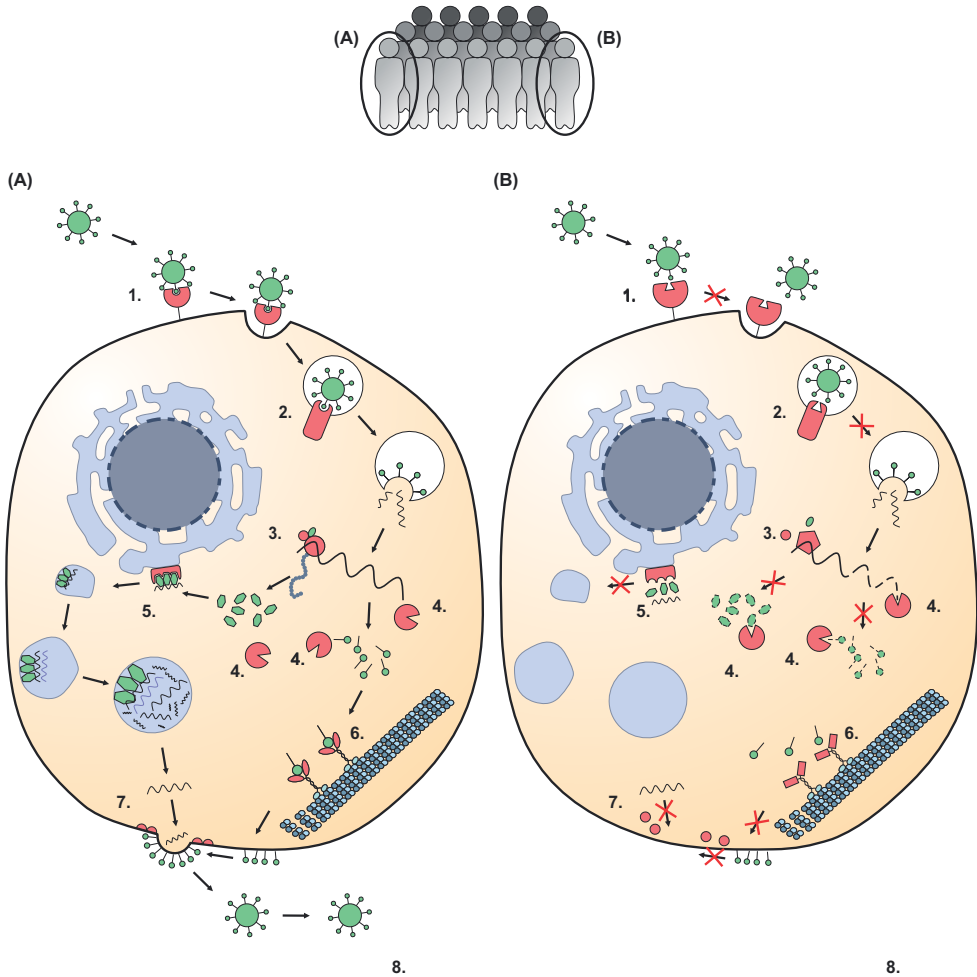
Model organisms offer alternative opportunities for unravelling the molecular mechanisms that are causal to individual differences of viral susceptibility. Here we review the use of model organisms to study the effect of genetic variation on viral infection. Advanced quantitative



genetic tools in model organisms allow for identifying polymorphisms that determine the viral susceptibility in natural populations. The quantitative genetic tools in model organisms provide ways to identify small-effect or rare polymorphisms with an effect on the viral susceptibility. Moreover, in order to mechanistically understand why individuals differ in viral susceptibility, model organisms provide an excellent platform for investigation because individual allelic differences can be studied via experimental manipulation. These fundamental insights can help to guide research in humans through the discovery of homologs or gene networks that underlie natural differences in susceptibility to viral infections.

*Polymorphisms in host factors that interact with viruses cause individual differences in viral susceptibility*

Viruses are obligate intracellular parasites that depend on their host for replication by exploiting various parts of the host cell machinery (Pillay 2009). At the same time, viruses need to evade the innate immune system of the host cell to prevent being sensed and eliminated. Viruses interact with host factors, such as cellular receptors and motor proteins, during their life cycle. Proviral host factors are necessary for viral replication, whereas antiviral host factors inhibit or block viral infection. Potentially every polymorphism in a gene encoding a host factor that interacts with a virus may determine individual viral susceptibility (Figure 1). Several polymorphisms in host factors were identified by human population studies and GWAS to affect the viral susceptibility during different stages in the viral life cycle. A polymorphism in the cellular co-receptor CCR5 prevents HIV-1 from entering the cell, making some individuals resistant against HIV-1 (Dean *et al.* 1996; Liu *et al.* 1996; Samson *et al.* 1996). The polymorphism in the RNA trafficking gene RPAIN is hypothesized to increase viral replication and is associated with severe pneumonia after Influenza A infection (Zúñiga *et al.* 2012). The antiviral host factor BST2 restricts viral egress of HIV-1 by tethering the virus to the cell (Neil *et al.* 2008) and polymorphisms in BST2 and the regulatory sequences of BST2 are associated with the progression of HIV-1 infection (Hancock *et al.* 2015; Laplana *et al.* 2013). Moreover, several polymorphisms in immune pathways associate with the viral susceptibility of humans. The highly polymorphic Human Leukocyte Antigen cluster (HLA) regulates the human adaptive immune response. Genetic variation in the HLA underlies susceptibility differences for viral infections such as HIV-1, Hepatitis B and C virus, Epstein-Barr virus and measles virus (McLaren *et al.* 2015; Ovsyannikova *et al.* 2005; Pittman *et al.* 2016; Salek-Ardakani *et al.* 2002). Furthermore, polymorphisms in the innate immune sensor MDA5 and in and around the cytokine IFN- $\lambda$ -3 are associated with Hepatitis C virus clearance and the responsiveness upon IFN treatment (Ge *et al.* 2009; Hoffmann *et al.* 2015; Thomas *et al.* 2009).



**Figure 1 Genetic polymorphisms can affect the viral life cycle in the cell leading to susceptible and a resistant individual** – A hypothetical viral life cycle (based on a positive stranded RNA viral life cycle) is shown for the cells of a susceptible (A) and a resistant (B) individual. Host factors are shown in red and viral factors are shown in green. A comparison between the viral life cycles of both cells illustrates several steps where individual polymorphic differences in host factors can affect the viral susceptibility. Step 1: in the susceptible cell the virus binds to the cellular receptor, whereas in the resistant cell the virus cannot enter due to polymorphic changes leading to insufficient binding capacity. CCR5 $\Delta$ 32 is a well-known polymorphism in a cellular co-receptor preventing HIV-1 entry (Dean et al. 1996; Liu et al. 1996; Samson et al. 1996). Step 2: in the susceptible cell the virus successfully uses an intracellular transporter, whereas in the resistant cell this is not the case due to genetic individual differences. Polymorphisms in the intracellular

*receptor NPC1 can prevent Ebola virus from being released into the host cell (Ndungo et al. 2016; Ng et al. 2015). Step 3: translation of the viral genome in the susceptible cell is successful, but not in the resistant cell. A polymorphism in a translation initiation factor is associated with resistance to Rice tungro spherical virus (Lee et al. 2010). Step 4: host immunity factors recognize the viral genome and proteins in the resistant cell, but natural genetic variation leads to failure to eliminate the virus in the susceptible cell. Multiple viral infections are affected by polymorphisms in the HLA region (McLaren et al. 2015; Ovsyannikova et al. 2005; Pittman et al. 2016; Salek-Ardakani et al. 2002). Step 5: viral proteins efficiently hijack the cellular machinery for genomic replication, whereas the virus in the resistant cell is unable to replicate due to genetic individual differences. Polymorphisms in the replication gene RPAIN have been associated with Influenza A virus replication (Zúñiga et al. 2012). Step 6: viral proteins are transported by the cellular motor proteins in the susceptible, but not in the resistant cell. Step 7: viral egress is facilitated by host factors in the susceptible, but not in the resistant cell. Polymorphisms in BST2 can prevent HIV-1 from exiting the host cell (Hancock et al. 2015; Laplana et al. 2013; Neil et al. 2008). Step 8: the virus is able to infect and replicate the susceptible individual, in contrast to the resistant individual.*

These examples of polymorphisms detected by GWAS illustrate the power of GWAS to detect genetic variants associated with viral susceptibility. However, GWAS explains a small fraction of the total variation observed, which is in part due to experimental limitations of human GWAS studies. When a polymorphism is rare and/or has a small-effect on the phenotype, the association will explain only a small part of the total phenotypic variation in the population and is therefore not detected by the statistical test (Manolio *et al.* 2009). Moreover, GWAS in humans has limited possibilities to detect the mechanisms underlying small-effect or rare polymorphisms in the examined population for technical and ethical considerations, e.g. genetic manipulations and experiments cannot be conducted.

### *Use of model organisms to unravel the interplay between host genetic variation and viral infection*





Quantitative genetic approaches in model organisms provide means to detect genetic variants and the underlying mechanism involved in viral susceptibility (Mashimo *et al.* 2002; Welton *et al.* 2005). Increased awareness concerning the importance of genetic variation in natural populations has prompted model organism researchers to study the mechanisms of genetic variation using segregating populations generated by parental crossings (Gasch *et al.* 2016). The mapping populations consist of genotyped inbred populations, each harboring different recombinations of the parental alleles. Subsequent phenotyping for viral susceptibility in the inbred strains can yield genetic variants including single nucleotide polymorphisms (SNPs) in coding and non-coding gene regions. Inbred populations with many allelic breakpoints increase the possibility for identification of small-effect or rare polymorphisms because of a high mapping resolution. As many pathways involved in viral

infection are conserved across species, the search for genetic variants in model organisms may identify host factors that function similar as their human homologs.

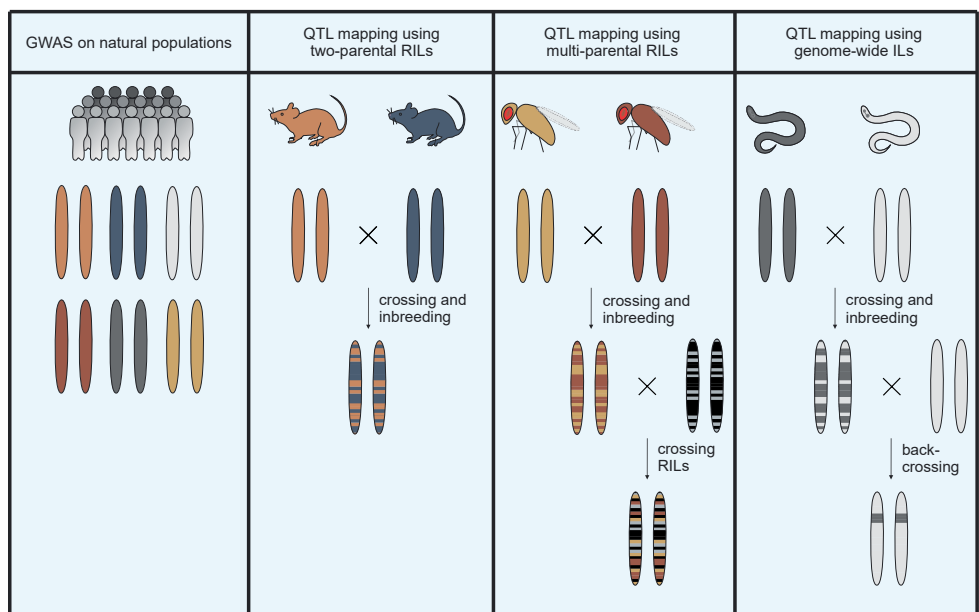
Several inbred populations derived from two parents have been created for model organisms which comprise inbred strains of wild isolates, Recombinant Inbred Line (RIL) and Introgression Line (IL) populations (Figure 2). RILs and ILs can be used for mapping Quantitative Trait Loci (QTL) associated with viral susceptibility. QTL mapping uses RILs and ILs derived from genetically divergent parents that differ in susceptibility to virus infection. The parents are crossed and the offspring is inbred to obtain a population of homozygous RILs, each having different genotypes. Once fully genotyped for genetic markers like SNPs, every individual RIL can be measured for viral susceptibility. QTL mapping statistically correlates viral susceptibility and the genotype of the RILs for every locus on the chromosome. Significant QTL peaks indicate which locus is likely determining the phenotype. In case the detected QTL are relatively broad and cover a large part of the chromosome harboring many candidate polymorphisms, genetic loci identified in QTL studies can be further fine mapped with ILs. ILs contain a single genetic fragment (the introgression) of one wild type strain in the complete genetic background of the other strain. Moreover, a causal relation between the phenotype and the introgression in the target region experimentally verifies the QTL. Next to two-parental RILs, multiparent RILs can be used to increase the mapping resolution (Iraqi *et al.* 2012; King *et al.* 2012b). These RILs are created after several rounds of crossing starting with multiple parents, increasing the genetic variation compared to two-parental crosses. A limitation of two-parental RILs is that they do not encompass the full diversity of allelic variation that exists in natural populations. Inclusion of multiple alleles allows for more precise mapping and identification of potential regulatory variants.

Mouse (*Mus musculus*), fruit fly (*Drosophila melanogaster*) and nematode (*Caenorhabditis elegans*) are major model organisms for genetic and molecular research, including virological research focused on pathogenesis, tissue tropism and (evasion of) immune responses. Below we summarize the quantitative genetic tools that are available for these three model organisms and describe how these tools have been used to identify polymorphisms involved in viral susceptibility (Figure 2). Furthermore, we illustrate how the studies in model organisms can guide detection of genetic variants associated with viral susceptibility in other species, including humans.

(A)

		Application of molecular tools	GWAS on natural populations	QTL mapping using two-parental RILs	QTL mapping using multi-parental RILs	QTL mapping using genome-wide ILs
<i>Mus musculus</i> Genome size: 2.5 Gb		✓	✓	✓	✓	✓
<i>Drosophila melanogaster</i> Genome size: 180 Mb		✓	✓	✓	✓	
<i>Caenorhabditis elegans</i> Genome size: 97 Mb		✓	✓	✓		✓
Human Genome size: 2.9 Gb			✓			

(B)



**Figure 2 Quantitative genetic tools in model organisms that can be used to study viral infection – A)** An overview of the tools that facilitate quantitative genetic studies on viral infection in mice, fruit flies and nematodes. A comparison is made with the possibilities for human research. **B)** The genetic composition of several types of quantitative genetic populations. GWAS populations contain individuals with different genetic backgrounds. RIL populations contain the genetic fragments of two strains that are crossed. Multiparent RIL populations contain genetic fragments from more than two parents, by crossing RILs that originate from distinct parental strains. IL populations contain a single genetic background from one parental strain in the full genome of the other parental strain. ILs are created by backcrossing RILs with one of the parental strains.



*Mus musculus*

Mice (*Mus musculus*) are widely used model organisms because these small mammals are relatively close related to humans (Chinwalla *et al.* 2002). Contrary to other small model invertebrate organisms, in mice the adaptive immune system can be studied. Mice can be infected with several human viruses, such as Influenza A virus and chikungunya virus (Gardner *et al.* 2010; Thangavel & Bouvier 2014). Moreover, either the mouse (immune system) or the virus can be genetically adapted to facilitate infection with additional human viruses, including Zika and HIV-1 (Guabiraba & Ryffel 2014; Lazear *et al.* 2016; Mestas & Hughes 2004; Shultz *et al.* 2012; Victor Garcia 2016). A collection of inbred mice populations is available to investigate genotype-phenotype effects. These populations include the regularly used multiparent RILs of the Collaborative Cross population and the Chromosome Substitution Strains which can be seen as an IL population with large introgressions (Buchner & Nadeau 2015). Moreover, RIL and IL populations are also custom made by researchers to answer specific questions.

Multiparent RIL mice of the Collaborative Cross population were infected with a mice-adapted strain of Ebola virus. Some mice strains were completely resistant, whereas others developed the lethal hemorrhagic fever characteristic for Ebola virus infection. Collaborative Cross strains with different phenotypes upon infection were crossed after which the viral susceptibility and transcriptional response of the F1 offspring was tested. This approach yielded the identification of two susceptibility loci. One of the loci could be identified in more detail and it was shown that the different susceptibilities to Ebola virus are likely due to distinct *Tie2* (also called *Tek*) polymorphisms (Rasmussen *et al.* 2014). *Tie2* is involved in sepsis upon infection with diverse pathogens and forms a target gene for therapeutics that may relieve Ebola virus infection (Ghosh *et al.* 2016; Han *et al.* 2016).

Chromosome Substitution Strains contain a chromosome from one parent in the full genome of the other parent; therefore found QTLs can be specifically attributed to a location (Nadeau *et al.* 2000). The Chromosome Substitution strains have been used to study susceptibility differences to the bacterial pathogen *Staphylococcus aureus* and did identify two causal polymorphisms (Ahn *et al.* 2010). A similar approach could be taken to study the effect of viral infection in this population.

Experiments combining molecular and quantitative genetic mapping techniques showed that polymorphisms in the gene *Mx1* control several viral infections in mice (Boon *et al.* 2009; Ferris *et al.* 2013; Horby *et al.* 2010; Nedelko *et al.* 2012; Thach *et al.* 2001; Welton *et al.* 2005). The functioning of *Mx1* against influenza A virus depends on the genetic background of the mice, indicating *Mx1* resistance may be regulated by other, interacting genes (Shin *et al.* 2015). Future studies in mice may show which molecular pathways underlie *Mx1* resistance in different genetic backgrounds. In humans the homolog *MxA* is also a restriction factor

of Influenza A virus (Turan *et al.* 2004; Xiao *et al.* 2013), however phenotypic variation in Influenza A susceptibility in humans has not been related to *MxA*. Mice experiments suggest that the genetic architecture underlying *MxA* resistance is complex, therefore future studies in humans could focus on investigating *MxA* polymorphisms in patients with severe influenza syndromes (Ciancanelli *et al.* 2016). A focused search in humans with severe syndromes may identify rare *MxA* polymorphisms or cases in which the *MxA* polymorphism in combination with the genetic background is deleterious.

Commercially available and custom made RILs were used in genetic mapping to reveal a susceptibility locus for West Nile virus in mice. The genetic locus was fine mapped using custom made ILs and contains a polymorphism in the gene *Oas1b* (or 2'-5'-*OAS1 L1*). *Oas1b* degrades viral RNAs, which explains differences in West Nile virus susceptibility (Mashimo *et al.* 2002; Perelygin *et al.* 2002). Subsequently, populations of susceptible humans were analyzed to find that polymorphisms in the homolog *OAS1* do indeed affect West Nile virus susceptibility in humans (Lim *et al.* 2009). Taken together, these studies illustrate the value of detecting a causal polymorphism in a mouse gene for translational analysis toward detection of causal polymorphisms in human populations.

### *Drosophila melanogaster*

The fruit fly *Drosophila melanogaster* is an important model for studying genetic variation of virus infection, mainly because natural populations can be collected relatively easy which results in the availability of genetically highly diverse populations. *D. melanogaster* can be infected by at least 30 viruses in nature and around 30% of *D. melanogaster* individuals in the wild carry a viral infection (Webster *et al.* 2015). Researchers using *D. melanogaster* can use the roughly 200 inbred lines of the *Drosophila melanogaster* Genetic Reference Population Lines for GWAS (Huang *et al.* 2014; Mackay *et al.* 2012) or the 1700 multiparent RILs of the *Drosophila* Synthetic Population Resource for high resolution QTL mapping (King *et al.* 2012a).

A GWAS in the *Drosophila melanogaster* Genetic Reference Population showed that common, large-effect polymorphisms explain most of the phenotypic variation in anti-viral responses against *Drosophila* Sigma virus and *Drosophila* C virus (Magwire *et al.* 2012). A subsequent QTL mapping using the same viruses in the *Drosophila* Synthetic Population Resource showed a similar overall trend of large-effect polymorphisms that determine the viral load. However, the QTL mapping technique increased the mapping resolution compared to the previously performed GWAS, therefore additional polymorphisms were identified. The additional polymorphisms included one in a rare, but major-effect gene named *Ge-1* (Cogni *et al.* 2016). A polymorphism in *Ge-1* also controls susceptibility towards a rhabdovirus as identified using a custom made RIL population. *Ge-1* functions as a bridge between two antiviral host factors

and the polymorphism disrupts the link between the two binding domains of Ge-1 (Cao *et al.* 2016). These studies illustrate that a rare major-effect gene may be missed by GWAS, but can be identified by QTL mapping.

One of the advantages of *Drosophila melanogaster* is that conclusions based on results obtained in laboratory populations can be investigated in wild populations. GWAS and QTL mapping both found that the *ref(2)P* polymorphism is the major determinant of viral susceptibility in populations in the lab. The function of *ref(2)P* in antiviral immunity links to the innate immunity of the Toll-signaling and autophagy pathways (Mussabekova *et al.* 2017). Field studies confirmed that polymorphisms in *ref(2)P* affect which flies become infected in the wild (Wilfert & Jiggins 2010), illustrating the use of *D. melanogaster* to pinpoint polymorphisms that define viral susceptibility in nature. Although the *ref(2)P* polymorphism itself may not hold potential for human therapeutics, studies in fruit flies can clarify how polymorphisms providing resistance spread through natural populations (Bangham *et al.* 2007), in a similar fashion as these polymorphisms may spread in the human population.

Moreover, *D. melanogaster* can also be infected with several human pathogens, such as Sindbis virus and West Nile virus (Chotkowski *et al.* 2008; Kemp *et al.* 2013; Sabin *et al.* 2009). These arboviruses are carried by mosquito vectors and both virus and vector can spread quickly due to increased globalization patterns (Gould *et al.* 2017). Therefore, the diseases that result from the infections are important threats to global health. Genetic mapping in one of the available *Drosophila* panels may unveil polymorphisms that alter the susceptibility of viral vectors and give further insights in the molecular basis of infection.

### *Caenorhabditis elegans*

The self-fertilizing hermaphroditic nematode *Caenorhabditis elegans* has recently become an important model for studying viral genetics. *C. elegans* does not suffer from inbreeding depression whereas males can be used for genetic exchange (Gaertner & Phillips 2010). Genetically diverse wild strains are available, and the overall genetic variation within the species is comparable to humans (Andersen *et al.* 2012; Cook *et al.* 2017). Existing genetic tools comprise several RIL populations and an IL population that covers the complete genome (Andersen *et al.* 2015; Doroszuk *et al.* 2009; Li *et al.* 2006; Rockman & Kruglyak 2009). ILs can be backcrossed with the parental strain to increase the mapping resolution in target areas (Bernstein & Rockman 2016). *C. elegans* can be infected with the human zoonotic Vesicular stomatitis Indiana virus and the Orsay virus that is *C. elegans* specific (Félix *et al.* 2011; Gammon *et al.* 2017; Geng *et al.* 2012).

Genetically diverse wild *C. elegans* strains showed different susceptibilities to the naturally infecting Orsay virus (Ashe *et al.* 2013; Félix *et al.* 2011; Sterken *et al.* 2014). A GWAS using a selection of wild strains located a susceptibility locus. Subsequently ILs specific for this

location were created by crossing a resistant and a susceptible strain. Experiments in the ILs showed that a *drh-1* polymorphism largely explains differences in viral susceptibility (Ashe *et al.* 2013). Mammalian *drh-1* homologs, called RIG-I genes, recognize viruses and trigger the anti-viral response (Schlee 2013). Therefore the function of *drh-1* gene was suggested to be conserved, even though the responding pathways differ (Ashe *et al.* 2013). Although this study did not identify a previously unknown gene, studies in *C. elegans* suggest that polymorphisms in *drh-1* homologs may underlie natural differences in viral susceptibility, something that could be investigated in human populations. Moreover, some strains that have the susceptible *drh-1* polymorphism are not susceptible themselves. Follow-up experiments could therefore provide additional insights in the role of the genetic background on the functioning of viral sensors like *drh-1*.

## Future perspectives

Studying viral infection in model organisms provides information that can guide and support research on viral infections in human populations. Current advances, such as the development of advanced multi-parent RILs, improve the mapping resolution and effectiveness of quantitative genetics tools in model organisms. The new mapping tools increase the chances of finding polymorphisms that cause differences in viral susceptibility. Human homologs of causal genes in model organisms are candidate genes that may define viral susceptibility of human populations as well.

The molecular mechanisms behind viral susceptibility differences can be found *in vivo* in model organisms using transcriptomics, proteomics and mutational screenings. There is a pleiotropy of techniques available in these models and some recent advances promise to make unravelling molecular mechanisms behind susceptibility differences even more straightforward. Here we highlight only a few promising techniques and suggest how these can be used to address individual differences in viral susceptibility. After identification of a candidate polymorphism homozygous recombination using CRISPR-Cas9 can be applied to change or insert a specific polymorphism in different genetic backgrounds. Therefore, the effect of a specific polymorphism can be tested contrary to other approaches using chemical mutagenesis or knockdown by RNAi. The effect of a polymorphism in a gene may be different than completely knocking-out or knocking-down the same gene. Moreover, advances in transcriptional studies, including RNA-seq or tissue-specific transcriptomics, provide better clues on the pathways involved in the viral susceptibility. Also increasing amounts of big data, including genome sequences and protein structures can be used to predict the effect of a polymorphism on the functioning of the host factor.

In conclusion, studying the effect of genetic variation on viral infections in model organisms can a) provide fundamental insights in the molecular and the genetic architecture of viral infection, b) identify unknown host factors involved in viral infection and c) provide candidate genes for human population studies that aim to identify which host factors control individual viral susceptibility.

## **Acknowledgments**

We thank the POP-group of the Laboratory of Nematology for valuable comments on the manuscript.



## Chapter 3

# Balancing selection of the Intracellular Pathogen Response in natural *Caenorhabditis elegans* populations

Lisa van Sluijs<sup>1,2</sup>, Kobus J. Bosman<sup>1</sup>, Frederik Pankok<sup>1</sup>, Tatiana Blokhina<sup>1</sup>, Joost A. G. Riksen<sup>1</sup>, Basten L. Snoek<sup>3</sup>, Gorben P. Pijlman<sup>2</sup>, Jan E. Kammenga<sup>1</sup>, Mark G. Sterken<sup>1,2</sup>

<sup>1</sup>Laboratory of Nematology, Wageningen University and Research, the Netherlands

<sup>2</sup>Laboratory of Virology, Wageningen University and Research, the Netherlands

<sup>3</sup>Theoretical Biology and Bioinformatics, Utrecht University, the Netherlands

Manuscript submitted; pre-print published on BioRxiv (2020), 579151



## Abstract

Genetic variation in host populations may lead to differential viral susceptibilities. Here, we investigate the role of natural genetic variation present for an antiviral pathway, the Intracellular Pathogen Response (IPR), underlying susceptibility to Orsay virus in the model organism *Caenorhabditis elegans*. The IPR involves transcriptional activity of 80 genes including the *pals*-genes. The *pals*-genes form an expanded gene family which hints they could be shaped by an evolutionary selective pressure. Here we examine the genetic variation in the *pals*-family for traces of selection and explore the molecular and phenotypic effects of having distinct *pals*-gene alleles. Genetic analysis of 330 global *C. elegans* strains reveals that genetic diversity within the IPR-related *pals*-genes can be categorized in a few haplotypes worldwide. Importantly, two key IPR regulators, *pals-22* and *pals-25*, are in a genomic region carrying signatures of balancing selection. Therefore, distinct *pals-22/pals-25* alleles have been maintained in *C. elegans* populations over time, which suggests different evolutionary strategies exist in IPR regulation. We investigated the IPR by infecting two *C. elegans* strains that represent distinct *pals-22/pals-25* haplotypes, N2 and CB4856, with Orsay virus to determine their susceptibility and transcriptional response to infection. Our data suggests that regulatory genetic variation underlies constant high activity of IPR genes in CB4856 which could determine the host transcriptional defense. We found that CB4856 shows initially lower viral susceptibility than N2. High basal IPR expression levels might help counteract viral infection directly, whereas N2-like strains that need to activate the IPR genes first may have a slower response. Nevertheless, most wild strains harbor N2-like alleles for the *pals*-genes. Our work provides evidence for balancing genetic selection of immunity genes in *C. elegans* and illustrates how this may shape the transcriptional defense against pathogens. The transcriptional and genetic data presented in this study therefore provide a novel perspective on the functional diversity that can develop within a main antiviral response in natural host populations.

## Introduction

Viral infections are a key element in all naturally occurring populations. Genetic variation can change host-virus interactions by altering coding sequences of protein products. Moreover, host-virus interactions can also be influenced by genetic variation due to altered gene copy numbers. Structural and regulatory genetic variation may both affect the viral susceptibility after infection, making some individuals within the population more resistant than others ([Chapter 2](#)) (Franco *et al.* 2013; Piasecka *et al.* 2018; Wang *et al.* 2018). Presence of viruses can thereby select for beneficial genetic variants to remain present in the population (Enard *et al.* 2016; Wilke & Sawyer 2016).

The nematode *Caenorhabditis elegans* and its natural pathogen Orsay virus (OrV) are used as a powerful genetic model system to study host-virus interactions (Félix *et al.* 2011). OrV is a positive-sense single-stranded RNA virus infecting *C. elegans* intestinal cells where it causes local disruptions of the cellular structures (Félix *et al.* 2011; Franz *et al.* 2013). Three antiviral responses are known of which RNA interference (RNAi) and uridylation both target viral RNA for degradation (Ashe *et al.* 2013; Félix *et al.* 2011; Le Pen *et al.* 2018; Sterken *et al.* 2014). The third response, the so-called Intracellular Pathogen Response (IPR), is thought to relieve proteotoxic stress from the infection by the OrV and other intracellular pathogens (Bakowski *et al.* 2014; Osman *et al.* 2018; Reddy *et al.* 2017, 2019). The 80 genes involved in the IPR pathway are controlled by *pals-22* and *pals-25* that function as a molecular switch between growth and antiviral defense. The gene *pals-22* promotes development and lifespan, whereas *pals-25* stimulates pathogen resistance (Reddy *et al.* 2017, 2019). Of the 80 IPR genes that become differentially expressed upon infection, 25 genes belong to the *pals*-gene family. Although the function of most *pals*-proteins remains opaque, PALS-22 and PALS-25 physically interact together and are likely to interact with additional PALS-proteins (Panek *et al.* 2020). The total *pals*-gene family contains 39 members mostly found in five genetic clusters on chromosome I, III, and V (Chen *et al.* 2017; Leyva-Díaz *et al.* 2017). Both the antiviral IPR and the antiviral RNAi pathway require presence of *drh-1* that likely functions as a viral sensor (Ashe *et al.* 2013; Sowa *et al.* 2019).

At present, natural populations of *C. elegans* have been isolated worldwide and catalogued into 330 isotypes maintained by the *C. elegans* Natural Diversity Resource (CeNDR) (Cook *et al.* 2017). The collection contains *C. elegans* nematodes from every continent except Antarctica and each isotype in the CeNDR collection has been sequenced. Previous research found that genetic variation in the gene *drh-1* determines susceptibility to the OrV (Ashe *et al.* 2013), yet most likely additional genetic variants can influence host-virus interactions in nature. The CeNDR database provides an ideal platform to investigate worldwide genetic variation and traces of evolutionary selection in antiviral genes in *C. elegans*.

Current studies investigating the IPR in *C. elegans* have focused on the reference genotype Bristol N2 (Panek *et al.* 2020; Reddy *et al.* 2017, 2019; Sowa *et al.* 2019). Here, we investigated genetic variation in the *pals*-gene family which is specifically expanded in *C. elegans* (humans and mice only contain a single *pals*-gene ortholog) (Leyva-Díaz *et al.* 2017). Expanded gene families often result from evolutionary selection, which strengthens the hypothesis that genetic variants in the *pals*-family could determine viral susceptibility (Thomas 2006). We set out to examine if the *pals*-genes experience selective pressure by analyzing the genetic variation in 330 wild strains in the CeNDR database. We found that only a few haplotypes occur worldwide for the *pals*-genes and that some are in regions that are of ancient origin. This indicates that different pools of *pals*-genes have been maintained in *C. elegans* populations: a hallmark of balancing selection. Genetic variation in the *pals*-gene family, and specifically in the IPR-regulators *pals*-22 and *pals*-25, may affect susceptibility to viral infection. This hypothesis is explored by infecting two well-studied strains, Bristol N2 and the Hawaiian strain CB4856, representing distinct *pals*-22/*pals*-25 haplotypes. Our data suggest that regulatory genetic variation in *pals*-22/*pals*-25 can determine (basal) IPR gene expression, illustrating that natural genetic variation in IPR genes may influence host-pathogen interactions in wild populations.



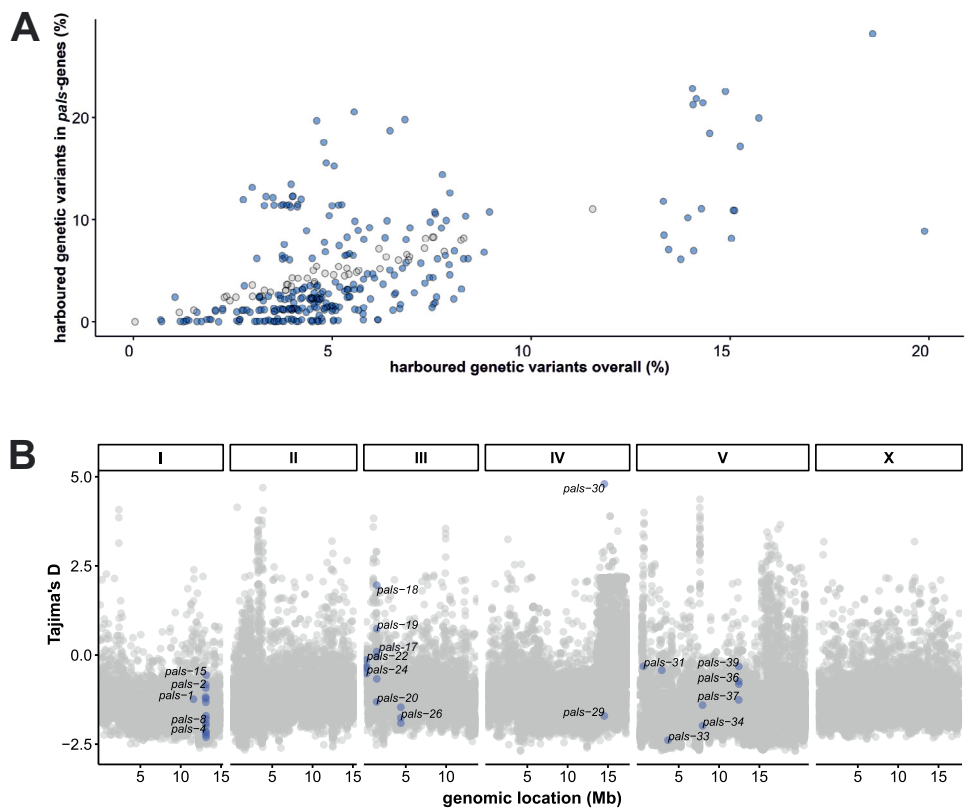
## Results

### *Global genetic variation in the pals-family is shaped by balancing selection*

The Intracellular Pathogen Response (IPR) counteracts viral infection in *Caenorhabditis elegans* and involves activity of 25 *pals*-genes (Reddy *et al.* 2017, 2019). To examine genetic variation in the *pals*-family genes, we investigated sequence information from the 330 wild strains from the CeNDR database (Cook *et al.* 2017). For each wild strain the genetic variation (compared to the reference strain N2) was summarized for genes in the *pals*-family and for all genes. For 48 wild strains genetic variation (defined by SNPs) in the *pals*-family was higher than expected compared to the overall genetic variation (chi-square test, FDR < 0.0001), but for 204 strains of the 330 analyzed strains the *pals*-gene family contained less variation than the overall genetic variation (chi-square test, FDR < 0.0001) (Figure 1A) (See Material and Methods for details). This indicated that while the *pals*-gene family is an expanded gene family, most wild strains contain relatively little genetic variation in the *pals*-genes compared to N2.

Populations from distinct geographical locations may encounter different selective pressures (Sivasundar & Hey 2005; Volkers *et al.* 2013). However, after mapping the amount of natural variation to the geographical location, no clear pattern could be found (Figure S1). Interestingly, some local strains show highly diverging levels of genetic diversity within the *pals*-family. For example, strain WN2002 was isolated in Wageningen (the Netherlands) and contains 3 times more genetic variation in the *pals*-family than the average of other genes. Strain WN2066 was isolated from the same compost heap as WN2002. Yet, compared to N2, WN2066 contains only 0.27% genetic variation in the *pals*-family set against an overall genetic variation of 2.67%. This shows local diversity for the locus that might be retained in the population due to differential microenvironmental pressures.

Next, we tested whether DNA sequence divergence was governed by genetic drift, or that selective forces were possibly acting on the *pals*-family. Although overall Tajima's D (TD) values in *C. elegans* populations are low as a result of overall low genetic diversity ( $TD_{\text{mean}} = -1.08$ ,  $TD_{\text{median}} = -1.12$ ) (Andersen *et al.* 2012), four *pals*-genes (*pals*-17, *pals*-18, *pals*-19, and *pals*-30) showed positive TD values that indicate either balancing selection acts on these genes or that there was a low frequency of rare alleles (Figure 1B). The most extreme example was *pals*-30 that had a TD value of 4.8: the highest value of all tested *C. elegans* genes. In total, 11 out of 39 *pals*-genes had values that fall within the 10% highest TD values for *C. elegans* ( $TD > -0.42$ ) and these top 10% genes included IPR regulators *pals*-22 and *pals*-25 (Table S1).

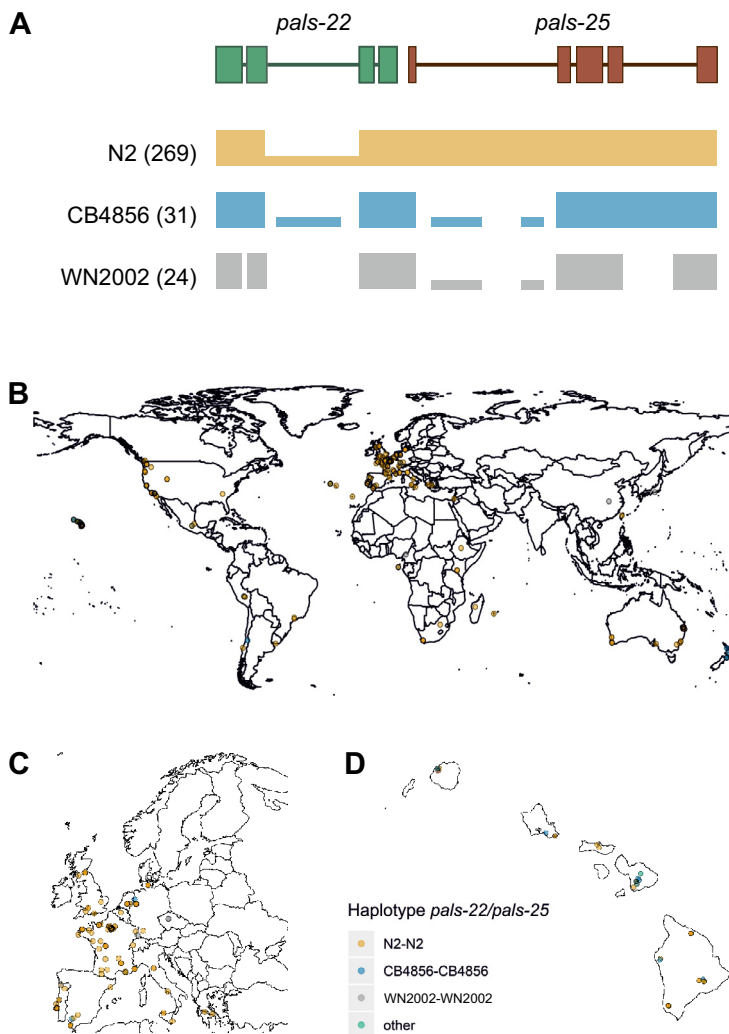


**Figure 1** Natural variation in the *C. elegans* *pals*-gene family worldwide – A) The percentage of genetic variants (defined by SNPs) in the *pals*-gene family compared to the overall natural variation for each of the 330 wild isotypes (Cook et al. 2017). The number of SNPs is relative to the reference strain N2. Blue dots indicate the amount of variation in the *pals*-genes is different than expected from the overall natural variation (Chi-square test, FDR < 0.0001). B) Tajima's D values per gene in the *C. elegans* genome calculated from sequence data of the 330 wild strains in the CeNDR database (Cook et al. 2017). Blue dots indicate Tajima's D values for *pals*-genes.

Subsequently, we delved into the genetic diversity for each *pals*-gene by investigating the number of variants per *pals*-gene (Figure S2, Table S1). Several *pals*-genes contained hardly any genetic variation and were therefore conserved on a worldwide scale. This conserved group contains the gene *pals-5* which acts downstream in the IPR (Reddy et al. 2017). Other *pals*-genes were highly variable, sometimes containing hundreds of polymorphisms (SNPs) in a single gene. Interestingly, for most genes in the diverse group, few alleles exist worldwide.

For example, three alleles were found for the gene *pals-25*: strains that harbor an N2-like allele, an allele containing ~30 polymorphisms (the well-studied Hawaiian strain CB4856 belongs to this group) or an allele containing ~95 polymorphisms (illustrated by the strain WN2002 from the Netherlands). In total, 19 out of 24 highly variable *pals*-genes show a clear grouping within 2 or 3 haplotypes suggesting that these haplotypes are actively maintained in the populations which supports that balancing selection could be acting on these genes (Figure S2, Table S1).

Manual inspection of the mapped reads in the 330 CeNDR strains showed evidence for extensive polymorphisms in the *pals-22/pals-25* locus that regulate the IPR transcriptional response (Reddy *et al.* 2017, 2019). In total, we found three major *pals-22/pals-25* haplotypes (N2-like, CB4856-like, and WN2002-like) that occur globally (Figure 2, Table S2). The highest local genetic diversity was found on the Hawaiian islands and Pacific region where *C. elegans* could originate from (Crombie *et al.* 2019). The genetic variation in the region where *pals-22* and *pals-25* are located is estimated to have diverged  $10^6$  generations ago (Thompson *et al.* 2015). Notably, *pals-22* and/or *pals-25* are both thought to have early stop codons in CB4856 and WN2002 that could change or disrupt their functioning (Cook *et al.* 2017; Leyva-Díaz *et al.* 2017). Yet, poor mapping to the reference genome in these highly diverse regions hampers the reliability of exact variant calling; in 24 wild strains most of the intron sequence was not covered by reads at all. The latter suggests that the genetic sequence in those strains is highly polymorphic. In conclusion, within the *pals*-gene family we found genes with either globally conserved or a few genetically distinct alleles. In particular, the *pals*-genes with a division into a few haplotypes show atypically high Tajima's D values when compared to other *C. elegans* genes. Together, our findings indicate that the *pals*-genes have been experiencing evolutionary pressure that resulted in long-term balancing selection of these genes.



**Figure 2 Worldwide haplotype diversity found for *pals-22* and *pals-25*** – A) Three distinct haplotypes were found at the *pals-22/pals-25* locus, here represented by an illustration of the read coverage at the locus. The most common haplotype (N2-like) is found in 269 stains and shows low coverage of the second intron of *pals-22*. The second common haplotype is CB4856-like and was found in 31 strains. For these strains coverage indicates strong structural variation in the introns of both genes, as well as larger insertions/deletions. Then, the WN2002-like haplotype was found in 28 strains and consists of very extensive structural variation at the locus, where almost the entire intron structure is not covered by reads. B) A geographical representation of the *pals-22/pals-25* haplotypes found worldwide. C) Zoomed in representation of Figure S2B of the strains collected in Europe. D) Zoomed in representation of Figure S2B of the strains collected on Hawaii.

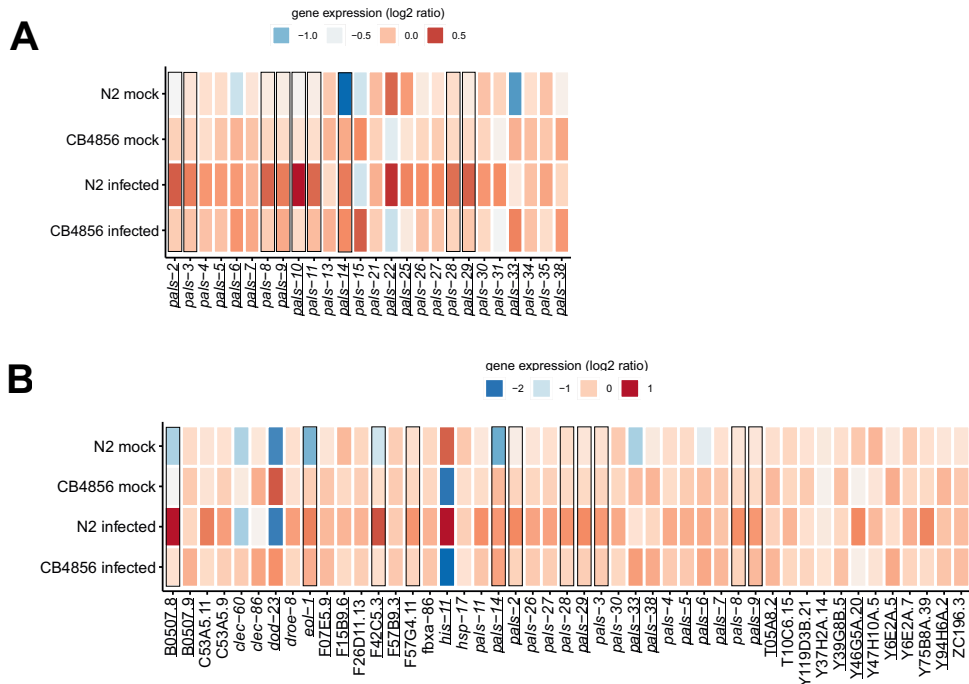
### *Basal pals-gene expression differs between N2 and CB4856*

Distinct *pals*-gene profiles may result in distinct IPR activity between wild isolates. To study whether the *pals*-22/*pals*-25 haplotypes (N2-like, 268 strains, CB4856-like, 31 strains, or WN2002-like, 24 strains) show differential gene expression for these two genes, we mined DNA and RNA hybridization data from (Volkers *et al.* 2013). We found that observed gene expression differences for WN2002 were mainly due to hybridization errors. However, for CB4856 hybridization was similar to N2 and shows *pals*-22 expression was significantly lower than in N2 (linear model,  $p < 0.001$ ) (Table S3). As the strains N2 and CB4856 are well-studied and their transcriptional responses to different types of stress have been described (Capra *et al.* 2008; Li *et al.* 2006; Snoek *et al.* 2017; Viñuela *et al.* 2012), we continued to characterize the IPR in these two strains. We measured their transcriptional profile under standard conditions and after exposure to the OrV. These nematodes were 56h old and received either a mock or OrV infection treatment when they were in the L2 stage.

First, we focused on transcriptional differences because of genetic variation between both wild types. Expression patterns of the full dataset were analyzed by means of a principal component analysis (PCA). Genotype explained the main difference in gene expression patterns (36.1%), which is in line with previous results, see for example (Capra *et al.* 2008; Li *et al.* 2006; Snoek *et al.* 2017; Volkers *et al.* 2013) (Figure S3). Among the 6383 genes (represented by 9379 spots) that were differentially expressed between N2 and CB4856 (under mock conditions) were 131 genes known to be involved in OrV infection (Table S4, S5A). These include twenty-three IPR genes of which twenty show higher expression in CB4856. Moreover, ten *pals*-genes showed higher basal expression in CB4856 compared to N2 (Figure 3A) (FDR < 0.05). For *pals*-22 and *pals*-25, expression is higher in N2 than in CB4856 (Figure 3A) (FDR < 0.05, Table S5A), which may determine gene expression of other IPR members (Figure 3B) (Reddy *et al.* 2019).

The strain CB4856 harbors much of the total genetic variation known to occur in nature for the genomic *pals*-clusters on chromosome III and V. Given that our experiment used microarrays, we accounted for hybridization errors and excluded 17 IPR genes and 11 *pals*-genes from the analysis based on probable hybridization differences (Table S6A-C). Hybridization was considered genuine when probes had at least 95% overlap between probe and genome sequence and similar hybridization in N2 and CB4856 (Table S6D). Furthermore, expression patterns measured by RT-qPCR confirmed the microarray data for *pals*-6, *pals*-14, and *pals*-22, although the slight difference in *pals*-25 expression on the microarrays was not replicated (Figure S4). Besides, lower expression of *pals*-22 in CB4856 compared to N2 was also observed in an RNA-seq study (Vu *et al.* 2015).





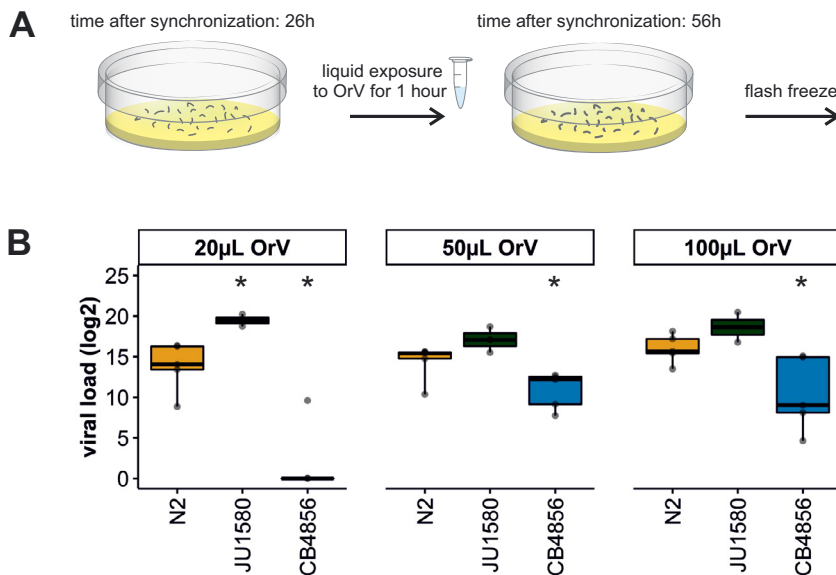
**Figure 3 Gene expression of IPR and pals-genes in *C. elegans* N2 and CB4856 under control and OrV-infected conditions** – A) Heat-map showing the  $\log_2$  intensities of pals-genes in N2 mock, N2 infected, CB4856 mock and CB4856 infected conditions B) Heatmap showing the expression of IPR genes in  $\log_2$  intensities in N2 mock, N2 infected, CB4856 mock and CB4856 infected conditions. Underlined genes showed significant (basal) expression differences based on genotype (FDR < 0.05), whereas squares indicated the genes where treatment- or the combination of treatment and genotype had a significant effect (FDR < 0.1) (Table S5).  $\log_2$  ratios are based on the average expression of the gene of interest in the overall dataset. Therefore, the  $\log_2$  ratios per experimental group indicate the deviation from the average value. Please note, a subset of the pals-genes, namely the pals-genes that are also IPR genes (defined in (Reddy et al. 2019)) are depicted twice. This allows for direct comparison to other pals-genes that do not become differentially expressed upon infection (like *pals-22* and *pals-25*) and to non-pals IPR genes.

Most of the genetically diverse *pals*-genes on chromosome III and V have previously been shown to display local regulation of gene expression (*cis*-quantitative trait locus; *cis*-eQTL) (Table S7). Moreover, at least 10 genes across different *pals*-clusters were regulated by genes elsewhere in the genome (*trans*-eQTL) (Table S7). The established IPR regulators *pals-22* and *pals-25* are likely candidates for this regulatory role. Most of the expression QTL were consistently found across multiple studies, environmental conditions and labs (Li et al. 2006, 2010; Rockman et al. 2010; Snoek et al. 2017; Sterken et al. 2014; Viñuela et al. 2010, 2012).

Therefore, we conclude that regulatory genetic variation of *pals*-genes could underlie observed differences in basal expression of IPR genes between the strains N2 and CB4856.

### *Stress-related phenotypes in the N2 and CB4856 strain*

Multiple (a)biotic stressors, including heat stress and viral exposure, activate the IPR pathway (Reddy *et al.* 2017). Differences in IPR and *pals*-gene activity between N2 and CB4856 may determine the phenotype of stressed nematodes which was investigated here. The response of CB4856 to the OrV has so far only been measured in one experiment. This research indicated that mixed-stage populations of N2 and CB4856 that are continuously exposed to OrV for four days have comparable viral susceptibility (Ashe *et al.* 2013). We could confirm these findings (Figure S5), but also found that young adult CB4856 nematodes collected for our transcriptional analysis were less susceptible than N2 nematodes (Figure 4). As a control, we took along the susceptible strain JU1580, which indeed was more susceptible than the strain N2 for both assays (Félix *et al.* 2011; Sterken *et al.* 2014) (Figure 4B, Figure S5).



**Figure 4 Viral susceptibility of *C. elegans* N2, CB4856 and JU1580 upon initial infection** – A) Strains were infected 26 hours post bleaching by liquid exposure to OrV for 1 hour. 30 hours after the exposure nematodes were collected and flash frozen before further processing. B) Viral loads (log<sub>2</sub>) as determined by RT-qPCR for the genotypes N2, CB4856 and JU1580 after exposure to 20, 50 or 100 µL OrV/500 µL infection solution (student t-test; \*  $p < 0.05$ ) (4 biological replicates containing roughly 400 nematodes per sample). Boxplots show the division of the data into four quantiles with the middle bar indicating the median.

We investigated the cause of difference in susceptibility in N2 and CB4856 nematodes 30h post infection. As earlier studies have shown that CB4856 nematodes feed at least as much as N2 nematodes (Balla *et al.* 2015; Wang *et al.* 2019), the initially lower susceptibility of CB4856 nematodes could not be explained by reduced food intake. Therefore, another explanation would be a different number of infected individuals or cells. Using Fluorescent in situ Hybridization (FISH) staining of OrV we could only visualize a few infected cells (infection detected in less than 1% of the animals), despite viral infection could be confirmed by RT-qPCR (Table S8). Contrary, almost half the nematodes showed fluorescence in a control sample of plate-infected JU1580 validating the sensitivity of the assay (Ashe *et al.* 2013; Frézal *et al.* 2019) (Table S8). Furthermore, we used FISH on plate-infected IPR reporter strains (*pals-5::GFP* and *F26F2.1::GFP* in an N2 background) and observed that most nematodes that showed OrV FISH signal also showed IPR gene expression (Figure S6). About 43% of the individual nematodes in the population showed IPR expression (for both reporter strains). But many individuals that showed IPR gene expression did not show OrV FISH signal (Figure S6). More specifically, 21% of the animals that showed *pals-5* expression did not show OrV stained cells (Table S8). For the *F26F2.1* reporter strain even 95% of the animals showed *F26F2.1* expression but no OrV stained cells (Table S8). Together, this suggests that the OrV does not reach high levels of infection in both N2 and CB4856 30h post infection, but low levels of OrV might already activate IPR members.

IPR genes also become differentially expressed upon prolonged heat stress providing thermotolerance in N2 (Panek *et al.* 2020; Reddy *et al.* 2017, 2019). Although the response of N2 and CB4856 to different temperatures has been well investigated, the role of IPR genes in heat stress has not been specifically compared between N2 and CB4856. Both strains have a different temperature response as CB4856 nematodes die more often when experiencing severe heat stress, yet surviving nematodes live longer than those under standard conditions (Jovic *et al.* 2019; Rodriguez *et al.* 2012). Additionally, CB4856 and N2 naturally prefer different environmental temperatures. Contrary to N2, where temperature preference can be conditioned, this preference is fixed in CB4856 (Anderson *et al.* 2007; Gaertner *et al.* 2012; Jurado *et al.* 2010).

### *The transcriptional response to IPR impulses in N2 and CB4856*

We proceeded by investigating the transcriptional IPR response to OrV infection and heat stress in N2 and CB4856. The transcriptional response of age-synchronized CB4856 and N2 to heat stress was previously described by (Jovic *et al.* 2017, 2019) and we re-analyzed the data with a focus on the *pals*-genes. The transcriptomes of heat shocked and recovering nematodes were compared to the control treatment. We found that IPR genes became differentially expressed after heat-shock and during recovery in both N2 and CB4856 (Figure S7B). The mean basal IPR

expression was higher in CB4856 than in N2 under control, heat shock and recovery conditions. Subsequently, we analyzed the transcriptomes of nematodes collected 30 hours post infection (Figure 3). As expected, based on the PCA relatively few genes did respond to the OrV in our experiment (Figure S3) (Table S5B). Gene expression analysis by a linear model showed that 27 genes (represented by 57 spots) were differentially expressed upon infection by OrV ( $FDR < 0.1$ ) (Figure S8A) and 18 genes (represented by 44 spots) were differentially expressed by a combination of both treatment and genotype ( $FDR < 0.1$ ) (Figure S8B). These two groups of genes were largely overlapping (Figure S8C) and most of these genes only respond to infection in the genotype N2 (Figure 3B). Many of the *pals*-gene family members became higher expressed after infection in N2, but not in the strain CB4856 (Figure 3A). Some of these difference between N2 and CB4856 are due to hybridization errors in CB4856 (Table SA-E), but also 11 genes with correctly aligned probes in CB4856 did not show a response (Figure 3A, B). In agreement with literature, the *pals*-genes in the cluster on chromosome III (*pals-17* until *pals-25*) were not differentially expressed after OrV infection (Leyva-Díaz *et al.* 2017; Reddy *et al.* 2017).

Compared to previous studies, which found 48 or 320 genes responding to OrV infection, fewer differentially expressed genes are found in this set-up which may be a result of shorter exposure to the OrV (1h in liquid here compared to constant exposure on the plate in other assays) (Chen *et al.* 2017; Sarkies *et al.* 2013). Because of the minor effect of OrV infection on transcriptional activity, a relaxed false discovery rate ( $FDR < 0.1$ ) was used to analyze the data. As all genes discovered using this threshold were IPR genes that are previously described by others, these were probably true positive hits (Chen *et al.* 2017; Sarkies *et al.* 2013).

Thus, N2 showed a weak, but expected IPR to OrV infection, however we did not detect increased expression of IPR genes in CB4856 nematodes. On the other hand, *pals*-genes could be activated by heat stress in CB4856. The lack of transcriptional response 30h post infection could result from early activation of IPR genes or the infection stress being too mild stress to trigger the IPR. We tested the first hypothesis by measuring IPR activity over a 30-hour time-course. This did not show evidence for an earlier IPR in CB4856 than N2 although we noticed that OrV responsive genes in CB4856 were more dynamic than in N2 (Figure S9A) (Table S5C). Because we measured higher viral loads for plate-infected CB4856 (Figure S5), we measured gene expression of *pals-6*, *pals-14*, *pals-22*, and *pals-25*, by RT-qPCR for these populations. We found that *pals-6* and *pals-14* were upregulated in CB4856, suggesting that higher pathogen pressure is sufficient to activate the IPR in this strain (Figure S9B).

## Discussion

### *IPR genes of the pals-family are under balancing selection*

We have studied the effect of genetic diversity in relation to natural viral infection of the nematode *C. elegans*. Our findings show that genetic variation in *C. elegans* affects the Intracellular Pathogen Response (IPR): a transcriptional response that counteracts pathogens by increased proteostasis and in which at least 27 *pals*-genes are involved (Reddy *et al.* 2017, 2019). The 39 members of the expanded *pals*-gene family are mostly conserved within the *C. elegans* species and the *pals*-genes for which divergent alleles do occur can be clustered into a few distinct haplotypes.

Population genetic analyses showed that IPR genes are experiencing selective pressure which could be a result of balancing selection, population bottlenecks or presence of rare genetic variants. We argue that balancing selection is the most likely cause for three reasons. First, *pals-22* and *pals-25* were experimentally validated to regulate the IPR and to balance growth and immunity (Reddy *et al.* 2019). Second, we observed that few major haplotypes occur for this gene-pair and most other *pals*-genes. Manual inspection of the *pals-22/pals-25* locus did not suggest presence of rare variants, rather the presence of a highly divergent region of ancient origin (Thompson *et al.* 2015). Third, presence of the *pals-22/pals-25* divergent region did not correlate with overall genetic variation, hence is unlikely to be the result of a bottleneck.

Besides *pals-22* and *pals-25*, multiple other *pals*-genes studied here show signs of balancing selection (high Tajima's D values) (Tajima 1989), in particular the genes on the first and second cluster on chromosome III (0.1 and 1.4Mb). In contrast, most of the genes in *C. elegans* show negative Tajima's D values due to a recent selective sweep affecting chromosome I, IV, V, and X. This greatly reduces the genetic variation within the species (Andersen *et al.* 2012). The *pals*-genes with relatively high Tajima's D values on chromosome III are located in a region that has diverged early in the natural history of *C. elegans* (Thompson *et al.* 2015). Despite this ancient divergence, few haplotypes occur for this region. Also *pals-22* and *pals-25* fall within this region and are therefore expected to generally show similar transcription patterns for the IPR pathway for most *C. elegans* strains. This is in line with our results that show only a minority of strains, including CB4856, carry distinct regulatory genetic variants.

Genetic variation in the *pals*-gene family regulates an evolutionary important transcriptional response to environmental stress, which include pathogens. Given the minor effect of OrV on fecundity (Ashe *et al.* 2013; Félix *et al.* 2011), it seems unlikely that the OrV is one of the pathogens that exert selection pressure underlying the balancing selection. However, immunity responses upon microsporidia and oomycete infection are also mediated by the



IPR (Bakowski *et al.* 2014; Osman *et al.* 2018; Reddy *et al.* 2017, 2019). As these pathogens are lethal (Zhang *et al.* 2016), we think that it is possible that these classes of pathogens underlie maintenance of different IPR haplotypes in natural populations. A recent example shows balancing selection in the plant genus *Capsella* also results in maintenance of ancestral genetic variation in immunity genes. The two *Capsella* species studied retained genetic variability at immunity loci, despite a recent population bottleneck and reproduction by selfing that together reduced overall genetic variation. Here, parallels can be drawn to *C. elegans*, a species that also mainly reproduces by selfing and has experienced loss of global genetic diversity as a result of a selective sweep (Andersen *et al.* 2012). Together, these studies show that within natural populations immunity-related genetic variation can be retained by balancing selection. Yet, also selective pressure of abiotic factors may result in balancing selection, for example in sunflower species that adapt to different environments (Todesco *et al.* 2020). As the IPR also responds to heat stress, abiotic factors may also contribute to balancing selection of the IPR.

### *Transcriptional activation of the IPR in two genetically diverse strains*

Although our data shows that the strain CB4856 has multiple IPR genes with higher basal expression than in the strain N2 it is unclear whether this is linked to its response to OrV infection. We only observe lower viral accumulation in CB4856 compared to N2 during the first 30 hours of infection. During this initial period of viral infection, we did not detect upregulation of IPR genes in CB4856 compared to its basal expression. Yet, after a longer period of viral exposure, CB4856 accumulates as much virus in the population as N2 and subsequently IPR genes were also upregulated in CB4856. Our results provide novel insights in natural genetic variation in *C. elegans* immunity genes and measure the transcriptional response upon Orsay virus in two strains with genetically distinct *pals* alleles.

CB4856 shows a high basal expression of multiple IPR genes probably due to regulatory genetic variation in the *pals*-genes. Longer exposure to OrV or heat-stress is sufficient to activate the IPR in both the N2 and CB4856 strain. Therefore, a higher threshold for viral accumulation or cellular stress may prevent IPR activation in CB4856. Transgene-based expression of OrV under promoters with different expression levels may help to uncover which level of OrV is necessary for IPR activation (Jiang *et al.* 2017; Sowa *et al.* 2019). We could also show that a difference in timing of the response was unlikely by performing a time-course experiment. Because Orsay virus infection does not infect all individuals within the population and its cellular tropism is limited to the intestinal cells (Félix *et al.* 2011; Franz *et al.* 2013), we cannot rule out there may be local changes in transcription in the intestinal cells or for certain individuals of CB4856. Transcriptional techniques, such as single-cell RNA-seq or TOMO-seq (Ebbing *et al.* 2018; Trapnell *et al.* 2017), that identify gene expression within infected cells could provide more details about the transcriptional response.

The role of the IPR in fitness upon OrV infection remains elusive. As findings in N2 show that there is a balance in spending energy on antiviral defense versus growth, it could be that laboratory selection on growth in the N2 strain has affected parts of the regulatory pathway thus trading-off constant high expression of IPR genes against traits beneficially under laboratory conditions (Reddy *et al.* 2019; Sterken *et al.* 2015). It is therefore not clear if the N2 strain that carries the worldwide most common allele, would be representative of the typical response of that allele. Still, having a plastic instead of a fixed IPR response may be a beneficial strategy in nature. In that case, the fixed genetic response of CB4856 could be explained by strong selection for stress resistance. The hypothesis that the IPR determines immunity and growth in diverse genetic backgrounds can be tested by comparing IPR gene expression, development, and pathogen susceptibility of additional wild-strains that have an ‘N2-like’, ‘CB4856-like’, or ‘WN2002-like’ haplotype for the IPR genes, in particular *pals-22* and *pals-25*.

#### *Are there alternative IPR strategies?*

Strains potentially harbor regulatory genetic variation tailored to specific environments. In a harsh environment constant activity of the IPR may be preferred over low expression. Finding out which environmental factor could explain the population genetic patterns within the *pals*-genes of the IPR will be challenging. The IPR pathway has been shown to respond to multiple environmental stressors including intestinal and epidermal pathogens, but also heat stress (Reddy *et al.* 2017, 2019). Despite the increasing amount of ecological data for both *C. elegans* (Cook *et al.* 2017) and its pathogens (Frézal *et al.* 2019; Richaud *et al.* 2018; Zhang *et al.* 2016), it is not yet sufficient to draw any firm conclusions whether co-occurrence of host and pathogen drives evolution within the *pals*-family. However, some evidence exists that host-pathogen interactions can affect the genotypic diversity at a population level. In Orsay (France), the location where Orsay virus is found, diversity in pathogen susceptibility potentially explains the maintenance of several minority genotypes. These minority genotypes are outcompeted in the absence of the intracellular pathogen *Nematocida parisii*, but perform better in the presence of the pathogen (Richaud *et al.* 2018). Perhaps this can also help explain our observations of divergent *pals22/pals-25* haplotypes in strains found at the same site. Moreover, experimental evolution experiments hold the potential to bridge the gap between the lab and the field by investigating if the presence of intracellular pathogens invokes any genetic and transcriptional changes within the *pals*-family (Gray & Cutter 2014; Teotonio *et al.* 2017).

The ‘slow’ transcriptional response in the CB4856 strain may have different biological reasons. As CB4856 populations showed lower viral loads, the lack of a transcriptional response could simply result from fewer infected nematodes or fewer infected cells per animal. Fluorescent staining of infected cells showed that only a few cells per nematode population (containing hundreds of individuals) could be visualized to be infected with the OrV, making a comparison between the two strains not trustworthy. Yet the number of plate-infected individuals that were stained was comparable to literature (Sarkies *et al.* 2013) and we could reliably detect the OrV by RT-qPCR in infected N2 and CB4856 populations. This indicates that the infection assay works properly and suggests that the viral levels per cell remained under the detection limit. Perhaps our data represents only the first cycle of viral replication as previous findings indicate that viral loads reached a plateau 30 hours post exposure with this method (Sterken *et al.* 2014). Later on, viral reproduction may increase further in line with observations that plate-infected N2 and CB4856 populations were equally infected. Thus, CB4856 nematodes may only be capable of limiting viral infection temporally.

This study provides insights into the natural context of the evolutionary conserved genetic and the plastic, transcriptional response after infection. We show that relatively little genetic diversity is found worldwide within clusters of *pals*-genes that regulate the IPR transcriptional response. In addition, the genetic diversity that exists is captured by only a few highly divergent haplotypes occurring worldwide. Therefore, we suggest that genes that function in the IPR transcriptional response could be under balancing selection, possibly from intracellular pathogens. Our results show that genetic variation within wild *C. elegans* strains could shape the basal expression of IPR genes and this may determine the transcriptional response after infection. Thus, this study provides new insights into the diversity of ways that hosts can develop both genetic and transcriptional responses to protect themselves from harmful infections.

## Methods

### *Nematode strains and culturing*

*C. elegans* strains N2 (Bristol) and CB4856 (Hawaii) were kept on 6-cm Nematode Growth Medium (NGM) dishes containing *Escherichia coli* strain OP50 as food source (Brenner 1974). The strains ERT54 (jyIs8[*pals-5p::GFP*, *myo-2p::mCherry*] X) and ERT71 (jyIs15[F26F2.1p::GFP; *myo-2::mCherry*]) were kind gifts from Emily Troemel (Bakowski *et al.* 2014; Reddy *et al.* 2017, 2019). Strains were kept in maintenance culture at 12°C and the standard growing temperature for experiments was 20°C. Fungal and bacterial infections were cleared by bleaching (Brenner 1974). The strains were cleared of males prior to the experiments by selecting L2 larvae and placing them individually in a well in a 12-wells plate at 20°C. Thereafter, the populations were screened for male offspring after 3 days and only the 100% hermaphrodite populations were transferred to fresh 9-cm NGM dishes containing *E. coli* OP50 and grown until starved.

### *Short-term Orsay virus infection assay (exposure in liquid)*

Orsay virus stocks were prepared according to the protocol described before (Félix *et al.* 2011). After bleaching, nematodes were infected using 20, 50, or 100µL Orsay virus/500µL infection solution as previously described (Sterken *et al.* 2014). Mock infections were performed by adding M9 buffer instead of Orsay virus stock (Brenner 1974). For each strain the maximum viral load was determined (4 biological replicates). The maximum viral load is the highest viral load that can be obtained for a strain and is reached when increasing amounts of virus do not significantly affect the viral load anymore (t-test,  $p > 0.05$ ). For N2 and JU1580 20µL of OrV sufficed to reach the maximum viral load. For CB4856 at least 50µL OrV was needed to maximize the viral load. Hence, using 50µL OrV/500µL infection solution the maximum viral load was obtained for all three strains which was therefore used in subsequent experiments (Figure 1B). Two virus stocks were used for these experiments: one for the first four biological replicates and one for the remaining four replicates.

The samples for the viral load and transcriptional analysis were infected in Eppendorf tubes with 50 µL Orsay virus/500µL infection solution 26 hours post bleaching (L2-stage) (8 biological replicates per treatment per genotype). The nematodes were collected 30 hours after infection. The samples for the transcriptional analysis of the time-series were infected with 50 µL Orsay virus/500µL infection solution at 40 hours post bleaching (L3-stage). These strains were infected in the L3 stage to obtain high RNA concentrations for microarray analysis also in the early samples. The nematodes were collected at time points: 1.5, 2, 3, 8, 10, 12, 20.5, 22, 24, 28, 30.5, or 32 hours post-infection (1 biological replicate per treatment

per genotype per time point). Viral loads of the samples were determined by RT-qPCR as described by (Sterken *et al.* 2014). A single Orsay virus stock was used for this experiment.

#### *Long-term Orsay virus infection assay (exposure on plate)*

Orsay virus stocks were obtained as described previously (Félix *et al.* 2011). The long-term exposure assay was based on previous experiments by (Ashe *et al.* 2013; Félix *et al.* 2011). Three young adult nematodes were placed on a plate with 20, 50, or 100µL Orsay virus that was added to the plate shortly before transfer. M9 was added to mock-treated plates instead of Orsay virus stock. Two days after incubation part of the population was transferred to a fresh plate to prevent starvation. Four days (96 hours) after placing the first nematodes, populations were collected for RNA isolation. Viral loads of the samples were determined by RT-qPCR as described by (Sterken *et al.* 2014). A single Orsay virus stock was used for this experiment.

#### *Fluorescent in situ hybridization (FISH) of infected nematodes*

Custom Stellaris FISH Probes were designed against OrV RNA1 by utilizing the Stellaris RNA FISH Probe Designer (Biosearch Technologies, Inc., Petaluma, CA) available online at [www.biosearchtech.com/stellarisdesigner](http://www.biosearchtech.com/stellarisdesigner). The mock-treated or infected nematodes were hybridized with the Stellaris RNA FISH Probe set labeled with CAL Fluor® Red 590 Dye (Biosearch Technologies, Inc.), following the manufacturer's instructions available online at [www.biosearchtech.com/stellarisprotocols](http://www.biosearchtech.com/stellarisprotocols) based on protocols by Raj *et al.* (Femino 1998; Raj & Tyagi 2010; Raj *et al.* 2008).

N2 and CB4856 populations were fixed 30h after infection (according to the short-term infection assay). Half of the nematodes were flash frozen to determine the viral load in the populations (Sterken *et al.* 2014) and the other half were used in the FISH procedure. Eight biological replicates were performed for this assay. The strains JU1580, ERT54 and ERT71 were mock-treated or OrV infected by chunking nematodes from a starved to a fresh plate and adding either 50µL M9 or OrV. These nematodes were fixed for FISH 48 hours post mock-treatment or infection. For the reporter strains (ERT54 and ERT71) three biological replicates were performed and JU1580 nematodes were infected once.

#### *RNA isolation*

The RNA of the samples in the transcriptional analysis (infected 26 hours post bleaching and collected 56 hours post bleaching) was isolated using Maxwell® 16 Tissue LEV Total RNA Purification Kit, Promega according to the manufacturer's instructions including two modifications. First, 10µL proteinase K was added to the samples (instead of 25µL). Second,



after the addition of proteinase K samples were incubated at 65°C for 10 minutes while shaking at 350 rpm. Quality and quantity of the RNA were measured using the NanoDrop-1000 spectrophotometer (Thermo Scientific, Wilmington DE, USA).

The RNA of the samples in the time series was isolated using the RNeasy Micro Kit from Qiagen (Hilden, Germany). The ‘Purification of Total RNA from Animal and Human Tissues’ protocol was followed, with a modified lysing procedure; frozen pellets were lysed in 150µl RLT buffer, 295µl RNase-free water, 800µg/ml proteinase K and 1% β-mercaptoethanol. The suspension was incubated at 55°C at 1000 rpm in a Thermomixer (Eppendorf, Hamburg, Germany) for 30 minutes or until the sample was clear. After this step the manufacturer’s protocol was followed. Quality and quantity of the RNA were measured using the NanoDrop-1000 spectrophotometer (Thermo Scientific, Wilmington DE, USA) and RNA integrity was determined by agarose gel electrophoresis (3µL of sample RNA on 1% agarose gel).

#### *cDNA synthesis, labelling and hybridization*

The ‘Two-Color Microarray-Based Gene Expression Analysis; Low Input Quick Amp Labeling’ -protocol, version 6.0 from Agilent (Agilent Technologies, Santa Clara, CA, USA) was followed, starting from step five. The *C. elegans* (V2) Gene Expression Microarray 4X44K slides, manufactured by Agilent were used.

#### *Data extraction and normalization*

The microarrays were scanned by an Agilent High Resolution C Scanner with the recommended settings. The data was extracted with Agilent Feature Extraction Software (version 10.7.1.1), following manufacturers’ guidelines. Normalization of the data was executed separately for the transcriptional response data (infected at 26 and collected at 56 hours post bleaching) and the transcriptional response of the time series. For normalization, R (version 3.3.1. x64) with the Limma package was used. The data was not background corrected before normalization (as recommended by (Zahurak *et al.* 2007)). Within-array normalization was done with the Loess method and between-array normalization was done with the Quantile method (Smyth & Speed 2003). The obtained single channel normalized intensities were  $\log_2$  transformed and the transcriptional response data (infected 26 hours post bleaching) was batch corrected for the two different virus stocks that were used for infection. The obtained (batch corrected)  $\log_2$  intensities were used for further analysis using the package ‘tidyverse’ (1.2.1) in R (3.3.1, x64).

The transcriptome datasets generated are deposited at ArrayExpress (E-MTAB-7573 and E-MTAB-7574). The data of the 12 N2 mock samples of the time series has previously been described (Snoek *et al.* 2015).

### Principal component analysis

A principal component analysis was conducted on the gene-expression data of the both the transcriptional response and the transcriptional response of the time series. For this

$$R_{i,j} = \log_2\left(\frac{y_{i,j}}{\bar{y}_i}\right)$$

purpose, the data was transformed to a  $\log_2$  ratio with the mean, using where  $R$  is the  $\log_2$  relative expression of spot  $i$  ( $i = 1, 2, \dots, 45220$ ) for sample  $j$ , and  $y$  is the intensity (not the  $\log_2$ -transformed intensity) of spot  $i$  for sample  $j$ . The principal component analyses were performed independently per experiment. The transformed data was used in a principal component analysis, where the first six axes that explain above 4.9% of the variance were further examined.

### Linear models

The  $\log_2$  intensity data of the nematodes that were mock-treated 26 hours post bleaching and collected 56 hours post bleaching was analyzed using the linear model

$$Y_i = G + \varepsilon$$

with  $Y$  being the  $\log_2$  normalized intensity of spot  $i$  ( $1, 2, \dots, 45220$ ).  $Y$  was explained over genotype ( $G$ ; either N2 or CB4856) and the error term  $\varepsilon$ . The significance threshold was determined by the *p.adjust* function, using the Benjamini & Hochberg correction ( $FDR < 0.05$ ) (Benjamini & Hochberg 1995). The analyzed dataset is part of the dataset containing mock-treated and OrV infected samples.

The  $\log_2$  intensity data of the nematodes that were either mock-treated or infected 26 hours post bleaching and collected 56 hours post bleaching was analyzed using the linear model

$$Y_i = G + T + G \times T + \varepsilon$$

with  $Y$  being the  $\log_2$  normalized intensity of spot  $i$  ( $1, 2, \dots, 45220$ ).  $Y$  was explained over genotype ( $G$ ; either N2 or CB4856), treatment ( $T$ , either infected or mock), the interaction between genotype and treatment and the error term  $\varepsilon$ . The significance threshold was determined by the *p.adjust* function, using the Benjamini & Hochberg correction ( $FDR < 0.1$  for  $T$  and  $G \times T$ ,  $FDR < 0.05$  for  $G$ ) (Benjamini & Hochberg 1995).

The  $\log_2$  intensity data for samples of the time series was analyzed using the linear model

$$Y_i = D + G + T + \varepsilon$$

with Y being the  $\log_2$  normalized intensity of spot i (1, 2, ..., 45220). Y was explained over development (D, time of isolation: 1.5, 2, 3, 8, 10, 12, 20.5, 22, 24, 28, 30.5, or 32 hours post-infection), genotype (G; either N2 or CB4856), treatment (T; either infected or mock) and the error term  $\varepsilon$ . The significance threshold was determined by the *p.adjust* function, using the Benjamini & Hochberg correction ( $FDR < 0.05$ ) (Benjamini & Hochberg 1995). For the samples in the timeseries a correlation coefficient (r) was obtained by calculating the slope of gene expression over time.

The  $\log_2$  intensity data of the nematodes that were exposed to heat shock (obtained from (Jovic *et al.* 2017)) was analyzed using the linear model

$$Y_i = G + T + G \times T + \varepsilon$$

with Y being the  $\log_2$  normalized intensity of spot i (1, 2, ..., 45220). Y was explained over genotype (G; either N2 or CB4856), treatment (T, either control, heat shock or recovery), the interaction between genotype and treatment and the error term  $\varepsilon$ . The significance threshold was determined by the *p.adjust* function, using the Benjamini & Hochberg correction ( $FDR < 0.05$ ) (Benjamini & Hochberg 1995).

### *Functional enrichment analysis*

Gene group enrichment analyses were performed using a hypergeometric test and several databases with annotations. The databases used were: the WS258 gene class annotations, the WS258 GO-annotation, anatomy terms, phenotypes, RNAi phenotypes, developmental stage expression, and disease related genes ([www.wormbase.org](http://www.wormbase.org)) (Lee *et al.* 2018; Stein *et al.* 2002); the MODENCODE release 32 transcription factor binding sites ([www.modencode.org](http://www.modencode.org)) (Gerstein *et al.* 2010), which were mapped to transcription start sites (as described by (Tepper *et al.* 2013)). Furthermore, a comparison with previously identified genes involved in OrV infection was made using a custom-made database (Table S4).

Enrichments were selected based on the following criteria: size of the category  $n > 3$ , size of the overlap  $n > 2$ . The overlap was tested using a hypergeometric test, of which the p-values were corrected for multiple testing using Bonferroni correction (as provided by *p.adjust* in R, 3.3.1, x64). Enrichments were calculated based on unique gene names, not on spots.

### Probe alignment

Probe sequences of the *pals*-genes and IPR-genes (*C. elegans* (V2) Gene Expression Microarray 4X44K slides, Agilent) were aligned to the genome sequence of CB4856 (PRJNA275000) using command-line BLAST, using blastn with standard settings (Blast command line application; v2.2.28) ([https://git.wur.nl/mark\\_sterken/Orsay\\_transcriptomics](https://git.wur.nl/mark_sterken/Orsay_transcriptomics)) (Altschup *et al.* 1990; Cook *et al.* 2017; Thompson *et al.* 2015). We also compared the *pals*-genes probes to a differential hybridization experiment, to see if differences in DNA sequence explain the mRNA-based hybridization differences. Therefore, we obtained data from Volkers *et al.*, 2013 (normalized data, E-MTAB-8126) and correlated the gene-expression differences with the hybridization differences (Volkers *et al.* 2013).

### Gene expression measurements by RT-qPCR

Gene expression measurements were performed on the cDNA of each of the 32 samples used in the N2 and CB4856 gene expression analysis of 30 hours of mock-treated or infection (8 biological replicates) and on the cDNA of N2 and CB4856 mock-treated or infected samples exposed to 50μL on plate (4 biological replicates). Gene expression was quantified by RT-qPCR using custom designed primers (*pals*-6 forward 5'-TGGGTTCTGGATCAAGCAAAT-3', *pals*-6 reverse 5'-TGTTCTAGAGCTGCCTGTCTCTG-3', *pals*-14 forward 5'-TCGGGAAAGCATCAATGAACTGC-3', *pals*-14 reverse 5'-TGTTGTGCCTCTCCTCTGCC-3', *pals*-22 forward 5'-TTTAAATCTTGAAAGTGACCGCTGGG-3', *pals*-22 reverse 5'-ACTCTCTGTTGTCTCTTGCAAAATT-3', *pals*-25 forward 5'-TGCAATCGAAGATTGGTGA-3', *pals*-25 reverse 5'-AAATTCTAACTTGCTCAGCATGGA-3') that overlap at least one exon-exon border to prevent unintended amplification of any remaining DNA. RT-qPCR was performed on the MyIQ using iQ SYBR Green Supermix (Biorad) and the recommended protocol. Gene expression in each sample was quantified for the gene of interest and two reference genes (Y37E3.8 and *rpl*-6) (Sterken *et al.* 2014) in duplo.

To determine the relative gene expression, we normalized the data as in (Sterken *et al.* 2014). In short, we normalized the *pals*-gene expression based on the two reference genes using

$$E = \frac{Q_G}{((Q_{rpl-6} / \bar{Q}_{rpl-6}) + (Q_{Y37E3.8} / \bar{Q}_{Y37E3.8}))}$$

where E is the normalized gene expression,  $Q_G$  is the expression of the gene of interest,  $Q_{rpl-6}$  is the expression of the reference gene *rpl*-6 and  $Q_{Y37E3.8}$  is the expression of the reference gene Y37E3.8.

### *Genetic variation analysis*

Genetic data on *C. elegans* wild strains were obtained from the CeDNR website (release 20180527) (Cook *et al.* 2017). The data was further processed using custom made scripts ([https://git.wur.nl/mark\\_sterken/Orsay\\_transcriptomics](https://git.wur.nl/mark_sterken/Orsay_transcriptomics)). In short, the number of polymorphisms in the *pals*-family within a strain was compared to the total number of natural polymorphisms found in that strain. The N2 strain was used as the reference strain. A chi-square test ( $FDR < 0.0001$ ) was used to determine whether strains showed less or more variation than expected within the *pals*-gene family compared the total natural variation observed. Next, we also manually inspected the *pals-22/pals-25* locus of each of the 330 isolates via the Variant Browser tool on the CeNDR website ([www.elegansvariation.org](http://www.elegansvariation.org)) (Cook *et al.* 2017). The *pals-22/pals-25* locus could be classified in three major groups based on structural variation observed in the bam-files.

The number of polymorphisms within the *pals*-gene family was further specified per gene. Tajima's D values were calculated per gene within the *C. elegans* genome using the PoPGenome package (Pfeifer *et al.* 2014). The number of polymorphisms within the *pals*-gene family were compared to the geographical origin of the strain obtained from the CeDNR database (Cook *et al.* 2017). The data were visualized using the packages 'maps' (3.3.0) and 'rworldmap' (1.3-6) using custom written scripts ([https://git.wur.nl/mark\\_sterken/Orsay\\_transcriptomics](https://git.wur.nl/mark_sterken/Orsay_transcriptomics)).

### *eQTL data analysis*

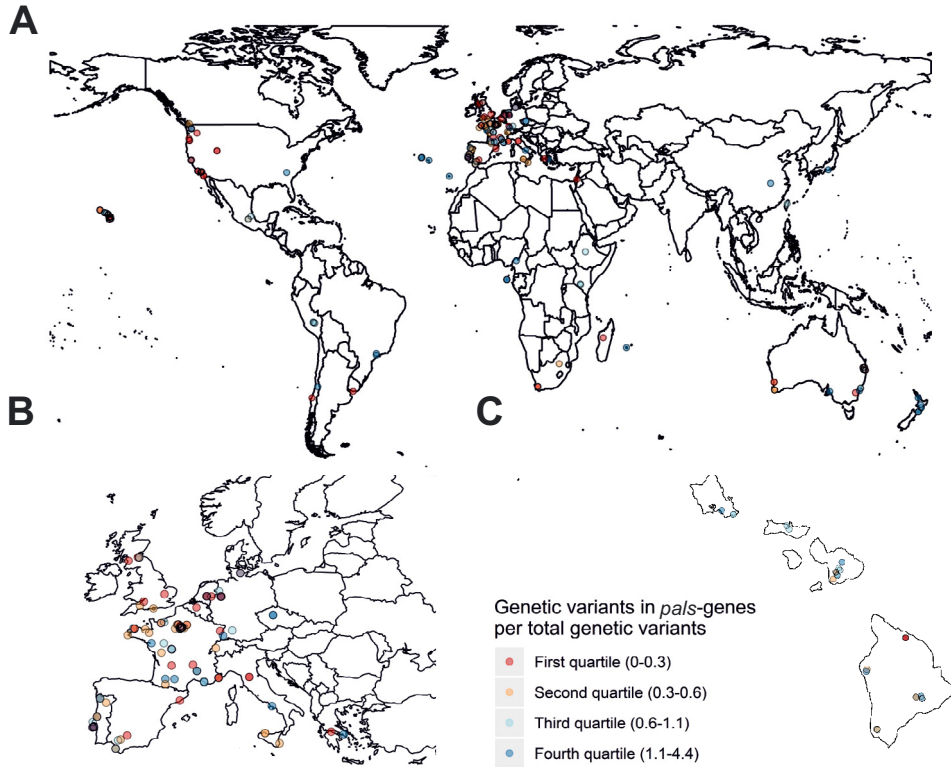
The eQTL data was mined from <http://www.bioinformatics.nl/EleQTL> (Snoek *et al.* 2020).

## **Acknowledgments**

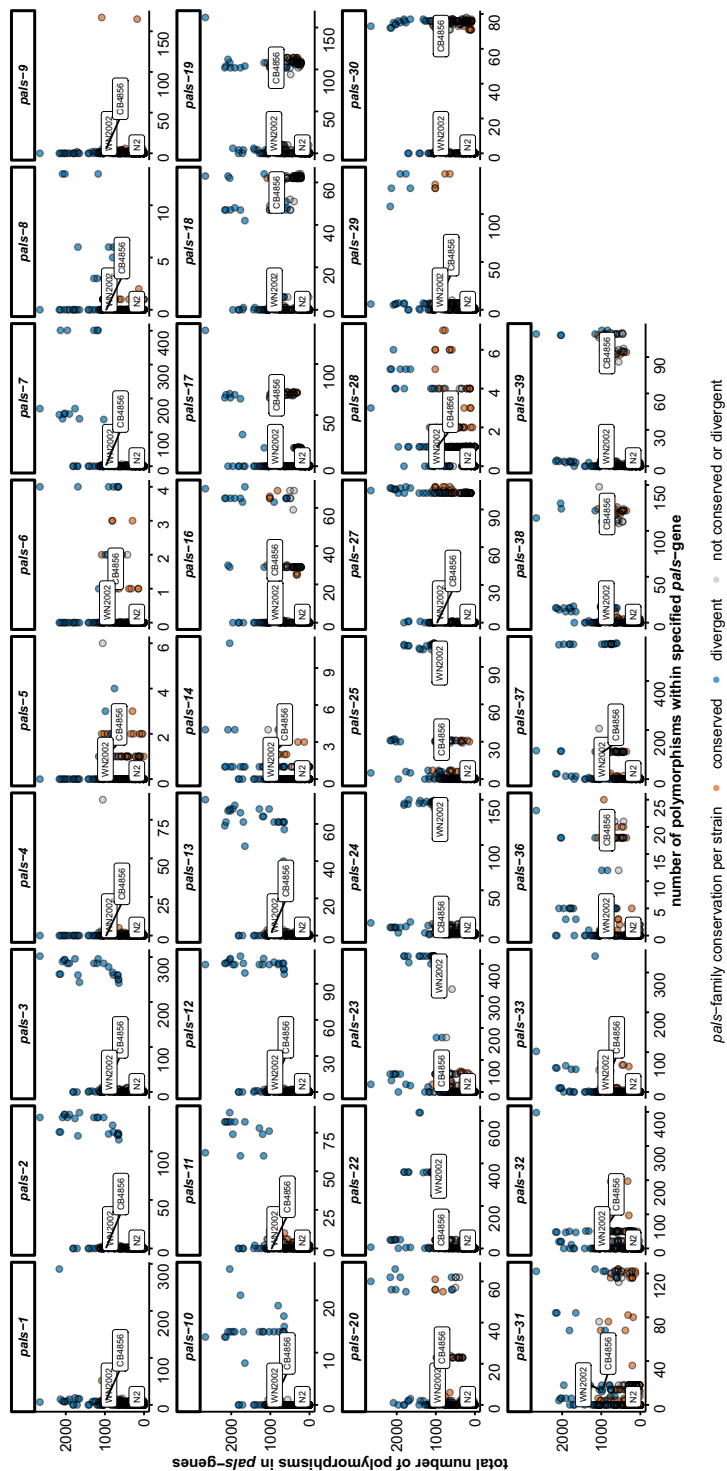
The authors want to thank Erik Andersen for hosting and sharing natural variation data on CeNDR and his advice on population genetic analyses. Moreover, we thank Emily Troemel for generously sharing the IPR reporter strains ERT54 and ERT71. The authors like to thank Sanne van Hamond, Sophie Vromans and Marèl Scholten for their RT-qPCR primer designs and efficiency tests. We thank WormBase for being a versatile and accessible resource.



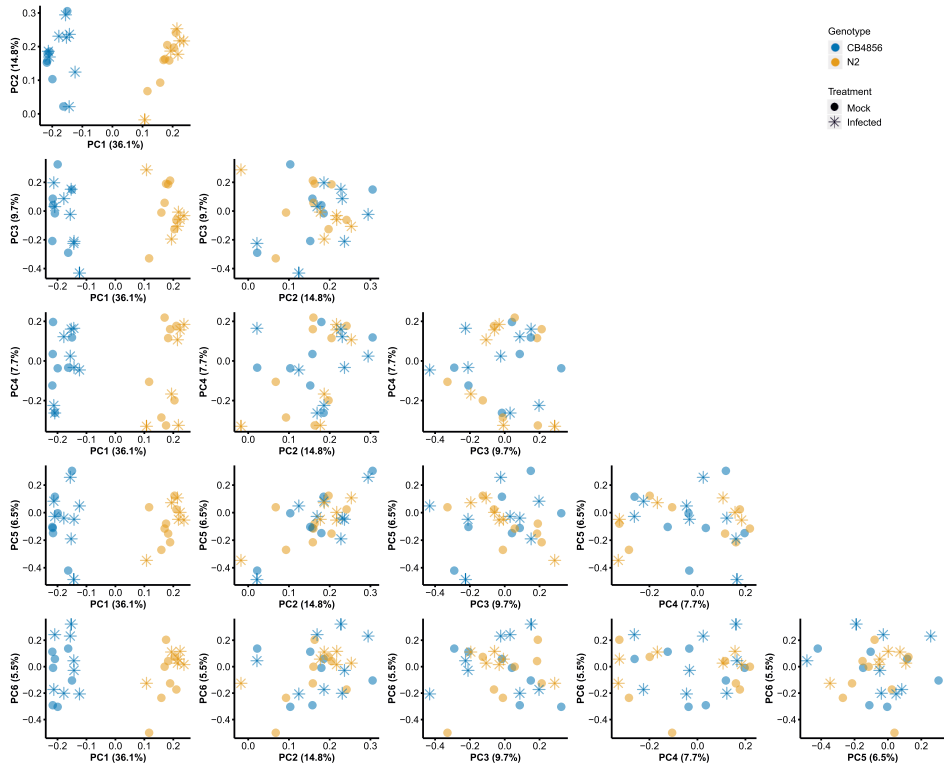
## Supplementary information



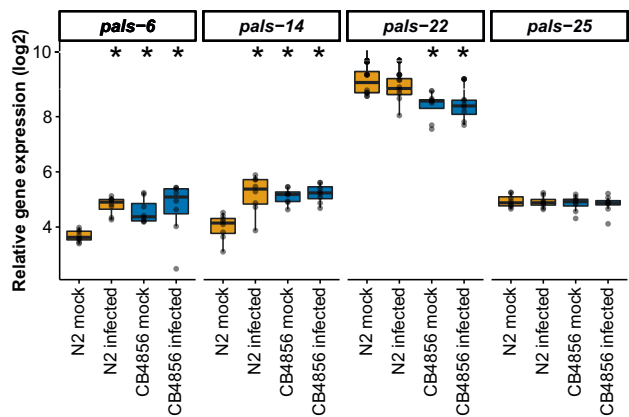
**Figure S1 Geographical distribution of natural variation within the *C. elegans* *pals*-gene family** – A) Location of CeNDR strains worldwide. The amount of natural variation (%) within the *pals*-pathway is indicated by the color of the dot. All natural strains have been grouped in a quantile (the first quartile exhibits least natural variation in the *pals*-pathway, the fourth exhibits most natural variation). B) Zoomed in representation of Figure S1A of the strains collected in Europe. C) Zoomed in representation of Figure S1A of the strains collected on Hawaii.



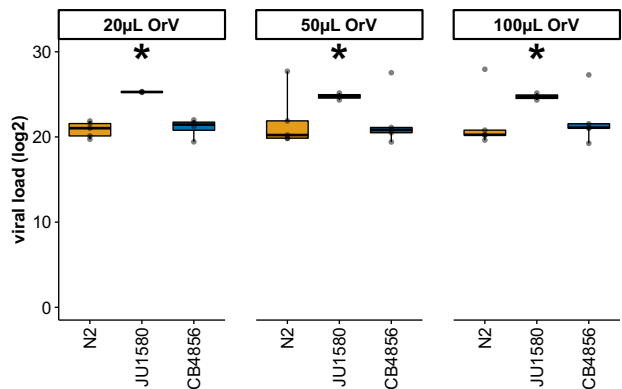
**Figure S2 Genetic variation per pals-gene per *C. elegans* strain** – The total number of SNPs within the pals-gene family is plotted against the number of known SNPs per pals-gene. Each dot represents a strain of the CeNDR database and the color of the strain indicates if the pals-family within this strain is depleted or enriched in polymorphisms as determined by the chi-square test (FDR < 0.0001).



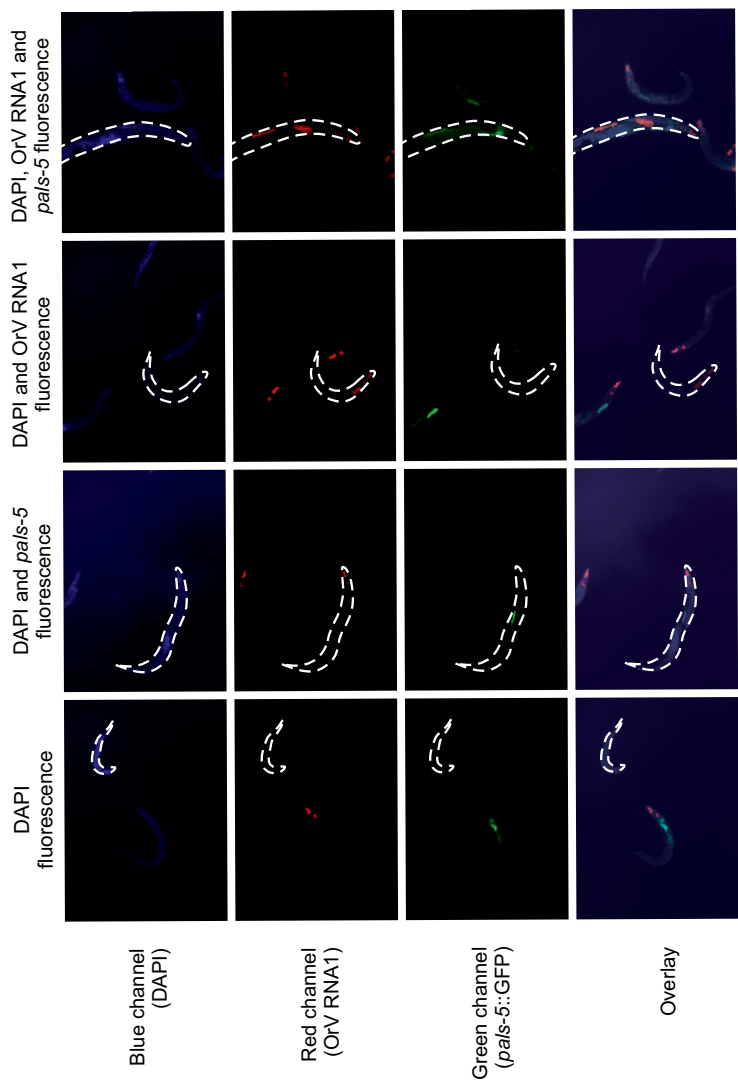
**Figure S3 Principal component analysis for gene expression in (un)infected *C. elegans* N2 and CB4856**  
 – Principal component analysis for the gene expression data obtained for the nematodes that were infected 26 hours post bleaching and collected 30 hours post infection. The six PC axes that explain at least 5% of the total variation are shown, numbered from 1 (explaining 36.1% of variance) to 6 (explaining 5.5% of variance). Each dot resembles the location of a sample on these axes. The genotype (N2 or CB4856) is indicated by color and the treatment (mock or OrV infection) is indicated by shape.



**Figure S4 Gene expression of *pals-6*, *pals-14*, *pals-22*, and *pals-25* in the 30 hours exposure assay determined via RT-qPCR** – Box-plots of relative gene expression patterns for *pals-6*, *pals-14*, *pals-22*, and *pals-25* which were determined using RT-qPCR. Measurements were performed on the same 32 samples that were used in the microarray analysis of (mock) infected N2 and CB4856. Each dot represents the expression within a sample. The expression of N2 mock was compared to N2 virus, CB4856 mock and CB4856 virus measurement per gene using a student t-test (\* $p < 0.05$ ).

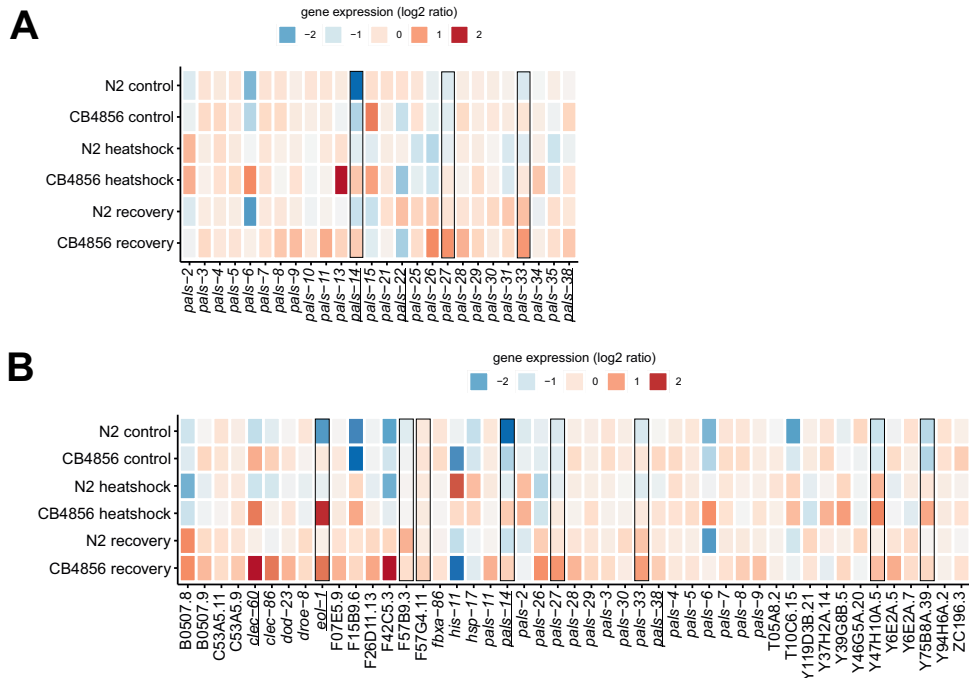


**Figure S5 Viral loads of *C. elegans* N2, CB4856, and JU1580 after 4 days of exposure** – Strains were infected with OrV by adding 20, 50, or 100µL OrV to an NGM plate with three young adults followed by 4 days of incubation (4 biological replicates). Viral susceptibility was determined by measuring viral loads (log<sub>2</sub>) by RT-qPCR (student t-test; \*  $p < 0.05$  compared to N2). Boxplots show the division of the data into four quantiles with the middle bar indicating the median.

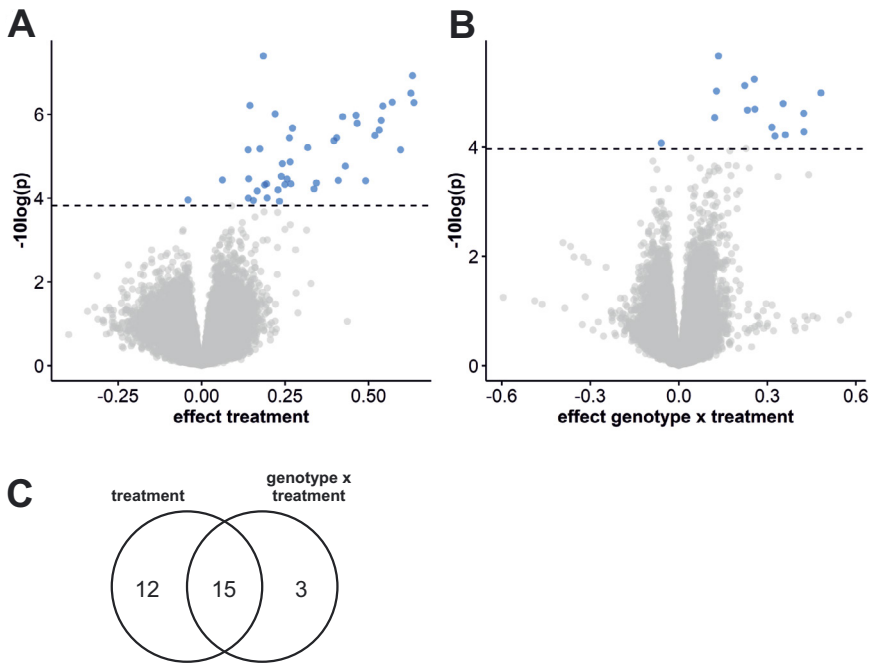


**Figure S6 Fluorescent in situ hybridization (FISH) of ERT54** – After OrV FISH staining plate-infected ERT54 nematodes, individuals expressing pals-5 and containing OrV infected cells were observed. These stainings could show in the same individual but were also observed separately. Different fluorescent channels show DAPI nuclear staining (blue), (intestinal) OrV RNA1 signal as stained by FISH (red), expression of the reporter pals-5::GFP (green). The exposure time in the red channel was fixed to 1s, for other channels the exposure times were set automatically. All nematodes show red fluorescence due to the myo-2::mCherry reporter in the pharynx.

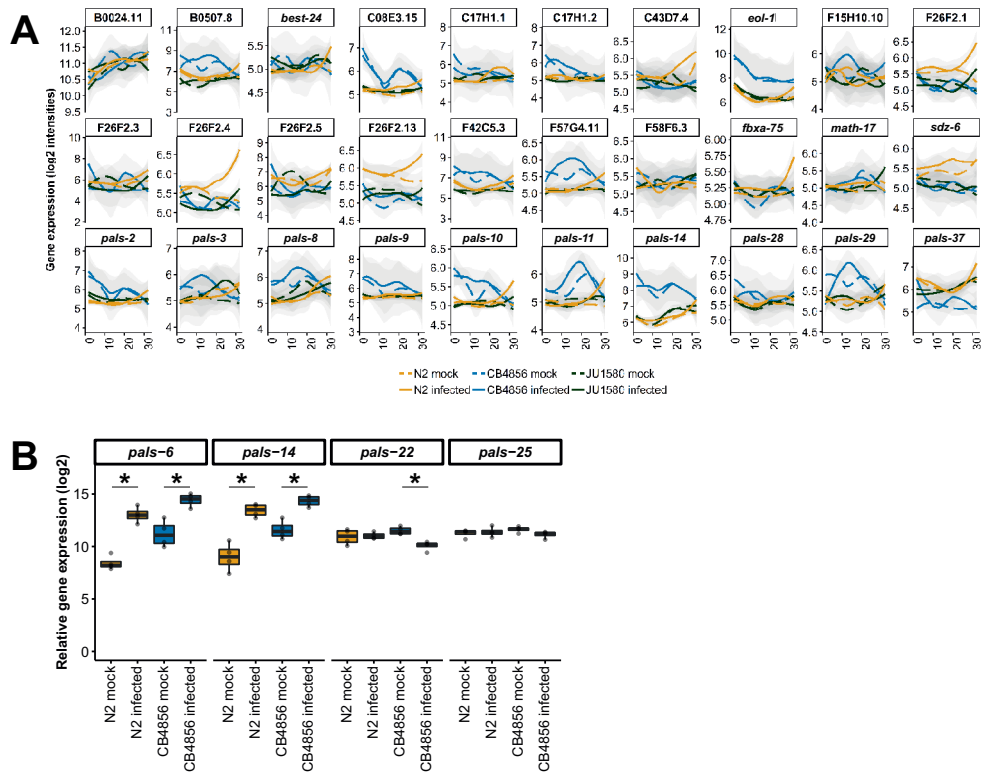




**Figure S7 Gene expression of pals-genes and IPR genes before and after heat-shock** – A) Heat-map showing the log<sub>2</sub> intensities of pals-genes in N2 control, N2 heat-shock, N2 recovery, CB4856 control, CB4856 heat-shock and CB4856 recovery conditions B) Heatmap showing the expression of IPR genes in log<sub>2</sub> intensities in N2 control, N2 heat-shock, N2 recovery, CB4856 control, CB4856 heat-shock and CB4856 recovery conditions. This dataset is described by (Jovic et al. 2019; Snoek et al. 2017). Underlined genes showed significant (basal) expression differences based on genotype (FDR < 0.05), whereas squares indicated the genes where treatment had a significant effect (FDR < 0.1). Log<sub>2</sub> ratios are based on the average expression of the gene of interest in the overall dataset. Therefore, the log<sub>2</sub> ratios per experimental group indicate the deviation from the average value. A subset of the pals-genes, namely the pals-genes that are also IPR genes (defined in (Reddy et al. 2019)) is depicted twice to facilitate direct comparison to non-IPR pals-genes and non-pals IPR genes.



**Figure S8 Differentially expressed genes upon OrV infection** – A) Volcano plot showing the effect of treatment (OrV infection) on global gene expression patterns. Microarray spots (blue) above the FDR threshold (dotted line) were considered differentially expressed. B) Volcano plot showing the effect of the interaction between treatment (OrV infection) and genotype (N2 or CB4856) on global gene expression patterns. Microarray spots (blue) above the FDR threshold (dotted line) were considered differentially expressed. C) Venn diagram indicating the number of genes that are differentially expressed and associated with an effect treatment (OrV infection) and/or genotype x treatment.



**Figure S9 Analysis of *C. elegans* expression patterns for OrV-response genes over time and long-term pathogenic exposure** – A) Gene expression patterns for N2 mock, N2 infected, CB4856 mock, CB4856 infected, JU1580 mock, and JU1580 infected nematodes of the 30 genes responding to OrV infection that were found in dataset of infected N2 and CB4856 nematodes (infected at 26 and collected at 56 hours post bleaching). The lines represent the fit through the data (by loess) and the grey area around a curve represents the 95% confidence interval, genotypes are indicated by colors and mock infection is indicated by dashed lines. B) Box-plots of relative gene expression patterns for pals-6, pals-14, pals-22, and pals-25 measured with RT-qPCR after exposure to OrV. Measurements were performed on the same 4 N2 and CB4856 samples that were obtained by continuous exposure to either mock-conditions or 50μL OrV. Each dot represents the expression within a sample. The expression of N2 mock was compared to N2 virus and the expression of CB4856 mock was compared to the CB4856 virus measurement per gene using a student *t*-test (\**p* < 0.05).

The full supplementary datasets can be found at <https://www.biorxiv.org/content/10.1101/579151v5> and <https://tinyurl.com/SL-PhD-thesis>. Here, the first four rows of the tables are shown to illustrate the contents.

**Table S1 Population genetic properties per *pals-gene*** – Conservation (genes are considered conserved when the vast majority of CeNDR strains contains < 25 polymorphisms), haplotype number and Tajima's D value per *pals-gene*. Genes that are upregulated upon infection in (Chen et al. 2017) are indicated.

Public name	Sequence name	Chromosome	Basepair position	Conserved	Haplotypes	Upregulated upon infection	Tajima D
pals-1	F15D3.8	I	11579848	yes	1	yes	-1.24
pals-2	C17H1.3	I	13104242	no	2	yes	-0.92
pals-3	C17H1.4	I	13110405	no	2	yes	-1.23

**Table S2 Haplotypes of *pals-22* and *pals-25* per CeNDR strain** – Haplotypes were manually inspected based for sequence coverage of *pals-22* and *pals-25* in the CeNDR database. The haplotype groups were noted for both genes per strain.

Strain	<i>pals-22</i>	<i>pals-25</i>	Haplotype
CB4856	CB4856	CB4856	CB4856-CB4856
CX11276	CB4856	CB4856	CB4856-CB4856
ECA191	CB4856	CB4856	CB4856-CB4856

**Table S3 Hybridization data from Volkers et al., 2013** - N2, WN2002, and CB4856 DNA (A) and RNA hybridization (B) to pals-22 and pals-25. The log<sub>2</sub> intensities for the probes are reported. Note that the samples were different for the two experiments. The full dataset was previously published in (Volkers et al. 2013) and can be found online (ArrayExpress E-MTAB-8126).

SpotID	Public name	Strain	Sample	log2 intensities
AGIWUR36367	pals-25	CB4856	12	13.64
AGIWUR36367	pals-25	N2	35	13.75
AGIWUR36367	pals-25	WN2002	34	13.24

SpotID	Public name	Strain	Sample	Batch	log2 intensities
AGIWUR36367	pals-25	CB4856	37	1	8.68
AGIWUR36367	pals-25	CB4856	61	2	8.50
AGIWUR36367	pals-25	N2	26	1	8.26

**Table S4 Genes involved during OrV infection** – Overview of *C. elegans* genes involved in OrV infection described in literature. Data is obtained from (Ashe et al. 2013; Bakowski et al. 2014; Chen et al. 2017; Félix et al. 2011; Frézal et al. 2019; Jiang et al. 2017; Le Pen et al. 2018; Reddy et al. 2017, 2019; Sandoval et al. 2019; Sarkies et al. 2013; Sowa et al. 2019; Sterken et al. 2014; Tanguy et al. 2017).

Sequence name	Public name	Paper	Type of evidence
F26F2.3		Chen BMC genomics 2017	RNA-seq in N2
F26F2.4		Chen BMC genomics 2017	RNA-seq in N2
F42C5.3		Chen BMC genomics 2017	RNA-seq in N2



**Table S5 Output linear models** – A) The linear correlation between genotype, treatment, the interaction between genotype and treatment and gene expression for the N2 and CB4856 30 hours post (mock) infection. B) The linear correlation between viral load and gene expression within infected samples per genotype for the N2 and CB4856 30 hours post (mock) infection. C) The linear correlation between development, genotype, treatment and gene expression for N2, JU1580 and CB4856 infected samples that were isolated 1.5–30 hours post infection.

SpotID	Gene	WBID	Term	Significance	Significance (adjusted)	Effect	Significance (FDR)
1			genotype	0.73	0.14	0.038	0.86
2			genotype	0.79	0.10	0.0065	0.89
3			genotype	0.21	0.69	0.029	0.42

SpotID	Gene	WBID	Term	Significance	Significance (adjusted)	Effect	Significance (FDR)
1			genotype	0.906434	0.043	0.0080	0.95
2			genotype	0.782822	0.17	-0.0067	0.88
3			genotype	0.52702	0.28	0.014	0.69

SpotID	Gene	WBID	Term	Significance	Significance (adjusted)	Effect	Significance (FDR)
1			genotype	0.205048	0.69	0.078	0.57
2			genotype	0.517155	0.29	-0.013	0.83
3			genotype	0.79833	0.098	-0.0044	0.94

**Table S6 Microarray probe binding to the CB4856 genome** – A) BLAST alignment of *C. elegans* CB4856 sequence (PJRNA275000) to microarray probes (*C. elegans* (V2) Gene Expression Microarray 4X44K slides) that detect the pals-genes. Probes are considered to align correctly when the there is more than 95% overlap on the correct chromosome. B) N2 and CB4856 DNA hybridization to the microarray probes genes that are differentially expressed based on genotype, treatment or their interaction. A ratio of 0 indicates there is absolutely no difference between N2 and CB4856 hybridization. The full dataset was previously published in (Volkers et al. 2013) and can be found online (ArrayExpress E-MTAB-8126). C) BLAST alignment of *C. elegans* CB4856 sequence (PJRNA275000) to microarray probes (*C. elegans* (V2) Gene Expression Microarray 4X44K slides) detect the IPR genes. Probes are considered to align correctly when the there is more than 95% overlap on the correct chromosome. D) BLAST alignment of *C. elegans* CB4856 sequence (PJRNA275000) to microarray probes (*C. elegans* (V2) Gene Expression Microarray 4X44K slides) that detect the differentially expressed genes from the linear model investigating the terms genotype, treatment and genotype x treatment. Probes are considered to align correctly when the there is more than 95% overlap on the correct chromosome. E) N2 and CB4856 DNA hybridization to the microarray probes genes that are differentially expressed based on genotype, treatment or their interaction. A ratio of 0 indicates there is absolutely no difference between N2 and CB4856 hybridization. The full dataset was previously published in (Volkers et al. 2013) and can be found online (ArrayExpress E-MTAB-8126).

SpotID	Term	Chromosome	Gene basepair start	Gene basepair end	Sequence name	Sequence	Chromosome blast	Chromosome correct	Percentage aligned in CB4856	Alignment correct	Probe correct
AGIWUR13	geno-type	IV	1545296	1566192	C34H4.5	CTTCCGGAGGAAAA- TCCTGCAAACAATTT- GTATATTTTGTGCGT- GTTATTTTCTCATTCC	IV	yes	100	yes	yes
AGIWUR17	geno-type					CCTAAATTTCTGATT- TTCAGAGTTTGAGAC- CGTTTCGATTCAAAC- CCCCACCGAACCCAA	V	no	100	yes	no
AGIWUR19	geno-type	V	18142849	18143543	Y37H2A.13	GGAATGATTGTTCG- CACAGAGGACATCAA- CTTGTCGATGCTTCA- TAGTTCTTTAATCTTG	V	yes	98.31	yes	yes

SpotID	Term	Chromosome	Gene basepair start	Gene basepair end	Sequence name	WBID	Public name	Sequence	Differential hybridization N2 and CB4856
AGIWUR13	genotype	IV	1545296	1566192	C34H4.5	WBGene00016427	C34H4.5	CTTCCGGAGGAAAATCCTGCAAA- CAATTTGTATATTTGTGCGTGT- TATTTTCTCATTCC	-0.037
AGIWUR17	genotype							CCTAAATTTCTGATTTTCAGAGTTT- GAGACCGTTTCGATTCAAACCCCCAC- CGAACCCAA	-0.54
AGIWUR19	genotype	V	18142849	18143543	Y37H2A.13	WBGene00045417	Y37H2A.13	GGAATGATTGTTTCGCACA- GAGGACATCAACTTGTTCGAT- GCTTCATAGTTCTTTAATCTTG	0.66

SpotID	Chromo- some	Gene basepair start	Gene basepair end	Sequence name	WBID	Public name	Sequence	Chro- mosome blast	Chro- mosome correct	Percentage aligned in CB4856	Align- ment correct	Probe correct
AGI-WUR10130	I	12109096	12112505	M01G12.9	WB- Gene00010822	M01G12.9	AATTAACGAAAAAGAATA- AAGGGCAACACGTTTCTTGAA- GAGTTCACGAAGCTTGTTCCA	I	yes	28.33	no	no
AGI-WUR1075	II	1619894	1621261	C08E3.10	WB- Gene00015602	fbxa-158	TTGGTCATGCTGAGATTCGCTCTA- AGGATATGAACATCAGCGAACT- GATGAGAGTCTTCG	II	yes	98.33	yes	yes
AGI-WUR10921	IV	14539762	14542602	T27E7.6	WB- Gene00012091	pals-29	TTTCAAGAAAATCAAATATTAT- GCTGACATGGCGAGCAAGTTGAA- CAACCAATTCTCTCA	IV	yes	100	yes	yes

SpotID	Chromo- some	Gene basepair start	Gene basepair end	Sequence name	WBID	Public name	Sequence	Percentage aligned in CB4856
AGI-WUR20103	I	11579861	11581986	F15D3.8	WB- Gene00008858	pals-1	ATTTTCATAGGCTTCGGTTGAAGGCTCCGGAGCGTGATATTCCTTCTTCGTC- TACTCTAA	0
AGI-WUR5622	I	13104253	13106298	C17H1.3	WB- Gene00007656	pals-2	ACGATGTTCGACTGAATTTTGTGAGCTGGAAGCAACTGGAAATTGTAGAAGCGGT- GGTTG	100
AGI-WUR17946	I	13104253	13106298	C17H1.3	WB- Gene00007656	pals-2	TGTTCGACTGAATTTTGTGAGCTGGAAGCAACTGGAAATTGTAGAAGCGGTGGTT- GAACA	100

SpotID	Chromo- some	Gene basepair start	Gene basepair end	Sequence name	WBID	Public name	Sequence	Differential hybridization N2 and CB4856
AGI-WUR20103	I	11579861	11581986	F15D3.8	WB- Gene00008858	pals-1	ATTTTCATAGGCTTCGGTTGAAGGCTCCGGAGCGTGATATTCCTTCTTCGTC- TACTCTAA	-0.28
AGI-WUR5622	I	13104253	13106298	C17H1.3	WB- Gene00007656	pals-2	ACGATGTTCGACTGAATTTTGTGAGCTGGAAGCAACTGGAAATTGTAGAAGCGGT- GGTTG	-0.23
AGI-WUR17946	I	13104253	13106298	C17H1.3	WB- Gene00007656	pals-2	TGTTCGACTGAATTTTGTGAGCTGGAAGCAACTGGAAATTGTAGAAGCGGTGGTT- GAACA	0.094


**Table S7 eQTL found for the pals-genes** – Summary of cis- and trans-eQTL found for the pals-genes in previous eQTL studies, data obtained from (Snoek et al. 2020).

Study	Trait	QTL chromo-some	QTL basepair	QTL marker	QTL signifi-cance	QTL ef-fect	QTL peak	QTL marker left		QTL basepair left	QTL marker right	QTL basepair right	Gene chromo-some	Gene basepair	Sequence name	Public name	QTL type	QTL R2
Li_2006_16C	12994	III	600003	234	6.08	0.335	234	233		10000	233	600003	III	132418	C29F9.1	pals-22	cis	0.32
Li_2006_16C	20738	III	600003	234	6.35	1.096	234	233		10000	233	600003	III	120492	C29F9.4	pals-24	cis	0.33
Li_2006_16C	22845	III	600003	234	8.06	0.771	234	233		10000	233	600003	III	124778	C29F9.3	pals-23	cis	0.41

**Table S8 Fluorescent signals observed in infected nematodes** – Summary of the phenotypes that were observed after OrV FISH staining in the strains N2, CB4856, ERT54, ERT71, and JU1580. The total number of inspected nematodes is shown and the number of nematodes showing either a green fluorescent signal (IPR reporter gene + GFP), red fluorescent signal (OrV RNA1.1) or both signals.

Strain	Assay	Biological replicates	Total amount of nematodes	Red fluorescent (OrV RNA1.1) nematodes		Green fluorescent (pals-5) nematodes	Green fluorescent (F26F2.1) nematodes	Nematodes showing both green and red fluorescence	IPR signal (%)	Infection signal (%)
N2	Liquid exposure OrV	8	2172	10		0	0	0		0.5
CB4856	Liquid exposure OrV	8	2235	4		0	0	0		0.2
ERT054	Plate exposure OrV	3	183	20		79	0	13	43.2	10.9





## Chapter 4

### Punctuated loci on chromosome IV determine natural variation in Orsay virus susceptibility of *Caenorhabditis elegans* strains Bristol N2 and Hawaiian CB4856

Mark G. Sterken<sup>1,2,4</sup>, Lisa van Sluijs<sup>1,2,4</sup>, Yiru Wang<sup>1</sup>, Wannisa Ritmahan<sup>1</sup>,  
Mitra Gultom<sup>1</sup>, Joost A.G. Riksen<sup>1</sup>, Rita J.M. Volkers<sup>1</sup>,  
L. Basten Snoek<sup>3</sup>, Gorben P. Pijlman<sup>2</sup>, Jan E. Kammenga<sup>1</sup>

<sup>1</sup>Laboratory of Nematology, Wageningen University and Research, the Netherlands

<sup>2</sup>Laboratory of Virology, Wageningen University and Research, the Netherlands

<sup>3</sup>Theoretical Biology and Bioinformatics, Utrecht University, the Netherlands

<sup>4</sup>These authors contributed equally



## Abstract

Host-pathogen interactions play a major role in evolutionary selection and shape natural genetic variation. Viral infection of *C. elegans* has shown that the genetically distinct strains, N2 and CB4856, are differentially susceptible to the Orsay virus (OrV). Here we report the dissection of the genetic architecture of susceptibility to OrV infection. We compare OrV infection in the relatively resistant wild type Hawaiian CB4856 strain to the more susceptible canonical Bristol N2 strain. To gain insight into the genetic architecture of viral susceptibility, 52 fully sequenced recombinant inbred lines (CB4856 x N2 RILs) were exposed to OrV. This led to the identification of two loci on chromosome IV associated with OrV resistance. To verify the two loci and gain additional insight into the genetic architecture controlling virus infection, introgression lines (ILs) that together cover chromosome IV, were exposed to OrV. Of the 27 ILs used, 17 had an CB4856 introgression in an N2 background and 10 had an N2 introgression in a CB4856 background. Infection of the ILs confirmed and fine mapped the locus underlying variation in OrV susceptibility and we found that a single nucleotide polymorphism in *cul-6* contributes to the difference in OrV susceptibility between N2 and CB4856. An allele swap experiment showed the strain CB4856 became more susceptible by having an N2 *cul-6* allele, although having the CB4856 *cul-6* allele did not increase resistance in N2. Additionally, we found that multiple strains with non-overlapping introgressions showed a distinct infection phenotype than the parental strain, indicating that there are punctuated locations on chromosome IV determining OrV susceptibility. Thus, our findings reveal the genetic complexity of OrV susceptibility in *C. elegans* and suggest that viral susceptibility is governed by multiple genes.



## Introduction

Genetic variation plays a major role in the arms race between pathogen and host (Enard *et al.* 2016; Obbard *et al.* 2006; Vasseur *et al.* 2011). The interaction between host genetic background and pathogen can shape natural variation by imposing a strong selection regime on the affected population. Host genetic variation plays a role in ongoing viral outbreaks as illustrated by studies that correlate outcome of infection with Hepatitis, HIV, Zika, Ebola and SARS-CoV-2 to the host's genetic background (Dean *et al.* 1996; Heim *et al.* 2016; Hou *et al.* 2020; Nguyen *et al.* 2020; Rasmussen *et al.* 2014; Yun *et al.* 2018). Studying host-virus interactions in model systems can uncover genetic networks determining viral susceptibility (Chapter 2).

The nematode *Caenorhabditis elegans* encounters a variety of pathogens in its natural habitat, including bacteria, microsporidia, oomycetes, and fungi (Schulenburg & Félix 2017). So far, only one virus has been discovered that naturally infects *C. elegans*: the Orsay virus (OrV) (Félix & Wang 2019). In the laboratory this pathogen can be easily maintained and used to study host-virus interactions (Félix & Wang 2019). Host-virus interaction studies focusing on the effect of host genetic variation are facilitated by the androdieious mode of replication by which *C. elegans* reproduces. This makes *C. elegans* a powerful model to investigate the effect of host genetic variation as populations can be both inbred and outcrossed.

Three cellular pathways are used by *C. elegans* to defend itself against viral infections. First, the RNAi response is a highly adaptive and diverse pathway that plays a role in many processes in an organism, for example in development and antiviral responses in invertebrates (Grishok & Mello 2002; Tabara *et al.* 1999). In OrV infection, it recognizes the double stranded RNA replication intermediate, which ultimately leads to the production of small interfering RNAs (siRNAs) that target the viral RNA for degradation (Ashe *et al.* 2013; Félix *et al.* 2011; Sarkies *et al.* 2013; Sterken *et al.* 2014; Tanguy *et al.* 2017). Mutants defective for various genes in the RNAi pathway display higher viral susceptibility upon infection (Ashe *et al.* 2013; Félix *et al.* 2011; Sterken *et al.* 2014; Tanguy *et al.* 2017). Second, the OrV can be targeted by a distinct mechanism known as viral uridylation (Le Pen *et al.* 2018). Uridylation, like RNAi, leads to degradation of viral RNAs although both antiviral defenses function independently of one another. Third, the Intracellular Pathogen Response (IPR) is involved in defense against viral, fungal and microsporidian infections. The IPR is regulated by the gene pair *pals-22* and *pals-25* that balance the nematode's physiological programs between growth and immunity (Reddy *et al.* 2019). Infections are counteracted by upregulating a range of 80 IPR genes that reduce proteotoxic stress (Panek *et al.* 2020; Reddy *et al.* 2017, 2019). For most IPR genes, their biochemical function is currently unknown, but IPR gene *cul-6* functions in the E3 ubiquitin ligase complex and protects against viral and microsporidian infection (Bakowski *et al.* 2014; Panek *et al.* 2020). Furthermore, the gene *drh-1* (encoding a RIG-I like protein) mediates the IPR response specifically upon OrV infection connecting IPR and RNAi pathways which both depend on this gene (Sowa *et al.* 2019).

Natural variation influences the susceptibility to OrV infections. Initially, it was observed that the natural *C. elegans* strain JU1580 is more susceptible to infection with OrV than the reference strain Bristol N2 (Félix *et al.* 2011). This difference has been linked to a natural polymorphism in *drh-1* affecting the anti-viral RNAi response (Ashe *et al.* 2013). In addition to the natural variation in the RNAi response, genetic variation also determines the Intracellular Pathogen Response (IPR) against OrV infection. The Hawaiian strain CB4856 had higher (basal) expression of multiple IPR genes than N2, potentially resulting in higher resistance to OrV infection observed in CB4856 (Chapter 3). However, the genetic and transcriptional networks leading to this difference have not been uncovered.

The CB4856 and N2 strain are very polymorphic, with more than 400,000 polymorphisms, including insertions/deletions and single nucleotide variants (Kim *et al.* 2019; Thompson *et al.* 2015). Over the last decade, both strains have been jointly used in many quantitative genetics studies in *C. elegans*, focused on traits like: aging, stress tolerance and pathogen avoidance (Doroszuk *et al.* 2009; Rodriguez *et al.* 2012; Snoek *et al.* 2020; Viñuela *et al.* 2012). Most of these studies have been conducted on one of the two available recombinant inbred line (RIL) panels (Li *et al.* 2006; Rockman & Kruglyak 2009) or on the introgression line (IL) population which contains fragments of CB4856 in a background of N2 (Doroszuk *et al.* 2009).

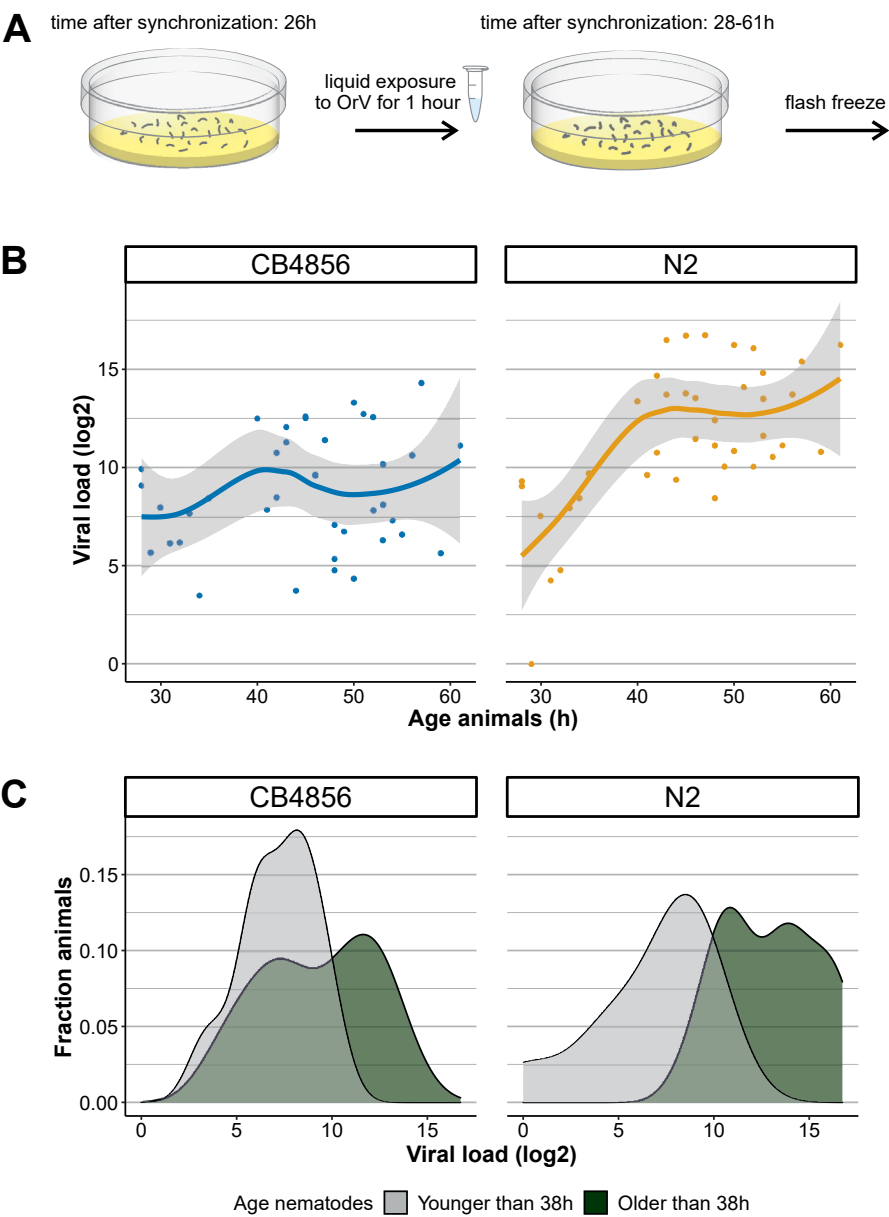
Here we set out to investigate genetic loci involved in the phenotypic differences between the Bristol N2 strain and the Hawaii CB4856 in response to OrV infection. Viral replication was characterized in N2 and CB4856 in a stage- and incubation time-dependent manner. Subsequently, we used inbred panels constructed from these strains to identify possible causal loci underlying the difference in viral susceptibility. We exposed a panel of 52 RILs to OrV and measured the viral load. We identified two QTL associated with differences in viral load on chromosome IV. Following-up, using a panel of 27 IL strains together covering the QTL location on chromosome IV led to the identification of 34 candidate genes involved in antiviral immunity. One of these candidate genes, the IPR gene *cul-6*, was tested for its role in OrV infection in the strains N2 and CB4856.

## Results

### *CB4856 displays resistance to OrV infection*

The two wild-type genotypes N2 and CB4856 respond differently to OrV infection: N2 shows higher viral loads than CB4856 30 hours after infection with OrV ([Chapter 3](#)). This difference could arise due to a slower developing infection, a difference in the stationary phase of the infection, or (molecular) differences at an individual level. The latter hypothesis was previously tested by visualizing OrV in infected individuals ([Chapter 3](#)), but individual infection levels were insufficient to explain differences between the strains. Therefore, we decided to focus on the kinetics of viral susceptibility. Infection kinetics were investigated by infecting both strains at an age of 26h (L2 stage) and measuring the viral load over 2-36 hours post infection (in 26-62 hour old animals) (Figure 1A). N2 developed a 3.2 units higher viral load than CB4856, in concordance with the observation in previous experiments ([Chapter 3](#)). Next to this, we observed that infection could be established in N2 more often than in CB4856 (76% (n = 121) versus 61% (n = 115) success rate).

Whereas the infection developed via a clear lag-phase in N2 during the first 12 hours, a large variation in viral loads was observed in the initial infection phase for CB4856 (Figure 1B). In this time series experiment, a significant amount of the variance was explained by genotype (ANOVA,  $p < 1 \cdot 10^{-4}$ ). We found that for some infected CB4856 populations the infection developed similar (but to lower level) compared to N2, however in other experiments the infection did not develop beyond levels reached in the lag-phase of the infection. Consequently, CB4856 populations that were 38h or older show either similar viral loads to populations that were younger (and thus shorter infected) or viral loads that reached a maximum (Figure 1C). On the other hand, N2 populations all reach higher loads after the lag-phase of infection was passed (Figure 1C). Therefore, the time passed since infection was also explaining variance in viral load (ANOVA,  $p < 1 \cdot 10^{-6}$ ). Constant exposure to OrV for four days resulted in similar viral loads between N2 and CB4856 ([Chapter 3](#)) (Ashe *et al.* 2013), thus suggesting that multiple rounds of viral replication are necessary to fully infect CB4856 populations. Together, these observations show that CB4856 initially develops a lower viral load and can control beginning infection better than N2.

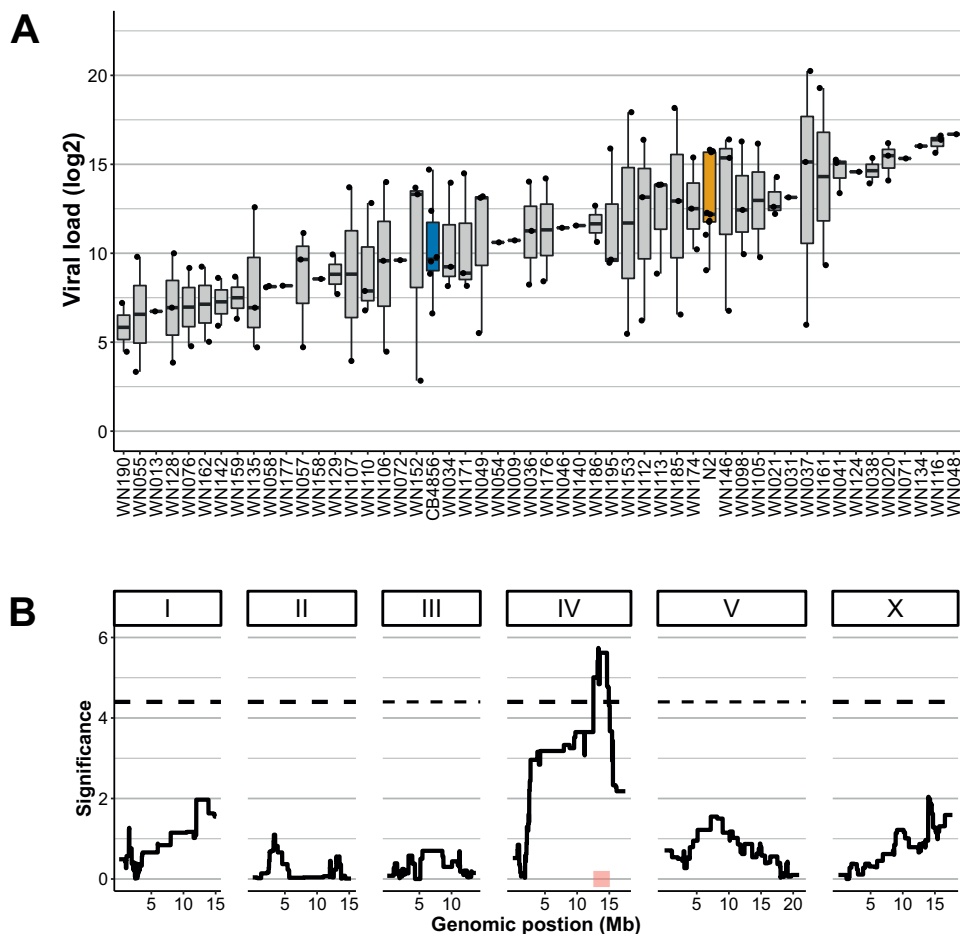


**Figure 1 Kinetics of OrV infection in N2 and CB4856** – A) Nematodes are infected by the OrV in liquid at the age of 26h hours (as in (Chapter 3)) before samples were washed of the plate 2-35h later and collected for viral load quantification. B) Development of OrV infection over time in N2 and CB4856 over a course of 33h. C) A density plot of the viral load measurements over time, divided in two groups: early infection (up to 12 hours post infection, grey) and late infection (after 12 hours post infection, green).

CB4856 nematodes are resistant to infection by the microsporidian *Nematocida parisii*, but only in the L1 stage (Balla *et al.* 2015). *N. parisii* shares its cellular tropism with OrV and both pathogens induce the same transcriptional response: the Intracellular Pathogen Response (IPR) (Chen *et al.* 2017; Reddy *et al.* 2017, 2019). Here, we tested if L1 CB4856 could have even higher resistance to OrV infection than the L2 animals we have infected before. Therefore, infection was compared in first (22-hour old) and second (28-hour old) larval stage animals (Figure S1A). N2 animals were infected in parallel for reference and the infection could develop for 30 hours after infection. We found for both genotypes that the viral loads were highly comparable between the L1 and in the L2 infected nematodes (Figure S1B). Thus, the relative resistance of CB4856 towards the OrV is not age-dependent, in contrast to resistance to the microsporidian *N. parisii*.

#### *A locus on chromosome IV links to resistance against OrV*

To find the causal genetic loci underlying the differences between N2 and CB4856 in viral load, recombinant inbred lines (RILs) constructed from a cross between these strains were infected with OrV (Figure S2A) (Li *et al.* 2006; Thompson *et al.* 2015). The RILs were infected in the L2 stage (at an age of 26h) and the infection was continued for 30h, after which the viral load was measured. The viral loads of the RILs followed a pattern of transgressive segregation, indicating that multiple genetic loci contribute to viral susceptibility (Figure 2A). We found a narrow-sense heritability ( $h^2$ , the fraction of trait variation explained by genotype) of 0.40 for the mean viral load (excluding populations that were not successfully infected). Linkage analysis for this trait identified a QTL on chromosome IV between 12.5 and 15.1Mb (Figure 3B) ( $R^2 = 0.37$ ). Besides performing a linkage analysis for the mean viral load of successfully infected populations, linkage analysis was performed for A) the median viral load (excluding unsuccessfully infected populations), B) the overall mean viral load (including unsuccessfully infected populations) and C) the minimum viral load observed for a strain (Figure S2A-C). Correlation analysis of the minimum viral load pointed towards an additional QTL location on chromosome IV between 2.6 and 2.8Mb. Thus, the QTL peak on the left side of chromosome IV could be linked to the success of infection, whereas the peak on the right side of chromosome IV was linked to the height of the viral load measured. Therefore, each locus may influence another biological aspect of OrV infection.



**Figure 2 OrV infections in a Recombinant Inbred Line panel with parental strains N2 and CB4856 –**  
A) Transgression plot of the viral loads of 52 RIL strains used for the infection assays. B) The QTL profile for mean viral load (excluding unsuccessful infections). The significant QTL peak is found at the end of chromosome IV at 13.3Mb (1.5 LOD-drop interval from 12.5-15.1Mb).

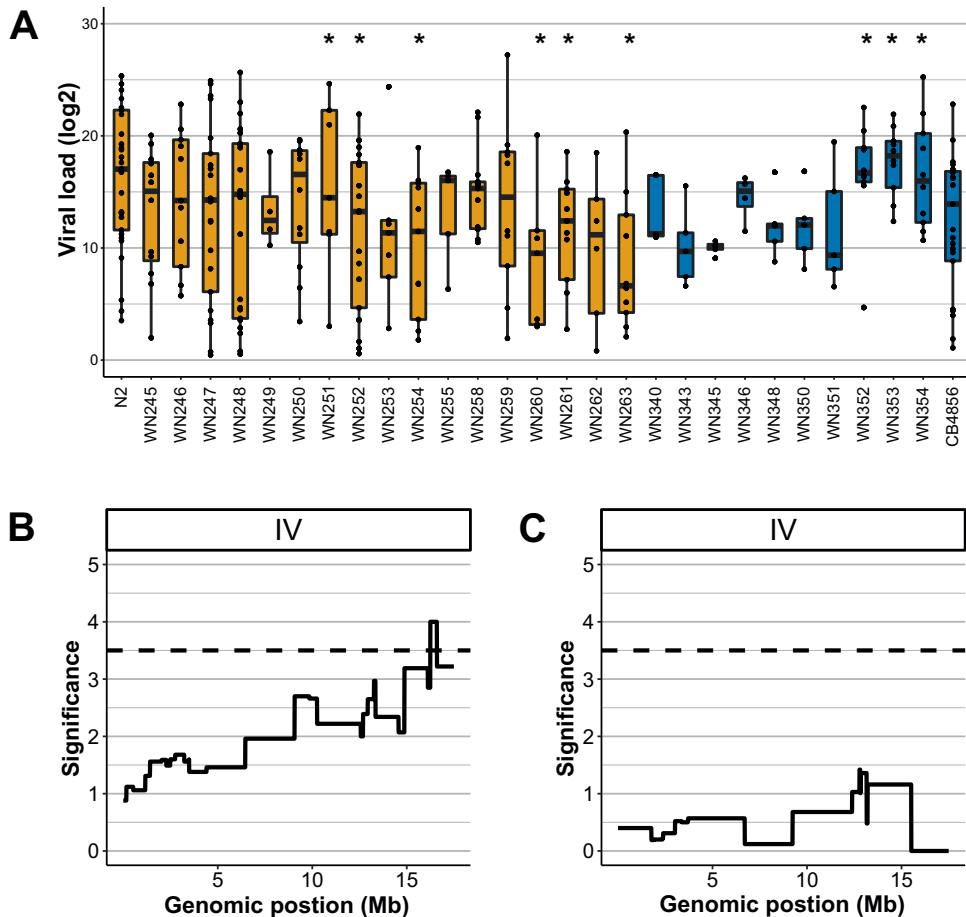
### Verification of the QTL locus by introgression lines

To experimentally verify the QTL peaks involved in the viral susceptibility difference between N2 and CB4856 introgression lines (ILs) were infected. ILs contain small fragments of one strain in the genetic background of another strain (Doroszuk *et al.* 2009). ILs that together cover chromosome IV were used and their viral loads were measured after infection. We used 10 ILs with a N2 fragment in the CB4856 background (IL<sub>CB4856</sub>) and 17 ILs with a CB4856



fragment in the N2 background (IL<sub>N2</sub>; Figure S2B). Of the 27 infected ILs 9 had a different viral load than the parental strain, demonstrating that presence of the introgression alters the viral susceptibility compared to the parent. We found that the IL<sub>CB4856</sub> strains WN352, WN353, and WN354 showed a phenotype distinct from the parental strain (two-sided t-test,  $p < 0.05$ ) overlapping the right QTL peak at 12.41-12.89Mb. In agreement, two IL<sub>N2</sub> strains covering this QTL were more resistant than N2 (WN252, WN254) (Figure 3A), but contrary three strains with the CB4856 fragment in the N2 background covering the same location did not show a lower viral load than the N2 strain (WN258, WN259, and WN261). In addition, IL<sub>N2</sub> strain WN263 with an introgression from 14.87-17.49Mb had a lower susceptibility than N2. Together, these results indicate that there are multiple loci genes underlying the susceptibility difference between N2 and CB4856 and these can interact with each other.

Linkage analysis on the IL<sub>N2</sub> panel showed the highest correlations for viral load and genetic background on the right side of chromosome IV with a QTL peak around 16Mb (Figure 3B), whereas the IL<sub>CB4856</sub> panel mapping did not show an effect of the introgression (Figure 3C). The resolution for IL mappings is relatively low compared to RIL mappings, because of fewer breakpoints in the population. Therefore, the IL peak mapped in the ILs could rely on the same genetic variation as the QTL peak mapping in the RIL panel. The left-sided QTL peak on chromosome IV found in the RIL panel could not be confirmed by IL mapping, with only two (WN251 and WN252) out of fourteen strains that cover this location having a different phenotype than the parental strain.

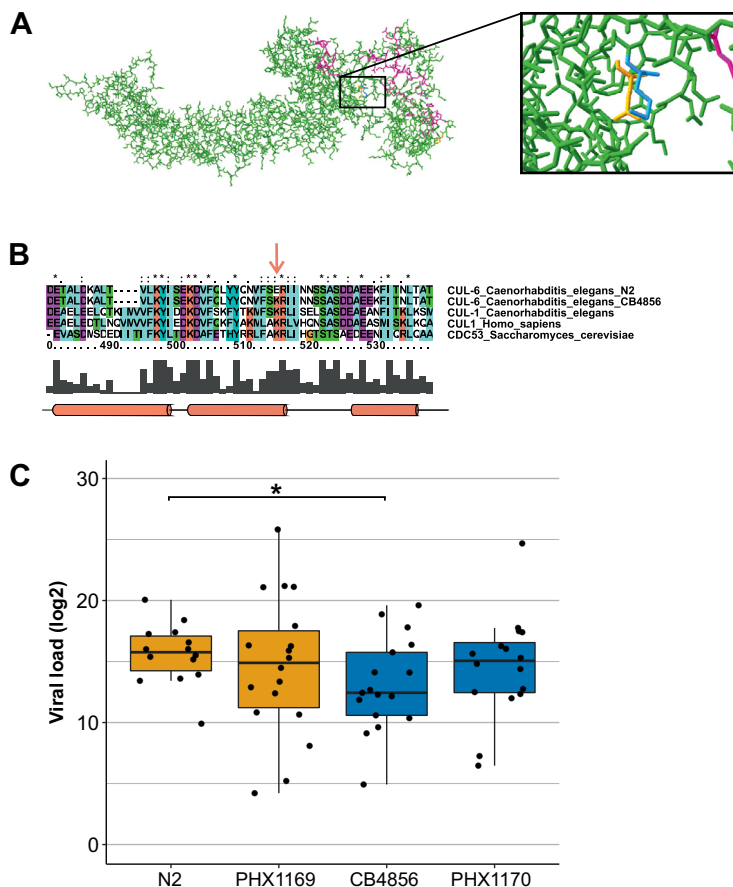


**Figure 3 OrV infections in two Introgression Line panels with parental strains N2 and CB4856 – A)** The viral loads of N2, CB4856 and 27 IL strains used for the infection assays. Of these, 17 strains have a CB4856 introgression in a N2 background (orange filled boxplots) and 10 have a N2 introgression in a CB4856 background (blue filled boxplots). An asterisk indicates that the strain is significantly different from its parental genetic background ( $p < 0.05$ , t-test). **B)** Linkage mapping profile for mean viral load (excluding unsuccessful infections) for the CB4856-in-N2 background panel. A significant peak is found on the right side of chromosome IV. **C)** Linkage mapping profile for mean viral load (excluding unsuccessful infections) for the N2-in-CB4856 background panel.

### *In search of causal genes underlying viral susceptibility differences between N2 and CB4856*

Linkage analysis in both RILs and ILs indicated that viral susceptibility differences between N2 and CB4856 were governed by multiple loci. We set out to see if we could identify polymorphic genes that determine the susceptibility difference between N2 and CB4856. We focus on the 12.41-12.89Mb region on chromosome IV, because this region was mapped in the RIL panel and supported by analysis of the ILs. This region contains 34 polymorphic genes of which 25 contain a non-synonymous change in the coding sequence (Supplementary Table S3). The candidate genes in this region have diverse functions, including genes with a known immune function against bacterial or viral infection. One these is the gene *cul-6*. This gene is regulated by the IPR and knockdown of *cul-6* increases the susceptibility to OrV in N2 nematodes (Bakowski *et al.* 2014; Reddy *et al.* 2017, 2019). The *cul-6* gene contains a single nucleotide polymorphism in the 428<sup>th</sup> amino acid changing a negatively charged glutamic acid in N2 into a positively charged lysine in CB4856 close to the RBX-1 binding site (Figure 4A). The amino acid at this position has been highly conserved from yeast to humans in the closely related CDC53 and CUL-1 proteins encoding a lysine in all cases (amino acid conservation between *C. elegans* CUL-1 and CUL-6 is 47%) (Figure 4B) (Zheng *et al.* 2002).

To test whether the *cul-6* polymorphism explains the difference in viral susceptibility between N2 and CB4856 we used CRISPR-Cas9 strains encoding the *cul-6* N2 allele in the CB4856 genetic background (PHX1170) and the *cul-6* CB4856 allele in the N2 genetic background (PHX1169). Based on the results of the introgression line analysis the strain N2-in-CB4856 strain was expected to be as susceptible as N2. We indeed observed a higher viral load for PHX1170 than for CB4856 (Figure S3). Although this CB4856-PHX1170 difference was not significant ( $p = 0.31$ ), the difference between N2 and PHX1170 was also not significant ( $p = 0.41$ ), indicating PHX1170 has a phenotype intermediate between N2 and CB4856. Furthermore, PHX1169 retained high viral susceptibility, nevertheless we noticed that the variance in measurements was higher than in the N2 strain. Together, this shows that the *cul-6* polymorphism contributes to viral susceptibility differences between N2 and CB4856, yet the effect size of this allele is modest. The resistant phenotype of CB4856 cannot be fully allocated to this allele, because it does not confer resistance in a susceptible background and CB4856 and PHX1170 were more similar than CB4856 and N2. This shows having the susceptible *cul-6* allele makes the strains vulnerable to infection but having the resistant allele does not protect strains with an otherwise susceptible background.



**Figure 4** The gene *cul-6* in CB4856 and N2 and its effect on viral susceptibility – A) Structure prediction of *C. elegans* CUL-6. The lysine present in the CB4856 allelic variant is shown in blue, the glutamic acid present in the N2 allelic variant in orange. The RBX-1 binding domain is shown in purple. B) Part of the sequence alignment between *Homo sapiens* CUL1, *Saccharomyces cerevisiae* CDC53, *C. elegans* CUL-1 and the *C. elegans* N2 and CB4856 allelic variants of CUL-6. The location of the N2 and CB4856 polymorphism is indicated with an arrow. The amino acid conservation is indicated by the grey bars at the bottom and by the annotations on top (single dot: weakly conserved, double dot: strongly conserved, asterisk: completely conserved). The gene *cul-6* contains a polymorphism between N2 and CB4856 at a conserved site. Colors are based on the amino acid properties and locations of alpha-helices are indicated by cylinders (Zheng et al. 2002). C) Viral susceptibility of N2, CB4856, PHX1169 (N2 genetic background carrying a CB4856 *cul-6* allele) and PHX1170 (CB4856 genetic background carrying a N2 allele) (asterisk indicating  $p < 0.05$ , *t*-test).

## Discussion

Here we have unraveled the genetic architecture of viral susceptibility in the *C. elegans* strains N2 and CB4856. We found two QTL peaks linking to susceptibility differences on chromosome IV and confirm the QTL on the right side of chromosome IV using a selection of introgression lines. Observations made for individual ILs show that multiple loci on chromosome IV contribute to viral susceptibility. When we zoomed in on the 12.4-12.9Mb region that likely contains a causal gene, we identified 34 polymorphic genes which are candidates to explain differences in viral susceptibility between N2 and CB4856. Allele swap experiments between one of this candidate genes, the IPR gene *cul-6*, indicated a single nucleotide polymorphism underlies susceptibilities differences. Nevertheless, other genetic loci contribute to the whole phenotypic variation between N2 and CB4856. These findings show that the genetic architecture of OrV susceptibility is a complex, polygenic trait.

### *The OrV infection phenotypes of CB4856 and N2*

Our experiments confirmed that there is a difference in early OrV infections between N2 and CB4856 (Chapter 3) and that this difference retains independently of age of infection. There are several possibilities for the observed differences in OrV infection between N2 and CB4856. Three of these possibilities are: i) individual nematodes in the CB4856 population are less likely to be infected, ii) in CB4856 a lower number of cells is infected, iii) or the infection topology in CB4856 is different from N2. Although we have addressed these questions in our previous study, we could not visualize OrV in enough individuals to state what may contribute to the observed viral susceptibility difference at a population level. Having a more sensitive method than Fluorescent *in situ* Hybridization (FISH) could help to answer this question. For example, having fluorescently labelled OrV might help to visualize lower infection levels. Additionally, a reporter strain could be made by integrating a fluorescently labelled copy of one of the IPR genes that responds similar to infection in N2 and CB4856. A non-microscopic approach may be found by developing a protocol for single-nematode RT-qPCRs to detect the OrV.

### *Chromosome IV is implicated in natural variation in OrV infection*

By exposing RILs and ILs to OrV infection, we identified a QTL on chromosome IV that is implicated in a lower viral load due to the CB4856 allele. A genome wide association study (GWAS) on OrV infection in *C. elegans* also implicated chromosome IV (Ashe *et al.* 2013), but unlike these authors, we did not find a peak near the *drh-1* locus. This was in line with expectations as only two polymorphisms are found in the introns between N2 and CB4856 for this gene (Thompson *et al.* 2015). Still, the more distal associations uncovered by the GWAS could potentially result from the same allelic variation as the QTL between 12.41-12.89Mb,

because the GWAS identified five peaks on chromosome IV which are located between 5 and 13Mb. Therefore, natural populations of *C. elegans* may carry similar genetic variants conferring OrV resistance as N2 and CB4856.

In our previous study investigating viral susceptibility differences between N2 and CB4856 we found that CB4856 has higher basal expression of IPR genes which we hypothesized may be caused by distinctive *pals-22/pals-25* expression patterns ([Chapter 3](#)). These genes, the respective repressor and activator of the IPR, are located adjacent to each other on the left hand of chromosome II (Reddy *et al.* 2019). eQTL studies indicated that local genetic variation (*cis*-eQTL) regulate expression of *pals-22* and *pals-25* (Li *et al.* 2006, 2010; Rockman *et al.* 2010; Snoek *et al.* 2017; Sterken *et al.* 2014; Viñuela *et al.* 2010, 2012). Nevertheless, we did not observe a link between natural genetic variation in viral susceptibility in N2 and CB4856 and the *pals-22/pals-25* locus on chromosome II. There may be many reasons for this (including a lack of causality of this locus), but our results show that we could only explain a minor fraction of the heritability by the QTL locations we found. This result is typical for QTL mappings of complex traits and suggests that additional loci contribute to the viral susceptibility difference between N2 and CB4856. These loci may have small effect sizes, interactions or are affected by a (currently unknown) environmental cause (Eichler *et al.* 2010).

### *Orsay virus susceptibility has a polygenic basis*

The QTL in the RIL panel and follow-up fine mapping in the ILs identified a relatively small locus containing 34 polymorphic genes contributing to the viral susceptibility towards OrV infection. We investigated the effect of a *cul-6* polymorphism and found that this SNP contributes to viral susceptibility. This allele functions in one direction by making the resistant CB4856 background susceptible when carrying the N2-allele. Other approaches may clarify how the *cul-6* polymorphism affects the functioning of the E3 ubiquitin ligase complex. For example, biochemical approaches such as ELISA or immunoblots could be used to investigate binding of the two *cul-6* protein variants to other E3 ubiquitin ligase complex members (Pollard 2010). Additionally, tagging *cul-6* with a fluorescent label could further reveal cellular functioning of both variants, especially in OrV infected cells that are also stained by FISH.

The phenotypic difference between N2 and CB4856 cannot be entirely explained by the *cul-6* allele alone. The 12.4-12.9Mb region also contains multiple other genes that may affect viral susceptibility. Some of these are transcriptionally activated by OrV infection, others have a more general or unknown cellular function. Besides, the 12.4-12.9Mb region specifically investigated here, we show that are multiple other loci and genes contributing to viral susceptibility on chromosome IV. The left side of chromosome IV appeared involved in determining the success of infection, but we could not verify this result in the ILs. This



may be because there is no correlation between phenotype and genotype, but it may also be because only a small fraction of infections fails, complicating studying this trait. Nevertheless, some ILs covering the left side of chromosome IV had a viral susceptibility distinct from the parent. Additionally, strain WN351 carries a susceptible introgression at the 12.4-12.9Mb locus but remained resistant. This strain has a large introgression also covering the left side of chromosome IV, where interacting genes may be located.

Our results reveal the complex genetic basis of OrV susceptibility. These results are in line with other studies mapping variation in viral susceptibility to the hosts genome (see for example (Al-Qahtani *et al.* 2013; Heim *et al.* 2016; McLaren *et al.* 2015; Nedelko *et al.* 2012; van Manen *et al.* 2012)). This may not be surprising as viruses use the hosts cellular machinery to replicate and hosts have multiple mechanisms to counteract viruses, therefore host-virus interactions will comprise many genetic interactions that can be affected by genetic variation ([Chapter 2](#)). Thus, future studies may aim to uncover genetic networks rather than a single gene to further enhance our understanding of natural variation in host-virus interactions.

## Material and methods

### *C. elegans strains and culturing*

*C. elegans* strains Bristol N2 and Hawaii CB4856 were used and strains derived from crosses between these two wild-type strains. In this paper 52 recombinant inbred lines, 17 introgression lines with an N2-background ( $IL_{N2}$ ) and 10 introgression lines with a CB4856 background ( $IL_{CB4856}$ ) covering chromosome IV were used (Supplementary Table S1, S2A) (Doroszuk *et al.* 2009; Li *et al.* 2006). All these genotypes have been confirmed by full-genome sequencing. The strains PHX1169 *cul-6(syb1169)* and PHX1170 *cul-6(syb1170)*, containing the *cul-6* CB4856 allele in a N2 background and the N2 *cul-6* allele in a CB4856 background respectively, have been created by CRISPR-Cas9 by SunyBiotech (<http://www.sunybiotech.com>) (Supplementary Text S1). These genotypes have been confirmed by PCR sequencing.

The nematodes were kept at 12°C between experiments on 6 cm NGM plates seeded with *E. coli* OP50. Bleaching was used to synchronize populations and to remove bacterial or fungal contaminations (Brenner 1974). Before experiments, a population without males was created by picking single worms in the L1/L2 stage and transferring hermaphrodite populations to fresh 9 cm NGM plates. New experiments were started by bleaching an egg-laying population grown at 20°C.

### *Orsay virus stock preparation*

Orsay virus stocks were generated by isolating OrV from a persistently infected JU1580 culture (Félix *et al.* 2011). Over 100 JU1580 populations were grown on 9cm NGM plates containing twice the usual amount of agar to prevent the nematodes from burying into the agar (34g/L). The nematodes were collected by washing the animals off the plate with M9 buffer and collecting the suspension in an Eppendorf tube. The suspension was flash frozen in liquid nitrogen to break the nematodes and slowly thawed at 4°C. The suspension was centrifuged for 5 minutes to pellet the bacteria and nematodes. The supernatant was collected and passed through a 0.2µm filter. The obtained virus stock was divided in aliquots, flash frozen in liquid nitrogen, and stored at -80°C until use. Specific infectivity of the virus stock was tested by serial dilution infections in *C. elegans* JU1580 (Sterken *et al.* 2014).

### *Infection experiments*

The infection assay was conducted as described in (Sterken *et al.* 2014). Populations were synchronized (t = 0 hours) and grown at 20°C on 9 cm NGM plates. Just before of infection, the strains were washed off the plate with M9 buffer and pelleted by centrifugation. The supernatant was removed and the strains were exposed to OrV in liquid for 1 hour. The

worms were washed 3 times with M9 and placed on a fresh 9 cm NGM plate. Infections were performed on animals in the L1 (22 hours post bleaching) or L2 (26 or 28 hours post bleaching as indicated in the text) stage.

For the replication kinetics experiments on N2 and CB4856, the animals were harvested 2-35 hours post infection. This experiment was conducted in 8 independent biological replicates, each evenly covering the time-series. For the viral load experiments on the RIL and IL panels and the *cul-6* allele swap strains, the animals were harvested 30 hours post infection. The experiment in the RIL panel was conducted on 3 independent biological replicates. The experiment in the IL panel was conducted on at least 5 independent biological replicates. The experiment using the *cul-6* allele swap strains was conducted on 21 independent biological replicates.

### *RNA isolation*

The RNA was isolated using a Maxwell® 16 AS2000 instrument with a Maxwell® 16 LEV simply RNA Tissue Kit (both Promega) following the recommended protocol, except the addition of 10 mg of proteinase K during the lysis step. The lysate was incubated in a Thermomixer (Eppendorf) for 10 minutes at 65°C at 1,000 rpm. After isolation the quality and quantity of the RNA was determined via NanoDrop-1000 (Thermo Scientific).

### *cDNA preparation and RT-qPCR*

cDNA was synthesized using the GoScript Reverse Transcriptase kit (Promega) following the recommended protocol with random hexanucleotides (Thermo Scientific) and 1µg of total RNA as starting material. The cDNA was quantified by RT-qPCR (MyIQ, Biorad) using Absolute QPCR SYBR Green Fluorescein Mixes (Thermo Scientific) or iQ SYBR Green Supermix (Biorad) following the recommended protocol. The samples were quantified using the primers described by (Sterken *et al.* 2014).

The qPCR data was processed using R (version 4.0.2), as described before (Sterken *et al.* 2014).

$$Q_{gene} = 2^{40 - Ct_{gene}}$$

In short, before normalization, the RT-qPCR measurements were transformed by where  $Q$  is the expression of the gene and  $Ct$  is the measured  $Ct$  value of the gene. The viral expression was normalized by the two reference genes, using the formula

$$E = \frac{Q_G}{((Q_{rpl-6} / \bar{Q}_{rpl-6}) + (Q_{Y37E3.8} / \bar{Q}_{Y37E3.8}))}$$

where  $E$  is the normalized viral load,  $Q_v$  is the expression of the viral RNA and  $Q_{rpl-6}$  is the expression of reference gene *rpl-6* and  $Q_{Y37E3.8}$  is the expression of reference gene Y37E3.8. All viral load data presented here was batch corrected per experiment for the different viral stocks used in that experiment. The average viral loads of N2 and CB4856 (excluding unsuccessful infections) were used for batch correction as these two strains were taken along in every experiment.

From the replicate measurements in the RIL panel, several traits could be derived for QTL mapping over the RIL population. The following parameters were derived: mean viral load, median viral load, and minimum viral load. We excluded the unsuccessful infections (as these could also arise due to technical failures) and unless indicated otherwise.

### *Quantitative trait locus mapping RIL population*

Single locus QTL mapping was done using a linear model (version 4.0.2) to explain viral load and derived traits over the markers by

$$E_i \sim x_{i,j} + \epsilon_{i,j}$$

where  $E$  is the viral load of RIL  $i$  (1, 2, ..., 52) and  $x$  is the marker of RIL  $i$  at location  $j$  (a set of 1152 sequenced markers was used (Supplementary Table S2)). For  $E$  the outcome of each replicate of the experiment was averaged over the three biological replicates.

For the RIL mapping the statistical threshold was determined via a permutation analysis, where the values measured for  $E$  were randomly distributed over the genotypes. The same model as for the mapping was used and this analysis was repeated 1,000 times. The 950th highest  $p$ -value was taken as the threshold  $p$ -value for a false discovery rate of 0.05.

### *Heritability and variance calculations*

The narrow-sense heritability's ( $h^2$ ) were calculated per investigated trait by REML (Grishok & Mello 2002; Kruijer *et al.* 2014; Speed *et al.* 2012) using the package 'heritability' in R (version 4.0.2). Significance was determined via 1,000 permutations where the values measured for E were randomly distributed over the genotypes.

The variation explained by a QTL peak was calculated by

$$V_{Explained} = \frac{R_{QTL}^2}{h^2}$$

where  $R^2$  is the  $R^2$  from the fit of the peak marker and  $h^2$  is the narrow-sense heritability of the trait.

### *Introgression line analysis*

The viral loads obtained for the introgression lines were analyzed individually against N2 and CB4856 via a two-sided t-test assuming unequal variance in R (version 4.0.2). Experiments where no virus was detected were excluded from the analysis. Moreover, we performed linkage mapping for the two IL panels separately using a linear model to explain viral load over the markers by

$$E_i \sim x_{i,j} + \epsilon_{i,j}$$

where E is the viral load of IL i (1, 2, ..., 10 or 17) and x is the marker of IL i at location j (a set of 1152 sequenced markers was used (Supplementary S2). Each IL was compared against the respective parental strain (N2 or CB4856). For E the outcome of each replicate of the experiment was averaged over the biological replicates. A significance threshold was drawn at  $-\log_{10}(p) > 3.5$  for analysis of the data.

### *Allele swap analysis*

Because we observed a high level of variance in the viral loads in N2 and CB4856 and the effect size of the QTL<sub>IV:12.41-12.89</sub> was small we used a high level of replication for the allele swap experiments by performing 21 biologically independent infections using three different virus stocks. Unsuccessful infections were excluded from the analysis and the batch corrected viral load data (based on virus stock as described above) was subsequently checked for outliers. Outliers were defined by 1.5 times the interquartile range plus or minus the third or first quartile respectively. After removal of the outliers (7% of the measurements), a t-test assuming unequal variances was performed to analyze the data.

### *Protein structure analysis*

Protein sequences from the human CUL1 (NCBI Reference Sequence: NP\_003583.2), *Saccharomyces cerevisiae* CDC53 (GenBank: CAA98702.1), *Drosophila melanogaster* CUL-1 (GenBank: AAD33676.1) and *C. elegans* CUL-1 (GenBank: AAC47120.1), CUL-6 N2 allelic variant (GenBank: CAB01230.1) and CUL-6 CB4856 allelic variant were aligned in ClustalX (version 2.1) using the standard settings (Larkin *et al.* 2007). A structural model for the N2 and CB4856 allelic variant was predicted using the human CUL1 protein structure as a template in the SWISS-MODEL ExPASy web server. The standard search parameters were used, based on the SWISS-MODEL template library (version 14/01/2015) and the protein data bank (version 09/01/2015) (Altschul *et al.* 1997; Benkert *et al.* 2011; Mariani *et al.* 2011; Peitsch 1997; Remmert *et al.* 2012; Šali & Blundell 1993). The obtained models for N2 and CB4856 CUL-6 were compared in SwissPDBViewer (v. 4.1.0) (Guex *et al.* 2009).

### *Script availability*

Scripts that were custom written in R (version 4.0.2) are openly available at [https://git.wur.nl/lisavansluijs/OrV\\_QTL](https://git.wur.nl/lisavansluijs/OrV_QTL).

## **Acknowledgments**

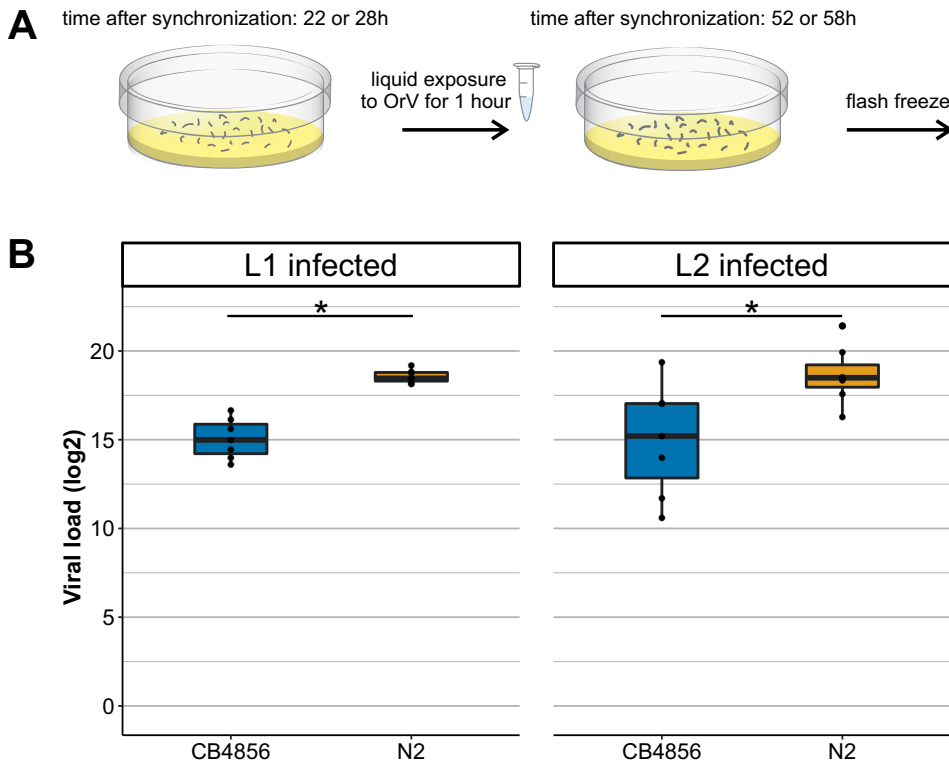
The authors want to thank all the people that contributed to the OrV research in N2 and CB4856: Kobus Bosman, Henrikje Smits, Jikke Daamen, Koen Semeijn, Yahya Zakaria Abdou Gaafar, Maarten Costerus, Yuqing Huang, Emma Lagae and Niels Vissers. We want to thank Marie-Anne Félix for providing us with the OrV. We thank Daniel Cook, Robyn Tanny, and Erik Andersen for help in sequencing IL<sub>CB4856</sub> strains.



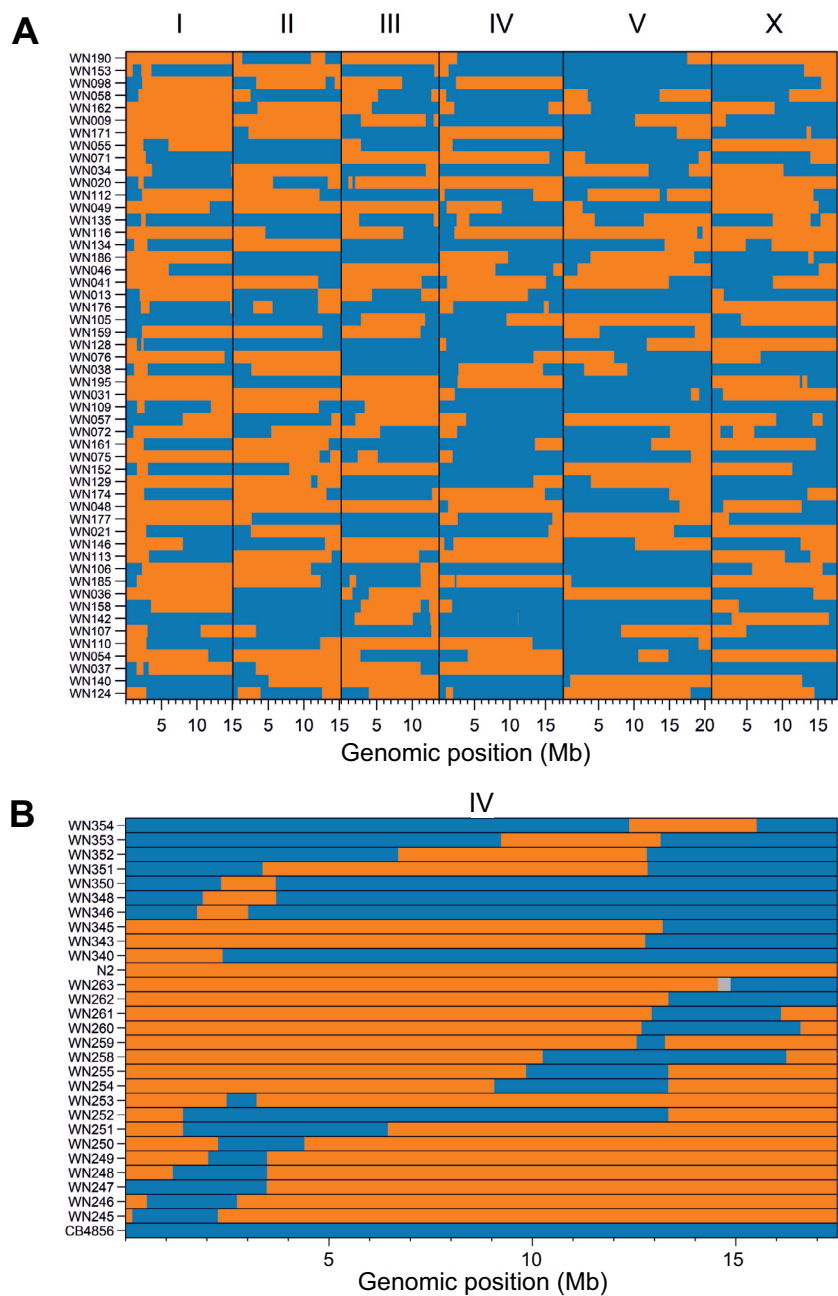
## Supplementary information

The supplementary text can be found at <https://tinyurl.com/SI-PhD-thesis>.

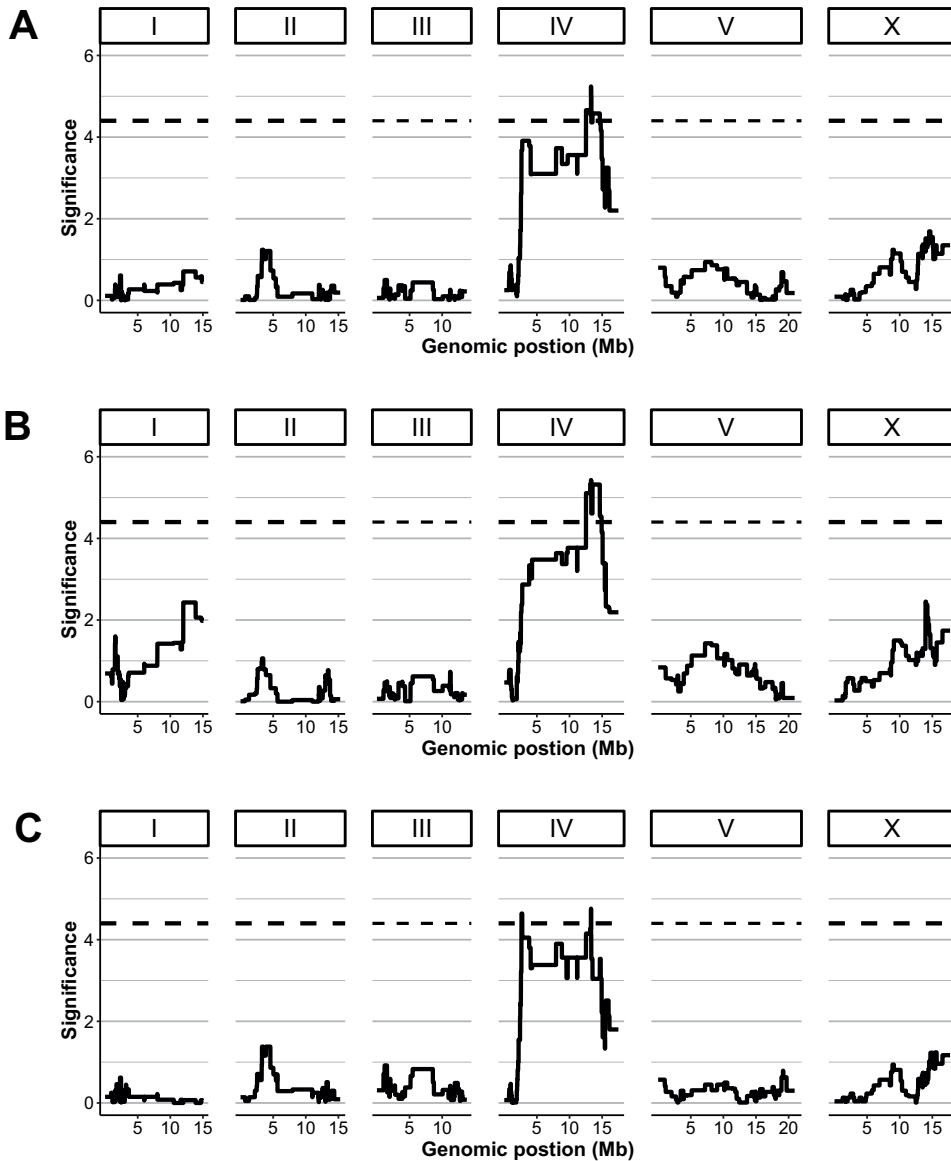
**Supplementary Text S1 Description of allele swap strains PHX1169 and PHX1170** – Genetic code of the strains PHX1169 and PHX1170 compared to their respective parental genotypes N2 and CB4856.



**Figure S1 Viral infection in L1 and L2 staged N2 and CB4856** – A) Viral infections were either started in the L1 stage (22h) or L2 stage (28h). Populations were exposed to the OrV in liquid and isolated 30 hours post infection for viral load measurements. B) Viral loads observed for N2 and CB4856 that were infected in the L1 or L2 stage. Genotype had a significant effect on viral susceptibility for both stages (asterisk indicating  $p < 0.05$ ,  $t$ -test), but viral susceptibility was not determined by infection age.



**Figure S2 Genetic maps of the RIL and IL strains** – A) Genetic map of the N2xCB4856 RIL panel. B) Genetic map of chromosome IV for the N2-in-CB4856 background (strain WN340-WN354) and CB4856-in-N2 background (strain WN245-WN263) ILs.



**Figure S3 QTL mapping for mean, median and minimum viral load** – A) The QTL profile for mean viral load (including unsuccessful infections). The significant QTL peak is found from 12.5-14.9Mb. B) The QTL profile for median viral load (excluding unsuccessful infections). The significant QTL peak is found from 11.2-15.1Mb. C) The QTL profile for minimum viral load (including unsuccessful infections). The significant QTL peaks are found from 2.6-2.8Mb and 11.2Mb.

The full supplementary tables and text can be found at <https://tinyurl.com/SI-PhD-thesis>. Here, the first four rows of the tables are shown to illustrate the contents.

**Supplementary Table S1 RIL and IL strains used in this study** – Strain names, previously used strain names (in other publications), strain type, genotype and previous publications mentioning the strains are given.

Strain	Strain alternative name	Strain type	Chromosome	Genotype	Position start	Position end	Paper
CB4856	cb4856	WT	I	CB4856	1	15072434	Thompson_2015_Genetics; this_paper
CB4856	cb4856	WT	II	CB4856	1	15279421	Thompson_2015_Genetics; this_paper
CB4856	cb4856	WT	III	CB4856	1	13783801	Thompson_2015_Genetics; this_paper

**Supplementary Table S2 Genetic map of RIL, IL strains and markers** – A) Strain genetic data used to perform genetic linkage mappings. A reduced number of columns is shown for the strain map: the full-size table contains 384 columns. B) Informative marker positions used to perform genetic linkage mappings.

Marker	CB4856	WN201	WN202	WN203	WN204	WN205	WN206	WN207	WN208	WN209
I1	-1	-1	1	1	1	1	1	1	1	1
I730000	-1	-1	1	1	1	1	1	1	1	1
I740000	-1	-1	1	1	1	1	1	1	1	1

Name	Chromosome	Position	Informative marker
I1	I	1	1
I730000	I	730000	1
I740000	I	740000	1

**Supplementary Table S3 Polymorphic genes between N2 and CB4856 at 12.41-12.89Mb on chromosome IV** – Selection of genes that have a polymorphism between N2 and CB4856 at the fine mapped QTL location. Gene name, location, strand, type of polymorphism, function according to WormBase and description in OrV literature are mentioned (format: first author, journal, year).

Gene sequence name	Gene public name	Genetic location	Gene strand	Synon- ymous substi- tutions	Non-syn- onymous substi- tutions	Exon dele- tions	Inframe insertions and/or deletions	Fully delet- ed	SNV in 5-prime	SNV in 3-prime	Function based on Worm- base overview	OrV related?
F08G5.2	ad-13	IV:12410480- 12412197	+	1	0	0	0	0	0	0	-	-
F08G5.6	irg-4	IV:12435804- 12438251	+	0	1	0	0	0	0	0	Involved in immune response against bacteria	Sarkies Genome Research 2013; Tanguy mBio 2017; Chen BMC Ge- nomics 2017
C25G4.1	clec- 185	IV:12438636- 12443012	+	0	1	0	2	0	0	0	-	-



## Chapter 5

# Virus infection modulates male sexual behavior in *Caenorhabditis elegans*

Lisa van Sluijs<sup>1,2</sup>, Jie Liu<sup>1</sup>, Mels Schrama<sup>1</sup>, Sanne van Hamond<sup>1</sup>,  
Sophie P. J. M. Vromans<sup>1</sup>, Marèl H. Scholten<sup>1</sup>, Nika Žibrat<sup>1</sup>, Joost A.G. Riksen<sup>1</sup>,  
Gorben P. Pijlman<sup>2</sup>, Mark G. Sterken<sup>1</sup>, Jan E. Kammenga<sup>1</sup>

<sup>1</sup>Laboratory of Nematology, Wageningen University and Research, the Netherlands

<sup>2</sup>Laboratory of Virology, Wageningen University and Research, the Netherlands

Manuscript submitted



## Abstract

Mating dynamics follow from natural selection on mate choice and individuals maximizing their reproductive success. Mate discrimination reveals itself by a plethora of behaviors and morphological characteristics, each of which can be affected by pathogens. A key question is how pathogens affect mate choice and outcrossing behavior. Here we investigated the effect of Orsay virus on the mating dynamics of the androdiecious (male and hermaphrodite) nematode *Caenorhabditis elegans*. We tested genetically distinct wild types and found that viral susceptibility differed between sexes in a genotype-dependent manner with males of reference strain N2 being more resistant than hermaphrodites. Males displayed a constitutively higher expression of Intracellular Pathogen Response (IPR) genes, whereas the antiviral RNAi response did not have increased activity in males. Subsequent monitoring of sex ratios over ten generations revealed that viral presence can change mating dynamics in isogenic populations. Sexual attraction assays showed that males prefer mating with uninfected rather than infected hermaphrodites. Together our results illustrate that viral infection can significantly affect male mating choice and suggest altered mating dynamics as a novel cause benefitting outcrossing under pathogenic stress conditions in *C. elegans*.



## Introduction

Sexual reproduction is the dominant reproductive strategy in the animal kingdom allowing for rapid adaptation to a(biotic) selective pressures (Butlin 2002; Lehtonen *et al.* 2012). Pathogens are main biotic drivers of evolution by forcing hosts to constantly adapt to the peril of infection and can benefit sexual over asexual reproduction (Bell 1982; Hamilton 1980; Jaenike 1978; Morran *et al.* 2011). Next to direct genetic selection, pathogens may interfere with mating systems in many other ways. For instance, by influencing sexual characteristics such as chemical cues, behavior and courtship. Such interferences have been recorded for a range of species showing that frogs, mice, flies and humans are less likely to mate with infected conspecifics (Kavaliers & Choleris 2018; Kiesecker *et al.* 1999). Furthermore, physiological, molecular and genetic differences between the sexes can underlie differences in pathogen susceptibility, thereby potentially favoring one sex over another in an infected population (Gipson *et al.* 2019; Klein & Flanagan 2016; Scully *et al.* 2020). Despite sexual reproduction would be most efficient when both sexes are equally present (Fisher 1930), skewed sex ratios are often observed in nature. Pathogen-induced interference in populations can help explain unbalanced sex ratios (Dyson 2012; Engelstädter & Hurst 2009; Klein & Flanagan 2016; Lynch *et al.* 2018; Masri *et al.* 2013; Morran *et al.* 2011).

The androdiecious (male and hermaphrodite) nematode *C. elegans* and its natural associated pathogens, provide a versatile model to study sex-dependent host-pathogen interactions (Cutter *et al.* 2019). *C. elegans* can reproduce both by outcrossing and self-fertilization leading to variable male-hermaphrodite ratios. Males (XO) arise only rarely after spontaneous X chromosome non-disjunction with an estimated frequency of 0.1-0.4% in the laboratory (Teotónio *et al.* 2006). Although successful outcrossing raises male frequencies as mated hermaphrodites produce 50% male and 50% hermaphrodite (XX) offspring, ineffective mating behavior typically dwindles male frequencies (Borne *et al.* 2017; Chasnov *et al.* 2007; Garcia *et al.* 2007; Kleemann & Basolo 2007; Palopoli *et al.* 2008; Teotónio *et al.* 2006). Yet, elevated outcrossing levels benefit bacterial-infected *C. elegans* populations among others by having genetically more resistant offspring (Lynch *et al.* 2018; Masri *et al.* 2013; Morran *et al.* 2011).

Associated intracellular pathogens of *C. elegans* include microsporidia, oomycetes and a virus (Félix *et al.* 2011; Osman *et al.* 2018; Zhang *et al.* 2016). These pathogens induce a molecular defense mechanism, called the Intracellular Pathogen Response (IPR), that is distinct from the response to bacterial infections (Chen *et al.* 2017; Panek *et al.* 2020; Reddy *et al.* 2017, 2019; Sowa *et al.* 2019). The IPR involves different expression of 80 IPR genes and is controlled by the IPR inhibitor *pals-22* and IPR activator *pals-25* (Reddy *et al.* 2019). Additionally, viral infection is counteracted by RNAi and uridylation leading to a degradation of viral RNA (Ashe *et al.* 2013; Coffman *et al.* 2017; Félix *et al.* 2011; Le Pen *et al.* 2018; Sterken *et al.* 2014;

Tanguy *et al.* 2017). Males differ largely from hermaphrodites in various molecular pathways, but it is currently unknown if nematodes show sex-dependent resistance against intracellular pathogens such as viruses and if stress caused by these pathogens affects mating dynamics.

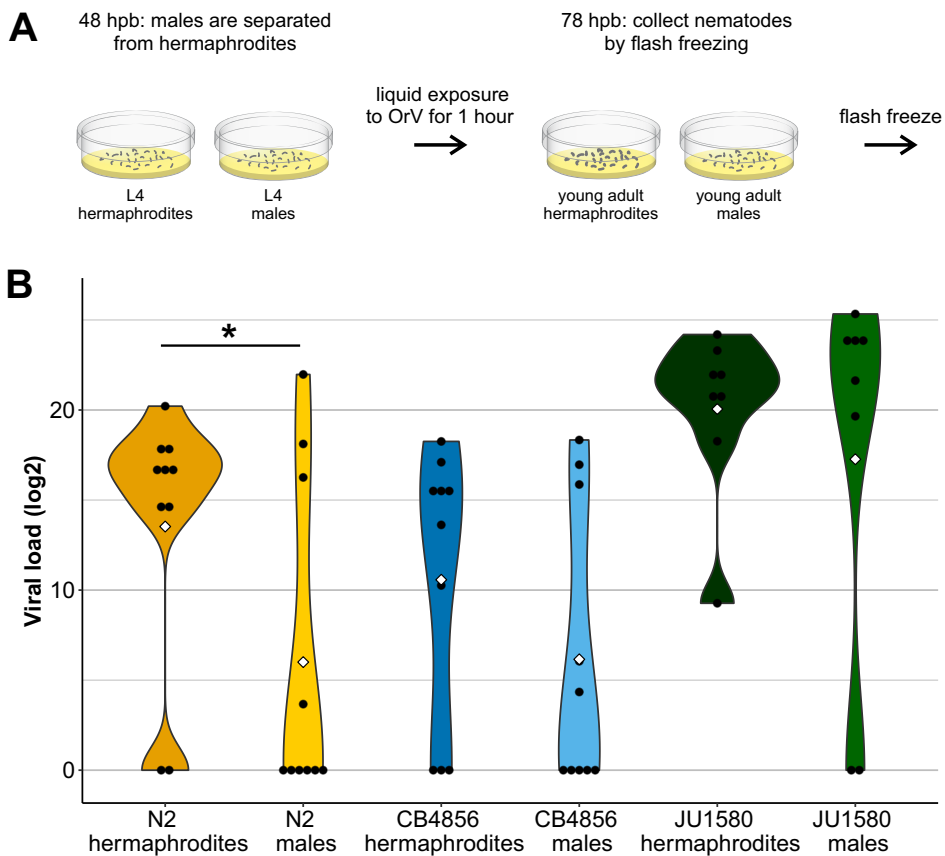
Here we studied infection responses for both sexes in different wild types of *C. elegans* to the naturally occurring intracellular pathogen Orsay virus (OrV). The non-lethal OrV is a positive-sense, single-stranded RNA virus (family *Nodaviridae*) that infects intestinal cells (Félix *et al.* 2011). Although relatively few individuals within populations become infected (up to 40% for the highly susceptible wild type JU1580), infected cells show severe morphological effects leading to fitness disadvantages (Ashe *et al.* 2013; Félix *et al.* 2011; Frézal *et al.* 2019). We found that viral infection affects male mating choice in *C. elegans* and suggest mating dynamics as a novel cause benefiting outcrossing under pathogenic stress conditions in *C. elegans*.

## Results

### *Viral susceptibility differs between sexes in a genotype-dependent manner*

Different genotypes of *C. elegans* hermaphrodites vary in susceptibility to the OrV (Ashe *et al.* 2013; Félix *et al.* 2011), but the susceptibility of *C. elegans* males has not been investigated. To compare the viral susceptibility of *C. elegans* males and hermaphrodites, both sexes were infected for three genetically distinct genotypes: N2, CB4856 and JU1580 (Figure 1A) (Dataset S1). We found that N2 males had lower viral loads than hermaphrodites (Figure 1B) (bootstrap,  $p = 0.047$ ). CB4856 males and hermaphrodites were often not successfully infected and had similar viral loads (Figure 1B) (bootstrap,  $p = 0.20$ ). JU1580 males and hermaphrodites are both highly susceptible to viral infection (Figure 1B) (bootstrap,  $p = 0.48$ ). Since the largest difference between males and hermaphrodites was seen for the N2 strain, OrV susceptibility was also tested in mixed male and hermaphrodite N2 populations. The mixed-sex populations contained around 30–40% of males, contrary to a hermaphrodite-only population. As expected, male presence had a tendency to lower the viral load in the population (t-test,  $p = 0.07$ ) (Figure S1) (Dataset S1). Notably, all mixed-sex populations were successfully infected, contrary to the male-only populations.

Since OrV is taken up whilst nematodes feed, sex-based differences in ingestion may affect viral loads. Therefore, we quantified both the rate and volume of food intake in L4 and young adult males and hermaphrodites (Figure S2A) (Dataset S2). The rate of food intake was measured by counting pumping rates of the pharynx, a neuromuscular feeding organ. The results show that for neither of the life stages nor the genotypes and sex affected food intake rate ( $p > 0.05$ ) (Figure S2B). To quantify the feeding volume, fluorescent beads were mixed with *E. coli* OP50 and fed to the nematodes (Bakowski *et al.* 2014). In all cases the mean fluorescent signal was at least as high in the males as in the hermaphrodites, showing that the average amount of ingestion within the nematodes was similar or males had higher ingestion (Figure S2C). The total amount of fluorescent signal was higher for adult hermaphrodites than for adult males, which corresponds to their larger body size (t-test,  $p < 0.05$ ) (Figure S2D). In conclusion, the ingestion rates cannot explain observed differences in viral susceptibility observed between sexes and strains.



**Figure 1 Sex-specific viral susceptibility** – A) Male and hermaphrodite nematodes were separated 48 hours post bleaching. Subsequently, single-sex populations were exposed to the OrV in liquid for 1 hour and grown on plates until 78 hours post bleaching. Then, nematode populations were collected and viral loads were obtained by RT-qPCR. B) Viral loads obtained for hermaphrodites and males of the strains N2, CB4856 and JU1580. Each dot represents a biological replicate ( $n = 10$  for N2 and CB4856,  $n = 8$  for JU1580). The white diamond shows the mean per combination of sex and strain. Statistically significant differences are indicated by an asterisk (bootstrap,  $p < 0.05$ ).

### Antiviral activity in males and hermaphrodites

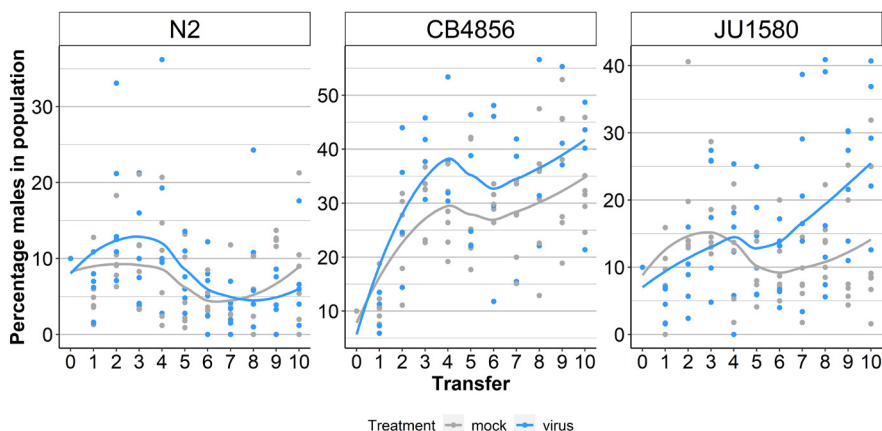
Males and hermaphrodites could have a different molecular response to OrV infection. Both the RNAi response and IPR were measured in males and hermaphrodites infected with OrV. Potent RNAi activity against the OrV resulting in viral siRNA production has been described (Ashe *et al.* 2013; Coffman *et al.* 2017). Small RNA sequencing of N2 males and hermaphrodites infected with OrV showed similar small RNA characteristics based on the observation of the main RNAi products: primary 23-nt long siRNAs (Figure S3A,B) and secondary 22-nt (antisense) siRNAs with a G at the 5'-end (Figure S3B). Additionally, we did not observe sex-based spatial clustering of 23-nt small RNAs on the viral genome for OrV RNA fragment 1 (encoding the RdRP). The siRNA hotspots on OrV fragment 2 differ between the sexes, but the number of siRNA reads are higher for hermaphrodites. The overall lower RNAi response in males does not explain their lower susceptibility to OrV infection (Figure S3C).

As the IPR is a transcriptional response, its activity was investigated by measuring gene expression in mock-treated and OrV-infected N2 adults. We selected samples with positively confirmed OrV infection by RT-qPCR. Male samples had a lower average viral load than hermaphrodite samples (Figure S4A). Adult males had lower expression of IPR regulators *pals-22* and *pals-25* than hermaphrodites (Figure S4B). Under mock conditions, expression differed for 26 IPR genes of which 24 were higher expressed in males (Figure S4B) (Dataset S3). OrV infection led to upregulation of 39 IPR genes for hermaphrodites and 10 IPR genes for males (Dataset S3). We also measured expression of the IPR regulators *pals-22* and *pals-25* and three IPR genes *pals-6*, *pals-14* and *eol-1* by RT-qPCR in untreated L4 nematodes (for N2, CB4856 and JU1580) to explore IPR expression upon viral exposure in our assay (Figure S5) (Dataset S3). We found that the IPR activator *pals-25* is higher expressed in males at this stage (linear model,  $p = 0.002$ ), suggesting other IPR genes might also be upregulated, matching previously collected data for N2 L4 males for the modENCODE project (Gerstein *et al.* 2010). Nevertheless, we did not detect increased expression for the limited set of (3/80) IPR genes we measured here (Figure S5). Together, our gene expression data collected for adults and the public modENCODE dataset for L4 nematodes indicate that IPR gene expression is constitutively high for males, which may reduce viral loads throughout the course of infection or potentially protect from OrV infection.

### *Viral infection changes mating dynamics*

The observed viral susceptibilities of males and hermaphrodites differed per strain (Figure 1). We asked whether this genotype-dependent viral susceptibility difference between the sexes could result in different mating dynamics per genotype. In particular, we hypothesized that N2 males, which are more resistant to the OrV, could become more dominant in infected populations. This hypothesis was tested by monitoring the male frequency in mock-treated and OrV infected populations for 10 generations. Starting populations contained 10% males of either N2, CB4856 or JU1580 nematodes in combination with hermaphrodites from the same genotype. Thus, genetic adaption of the nematodes via outcrossing to the OrV was not possible in this single-genotype set-up.

After observing the nematodes for 10 generations the frequency of N2 males did not increase in infected compared to mock infected populations which contradicted our previously stated hypothesis (mixed linear model,  $p = 0.52$ ) (Figure 2) (Dataset S4). However, we found that for CB4856 more males were present in infected populations (mixed linear model,  $p = 0.03$ ) (Figure 2), despite CB4856 males not showing a higher resistance to OrV than the hermaphrodites (Figure 1). For JU1580 viral presence increased male presence over time, contrary to male ratios in healthy populations (mixed linear model,  $p = 0.01$ ). After ten generations, JU1580 males made-up 25% of the infected population, compared to 14% in mock-infected conditions and CB4856 males were 5% more prevalent in infected than uninfected populations. Thus, we concluded that viral presence can change mating dynamics although it was not clear in which way male resistance contributes to this phenomenon.



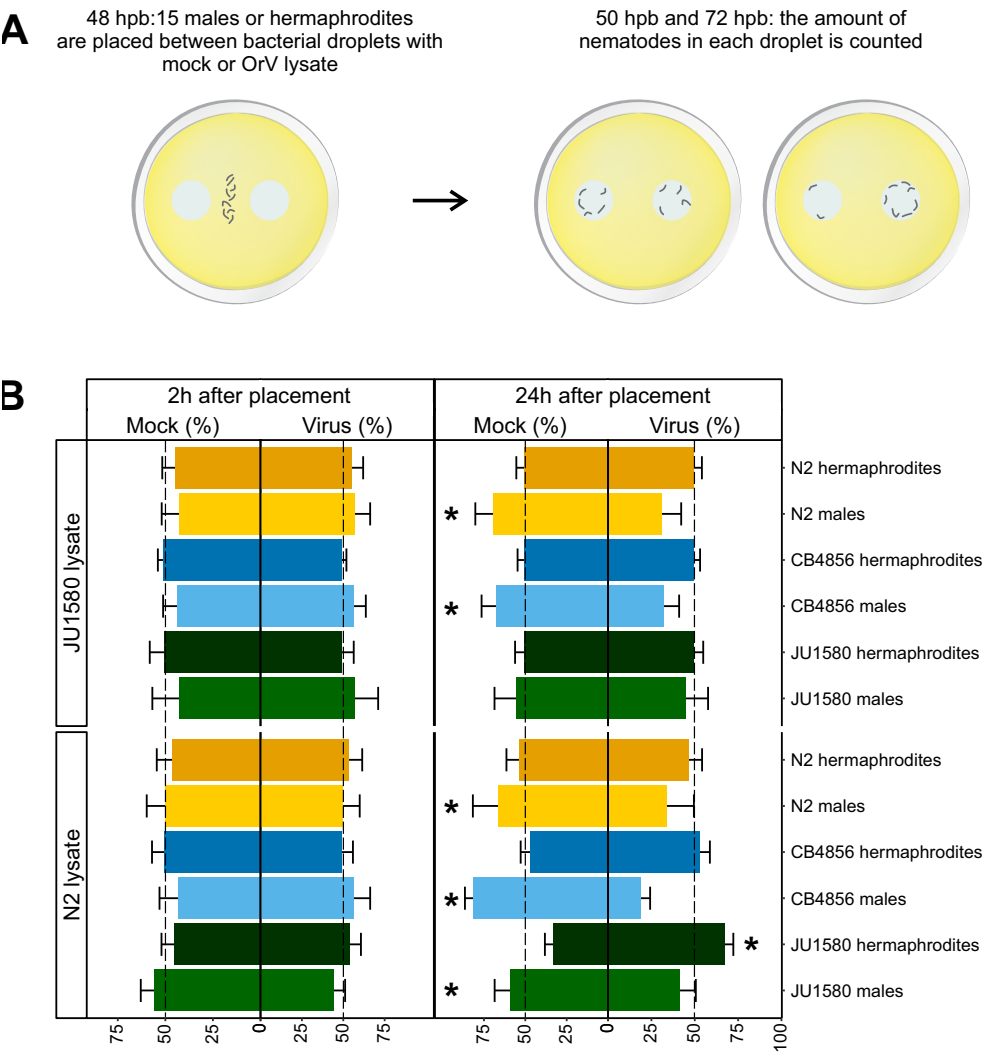
**Figure 2 Male frequency over ten generations** – The percentage of males in isogenic N2, CB4856 and JU1580 populations over 10 generations (30 days). Counts from mock populations are shown in grey, counts in OrV infected populations in blue. Each dot represents a technical replicate (plate with nematodes) and the technical replicates are equally divided over three biological replicates (having a different start date).

### *Males prefer healthy over infected hermaphrodites modulating sex*

To investigate if male *C. elegans* nematodes showed a preference for healthy hermaphrodites, a mating choice assay was designed. Healthy 48h-old males or hermaphrodites of the strains N2, CB4856 and JU1580 were given a choice between the lysate of mock-treated or the lysate of OrV-infected hermaphrodite populations (Dataset S5). Two types of lysate were used: lysates obtained from N2 or JU1580 nematodes. The number of nematodes on each spot was counted at 2 and 24 hours after placing them on the plate (nematode age is 50 and 72 hours respectively) (Figure 3A). Our results show that after 2 hours males and hermaphrodites do not show a preference for either the mock or virus spot on the plate (Figure 3B). However, after 24 hours adult males of all three genotypes display a significant preference for the N2-based mock lysate over the lysate of infected nematodes (Figure 3B). For the JU1580-based lysate the same trend was observed but was not significant for JU1580 males (Figure 3). Hermaphrodites did not show a preference for the mock lysate (Figure 3). Together, these results suggested that adult males were more attracted to healthy than infected hermaphrodites.

Males may distinguish infected- and mock-treated lysates by pheromones released by the hermaphrodites. Therefore, we used RB859 *daf-22* mutants, that do not excrete any male-attracting pheromones. We obtained the lysate of mock-treated and OrV-infected RB859 mutants for use in our choice assay (Supplementary Figure S6) (Dataset S5) (Von Reuss *et al.* 2012). Interestingly, the choice that males make between the lysates appears independent from pheromone signals, because we observed that adult males also favor uninfected over infected pheromone-free lysate of *daf-22* nematodes (chi-square test,  $p < 0.001$ ) (Figure S6). Additionally, also young adult hermaphrodites choose the mock lysate in this experiment (chi-square test,  $p < 0.001$ ), yet with a weaker preference than the males. Thus, males appear to have another way to distinguish uninfected from infected lysates.

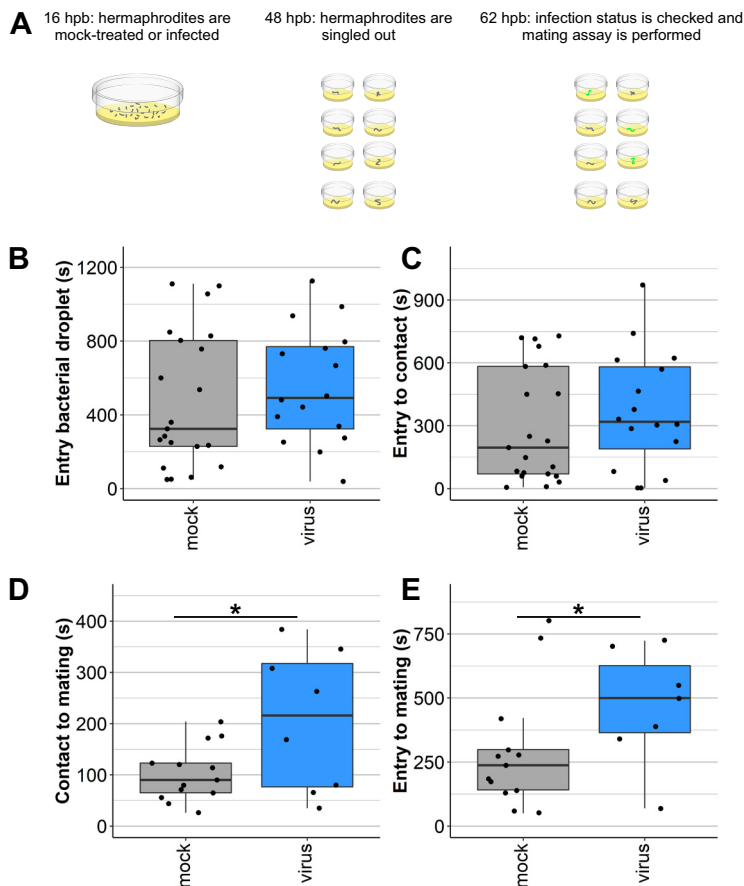




**Figure 3 Choice assay between mock- and OrV-infected nematode lysate** – A) Males and hermaphrodite of the genotypes CB4856, JU1580 and N2 were allowed to choose between a bacterial droplet containing the lysate of either a mock-treated or infected population of mixed-stage hermaphrodites. Fifteen male or hermaphrodite nematodes were placed in-between these droplets and were observed at 2 and 24h after placement. B) The percentage of nematodes that were observed in the mock or OrV droplet on the plate at 2 or 24 hours after placement. Only nematodes that were in either the mock or OrV droplet were counted. Mock and OrV lysates were made by lysing either N2 or JU1580 populations. The dotted line indicates where nematodes would not prefer either of the spots. Error bars indicate standard error of the mean. Nematodes that showed a significant preference for one of the spots are indicated with an asterisk (test,  $p < 0.05$ ).

To investigate if the preference for the lysate of healthy hermaphrodites translates into mating behavior, we performed a mating assay. Males were placed with a single (un)infected, young adult hermaphrodite (Figure 4A). Subsequently, mating behavior of N2 males towards mock- or OrV-infected nematodes was quantified by filming their movements for 20 minutes (Movie S1) (Dataset S6). We recorded the time it took for males to 1) enter the bacterial droplet, 2) touch the hermaphrodite and 3) mate (Figure 4B-E). The videos show that when the hermaphrodite was infected, males took longer between the first contact and mating (t-test,  $p = 0.023$ ) (Figure 4D). The overall process from entering the bacterial droplet to mating took 45% longer for infected than uninfected nematodes (t-test,  $p = 0.003$ ) (Figure 4E).

Behavioral differences between healthy and infected hermaphrodites could determine mating efficiency. Therefore, hermaphrodite movements were measured before and after contact with the male (Dataset S6). OrV infected hermaphrodites moved faster than healthy hermaphrodites even before contact with the male (t-test,  $p = 0.04$ ) (Figure S7A). Nevertheless, the speed of movement did not appear to (strongly) determine effective mating in our assay, because 1) both fast and slower moving nematodes were mated and 2) there was only a weak correlation between velocity and time from the first contact to mating (linear model,  $R^2 = 0.15$ ,  $p = 0.07$ ) (Figure S7B). Hermaphrodites could also avoid male mating attempts by crawling away after contact with the male. We investigated the change in speed after male contact, but did not observe a stronger avoidance response for infected than in healthy hermaphrodites (t-test,  $p = 0.63$ ) (Figure S7C). Concluding, both attraction assays performed here indicate that male *C. elegans* nematodes would sooner mate with uninfected over OrV infected hermaphrodites. In cases where OrV infection affects the progeny production of hermaphrodites this altered mating behavior could determine male percentages and outcrossing in infected populations.



**Figure 4 Mating assay with mock-treated and OrV-infected hermaphrodites** – A) For the mating assay hermaphrodite populations of the reporter strain ERT54 (*pals-5::GFP*) were either mock-treated or infected with the OrV. Once the hermaphrodites were 48 hours old they were placed on individual plates for 14 hours. Then, nematodes were checked for expression of GFP indicating successful infection. Successfully infected individuals and mock-treated individuals were placed in a camera set-up. A young adult male was added at a set distance and mating behavior was observed for 20 minutes. B) The time recorded for the male to enter the bacterial droplet. C) The time recorded between the male entering the bacterial droplet and the first physical contact between male and hermaphrodite. D) The time recorded between the first contact and mating. E) The time recorded between the male entering the bacterial droplet and mating. For B-E statistically significant samples are indicated with an asterisk (t-test,  $p < 0.05$ ). The number of technical replicates (nematodes filmed) was  $n = 64$  for mock-treated hermaphrodites and  $n = 66$  for OrV infected hermaphrodites. The technical replicates were divided over 5 biological replicates (different days). Each dot represents a technical replicate and when a certain behavior was not observed in the 20-minute timeframe less dots than technical replicates are shown in the graph.

## Discussion

Pathogens have various manners to perturb mating system dynamics. Unravelling the mechanisms by which pathogens interfere in the mating dynamics of their hosts provides insight in the way pathogens affect the world around them. Here, we studied the differences in OrV infection of males versus hermaphrodites of the nematode *C. elegans*. We found that males of the strain N2 were less susceptible than hermaphrodites and that males had higher IPR activity. Viral infection changed the mating dynamics in infected population. Infected hermaphrodites were less attractive partners for males and viral presence can increase male presence in a population. Our findings provide a novel perspective for how pathogens shape mating dynamics in isogenic populations of *C. elegans*.

*C. elegans*-pathogen co-evolution experiments illustrate that multiple factors contribute to outcrossing in *C. elegans* (Lopes *et al.* 2008; Lynch *et al.* 2018; Morran *et al.* 2009, 2011). Mated hermaphrodites produce 50% male offspring and unmated hermaphrodites only have 0.2% male offspring (Anderson *et al.* 2010; Cutter *et al.* 2019), therefore an increase in *C. elegans* males implies outcrossing takes place in that population. Presence of the bacterial pathogen *Serratia marcescens* increases male frequencies in *C. elegans* populations and their outcrossed offspring is more fit to counteract infection than offspring from selfing hermaphrodites (Morran *et al.* 2011). On the other hand, the bacterial pathogen *Bacillus thuringiensis* decreased outcrossing in *C. elegans* populations. Interestingly, males suffer disproportionally from infection to *B. thuringiensis*, yet stayed stably present at rate of about 10%. This suggests that despite the direct disadvantage of susceptible males, the overall populations still indirectly benefits from outcrossing (Masri *et al.* 2013). However, male frequencies can also increase under stress in isogenic populations that cannot benefit from fitness-increasing genetic recombination (Lynch *et al.* 2018; Morran *et al.* 2009). In one of these cases, increased male presence is linked to their higher resistance to starvation-induced stress. Our data demonstrates that males prefer mating with healthy (instead of OrV infected) hermaphrodites. In highly susceptible populations OrV infection can reduce offspring numbers, but not all individuals in the population become infected (Ashe *et al.* 2013). Our findings advocate males would mostly mate with healthy hermaphrodites within infected populations. These healthy mated hermaphrodites have the highest offspring numbers and produce 50% males. On the other hand, unmated hermaphrodites with lower brood sizes produce only hermaphrodites. Over time, this would lead to an increase in male presence in infected populations which may explain the observations we made in JU1580 and CB4856 populations.

Mating behavior depends on more than just the presence of a pathogen. In general, *C. elegans* males are characterized by their inefficient mating behavior and N2 males even belong to the poorest outcrossers for this species (Garcia *et al.* 2007; Wegewitz *et al.* 2008). In contrast, CB4856 nematodes exhibit relatively efficient mating behavior (Wegewitz *et al.* 2008). Even though our results suggest that males may become more prevalent in infected populations by selecting the best partners, they will need to mate efficiently, or else male numbers will dwindle quickly. Thus, the inefficient mating behavior of N2 males may help to explain why viral presence did not lead to an increase in males in infected N2 populations. Additionally, the mating behavior of hermaphrodites may also change when they are infected. Infected individuals for many species behave differentially which can lead to divergent mating frequencies (Beltran-Bech & Richard 2014; Burand *et al.* 2005; Paciência *et al.* 2019). Under standard conditions *C. elegans* hermaphrodites can avoid (costly) mating among other by keeping on moving during the mating attempts of the male (Chasnov & Chow 2002; Garcia *et al.* 2007; Woodruff *et al.* 2014). Non-favorable mating conditions could have occurred more often for infected hermaphrodites since we have observed that they moved quicker even before the male has made contact with them. Although we could not conclude that hermaphrodite behavior affects mating from our data, we can also not fully reject this hypothesis.

Molecular cues are commonly used to distinguish infection from non-infected potential partners (Beltran-Bech & Richard 2014). *C. elegans* males are normally attracted by hermaphrodite-produced ascarosides (Chasnov *et al.* 2007; Pungaliya *et al.* 2009; Srinivasan *et al.* 2008). However, our choice experiment using an ascaroside-lacking *daf-22* mutant (Von Reuss *et al.* 2012) suggests that ascaroside signals do not play an essential role for males to distinguish healthy from infected hermaphrodites. In addition, publicly available transcriptional data shows that the main genes in the ascaroside-producing pathway (*acox-1*, *maoc-1*, *dhs-28* and *daf-22*) remain equally expressed in infected nematodes ([Chapter 3](#)) (Chen *et al.* 2017; Sarkies *et al.* 2013). Hence, other ways of (chemical) communication might contribute to the mating preference for uninfected hermaphrodites by *C. elegans* males. For example, *C. elegans* can recognize a small RNA to avoid a pathogenic bacterium (Kaletsky *et al.* 2020), which illustrates another way by which infected nematodes may be recognized.

Our findings indicate that sex-based susceptibility differences occur in a genotype-dependent manner based on the viral loads of three genetically distinct strains. Of note, the viral loads in hermaphrodite populations are likely underestimations since these populations contained the first eggs at the moment of collection. Eggs cannot be infected by the OrV (Félix *et al.* 2011), but embryos do express the reference genes that are used to normalize the viral expression. Previous studies indicated IPR gene expression is enhanced in L4 stage males (Gerstein *et al.* 2014; Leyva-Díaz *et al.* 2017) and we found that some IPR genes are also higher expressed in adult males. Notably, males are also more resistant to an intracellular fungus, although

the connection with the IPR has not been studied (Van Den Berg *et al.* 2006). Not all genes in the IPR network were more active in virus infected males than hermaphrodites. The IPR upregulates a large network of 80 genes against a broad range of stressors (Chen *et al.* 2017; Reddy *et al.* 2017, 2019), but the biochemical reaction to counteract the stress appears more specific. For example, the IPR proteins that promote thermotolerance do not enhance pathogen resistance (Panek *et al.* 2020) and *drh-1* is only required for induction of the antiviral IPR (Sowa *et al.* 2019). Furthermore, the IPR genes in JU1580 and CB4856 hermaphrodites respond different to OrV infection than in N2 ([Chapter 3](#)) (Sarkies *et al.* 2013); this may relate to similar viral susceptibilities for JU1580 and CB4856 between males and hermaphrodites. Thus, a more thorough understanding of the IPR is necessary to draw firm conclusions about IPR involvement in sex-specific viral susceptibility.

The importance of *C. elegans* males remains a topic of debate with most studies indicating their disadvantages and only a few providing them with a potential ecological role (Cutter *et al.* 2019). *C. elegans* males are typically rare in nature, further contributing to the idea that they may be less relevant for the species (Barrière & Félix 2005; Félix & Duveau 2012; Richaud *et al.* 2018; Sivasundar & Hey 2005). Nevertheless, males may prove beneficial under stress conditions, including viral infections, and could thereby play a supporting role in the natural history of *C. elegans*. Here, we have shown how the presence of an intracellular pathogen shifts mating behavior and shapes flexible outcrossing in this species.

## Materials and methods

### *Caenorhabditis elegans strains and culturing*

*C. elegans* wild isolate strains N2, CB4856 and JU1580 were used. Mutant strain RB859 *daf-22(ok693)* was obtained from the *Caenorhabditis* Genetic Center. The strain ERT54 (*jyIs8[pals-5p::GFP, myo-2p::mCherry]* X) was a kind gift from Emily Troemel (Reddy *et al.* 2017, 2019). The strains were grown on nematode growth medium (NGM) plates seeded with *Escherichia coli* OP50 as food source (Brenner 1974). Male populations were maintained weekly by crossing adult males with L4 hermaphrodites in a 3:1 ratio. Experiments were started by transferring a starved population to a fresh NGM dish followed by bleaching after the adults were egg-laying (3 days after transfer at 20°C, 4 days after transfer at 16°C) (Brenner 1974). All experiments were performed at 20°C unless indicated otherwise.

### *Orsay virus stock and mock lysate*

Standard used Orsay virus (OrV) stock was prepared as previously described (Félix *et al.* 2011). One modification was made, namely to use modified NGM medium with 34g/L agar (mNGM) plates to grow the nematodes as this prevented the nematode from burrowing into the agar. Briefly, persistently infected JU1580 starved nematodes were washed of 100 9cm diameter mNGM plates and flash-frozen in liquid nitrogen. The lysate was collected and filtered through a 0.2µm filter and stored at -80°C until used. The infectivity of the OrV stock was tested by a dose-response experiment in JU1580 (Sterken *et al.* 2014).

For the choice-assay experiments additional Orsay virus and mock stocks were created. Mock stocks were used in these experiments, because then both mock and OrV stocks contain nematode lysate that was expected to affect attraction of males. Mock and OrV stocks were made in parallel to minimize batch effects by the nematodes used. All animals were grown at 16°C. First, Orsay virus JU1580 mock stock was prepared by bleaching a persistently infected JU1580 population to remove the OrV. After the bleached population was starved it was transferred to 100 new 9cm mNGM plates. Subsequently, the same protocol as for obtaining the OrV stock was followed (Félix *et al.* 2011; Sterken *et al.* 2014). Second, N2- and RB859-based mock and OrV stocks were made by infecting and lysing N2 or RB859 nematodes instead of JU1580 nematodes. To infect the N2 or RB859 nematodes 100µL of previously obtained JU1580 OrV stock was added to a 9cm mNGM plate with a proliferating N2 population. Once this population was starved, the plate was divided over 100 fresh 9cm mNGM plates of which the populations were flash-frozen upon starvation. Afterwards the standard protocol was followed (Félix *et al.* 2011; Sterken *et al.* 2014). Third, young adult mock and OrV stocks were made by infecting N2 or RB589 L1 nematodes of collected from



10 9cm mNGM plates (24 hours post bleaching) with either 200 $\mu$ L RB589 mock or OrV stock (obtained from starved populations as described above) according to the protocol described by (Sterken *et al.* 2014). Populations were washed five times with M9 (instead of the standard three times) after exposure to the mock or OrV stock for 1h. Populations were collected when the first eggs appeared on the plate (90 hours post bleaching for N2 nematodes, 104 hours post bleaching for RB589 nematodes). The ‘young adult’ N2 and RB589 stocks were used in the *daf-22* choice assay experiment.

### *Orsay virus infection experiments*

For the single-sex experiments male and hermaphrodite nematodes were separated ~44 hours post bleaching by placing them on fresh plates. After 200 nematodes were handpicked per plate the populations were washed into an Eppendorf tube and infected 48 hours post bleaching as described previously (Sterken *et al.* 2014). Mock infections were performed by adding M9 instead of OrV unless indicated that a mock lysate (lysate of uninfected nematodes) was used. The experiment was performed in technical duplicate, therefore every sample contained ~400 nematodes. The samples for small RNA (sRNA) sequencing contained ~600 nematodes. However, for some experiments, not enough males (less than 400) were present on the plates to pick males from and in these cases a total of 200 male nematodes were picked per sample (and also for the hermaphrodite samples within the same experimental batch). Infected populations were collected 30 hours post infection (78 hours post bleaching). Viral loads were obtained as described by (Sterken *et al.* 2014).

As the viral loads for single-sex infections did not follow normal distributions and some samples were zero-inflated, a nonparametric bootstrap approach was used to estimate sampling distributions and perform statistical tests (Kulesa *et al.* 2015). Bootstrap samples were drawn 10,000 times and used to calculate the bootstrap test statistic (the simulated difference in mean viral load of males and hermaphrodites). Each bootstrap sample was drawn by randomly selecting a set of *n* experimentally observed viral loads (*n* being the number of observations made for both males and hermaphrodites; 20 for N2 and CB4856, 16 for JU1580) from the full dataset with replacement (meaning that the same observed viral load can be drawn multiple times). Every bootstrap sample draws half of the observations from male samples and half of the observations from hermaphrodite samples. The bootstrapped p-value was calculated by dividing the number of bootstrapped test statistics that were greater than the experimentally observed test statistic by the total number of bootstrap test statistics of 10,000. The bootstrap analysis has been performed in R (version 4.0.2).

For the mixed-sex experiments, 30 L4 hermaphrodites and 90 adult males were transferred to a plate that was incubated overnight. Hermaphrodites in absence of males were picked as a control. Subsequently, 30 male- or self-fertilized nematodes were transferred to a new

plate where they were allowed to lay eggs for 6 hours before the hermaphrodites were removed from the plate. The resulting eggs were incubated for 20 hours after which they were infected with the OrV as previously described (Sterken *et al.* 2014). Nematodes were collected as young adults 30 hours post infection (56 hours post start of egg-laying). Viral loads were obtained as described by (Sterken *et al.* 2014) and mixed and hermaphrodite-only samples were compared with a student t-test with equal variances.

### *RNA isolation*

RNA isolation was performed using the Maxwell<sup>®</sup> 16 AS2000 instrument with the Maxwell<sup>®</sup> 16 LEV simply RNA Tissue Kit (Promega). The recommended protocol was followed, except for adding 10mg of proteinase K per sample after addition of the lysis buffer. The lysis was performed in a shaker for 10 minutes at 65°C (1000rpm) (Eppendorf). RNA sample quantity and quality were measured by using NanoDrop (Thermo Scientific).

### *Gene expression measurements by RT-qPCR*

cDNA was made from 1µg of RNA using the GoScript Reverse Transcriptase kit (Promega) and following the recommended protocol with random hexamers (Thermo Scientific). Gene-expression measurements were performed on cDNA of ~100 untreated male or hermaphrodite nematodes (collected 48 hours post bleaching). Samples were collected for 5 biological replicates. Gene expression was quantified by RT-qPCR using custom designed primers (Dataset S7) that overlap at least one exon-exon border to prevent unintended amplification of any remaining DNA. RT-qPCR was performed on the MyIQ using iQ SYBR Green Supermix (Biorad) and the recommended protocol. Primer efficiencies were checked by testing dilution ranges for N2 and JU1580 male populations. Correct primer annealing to the CB4856 genome was tested *in silico* (Thompson *et al.* 2015). Gene expression in each sample was quantified for the gene of interest and two reference genes (Y37E3.8 and *rpl-6*) in technical duplicate. Gene expression was determined according to (Sterken *et al.* 2014) and the effect of sex and strain were determined per gene via a linear model

$$Y = S + G + S \times G + \varepsilon$$

with  $Y$  being the  $\log_2$  gene expression.  $Y$  was explained over sex ( $S$ ; male or hermaphrodite), genotype ( $G$ ; N2, CB4856 or JU1580), the interaction between sex and strain and the error term  $\varepsilon$ . When gene expression for a certain combination of gene and sample remained under the detection limit by RT-qPCR ( $C_t > 40$ ) that measurement was excluded from the analysis.

### Whole-genome gene expression by microarray

For whole-genome gene-expression measurements, microarrays were used on 8 biological replicates of N2 mock-treated or OrV-infected males and hermaphrodites (obtained via the ‘single-sex’ infection protocol). Only samples where the OrV could be detected by RT-qPCR were used for this analysis. Gene Expression Microarray 4X44K *C. elegans* V2 slides were used for the microarrays (Agilent Technologies, Santa Clara, CA, USA). Microarrays were performed based on the ‘Two-Color Microarray-Based Gene Expression Analysis; Low Input Quick Amp Labeling’ protocol, version 6.0 from Agilent (Agilent Technologies, Santa Clara, CA, USA), starting from step five. The microarrays were scanned (Agilent High Resolution C Scanner) and extracted using the Agilent Feature Extraction Software (version 10.7.1.1) with the recommended settings. Normalization was performed separately per sex using the Limma package in R (version 4.0.2). The Loess method was used for within-array normalization and the Quantile method for between array normalization. The obtained single channel normalized intensities were  $\log_2$  transformed. The  $\log_2$  intensities were used in further analysis using the package Tidyverse (1.3.0) in R (version 4.0.2). Gene expression of control genes that should either similar (*rpl-6*) or differentially (*tra-1*, *fem-3*) expressed between the sexes were checked to validate the data.

Basal gene expression differences in IPR gene expression were determined by selecting transcriptional data of the mock-treated samples ( $\log_2$  intensities) and running the following linear model

$$Y_i = S + \varepsilon$$

where  $Y$  is the  $\log_2$  normalized intensity of spot  $i$  (1, 2, ..., 45220) was explained over sex ( $S$ , male or hermaphrodite), and the error term  $\varepsilon$ . A significance threshold was by the *p.adjust* function, using the Benjamini & Hochberg correction ( $FDR < 0.05$ ) (Benjamini & Hochberg 1995).

The effect of treatment on the IPR gene response was analyzed by selecting transcriptional data ( $\log_2$  intensities) for the 80 IPR genes (as described by (Reddy *et al.* 2019)) and running the following linear model per sex

$$Y_i = T + \varepsilon$$

where  $Y$  is the  $\log_2$  normalized intensity of spot  $i$  (1, 2, ..., 109) was explained over treatment ( $T$ , mock or infected), and the error term  $\varepsilon$ . A common significance threshold was determined after combining the datasets of both sexes and using the *p.adjust* function with the Benjamini & Hochberg correction ( $FDR < 0.05$ ) (Benjamini & Hochberg 1995).

### *Small RNA sequencing*

Samples for small RNA sequencing contained at least 1µg of RNA with a concentration > 50ng/µL. Samples were sequenced by DNBseq™ Small RNA sequencing (BGI) using either a 5'-dependent or 5'-independent protocol. The protocol type refers to ligation of the 5'-adaptor. The 5'-dependent protocol detects only monophosphorylated small RNA strands, whereas the 5'-independent protocol also allows detection of small RNA strands that had 5'-triphosphate or 5'-capped modification. The read data was aligned to the Orsay virus genome (GenBank: HM030970.2 and HM030971.2) using Bowtie2 via the public server at usegalaxy.org to map the data (Afgan *et al.* 2018; Langmead & Salzberg 2013). Reads that aligned to the OrV genome were subsequently analyzed in R (version 4.0.2). Reads with a mapping quality lower than 40 were excluded from analysis.

### *Pumping rate counts*

Pumping rates were counted for L4 (48 hours post bleaching) and young adult (72 hours post bleaching) N2, CB4856 and JU1580 nematodes using a Leitz Greenough microscope and a counter. The pumping rates were counted twice for 20s per nematodes and the average of both measurements was used in further analysis. Pumping rate counting was performed at room temperature (~20°C) and the experiment was performed over 5 independent biological experiments counting 6 nematodes of each genotype/sex per experiment.

### *Fluorescent bead accumulation*

To estimate the volume of food intake nematodes were either exposed to the fluorescent beads in liquid or whilst feeding on the plate. The plate feeding assay was performed as previously described (Bakowski *et al.* 2014). In short, nematodes were incubated on the plate for 30 minutes at 25°C before feeding was halted by placing the plates on ice. The NGM plates were incubated with a mixture of 30 µL *E. coli* OP50 and 5µL of Fluoresbrite® Polychromatic Red Microspheres (1.00µm particles,  $4.55 \cdot 10^{10}$  particles/ml) (Polysciences, Inc.). The liquid exposed nematodes were incubated in a mixture of 30µL *E. coli* OP50, 5µL of Fluoresbrite® Polychromatic Red Microspheres (1.00µm particles,  $4.55 \cdot 10^{10}$  particles/ml) (Polysciences, Inc.), 105µL 0.1% NGM and 15µL OrV stock for 1h at room temperature to mimic the OrV infection assay. Then nematodes were washed once with 0.25mM levamisole in M9 to stop the feeding and remove the surplus of fluorescent beads. For both assays nematodes were fixed in 0.25mM levamisole in M9 and nematodes were imaged using the Axio Observer Z1m inverted microscope (Zeiss). The exposure time for the red fluorescent channel ( $\lambda=453$  nm) was fixed at 440ms. These experiments were performed independently for 5 times. Each biological replicate contained between 10-50 nematodes for imaging.

Images were analyzed using the ImageJ software (v1.51f, National Institutes of Health). The total surface area of the nematode was selected by using the Wand Tool on the brightfield image. Images for which the surface area of the nematode could not be selected due to nearby presence of, for example, a bubble were excluded from the analysis. Subsequently, the area size and minimum-, maximum- and mean fluorescence were determined for the selected area. The data was tested for normality by a Kruskal-Wallis test and subsequently analyzed by the two-sample Wilcoxon signed-rank test in R (version 4.0.2).

### *Male frequency counts*

Before the start of the experiment 30 male and 10 hermaphrodite L4 nematodes were crossed to obtain isogenic male populations for each genotype (N2, CB4856 and JU1580) and grown until starved at 20°C. From the same plate 10 hermaphrodites were picked that were also grown to starvation. Starved plates were transferred onto 10 9cm mNGM plates per genotype/sex combination at incubated at 20°C. Three days later nematodes were bleached and again incubated at 20°C. When nematodes reached the L4 stage (48 hours post bleaching) 15 males (from the mixed male-hermaphrodite plates) and 135 hermaphrodites (from the hermaphrodite only plates) per genotype/sex combination were placed onto a new plate. For each genotype/sex combination 4 plates were made of which 2 were mock infected by adding 100µL mock stock lysate and 2 were OrV infected by adding 100µL OrV stock lysate. The experiment was performed in biological triplicate resulting in six technical replicates per genotype/treatment combination. After the preparation of the plates the experiment was blinded by removing the name tags and replacing them for a number by a colleague not involved in the experiment. The plates were incubated at 20°C throughout the experiment.

Every 3 days the number of adult males and total adult nematodes was counted using a dissecting microscope. Counting started 6 days after bleaching and continued for 10 transfers. Plates were divided into 8 parts with the use of a pen and parts were counted till a total of at least 100 adult individuals were observed. After counting, the nematodes were washed off the plate in 2mL M9 and for each population approximately 100 nematodes were transferred by pipetting to a fresh mNGM plate. The remaining nematodes were pelleted and flash frozen for RNA isolation after transfer number 1, 5 and 10 and these samples were checked for OrV presence as a control. Based on this analysis two technical replicates of OrV-infected CB4856 were excluded from the analysis as they appeared to have cleared the infection.

After the counting, the samples were unblinded and analyzed using a pre-written R script. Statistical analysis was performed using a linear mixed effect model for each genotype separately, explaining the observed male frequency over the fixed effects transfer, treatment and interaction between transfer and treatment. Replicate was used as a random variable. The analysis was performed in R (version 4.0.2) using the packages ‘lmerTest’ and ‘lme4’ and the variance was analyzed using Satterthwaite’s method (Bates *et al.* 2015; Kuznetsova *et al.* 2017; Satterthwaite 1941).

### *Choice assay*

Choice assay experiments were based on chemical and bacterial choice assays described previously (Bargmann *et al.* 1993; Zhang *et al.* 2005). Two droplets of 15µL *E. coli* OP50 were placed on a 6cm NGM dish after which the plates were incubated at room temperature for two days (~20°C). Before the start of the assay either 15µL OrV stock or 15µL mock stock lysate was pipetted on top of the droplets. After the plates dried 15 N2, CB4856 or JU1580 L4 males or hermaphrodites (48 hours post bleaching) were placed in the middle between the droplets. After 2h and 24h the number of nematodes in each droplet was noted. Data was collected for five biological replicates. Per biological replicate the behavior of 45 nematodes for each genotype/sex combination was tested using 3 separate plates with 15 nematodes each. The data in this experiment was analyzed by combining the total counts of nematodes in the mock or OrV droplet and performing a Chi-square test in R (version 4.0.2).

### *Mating assay*

N2 male and ERT54 populations were synchronized by bleaching. 16 hours post bleaching either 50 µL OrV stock or M9 was added to an ERT54 population on a 9m NGM plate. 48 hours post bleaching N2 males were placed on 9cm NGM plate without hermaphrodites. At the same time, (un)infected ERT54 nematodes were individually transferred to 3cm NGM plates. Next, young adult ERT54 nematodes (62hpb) were inspected for *pals-5::GFP* expression under an Axio Observer Z1m inverted microscope (Zeiss). Animals with high intestinal *pals-5::GFP* expression were selected as positively infected nematodes, whereas animals without this expression were not used in the assay. Subsequently, N2 males from the male-only plate were added to either a mock-treated or an OrV-infected ERT54 individual. Males were placed at 0.9cm distance from the middle of the bacterial food droplet. Directly after adding the male to the plate filming for 20 minutes started. Movies captured 1 frame per second and were made at room temperature (~20°C) using 4 USB cameras simultaneously (Conrad, cat. nr. 191341–62) using the video capturing program VirtualDub (version 1.10.2) and ImageJ (version 1.52) (Schindelin *et al.* 2012; Schneider *et al.* 2012). Filming was performed using 62-66 hour old nematodes that did not start egg-laying yet.



Movies were inspected manually and for each of the movies three characteristics were reported: 1) the time it took before the male entered the bacterial droplet, 2) the time it took before the male first touched the hermaphrodite, 3) the time it took before mating started. A mating was noted when the males tail stayed attached to the hermaphrodites vulva for at least 3 frames and mating attempts stopped afterwards (based results by (LeBoeuf *et al.* 2014)). The time the mating started was noted as the time to mating. Statistical analysis of the data was performed with a student t-test assuming equal variances.

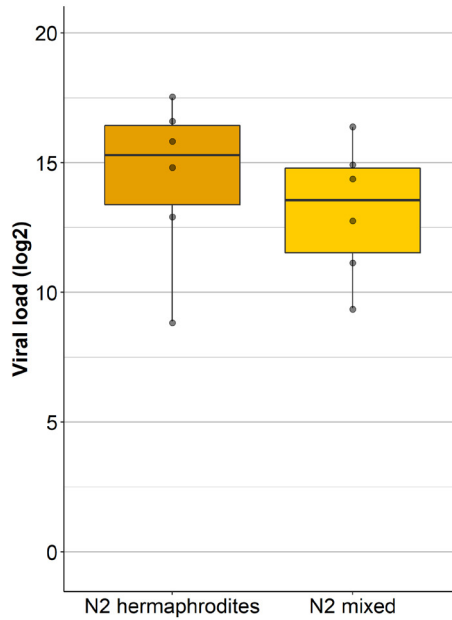
Hermaphrodite movements were investigated for 30 seconds before and 90 seconds after the first contact between male and hermaphrodite. A longer timeframe to measure the velocity of hermaphrodites after contact with the male was chosen, because males did often not directly start mating attempts and thus it may take longer than 30 seconds to see an effect of male mating attempts on the hermaphrodite velocity. Tracking was performed using the plugin 'Manual Tracking' in ImageJ (version 1.52) (Schindelin *et al.* 2012; Schneider *et al.* 2012). The movement data was subsequently analyzed in R (version 4.0.2). The velocity of the nematodes was compared using student t-test with equal variances and the correlation between 1) the time passed between first contact of the nematodes and mating and 2) the average velocity of the hermaphrodite was calculated by a general linear model.

### *Data availability*

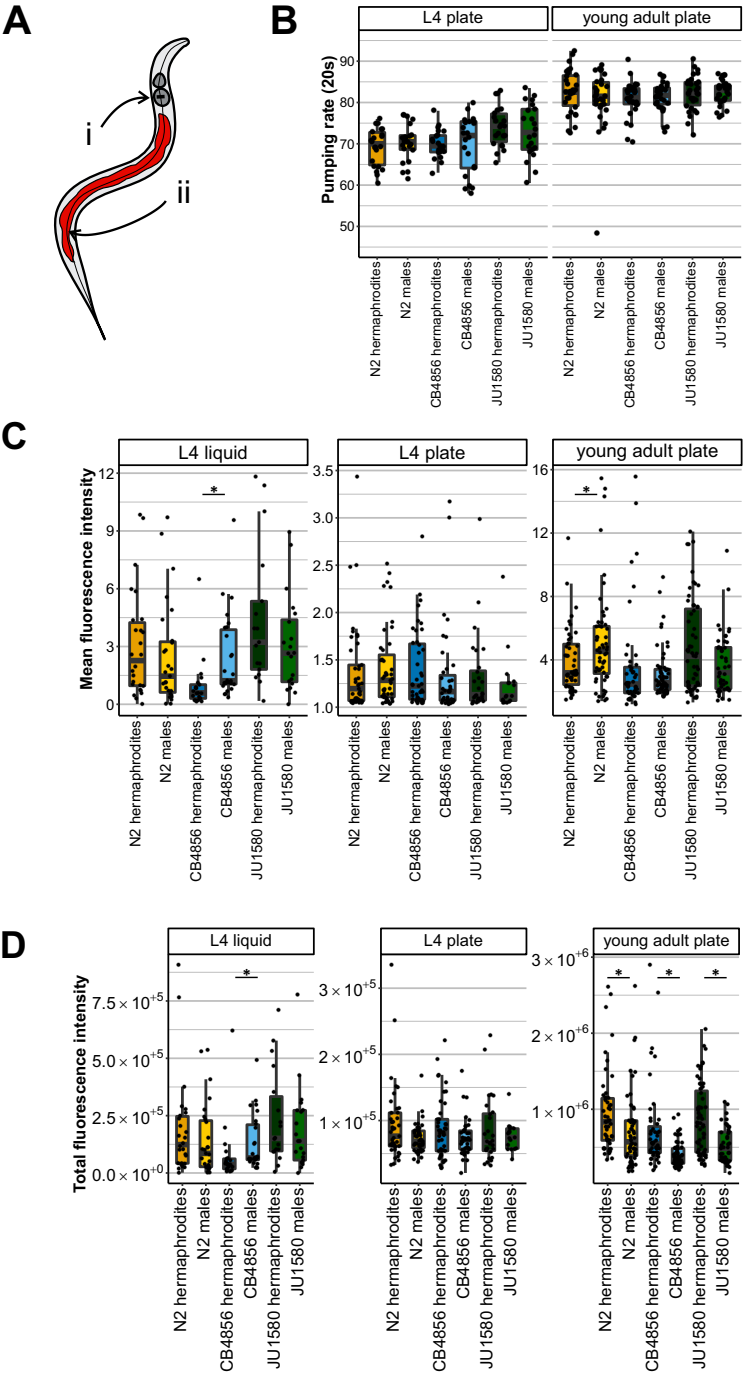
All custom written scripts in R can be accessed via [https://git.wur.nl/lisavansluijs/Male\\_hermaphrodite\\_viral\\_susceptibility](https://git.wur.nl/lisavansluijs/Male_hermaphrodite_viral_susceptibility). Microarray datasets are deposited at ArrayExpress (E-MTAB-9561).

## **Acknowledgments**

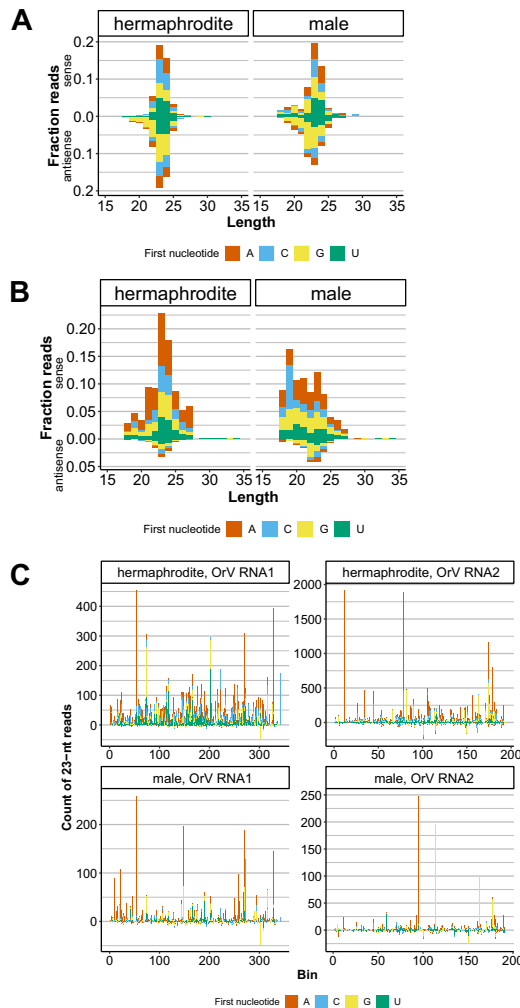
The authors thank Emily Troemel for sending us reporter strains and advice on food intake assays. We also thank Erik Andersen, Frank Schroeder and Marie-Anne Félix for advice about choice assay and mating experiments. Giel Göertz is thanked for sharing his knowledge about (processing) small RNA sequencing data. Katharina Jovic is thanked for advice about the VirtualDub script and camera set-up to simultaneously film nematode movements. Furthermore, we thank the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440), for providing strains.

**Supplementary information**

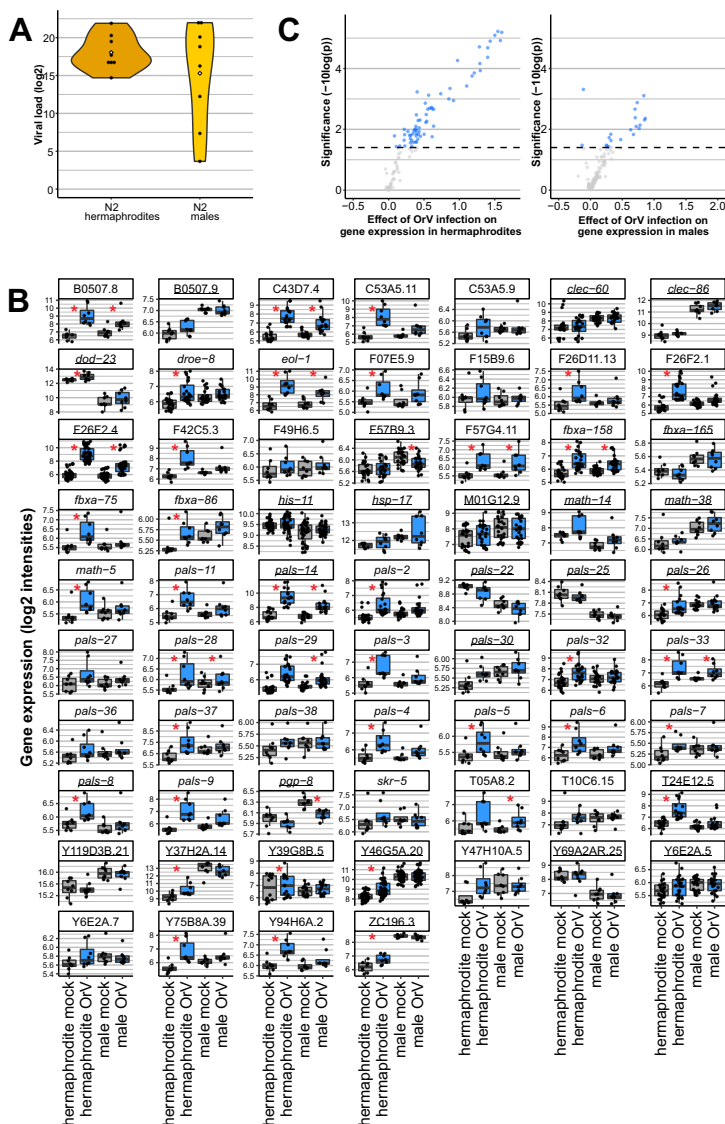
**Figure S1 Viral susceptibility of hermaphrodite and mixed-sex populations** – The viral susceptibility of hermaphrodite populations and mixed male and hermaphrodite populations 56 hours post bleaching (30 hours post infection). Infected populations were obtained by infecting the offspring of mated or selfing hermaphrodites.



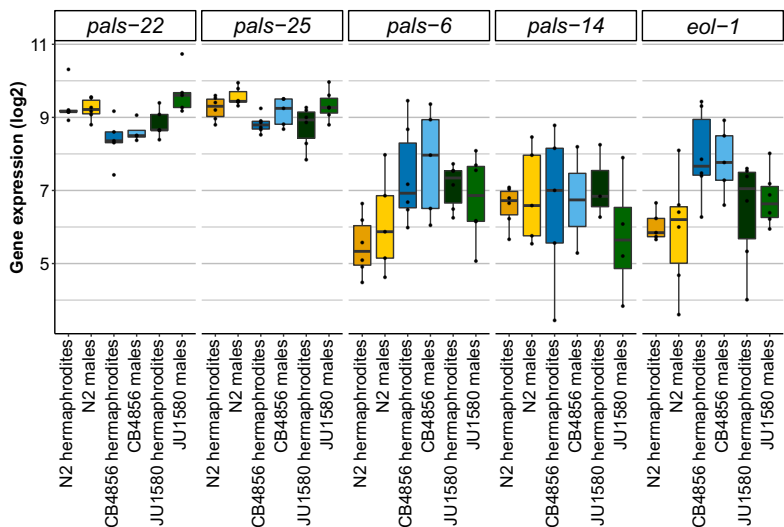
**Figure S2 Food intake quantification for males and hermaphrodites** – A) Food intake in males and hermaphrodites of the strains N2, CB4856 and JU1580 was quantified by i) measuring pumping rates by observing pharyngeal movements and ii) measuring accumulation of red fluorescent beads. The accumulation of red fluorescent beads was observed for L4 nematodes feeding in liquid (to mimic the viral exposure conditions) and L4 and young adult nematodes on the plate (to mimic incubation conditions). Pumping rates were counted for L4 and young adult nematodes feeding on the plate only. B) Pumping rates for N2, CB4856 and JU1580 males and hermaphrodites. C) Mean fluorescence measured for N2, CB4856 and JU1580 males and hermaphrodites. D) Total fluorescence measured for N2, CB4856 and JU1580 males and hermaphrodites. Statistical differences between the sexes are indicated with an asterisk ( $p < 0.05$ ). Each dot represents a technical replicate (observed nematode) which were divided over 5 biological replicates.



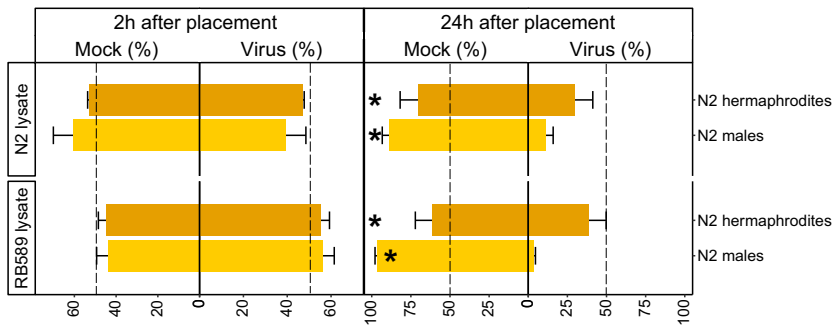
**Figure S3 Small RNA sequencing data for infected *N2* males and hermaphrodites** – Small RNA profile for *N2* males and hermaphrodites showing small RNA between 15 and 35 nucleotides in length that match the OrV genome. The fraction of sense reads is depicted on the positive y-axis, whereas antisense reads are shown on the negative y-axis. A) Small RNA sequencing profiles for samples that were sequenced in a 5'-dependent manner; only detecting monophosphorylated small RNAs. This data does not show secondary siRNAs as these are triphosphorylated in *C. elegans*. For these two samples males displayed a lower viral load than hermaphrodites. B) Small RNA sequencing data profiles for samples sequenced in a 5'-independent manner, which allowed for detection of triphosphorylated siRNAs. For these two samples males and hermaphrodites had comparable viral loads. C) Small RNA mapped to the OrV genome. These samples show clustering of 23-nt RNAs (primary siRNAs) that were obtained via 5'-independent sequencing. For these two samples males and hermaphrodites had comparable viral loads.



**Figure S4 Transcriptional response of N2 males and hermaphrodites to the Orsay virus by microarray**  
 – A) Viral loads of the samples that were analyzed by microarray. B) The effect of OrV infection on IPR genes in hermaphrodites where 39 genes (represented by 65 spots) are differentially expressed upon infection and the effect of OrV infection on IPR genes in males where 10 genes (represented by 20 spots) are differentially expressed upon infection. For both volcano plots the effect sizes are shown on the x-axis and the significance per spot is shown on the y-axis. The dotted line represents the statistical threshold (FDR = 0.05). C) Gene expression (log<sub>2</sub> intensities) of the IPR genes for mock-treated or OrV infected males and hermaphrodites. Underlined genes were differentially expressed between males and hermaphrodites under mock conditions, whereas red asterisks indicate different expression upon OrV infection per sex.

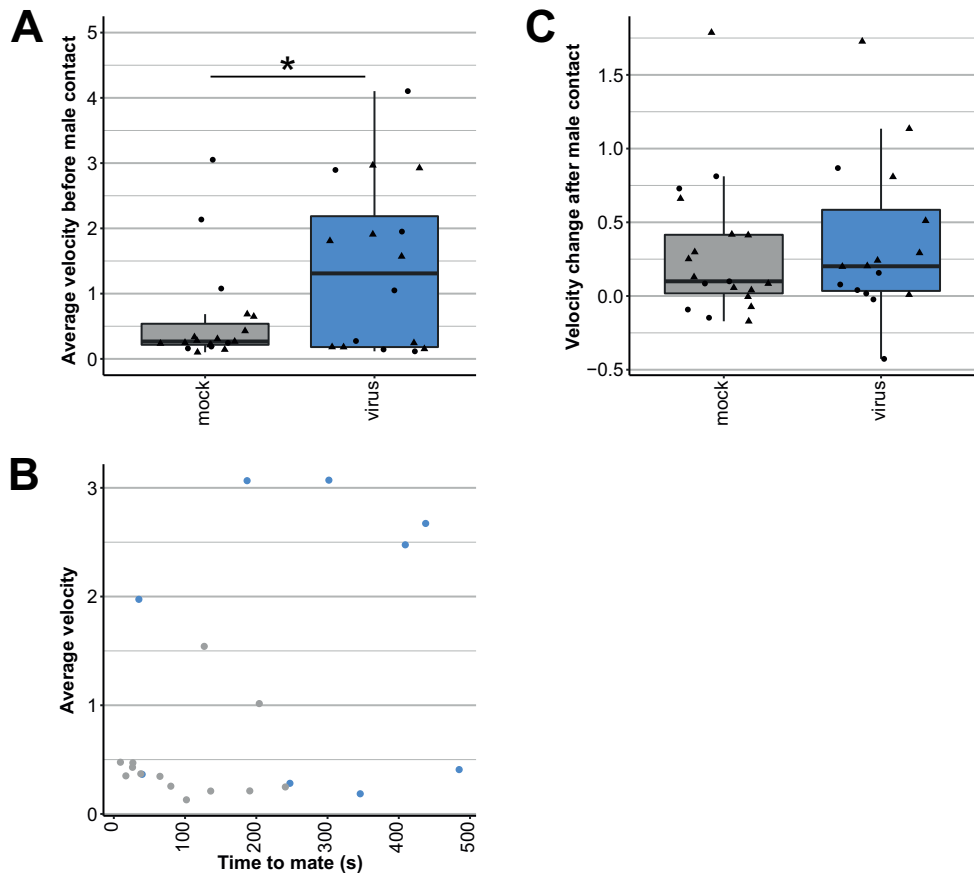


**Figure S5** Transcriptional response of N2, JU1580 and CB4856 males and hermaphrodites to the Orsay virus by RT-qPCR – Gene expression of pals-22, pals-25, pals-6, pals-14 and eol-1 for untreated L4 N2, JU1580 and CB4856 males and hermaphrodites as determined by RT-qPCR. The data represents 5 biological replicates, but samples that expressed the gene of interest under the RT-qPCR detection limit were excluded from the analysis.



**Figure S6** Choice assay between mock and OrV lysate of N2 and *daf-22* hermaphrodites – Choice assay where nematodes choose between the lysate of mock-treated or OrV infected young adult nematodes of either the genotype N2 or RB589 (*daf-22* mutant). The percentage of males and hermaphrodites that was observed in the mock or OrV lysate spot after 2 and 24 hours is shown. Error bars indicate standard error of the mean. Nematodes that showed a significant preference for one of the spots are indicated with an asterisk (test,  $p < 0.05$ ).





**Figure S7 Hermaphrodite movements** – A) Average velocity of the hermaphrodites during the 30 seconds before the first contact with the male. Velocities are depicted in arbitrary units. Each dot shows the velocity of a hermaphrodite and the shape of the dot (triangle = mated, dot = unmated) indicates if the hermaphrodite was mated before the end of the assay (20 minutes after placement of the male). The asterisk indicates a significant difference (t-test,  $p < 0.05$ ). B) Relationship between hermaphrodite velocity and mating times. Each dot represents the data from a mated nematode. The x-axis shows the time it took between the first contact between male and hermaphrodite and the time before the mating took place. The y-axis displays the average velocity of the hermaphrodites during 30 seconds before and 90 seconds after the first contact with the male. C) Change in velocity after the male and hermaphrodite make contact. The change is calculated by dividing the average velocity (measured over 30s) before and after (measured over 90s) the contact. The shape of the dot (triangle = mated, dot = unmated) shows if the hermaphrodite was mated before the end of the assay (20 minutes after placement of the male).

The supplementary movie can be found at <https://tinyurl.com/SI-PhD-thesis>.

**Movie S1 Example mating assay movie** – Example of a movie made for the mating assay. Males were placed at a set distance outside a bacterial droplet containing an infected or mock-treated young adult hermaphrodite. The recording time per movie was 20 minutes filming with 1 frame per second (this movie is accelerated 16 times). Different timepoints were manually noted: the timepoint the male enters the bacterial droplet, the timepoint the male touches the hermaphrodites for the first time and the timepoint a mating starts.

The full supplementary tables can be found at <https://tinyurl.com/SI-PhD-thesis>. Here, the first four rows of the dataset are shown to illustrate the contents.

**Supplementary Table S1 Viral loads** – A) Viral loads of both sexes of N2, CB4856 and JU1580. B) Viral loads of mixed-sex populations for N2.

Sample qPCR number	Strain	Sex	Treatment	Viral load
9	N2	h	OrV infected	16.3
10	N2	m	OrV infected	0.0
11	CB4856	h	OrV infected	0.0

Sample qPCR number	Strain	Sex	Treatment	Viral load
13	N2	hermaphrodite only	OrV infected	8.9
15	N2	mixed-sex	OrV infected	9.4
19	N2	hermaphrodite only	OrV infected	15.8

**Supplementary Table S2 Food intake** – A) Food intake as measured by pumping rate in L4 and young adult nematodes for males and hermaphrodites. B) Food intake measured by feeding fluorescent beads in liquid to L4 nematodes for both sexes. C) Food intake measured by feeding fluorescent beads on plates to L4 nematodes for both sexes. D) Food intake measured by feeding fluorescent beads on plates to young adult nematodes for both sexes.

Measurement	Strain	Sex	Stage	Date	Pumping rate 1	Pumping rate 2
1	CB4856	m	L4	20171206	64	64
2	CB4856	m	L4	20171206	61	63
3	CB4856	m	L4	20171206	64	54

Measurement	Strain	Sex	Experiment date	Stage nematodes	Treatment	Image number	Area number measured	Area	Mean	Min	Max	Median
1	CB4856	h	20190123	L4	liquid_1h_20C	18	1	79634	1.578	0	94	0
2	CB4856	h	20190123	L4	liquid_1h_20C	20	2	85481	2.316	0	88	0
3	CB4856	h	20190123	L4	liquid_1h_20C	21	3	65235	0.853	0	82	0

Measurement	Strain	Sex	Experiment date	Stage nematodes	Treatment	Image number	Area number measured	Area	Mean	Min	Max	Median
1	CB4856	h	20181105	L4	plate_30min_25C	1	1	75586	1.053	0	6	1
2	CB4856	h	20181105	L4	plate_30min_25C	2	1	118456	1.064	0	16	1
3	CB4856	h	20181105	L4	plate_30min_25C	3	2	76525	1.688	0	43	1

Measurement	Strain	Sex	Experiment date	Stage nematodes	Treatment	Image number	Area number measured	Area	Mean	Min	Max	Median
1	CB4856	h	20190214	YA	plate_30min_25C	35	1	186104	2.621	0	91	2
2	CB4856	h	20190214	YA	plate_30min_25C	36	2	196555	1.514	0	53	1
3	CB4856	h	20190214	YA	plate_30min_25C	38	3	216837	4.933	0	94	2

**Supplementary Table S3 Gene expression data** – A) The viral loads of the microarray samples. B) IPR gene expression under mock-conditions as determined by whole-transcriptome microarrays (young adult nematodes. C) IPR gene expression upon infection as determined by whole-transcriptome microarrays young adult nematodes. D) Gene expression of IPR genes as determined by RT-qPCR in untreated L4 nematodes.

Sample qPCR number	Strain	Sex	Treatment	Viral load
19	N2	m	OrV infected	3.7
20	N2	h	OrV infected	14.7
31	N2	m	OrV infected	22.0

SpotID	Gene	WBID	Term	Significance	Significance (FDR)	Effect	Significant
1			sex	0.31	0.51	-0.31	no
2			sex	0.63	0.20	-0.030	no
3			sex	0.31	0.51	-0.037	no

SpotID	Gene	WBID	Sex	Term	Significance	Significance (FDR)	Effect	Significant
659	B0284.1	WB-Gene00007131	h	treatment	0.026	1.58	0.39	yes
1075	C08E3.10	WB-Gene00015602	h	treatment	0.0089	2.051	0.43	yes
1526	Y46G5A.20	WB-Gene00012910	h	treatment	0.016	1.79	0.39	yes

Measurement	Strain	Sex	Gene	Gene expression (log2)
1	N2	h	pals-22	9.15
2	N2	m	pals-22	9.55
3	CB4856	h	pals-22	8.60

**Supplementary Table S4 Male frequencies** – Contains data about the percentage of males observed in mock-treated and OrV infected populations of N2, CB4856 and JU1580 over 10 transfers.

Sample number	Sample	Date	Transfer	Males	Total	Percentage
1	N2 mock	27-4-2019	0	15	150	10.0
2	CB4856 virus	27-4-2019	0	15	150	10.0
3	CB4856 mock	27-4-2019	0	15	150	10.0

**Supplementary Table S5 Choice assay** – Contains observations for the choice assays performed on the lysate of N2, JU1580 and RB859 nematodes. The number of nematodes in either the mock or OrV droplet is noted for 2 and 24 hours after placement. A) Observations for the JU1580 mock and OrV lysate. B) Observations for the N2 mock and OrV lysate. C) Observations for the N2 and RB859 mock and OrV lysate.

Strain	Sex	Timepoint	Technical replicate	Biological replicate	Mock	Virus
N2	h	2	1	1	4	7
N2	h	2	2	1	3	3
N2	h	2	3	1	6	7

Strain	Sex	Timepoint	Technical replicate	Biological replicate	Mock	Virus
N2	h	2	1	0	0	5
N2	h	2	2	0	1	1
N2	h	2	3	0	0	2

Strain	Lysate	Sex	Timepoint	Technical replicate	Biological replicate	Mock	Virus
N2	N2	h	2	1	1	8	0
N2	N2	h	2	2	1	2	4
N2	N2	h	2	3	1	1	5

**Supplementary Table S6 Mating assay** – A) Dataset showing the observed male behavior towards mock-treated or infected hermaphrodites. B) Dataset summarizing the movements of hermaphrodites.

Movie number	Movie ID	Strain male	Strain hermaphrodite	Treatment hermaphrodite	Male enters droplet (s)	First contact (s)	Mating (s)
1	20190905.1_N2m_ERT54h-mock treated_01	N2	ERT54	mock treated	NA	NA	NA
2	20190905.2_N2m_ERT54h-mock treated_01	N2	ERT54	mock treated	51	107	234
3	20190905.3_N2m_ERT54h-mock treated_01	N2	ERT54	mock treated	357	NA	NA

Movie id	Treatment	Phase	Average velocity	Change speed	Mating (s)	Mated	Time between first contact and mating (s)
20190905.10	OrV infected	after contact with male	3.54	1.73	476	yes	438
20190905.10	OrV infected	before contact with male	1.81	1.73	476	yes	438
20190905.12	OrV infected	after contact with male	3.17	0.20	842	yes	187

**Supplementary Table S7 Primers** – Overview of the RT-qPCR primers used in this study.

Gene	Organism	Primer forward	Primer reverse	Tm
pals-22	C. elegans	TTTAAATCTTGAAAGTGACCGCTGGG	ACTCTCTGTGTGCTTGCAAAATT	62
pals-25	C. elegans	TGCAATCCGAAGATTGGTGA	AAATCTAACTTGTCTCAGCATGGA	62
pals-6	C. elegans	TGGGTTCGGATCAAGCAAAT	TGTTCTAGAGTGCTGCTCTCTG	62





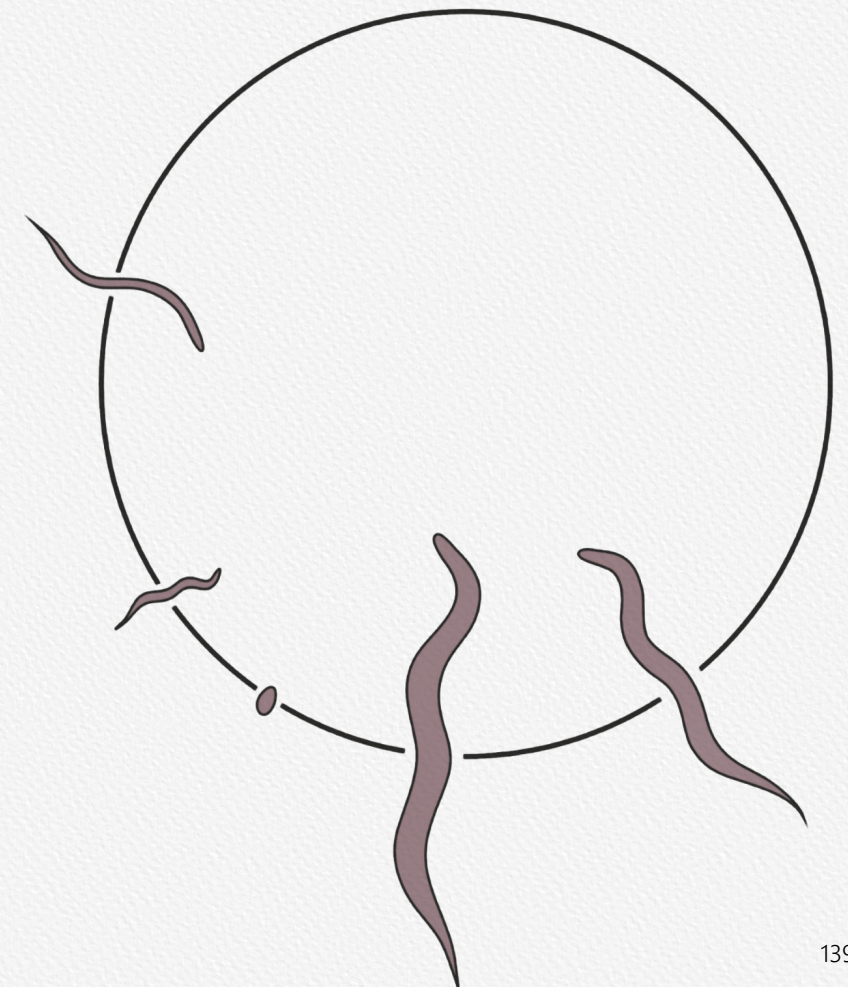
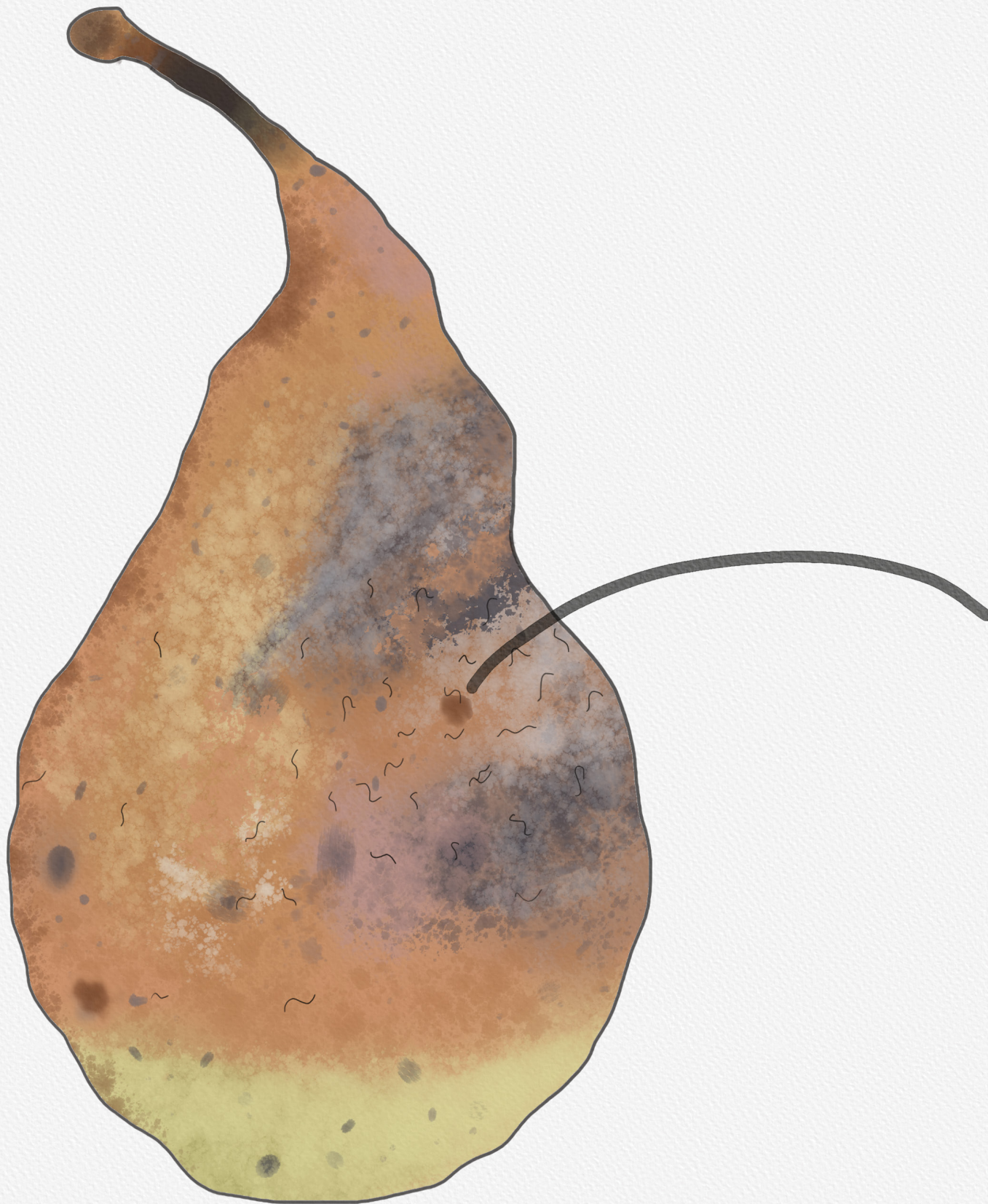
## Chapter 6

### General discussion

Lisa van Sluijs<sup>1,2</sup>

<sup>1</sup>Laboratory of Nematology, Wageningen University and Research, the Netherlands

<sup>2</sup>Laboratory of Virology, Wageningen University and Research, the Netherlands



*Caenorhabditis elegans* has kept a central position in biology since its introduction as a genetic model by Sydney Brenner in 1974 (Ankeny 2001; Corsi *et al.* 2015). Currently, five scientific papers are published every day mentioning the nematode. Most of these publications study the molecular functioning of genes and only a minority regards the natural life of *C. elegans*. Yet, the ecology of the nematode can help explain the biological function of genes, in particular of those with a currently unassigned role. On top of that, most studies consider the strain Bristol N2 as the genetic and functional standard, but wild *C. elegans* strains reveal additional genetic architectures and functioning protein variants (Cook *et al.* 2017; Gao *et al.* 2018; Lee *et al.* 2020; Sterken *et al.* 2015; Volkers *et al.* 2013). For example, the use of wild strains was key to study molecular mechanisms behind epigenetic inheritance, pheromone sensing and telomere length (Cook *et al.* 2016; Frézal *et al.* 2018; Lee *et al.* 2019).

In this thesis I investigated host-virus interactions in *C. elegans*. Viruses are ubiquitous obligatory parasites and one of nature's strongest selective forces. Viral presence changes the genetic composition of host loci involved in virus interaction (both immune- and non-immune genes) (Enard & Petrov 2020; Enard *et al.* 2016). Bioinformatic analyses of sequence data from wild populations revealed that antipathogenic genetic variation in the Intracellular Pathogen Response (IPR) has been maintained since the time that an outcrossing ancestor of *C. elegans* still existed ([Chapter 3](#)). The genetic variation present in the IPR between the *C. elegans* strains N2 and CB4856 underlies transcriptional differences in activity of this pathway ([Chapter 3](#)). Additionally, genetic variation on chromosome IV between the N2 and CB4856 strains contributes to the success of infection and the maximum viral load reached in infected populations ([Chapter 4](#)). The IPR also appears to determine viral susceptibilities between males and hermaphrodites, as males with a highly active IPR are less susceptible for infection ([Chapter 5](#)). Viral infection in male and mixed-sex populations not only revealed a sexual dimorphism in viral susceptibility, but also uncovered that mating dynamics changed under pathogenic conditions. Males preferred mating with healthy hermaphrodites and the proportion of males increased in infected populations of some strains. Together, the findings in this thesis illustrate mechanisms by which genetic and sex variation determine viral susceptibility and how the host-virus arms race has shaped parts of the *C. elegans* genome. In this chapter, I will place these insights into a broader perspective and provide my view on relevant follow-up experiments.

### *Linking genetic variation to antiviral defense in Caenorhabditis elegans*

*C. elegans* may have emerged as a model organism almost 50 years ago, large-scale efforts to understand the ecology of the organism only started 30 years later (Felix & Braendle 2010). Around the same time *C. elegans* became used more often in quantitative genetic studies that were in search for the genetic variants that underlie phenotypic variation (Gaertner & Phillips 2010). Nevertheless, these early studies were hampered by the relatively low genetic diversity in

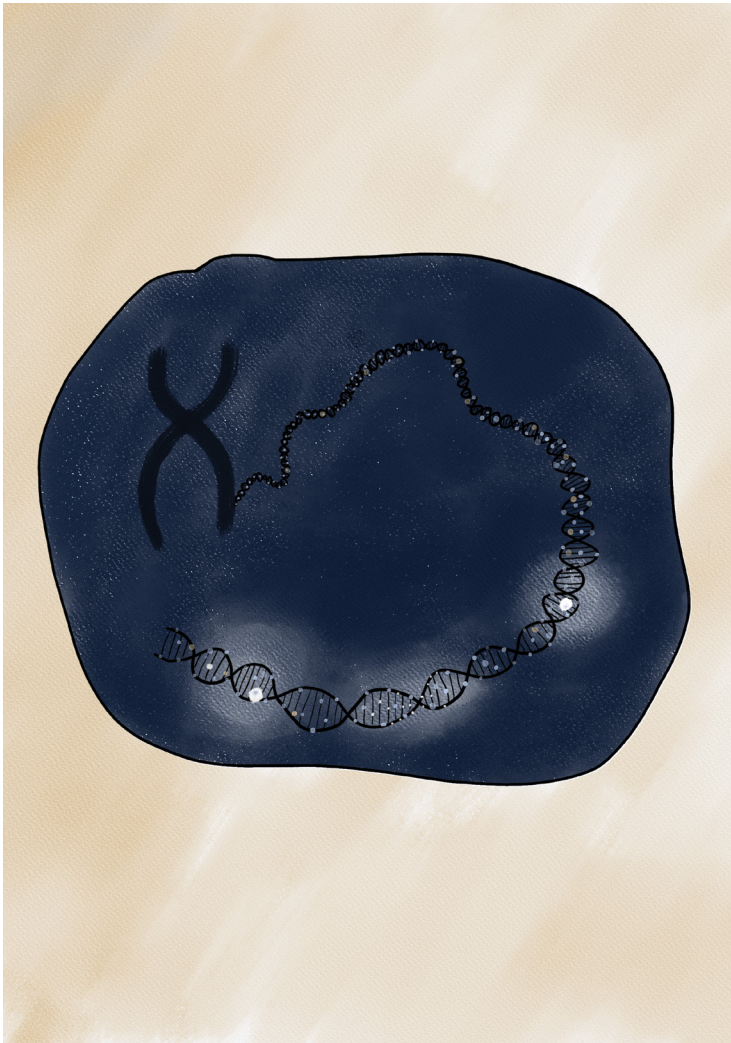


the strains available at that time. Since then, increased ecological knowledge about *C. elegans* has driven the collection of genetically more diverse strains (Cook *et al.* 2017; Frézal & Félix 2015). Genetic tools and resources to study these strains are now rapidly developed and shared by the scientific community. Classical RIL and IL panels are described in [Chapter 2](#), but also multiparent RILs have recently been established. The first multiparent mapping panel was created by crossing 16 parental strains, followed by a regime of experimental evolution for 50 generations (under control or increased salt conditions) (Noble *et al.* 2017). The second panel was made by crossing four parents under standard lab conditions (Snoek *et al.* 2019). The *Caenorhabditis elegans* Natural Diversity Resource (CeNDR) now facilitates the central organization of sampling efforts by collecting, maintaining, sequencing and distributing wild strains (Cook *et al.* 2017). Additionally, computational tools became increasingly user-friendly and accessible over the years. CeNDR has a tool for GWAS mappings and WormQTL facilitates comparative analysis of publicly available (e)QTL datasets (Cook *et al.* 2017; Snoek *et al.* 2020). These tools enhance understanding of the *C. elegans* genome.

In this thesis I made use of the novel tools to detect traces of viral selection in the *C. elegans* genome. In 2012, population genetic modelling analyses showed that much of the genetic diversity in *C. elegans* was eliminated by recent selective sweeps covering four of the six chromosomes (Andersen *et al.* 2012). Large regions of strong linkage disequilibrium (especially in the center of the chromosomes) underlie a further lack of genetic variability (Rockman *et al.* 2010). Nevertheless, de novo assembly of the CB4856 genome identified 61 regions that contain elevated levels of genetic variability (Thompson *et al.* 2015). These highly variable regions are undergoing long-term balancing selection and are enriched with ecologically relevant genes (Lee *et al.* 2020; Thompson *et al.* 2015). They cover among other genetic variations in two chemosensory genes and the two natural haplotypes identified likely represent different foraging strategies (Greene *et al.* 2016a,b; Thompson *et al.* 2015). The findings presented in this thesis provide additional evidence that balancing selection shaped the *C. elegans* genome. Using sequence information from CeNDR, the *pals*-genes, core genes within the IPR, were found to be under balancing selection. Using WormQTL it became clear that many *pals*-genes are transcriptionally regulated by local or distant genetic variation ([Chapter 3](#)) (Cook *et al.* 2017; Snoek *et al.* 2020). An intertwined network like this may be less prone to evolutionary change (Alvarez-Ponce *et al.* 2017; Fraser 2002; Helsen *et al.* 2019; Koubkova-Yu *et al.* 2018), but when a change occurs, functional diversity can be generated (Helsen *et al.* 2019; Koubkova-Yu *et al.* 2018). Indeed, limited genetic variations are found within the IPR and distinct IPR regulator (*pals*-22/*pals*-25) haplotypes may determine transcriptionally distinct IPR profiles for two strains N2 and CB4856 ([Chapter 3](#)). A third haplotype (WN2002) has not been investigated, but could uncover an extreme phenotype, because this strain has an even more genetically distinct *pals*-22/*pals*-25 locus compared to N2. Together these recent findings reveal how environmentally important genetic variation can be maintained for a species that mainly reproduces by self-fertilization.

Based on the finding that N2 and CB4856 have different viral susceptibilities, a QTL mapping approach was taken to uncover the causal genetic variation ([Chapter 4](#)). This study revealed that chromosome IV contains genetic variants that determine OrV susceptibility. Although a region was fine mapped and IPR gene *cul-6* was found to contribute to the difference between N2 and CB4856, it only explained part of the observed phenotype. Like many other studies, [Chapter 4](#) focuses on a single genetic variant encoding a protein change. Yet complex traits are proposedly affected by many, if not most, genes across the genome (Boyle *et al.* 2017; Consortium 2009; Fernández-Tajes *et al.* 2019; Loh *et al.* 2015; Shi *et al.* 2016; Visscher *et al.* 2006). According to the latter omnigenic model, core genes (for example protein encoding variants) can directly affect a trait, but these are influenced by trans-regulatory effects of omnipresent peripheral genes expressed in the same tissue (Liu *et al.* 2019). The viral susceptibilities of the N2 and CB4856 ILs follow this reasoning: they pointed towards multiple loci affecting viral susceptibility ([Chapter 4](#)). Perhaps if ILs covering other chromosomes were tested for viral susceptibility, more contributing loci would have been found.

Unraveling the full genetic basis of complex traits is challenging. Large-effect size loci can be relatively well identified by genetic mapping, but often still explain little of the overall phenotypic variation (Figure 1) (Manolio *et al.* 2009; Rockman 2012). Moreover, closely located modest effect loci may add up in the linkage analysis to show a single QTL peak and the many genes with a relative small effect (peripheral genes) are difficult to identify in genetic linkage studies (Figure 1) (Boyle *et al.* 2017; Liu *et al.* 2019; Rockman 2012). Furthermore, extensive genetic linkage disequilibrium in *C. elegans* can complicate pinpointing causal loci (Gaertner & Phillips 2010). [Chapter 3](#) suggests that the IPR regulators *pals-22* and *pals-25* could determine different transcriptional activity of the IPR in N2 and CB4856. *Pals-22* and *pals-25* are trans-acting master regulators that control at least 80 IPR genes across the genome (Reddy *et al.* 2019), but these were not detected by the QTL mapping here ([Chapter 4](#)). To resolve if *pals-22* and *pals-25* determine IPR activity and viral susceptibility, an gene expression QTL mapping (eQTL) approach could provide more information, because it also indicates target genes of regulators (Liu *et al.* 2019). Yet, a QTL approach mapping viral susceptibility was chosen here because eQTL mapping would require higher replication of viral infection in the RILs (due to the small effect of viral infection on the complete transcriptome; ([Chapter 3](#))). Also, eQTL mapping has higher costs and core IPR genes with sufficient effect sizes (like *cul-6*) can still be detected by QTL mapping ([Chapter 4](#)).



**Figure 1 Genetic linkage studies** – Genetic linkage studies aim to link causal genetic variants to phenotypic differences. Evidence is accumulating that for complex traits most genes in the genome will contribute to the phenotypic variation. Pinpointing these genes presents a challenge. Genes with a large effect size on the phenotype, such as *drh-1* on Orsay virus susceptibility, can give a bright signal (a major GWAS or QTL peak). But genes with a smaller effect remain unnoticed (here represented by the weak spots), because they produce a signal below the detection limit (perhaps genetic variants of IPR genes fall into this category). Still, together all these weak signals can contribute much to the overall trait. Finally, when multiple small or medium effect size alleles group closely together, this may result in a major GWAS or QTL peak. Yet, when trying to dissect the signal it is found it consists out of multiple signals (which could be ongoing with genes contributing to viral susceptibility on chromosome IV).

### *How do IPR haplotypes perform under natural conditions?*

A remaining question is how the different IPR haplotypes perform in nature. Because the three *pals-22/pals-25* haplotypes are found worldwide this suggests active maintenance throughout the natural history of *C. elegans* (Chapter 3). Most strains in the CeNDR database carry an N2-like *pals-22/pals-25* haplotype (83%). Only few strains have a CB4856 (10%) or WN2002 (7%) haplotype at this locus, despite our results indicate that CB4856-like strains could have a more constantly active IPR (Chapter 3). The N2-haplotype bias may occur because strains that invest in immunity cannot invest these resources in growth (Reddy *et al.* 2019). Hence the strains with a robust IPR may only increase in presence under non-favorable conditions when an active IPR provides an evolutionary advantage. It should be noted though that so far these fitness effects were only studied in N2 mutants and not in other genetic backgrounds (Reddy *et al.* 2019). Still, under optimal conditions a growth optimized IPR may be more advantageous than an antipathogenic IPR. Indeed, strains that are in minority in nature can be specialized for non-favorable conditions. This is illustrated by the observation that minority strains sampled from an orchard in France were better competitors under pathogenic conditions than the commonly sampled genotype. Yet under standard laboratory conditions, the minority strains were always outcompeted (Richaud *et al.* 2018). Because only one of these strains was sequenced, their IPR genotypes cannot be directly linked to these observations. Furthermore, the *pals-22/pals-25* N2-haplotype bias may also (partly) result from an overall N2-like strain bias in the CeNDR database. Strains from Europe, historically the most widely sampled continent, are often genetically similar to N2 (Schulenburg & Félix 2017). Future studies that investigate the IPR in a broader selection of wild strains are thus necessary.

Wild strains are often sampled from locations where *C. elegans* was previously found (hence these sites may have beneficial environments) and during fall when food seems most abundant (Schulenburg & Félix 2017). Sampling under less favorable conditions could gain insight in the evolutionary dynamics in natural populations, but can come with the challenge of not finding *C. elegans*. To illustrate this, in our lab, we have performed spring sampling trips to locations where we consistently find *C. elegans* in fall. Yet, in spring we could not find a single *C. elegans* individual on substrates (rotting plants, fruits and vegetables) or vectors (slugs, snails or isopods). More regular sampling for over 1.5 years in Germany also showed a lack of *C. elegans* during winter and spring (Petersen *et al.* 2014) and to date it is unknown where and how *C. elegans* hibernates during these seasons. Another way to collect strains that were exposed to divergent conditions is to sample in different locations. In particular, collecting *C. elegans* in the Pacific region appears promising. The Pacific holds the world's most diverse *C. elegans* strains and over the years it appeared that many of these were not affected by the selective sweep that reduced genetic variation elsewhere (Andersen *et al.* 2012; Crombie

*et al.* 2019). Naturally, this sparked enthusiasm among *C. elegans* researchers to uncover population structures in this region. A large sampling effort on Hawaii revealed that this may be the birthplace of the species (Crombie *et al.* 2019). The continuing sampling trips made by *C. elegans* researchers, but also citizen scientists (contact the Laboratory of Nematology at Wageningen University and/or see [elegansvariation.org/outreach](http://elegansvariation.org/outreach) if you like to give this a try), will continue to expand our knowledge about *C. elegans* in the wild.

### *Males to counteract pathogenic stress*

To assess the evolutionary impact of genetic variation in a species, it is necessary to consider its natural life cycle ([Chapter 1](#)). An important aspect of life for many animals is sexual reproduction via male-female mating. However, *C. elegans* has a more peculiar lifecycle. *C. elegans* mainly reproduces via self-fertilizing hermaphrodites and in nature males are rarely found (maximum 1% of the individuals) (Barrière & Félix 2005; Richaud *et al.* 2018; Sivasundar & Hey 2005). To illustrate, I have never come across a wild *C. elegans* male for over five years of sampling (at least once a year) around Wageningen. On top of that, males exhibit clumsy sexual behavior and hermaphrodites tend to avoid mating attempts further limiting sexual reproduction (Garcia *et al.* 2007). Even so, under unfavorable conditions male numbers can increase (Anderson *et al.* 2010; Kleemann & Basolo 2007; Lopes *et al.* 2008; Lynch *et al.* 2018; Masri *et al.* 2013; Morran *et al.* 2011). Interestingly, the single study that found relatively high male frequencies sampled soil (Sivasundar & Hey 2005): a substrate that later turned out non-optimal for *C. elegans*. Moreover, the few males sampled by Richaud *et al.* may have been found under non-optimal conditions too, because during this particular sampling trip proliferating populations were more scarce than for the other two excursions to the same place (Richaud *et al.* 2018). There is also genetic evidence that marks the (historical) presence of *C. elegans* males. The chromosome arms of *C. elegans* are more likely to recombine than the chromosome tips or centers during meiosis. The chromosome arms contain higher levels of genetic variation which indicates that outcrossing took place (Rockman & Kruglyak 2009). Moreover, albeit some genes involved in sexual competition have eroded to less-functional variants (Fierst *et al.* 2015; Noble *et al.* 2015; Thomas *et al.* 2012; Yin & Haag 2019), most male-specific functions have been conserved (Cutter *et al.* 2019).

The question remains if *C. elegans* males still have an ecologically relevant role (Chasnov 2013; Chasnov & Chow 2002; Cutter *et al.* 2019). Sex is costly, because males cannot produce offspring themselves, it halves genetic contributions from the parent to the offspring and sexual behavior takes energy and comes with risks (Gibson *et al.* 2017; Maynard Smith 1978). On the other hand, sex can prevent accumulation and fixation of deleterious mutations and provide efficient ways of natural selection (Chelo *et al.* 2019; Cutter 2005; Kondrashov 1984, 1985; Van Valen 1973). For *C. elegans*, the latter has proven essential under pathogenic



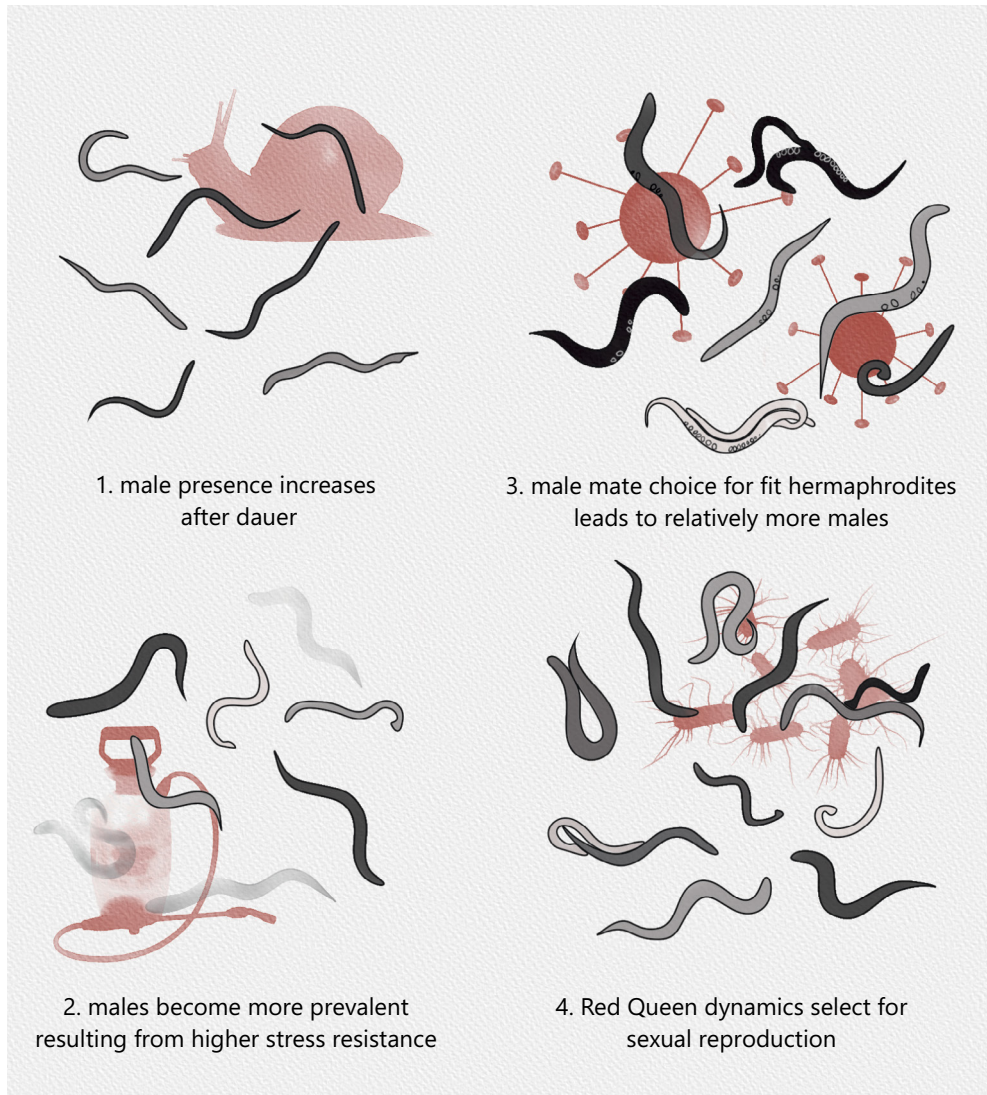
conditions. The Red Queen Hypothesis (RQH) predicts that co-evolving pathogens generate persistently dynamic selective conditions that favor outcrossing of the host (Van Valen 1973). Observations in *C. elegans* support this theory and show that outcrossed offspring with increased fitness towards the pathogen is essential for survival of the population (Masri *et al.* 2013; Morran *et al.* 2011; Slowinski *et al.* 2016). The starting populations in these experiments were genetically diverse (either resulting from mutagenesis or imposed outcrossing) and thus had enough standing genetic variation to generate fitter offspring after outcrossing. However, natural populations of *C. elegans* typically harbor little genetic variation and many sampled substrates reveal presence of a single genotype (Barrière & Félix 2007; Frézal & Félix 2015; Richaud *et al.* 2018). Therefore, I argue that (rare) male presence in the wild will often be driven by other factors than Red Queen dynamics.

The proportion of males in *C. elegans* populations can increase under (at least) three conditions that are ecologically relevant. First, males can become more prevalent after passing the dauer state. For some strains, males have higher dauer-survival chances and post-dauer mating is more efficient (Morran *et al.* 2009). Because natural populations of *C. elegans* are often found in the dauer state (Barrière & Félix 2005, 2007; Cook *et al.* 2017; Frézal & Félix 2015; Richaud *et al.* 2018), dauer-induced male presence may occur frequently (Figure 2). Second, adult males can be more stress resistant to pesticides and pathogens than hermaphrodites (Figure 2) ([Chapter 5](#)) (Lopes *et al.* 2008; Tan *et al.* 1999; Van Den Berg *et al.* 2006). If hermaphrodites perish more often than males under stress, males can become comparatively more common in the population. Yet little is known about sexual dimorphism in resistance to natural stressors. Third, I found that males are likely to select the fittest hermaphrodites within a stress-affected population ([Chapter 5](#)). Even within isogenic populations, not every individual is equally impaired by stress. Pathogens do not always infect every individual in an isogenic population ([Chapter 3](#)) (Ashe *et al.* 2013; Frézal *et al.* 2019) and individuals within isogenic populations differ in fitness and behavior (Perez *et al.* 2017; Stern *et al.* 2017). When males pick the fittest hermaphrodites (that produce most offspring) to mate with, the result is an increase in males, because mated hermaphrodites produce 50% males and unmated hermaphrodites (virtually) 100% hermaphrodites (Figure 2). Male mate choice for healthy hermaphrodites was only shown under OrV infection but could possibly be more common.

Importantly, the three reasons mentioned above can increase male proportions in isogenic or low-polymorphic populations that are often found in nature (Barrière & Félix 2007; Frézal & Félix 2015; Richaud *et al.* 2018). In isogenic populations mating will result in biparental offspring with the same genotype. This prevents fitness loss due to outbreeding depression which is one of the suggested reasons for male absence (Chasnov 2013; Dolgin *et al.* 2007). Yet, a fourth reason for increased male frequencies can be found in genetically variable populations under constant pathogen exposure where Red Queen dynamics could contribute to male maintenance

(Lynch *et al.* 2018; Masri *et al.* 2013; Morran *et al.* 2011) (Figure 2). Red Queen dynamics may be relatively uncommon in natural *C. elegans* populations, because these are often short-lived and genetically uniform (Frézal & Félix 2015). Increased male presence based on mate choice, survival chances and Red Queen dynamics depends on presence of stress; once gone there will be no (in)direct selection for males anymore so their numbers will drop. Together, this could underlie effective mating dynamics where (costly) sexual reproduction only takes place when needed, whilst when circumstances permit asexual reproduction becomes prevailing.

To gain better understanding about the effect of pathogens on mating dynamics additional data could be collected. Here, a link between male mate choice and male frequencies was presumed, but not experimentally proven (Chapter 5). An experiment that proves mate choice controls male frequencies in Orsay virus infected populations will be practically challenging, because it would require following the movements of the nematodes for generations. Moreover, the Orsay virus itself cannot be directly visualized in living animals and the currently available OrV reporter strains make use of GFP tagged IPR genes in an N2 background (Bakowski *et al.* 2014; Tanguy *et al.* 2017). These GFP-tagged IPR constructs could be inserted in other strains too, but because these IPR genes are differentially expressed in distinct genetic backgrounds and in males, this may not present a one-fits-all solution (Chapter 3, 5) (Sarkies *et al.* 2013). Therefore, ecological modelling could help to ratify my interpretations of experimental data (Ashby & Boots 2015; Campbell *et al.* 2017; Tybur & Gangestad 2011). New experiments could also use the more potent OrV isolate JUv2572 (Félix & Wang 2019; Frézal *et al.* 2019). This OrV isolate infects more individuals and could therefore lead to different male frequencies if mate choice matters. Furthermore, data could be collected in a complementary model system: that of the nematode *Caenorhabditis briggsae* and the viruses Le Blanc, Mélnik and Santeuil (Félix & Wang 2019; Félix *et al.* 2011; Franz *et al.* 2012; Frézal *et al.* 2019). Like OrV, these viruses belong to the nodaviruses and *C. briggsae* also reproduces via a male-hermaphrodite system. Using this system, it could be tested if pathogen-induced changes in mating dynamics are conserved between nematode species.

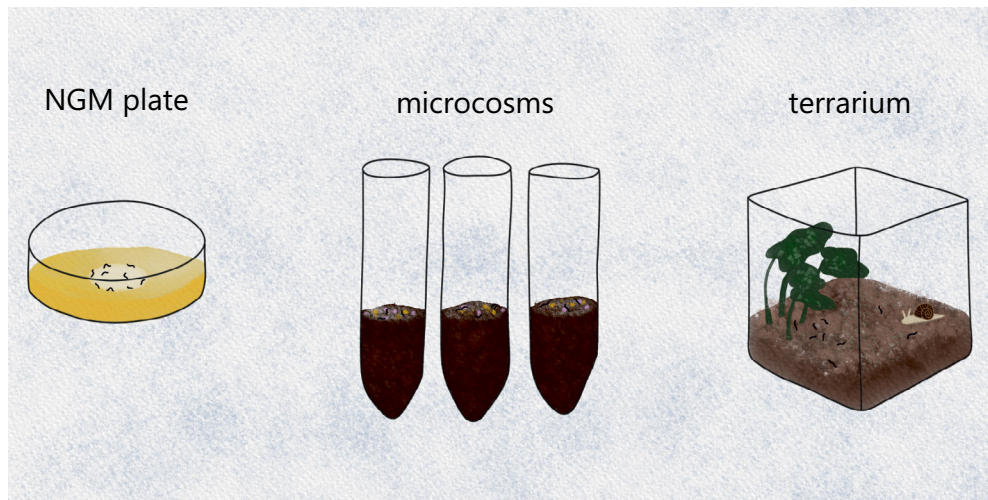


**Figure 2 Natural causes for male increase** – Male frequencies could rise in nature for four reasons. First, after passing through dauer (a common stage of natural life and often found on vectors as snails) *C. elegans* populations can have an increased presence of males. Second, males can be more resistant to lethal threats (for example a pesticide) and after hermaphrodites perish from the population relatively more males are present. Third, males can mate selectively with hermaphrodites that produce most offspring (for example the ones less affected by viral infection). Because these mated hermaphrodites will produce high numbers of offspring, of which 50% will be males, hermaphrodite frequencies are lowered. Fourth, when pathogens (for example pathogenic bacteria) co-evolve in a genetically diverse *C. elegans* population, constant co-adaption via Red Queen dynamics can underlie male presence.

## Laboratory insights into natural host-virus interactions

In this discussion and in my experimental work (Chapter 3, 5), I base my arguments on a combination of data from laboratory experiments and field studies. The conclusions are likely oversimplified because of the restricted, non-natural laboratory environment. Researchers typically ease some difficult aspects of *C. elegans* natural life in laboratory experiments by providing endless food and transferring the nematodes by picking or chunking them (replacing the natural role of isopods, slugs and snails). The oxygenic NGM plates *C. elegans* is usually grown on are detrimental for many wild isolates (Chang *et al.* 2006; Gray *et al.* 2004; Rogers *et al.* 2006; Sterken *et al.* 2015) and also prevent *C. elegans* from displaying natural dispersal behavior (Lee *et al.* 2012, 2017). Moreover, the sustained laboratory diet of a non-naturally associated *E. coli* OP50 determines the nematode's fitness (Berg *et al.* 2016; Dirksen *et al.* 2016, 2020; Samuel *et al.* 2016). Orsay virus infection can change the microbiome of *C. elegans* that were fed a variety of bacteria (Guo *et al.* 2017) which would lead to indirect fitness effects of viral infection in nature. Diet can also affect the IPR: *pals-22* mutant animals grown on HT115, another strain of *E. coli*, lacked increased thermotolerance (Panek *et al.* 2020).

The main challenge in trying to study the ecology of a microscopic animal is that they cannot be directly observed in their natural habitat. Yet, trying to provide a more-natural laboratory environment for *C. elegans* is within reach. CeMbio is a selection of bacteria that mimic the natural *C. elegans* microbiome providing all necessary nutrients for healthy growth (Dirksen *et al.* 2020). To retain the benefits of controlled laboratory experiments a complete molecular toolkit, that includes genome sequences, PCR primers and metabolic networks, has been developed for these bacteria (Dirksen *et al.* 2020). Furthermore, the growth substrate could mimic more-natural circumstances using microcosms or worm balls (Figure 3). Microcosms use soil as a substrate and nematodes are harvested at the end of the experiment using filter paper (Bååth *et al.* 1981; Franco *et al.* 2017; Gingold *et al.* 2013; Kenney *et al.* 2006). Worm balls lack the edges of normal plates and enable continuous movement of the nematode (Schulte *et al.* 2012). Moreover, applying technologies like 3D-printing of substrate and using a natural day-and-night temperature regime would contribute to naturalization of laboratory experiments. Undoubtedly these techniques would be more realistic than the flat NGM plates used most often, nevertheless the *C. elegans* community would benefit even more from larger set-ups that allow for targeted semi-wild studies. For example, by using a terrarium-like set-up that can be placed outside to perform seasonal field studies (Figure 3). A set-up like this would for example allow to perform competition experiments between the same strains under different semi-natural circumstances (for example by adding a pathogen to part of the 'terraria'). Additionally, this type of set-up may reveal more aspects of the *C. elegans* life cycle, for example where they reside during winter.



**Figure 3 Different *C. elegans* culture methods** – On the left a standard NGM (Nematode Growth Medium) plate containing a layer of the typical lab-fed bacterium *E. coli* OP50 is shown. In the middle microcosms are depicted: these contain soil and can be inoculated with bacteria like the CeMbio selection. Finally, to the right a small ecosystem is resembled in a terrarium like setting. Using this set-up has not been described for *C. elegans*, but may help close the knowledge gap between field studies and laboratory experiments.

The data provided in this thesis revealed a broad spectrum of insights into host-virus interactions. The effects of Orsay virus infection are relatively mild compared to that of other intracellular pathogens such as microsporidia that trigger the same defense pathway and appear to be more common in nature (Frézal *et al.* 2019; Zhang *et al.* 2016). Therefore, the genetic hallmarks I have related to viral infection might also be linked to presence of other intracellular pathogens like microsporidia (but also oomycetes or fungi). By performing microsporidian infections in males, the role of the IPR in determining fitness differences between the sexes could be further understood. Based on this thesis, I would also suggest studying ecologically relevant traits in the laboratory in a semi-natural set-up including a natural diet. Studying viral infections in a ‘terrarium’ would help considering the trade-offs between fitness and immunity in a natural environment and could be used to determine the role of different IPR haplotypes. Despite the remaining questions this thesis provides a novel understanding of how genetic and sex variation determine *C. elegans*-Orsay virus interactions. Therefore, this thesis provides a basis for further uncovering the natural host-pathogen ecology of this unique nematode species.





## Appendices

References

English summary

Nederlandse samenvatting

Acknowledgments

About the author

Publication list

PE&RC Training and Education Statement



## References

- Afgan E, Baker D, Batut B, Van Den Beek M, Bouvier D, *et al.* 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46:W537–44
- Ahmad S, Hur S. 2015. Helicases in Antiviral Immunity: Dual Properties as Sensors and Effectors. *Trends Biochem. Sci.* 40(10):576–85
- Ahn SH, Deshmukh H, Johnson N, Cowell LG, Rude TH, *et al.* 2010. Two genes on A/J chromosome 18 are associated with susceptibility to *Staphylococcus aureus* infection by combined microarray and QTL analyses. *PLoS Pathog.* 6(9):e1001088
- Al-Qahtani A, Khalak HG, Alkuraya FS, Al-Hamoudy W, Alswat K, *et al.* 2013. Genome-wide association study of chronic hepatitis B virus infection reveals a novel candidate risk allele on 11q22.3. *J Med Genet.* 50(11):725–32
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, *et al.* 1997. Gapped BLAST and PSI-BLAST : a new generation of protein database search programs. *Nucleic Acids Res.* 25(17):3389–3402
- Altschup SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic Local Alignment Search Tool. *J. Mol. Biol.* 215(3):403–10
- Alvarez-Ponce D, Feyertag F, Chakraborty S. 2017. Position matters: Network centrality considerably impacts rates of protein evolution in the human protein-protein interaction network. *Genome Biol. Evol.* 9(6):1742–56
- Andersen EC, Gerke JP, Shapiro JA, Crissman JR, Ghosh R, *et al.* 2012. Chromosome-scale selective sweeps shape *Caenorhabditis elegans* genomic diversity. *Nat. Genet.* 44(3):285–90
- Andersen EC, Shimko TC, Crissman JR, Ghosh R, Bloom JS, *et al.* 2015. A Powerful New Quantitative Genetics Platform, Combining *Caenorhabditis elegans* High-Throughput Fitness Assays with a Large Collection of Recombinant Strains. *G3 Genes[Genomes]Genetics.* 5(5):911–20
- Anderson JL, Albergotti L, Proulx S, Peden C, Huey RB, Phillips PC. 2007. Thermal preference of *Caenorhabditis elegans*: A null model and empirical tests. *J. Exp. Biol.* 210(17):3107–16
- Anderson JL, Morran LT, Phillips PC. 2010. Outcrossing and the Maintenance of Males within *C. elegans* Populations. *J. Hered.* 101(Supplement 1):S62–74
- Ankeny RA. 2001. The natural history of *Caenorhabditis elegans* research. *Nat. Rev. Genet.* 2(6):474–79
- Ashby B, Boots M. 2015. Coevolution of parasite virulence and host mating strategies. *Proc. Natl. Acad. Sci.* 112(43):13290–95
- Ashe A, Bécicard T, Le Pen J, Sarkies P, Frézal L, *et al.* 2013. A deletion polymorphism in the *Caenorhabditis elegans* RIG-I homolog disables viral RNA dicing and antiviral immunity. *Elife.* 2:e00994
- Bååth E, Lohm U, Lundgren B, Rosswall T, Söderström B, *et al.* 1981. Impact of Microbial-Feeding Animals on Total Soil Activity and Nitrogen Dynamics: A Soil Microcosm Experiment. *Oikos.* 37(3):257
- Bakowski MA, Desjardins CA, Smelkinson MG, Dunbar TA, Lopez-Moyado IF, *et al.* 2014. Ubiquitin-Mediated Response to Microsporidia and Virus Infection in *C. elegans*. *PLoS Pathog.* 10(6):e1004200
- Balla KM, Andersen EC, Kruglyak L, Troemel ER. 2015. A Wild *C. Elegans* Strain Has Enhanced Epithelial Immunity to a Natural Microsporidian Parasite. *PLoS Pathog.* 11(2):e1004583
- Bangham J, Obbard DJ, Kim K-W, Haddrill PR, Jiggins FM. 2007. The age and evolution of an antiviral resistance mutation in *Drosophila melanogaster*. *Proc. R. Soc. B Biol. Sci.* 274(1621):2027–34
- Bargmann CI, Hartweg E, Horvitz HR. 1993. Odorant-selective genes and neurons mediate olfaction in *C. elegans*. *Cell.* 74(3):515–27

- Barrière A, Félix MA. 2005. High local genetic diversity and low outcrossing rate in *Caenorhabditis elegans* natural populations. *Curr. Biol.* 15(13):1176–84
- Barrière A, Félix M-A. 2007. Temporal Dynamics and Linkage Disequilibrium in Natural *Caenorhabditis elegans* Populations. *Genetics*. 176(2):999–1011
- Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* 67(1):1–48
- Baugh LR. 2013. To grow or not to grow: Nutritional control of development during *Caenorhabditis elegans* L1 Arrest. *Genetics*. 194(3):539–55
- Bell G. 1982. *The Masterpiece of Nature : The Evolution and Genetics of Sexuality*. Springer International Publishing
- Beltran-Bech S, Richard FJ. 2014. Impact of infection on mate choice. *Anim. Behav.* 90:159–70
- Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc.* 57(1):289–300
- Benkert P, Biasini M, Schwede T. 2011. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*. 27(3):343–50
- Berg M, Stenuit B, Ho J, Wang A, Parke C, *et al.* 2016. Assembly of the *Caenorhabditis elegans* gut microbiota from diverse soil microbial environments. *ISME J.* 10(8):1998–2009
- Bernstein MR, Rockman M V. 2016. Fine-Scale Crossover Rate Variation on the *Caenorhabditis elegans* X Chromosome. *G3 Genes|Genomes|Genetics*. 6(6):1767–76
- Bloom JS, Ehrenreich IM, Loo WT, Lite T-LV, Kruglyak L. 2013. Finding the sources of missing heritability in a yeast cross. *Nature*. 494(7436):234–37
- Boon ACM, deBeauchamp J, Hollmann A, Luke J, Kotb M, *et al.* 2009. Host Genetic Variation Affects Resistance to Infection with a Highly Pathogenic H5N1 Influenza A Virus in Mice. *J. Virol.* 83(20):10417–26
- Borne F, Kasimatis KR, Phillips PC. 2017. Quantifying male and female pheromone-based mate choice in *Caenorhabditis* nematodes using a novel microfluidic technique. *PLoS One*. 12(12):1–14
- Boyle EA, Li YI, Pritchard JK. 2017. An expanded view of complex traits: from polygenic to omnigenic. *Cell*. 169(7):1177–1186
- Brenner S. 1974. The Genetics of *Caenorhabditis elegans*. *Genetics*. 77(1):71–94
- Brodin P, Jojic V, Gao T, Bhattacharya S, Angel CJL, *et al.* 2015. Variation in the human immune system is largely driven by non-heritable influences. *Cell*. 160(1–2):37–47
- Buchner DA, Nadeau JH. 2015. Contrasting genetic architectures in different mouse reference populations used for studying complex traits. *Genome Res.* 25(6):775–91
- Buniello A, Macarthur JAL, Cerezo M, Harris LW, Hayhurst J, *et al.* 2019. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* 47(D1):D1005–12
- Burand JP, Tan W, Kim W, Nojima S, Roelofs W. 2005. Infection with the insect virus Hz-2v alters mating behavior and pheromone production in female *Helicoverpa zea* moths. *J. Insect Sci.* 5(6):1–6
- Butlin R. 2002. The costs and benefits of sex: New insights from old asexual lineages. *Nat. Rev. Genet.* 3(4):311–17
- Campbell LJ, Head ML, Wilfert L, Griffiths AGF. 2017. An ecological role for assortative mating under infection? *Conserv. Genet.* 18(5):983–94

- Cao C, Magwire MM, Bayer F, Jiggins FM. 2016. A Polymorphism in the Processing Body Component Ge-1 Controls Resistance to a Naturally Occurring Rhabdovirus in *Drosophila*. *PLoS Pathog.* 12(1):1–21
- Capra EJ, Skrovanek SM, Kruglyak L. 2008. Comparative developmental expression profiling of two *C. elegans* isolates. *PLoS One.* 3(12):e4055
- Chalfie M, Tu Y, Euskirchen G, Ward W, Prasher D. 1994. Green fluorescent protein as a marker for gene expression. *Science.* 263(5148):802–5
- Chang AJ, Chronis N, Karow DS, Marletta MA, Bargmann CI. 2006. A distributed chemosensory circuit for oxygen preference in *C. elegans*. *PLoS Biol.* 4(9):1588–1602
- Chapman SJ, Hill AVS. 2012. Human genetic susceptibility to infectious disease. *Nat. Rev. Genet.* 13(3):175–88
- Charlier J, Höglund J, Morgan ER, Geldhof P, Vercruysse J, Claerebout E. 2020. Biology and Epidemiology of Gastrointestinal Nematodes in Cattle. *Vet. Clin. North Am. - Food Anim. Pract.* 36(1):1–15
- Chasnov JR. 2013. The evolutionary role of males in *C. elegans*. *Worm.* 2(1):e21146
- Chasnov JR, Chow KL. 2002. Why are there males in the hermaphroditic species. *Genetics.* 160(March):983–94
- Chasnov JR, So WK, Chan CM, Chow KL. 2007. The species, sex, and stage specificity of a *Caenorhabditis* sex pheromone. *Proc. Natl. Acad. Sci.* 104(16):6730–35
- Chelo IM, Afonso B, Carvalho S, Theologidis I, Goy C, *et al.* 2019. Partial selfing can reduce genetic loads while maintaining diversity during experimental evolution. *G3 Genes|Genomes|Genetics.* 9(9):2811–21
- Chen K, Franz CJ, Jiang H, Jiang Y, Wang D. 2017. An evolutionarily conserved transcriptional response to viral infection in *Caenorhabditis* nematodes. *BMC Genomics.* 18(1):303
- Chen X, Feng X, Guang S. 2016. Targeted genome engineering in *Caenorhabditis elegans*. *Cell Biosci.* 6(1):1–10
- Chinwalla AT, Cook LL, Delehaunty KD, Fewell GA, Fulton LA, *et al.* 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature.* 420(6915):520–62
- Chotkowski HL, Ciota AT, Jia Y, Puig-Basagoiti F, Kramer LD, *et al.* 2008. West Nile virus infection of *Drosophila melanogaster* induces a protective RNAi response. *Virology.* 377(1):197–206
- Ciancanelli MJ, Abel L, Zhang S-Y, Casanova J-L. 2016. Host genetics of severe influenza: from mouse Mx1 to human IRF7. *Curr. Opin. Immunol.* 38:109–20
- Coffman SR, Lu J, Guo X, Zhong J, Jiang H, *et al.* 2017. *Caenorhabditis elegans* RIG-I Homolog Mediates Antiviral RNA Interference Downstream of Dicer-Dependent Biogenesis of Viral Small Interfering RNAs. *MBio.* 8(2):1–15
- Cogni R, Cao C, Day JP, Bridson C, Jiggins FM. 2016. The genetic architecture of resistance to virus infection in *Drosophila*. *Mol. Ecol.* 25(20):5228–41
- Consortium IS. 2009. Common polygenic variation contributes to risk of schizophrenia that overlaps with bipolar disorder International. *Nature.* 460(7256):748–752
- Cook DE, Zdraljevic S, Roberts JP, Andersen EC. 2017. CeNDR, the *Caenorhabditis elegans* natural diversity resource. *Nucleic Acids Res.* 45(D1):D650–57
- Cook DE, Zdraljevic S, Tanny RE, Seo B, Riccardi DD, *et al.* 2016. The Genetic Basis of Natural Variation in *Caenorhabditis elegans* Telomere Length. *Genetics.* 204(1):371–83
- Cook SJ, Jarrell TA, Brittin CA, Wang Y, Bloniarz AE, *et al.* 2019. Whole-animal connectomes of both *Caenorhabditis elegans* sexes. *Nature.* 571(7763):63–71

- Corsi AK, Wightman B, Chalfie M. 2015. A transparent window into biology: A primer on *Caenorhabditis elegans*. *Genetics*. 200(2):387–407
- Crombie TA, Zdravljjevic S, Cook DE, Tanny RE, Brady SC, *et al.* 2019. Deep sampling of Hawaiian *Caenorhabditis elegans* reveals high genetic diversity and admixture with global populations. *Elife*. 8:e50465
- Cui Y, McBride SJ, Boyd WA, Alper S, Freedman JH. 2007. Toxicogenomic analysis of *Caenorhabditis elegans* reveals novel genes and pathways involved in the resistance to cadmium toxicity. *Genome Biol*. 8(6):1–15
- Cutter AD. 2005. Mutation and the experimental evolution of outcrossing in *Caenorhabditis elegans*. *J. Evol. Biol.* 18(1):27–34
- Cutter AD, Morran LT, Phillips PC. 2019. Males, Outcrossing, and Sexual Selection in *Caenorhabditis* Nematodes. *Genetics*. 213(1):27–57
- Dang TN, Naka I, Sa-Ngasang A, Anantapreecha S, Chanama S, *et al.* 2014. A replication study confirms the association of GWAS-identified SNPs at MICB and PLCE1 in Thai patients with dengue shock syndrome. *BMC Med. Genet.* 15(1):58
- Dean M, Carrington M, Winkler C, Huttley GA, Smith MW, *et al.* 1996. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. *Science*. 273(5283):1856–62
- Diaz SA, Viney M. 2015. The evolution of plasticity of dauer larva developmental arrest in the nematode *Caenorhabditis elegans*. *Ecol. Evol.* 5(6):1343–53
- Dirksen P, Assié A, Zimmermann J, Zhang F, Tietje A-M, *et al.* 2020. CeMbio - The *Caenorhabditis elegans* Microbiome Resource. *G3 Genes|Genomes|Genetics*. 10(September):g3.401309.2020
- Dirksen P, Marsh SA, Braker I, Heitland N, Wagner S, *et al.* 2016. The native microbiome of the nematode *Caenorhabditis elegans*: Gateway to a new host-microbiome model. *BMC Biol.* 14(1):1–16
- Dolgin ES, Charlesworth B, Baird SE, Cutter AD. 2007. Inbreeding and outbreeding depression in *Caenorhabditis* nematodes. *Evolution*. 61(6):1339–52
- Doroszuk A, Snoek LB, Fradin E, Riksen J, Kammenga J. 2009. A genome-wide library of CB4856/N2 introgression lines of CB4856/N2 introgression lines of *Caenorhabditis elegans*. *Nucleic Acids Res.* 37(16):1–16
- Dupuis S, Jouanguy E, Al-Hajjar S, Fieschi C, Al-Mohsen IZ, *et al.* 2003. Impaired response to interferon- $\alpha/\beta$  and lethal viral disease in human STAT1 deficiency. *Nat. Genet.* 33(3):388–91
- Dyson T. 2012. Causes and Consequences of Skewed Sex Ratios. *Annu. Rev. Sociol.* 38(1):443–61
- Ebbing A, Vértessy Á, Betist MC, Spanjaard B, Junker JP, *et al.* 2018. Spatial Transcriptomics of *C. elegans* Males and Hermaphrodites Identifies Sex-Specific Differences in Gene Expression Patterns. *Dev. Cell*. 47(6):801–813.e6
- Eichler EE, Flint J, Gibson G, Kong A, Leal SM, *et al.* 2010. Missing heritability and strategies for finding the underlying causes of complex disease. *Nat. Rev. Genet.* 11(6):446–50
- Enard D, Cai L, Gwennap C, Petrov DA. 2016. Viruses are a dominant driver of protein adaptation in mammals. *Elife*. 5:e12469
- Enard D, Petrov DA. 2020. Ancient RNA virus epidemics through the lens of recent adaptation in human genomes. *Philos. Trans. R. Soc. B Biol. Sci.* 375(1812):20190575
- Engelstädter J, Hurst GDD. 2009. The Ecology and Evolution of Microbes that Manipulate Host Reproduction. *Annu. Rev. Ecol. Evol. Syst.* 40(1):127–49
- Evans KS, Zhao Y, Brady SC, Long L, McGrath PT, Andersen EC. 2017. Correlations of genotype with climate parameters suggest *Caenorhabditis elegans* niche adaptations. *G3 Genes|Genomes|Genetics*. 7(1):289–98

- Farboud B. 2017. Targeted genome editing in *Caenorhabditis elegans* using CRISPR/Cas9. *Wiley Interdiscip. Rev. Dev. Biol.* 6(6):14–16
- Félix M-A, Duveau F. 2012. Population dynamics and habitat sharing of natural populations of *Caenorhabditis elegans* and *C. briggsae*. *BMC Biol.* 10(1):59
- Félix M-A, Wang D. 2019. Natural Viruses of *Caenorhabditis* Nematodes. *Annu. Rev. Genet.* 53:4.1-4.14
- Felix MA, Braendle C. 2010. The natural history of *Caenorhabditis elegans*. *Curr. Biol.* 20(22):R965–69
- Félix MA, Ashe A, Piffaretti J, Wu G, Nuez I, *et al.* 2011. Natural and experimental infection of *Caenorhabditis* nematodes by novel viruses related to nodaviruses. *PLoS Biol.* 9(1):e1000586
- Femino AM. 1998. Visualization of Single RNA Transcripts in Situ. *Science.* 280(5363):585–90
- Fernández-Tajes J, Gaulton KJ, Van De Bunt M, Torres J, Thurner M, *et al.* 2019. Developing a network view of type 2 diabetes risk pathways through integration of genetic, genomic and functional data. *Genome Med.* 11(1):1–14
- Ferris MT, Aylor DL, Bottomly D, Whitmore AC, Aicher LD, *et al.* 2013. Modeling Host Genetic Regulation of Influenza Pathogenesis in the Collaborative Cross. *PLoS Pathog.* 9(2):e1003196
- Fierst JL, Willis JH, Thomas CG, Wang W, Reynolds RM, *et al.* 2015. Reproductive Mode and the Evolution of Genome Size and Structure in *Caenorhabditis* Nematodes. *PLoS Genet.* 11(6):1–25
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature.* 391(6669):806–11
- Fisher RA. 1930. *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford
- Franco ALC, Knox MA, Andriuzzi WS, de Tomasel CM, Sala OE, Wall DH. 2017. Nematode exclusion and recovery in experimental soil microcosms. *Soil Biol. Biochem.* 108:78–83
- Franco LM, Bucasas KL, Wells JM, Niño D, Wang X, *et al.* 2013. Integrative genomic analysis of the human immune response to influenza vaccination. *Elife.* 2:e.00299
- Franz CJ, Renshaw H, Frezal L, Jiang Y, Félix M-A, Wang D. 2013. Orsay, Santeuil and Le Blanc viruses primarily infect intestinal cells in *Caenorhabditis* nematodes. *Virology.* 448:255–64
- Franz CJ, Zhao G, Felix M-A, Wang D. 2012. Complete Genome Sequence of Le Blanc Virus, a Third *Caenorhabditis* Nematode-Infecting Virus. *J. Virol.* 86(21):11940–11940
- Fraser HB. 2002. Evolutionary Rate in the Protein Interaction Network. *Science.* 296(5568):750–52
- Frézal L, Demoinet E, Braendle C, Miska E, Félix MA. 2018. Natural Genetic Variation in a Multigenerational Phenotype in *C. elegans*. *Curr. Biol.* 28(16):2588-2596.e8
- Frézal L, Félix MA. 2015. *C. elegans* outside the Petri dish. *Elife.* 2015(4):e05849
- Frézal L, Jung H, Tahan S, Wang D, Félix M-A. 2019. Noda-Like RNA Viruses Infecting *Caenorhabditis* Nematodes: Sympatry, Diversity, and Reassortment. *J. Virol.* 93(21):
- Gaertner BE, Parmenter MD, Rockman M V, Kruglyak L, Phillips PC. 2012. More than the sum of its parts: A complex epistatic network underlies natural variation in thermal preference behavior in *Caenorhabditis elegans*. *Genetics.* 192(4):1533–42
- Gaertner BE, Phillips PC. 2010. *Caenorhabditis elegans* as a platform for molecular quantitative genetics and the systems biology of natural variation. *Genet. Res.* 92(5–6):331–48

- Gammon DB, Ishidate T, Li L, Gu W, Silverman N, Mello CC. 2017. The Antiviral RNA Interference Response Provides Resistance to Lethal Arbovirus Infection and Vertical Transmission in *Caenorhabditis elegans*. *Curr. Biol.* 27(6):795–806
- Gao AW, Sterken MG, de Bos J uit, van Creijl J, Kamble R, *et al.* 2018. Natural genetic variation in *C. elegans* identified genomic loci controlling metabolite levels. *Genome Res.* 28(9):1296–1308
- Garcia LR, LeBoeuf B, Koo P. 2007. Diversity in mating behavior of hermaphroditic and male-female *Caenorhabditis* nematodes. *Genetics.* 175(4):1761–71
- Gardner J, Anraku I, Le TT, Larcher T, Major L, *et al.* 2010. Chikungunya Virus Arthritis in Adult Wild-Type Mice. *J. Virol.* 84(16):8021–32
- Gasch AP, Payseur BA, Pool JE. 2016. The Power of Natural Variation for Model Organism Biology. *Trends Genet.* 32(3):147–54
- Ge D, Fellay J, Thompson AJ, Simon JS, Shianna K V., *et al.* 2009. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature.* 461(7262):399–401
- Geng X, Harry BL, Zhou Q, Skeen-Gaar RR, Ge X, *et al.* 2012. Hepatitis B virus X protein targets the Bcl-2 protein CED-9 to induce intracellular Ca<sup>2+</sup> increase and cell death in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* 109(45):18465–70
- Gerstein MB, Lu ZJ, Van Nostrand EL, Cheng C, Arshinoff BI, *et al.* 2010. Integrative Analysis of the *Caenorhabditis elegans* Genome by the modENCODE Project. *Science.* 330(6012):1775 LP – 1787
- Gerstein MB, Rozowsky J, Yan K-K, Wang D, Cheng C, *et al.* 2014. Comparative analysis of the transcriptome across distant species. *Nature.* 512(7515):445–48
- Ghosh CC, David S, Zhang R, Berghelli A, Milam K, *et al.* 2016. Gene control of tyrosine kinase *TIE2* and vascular manifestations of infections. *Proc. Natl. Acad. Sci.* 113(9):2472–77
- Gibson AK, Delph LF, Lively CM. 2017. The two-fold cost of sex: Experimental evidence from a natural system. *Evol. Lett.* 1(1):6–15
- Gingold R, Moens T, Rocha-Olivares A. 2013. Assessing the Response of Nematode Communities to Climate Change-Driven Warming: A Microcosm Experiment. *PLoS One.* 8(6):e66653
- Gipson SAY, Jimenez L, Hall MD. 2019. Host sexual dimorphism affects the outcome of within-host pathogen competition. *Evolution.* 73(7):1443–55
- Gould E, Pettersson J, Higgs S, Charrel R, de Lamballerie X. 2017. Emerging arboviruses: Why today? *One Health.* 4(July):1–13
- Gray JC, Cutter AD. 2014. Mainstreaming *Caenorhabditis elegans* in experimental evolution. *Proc. R. Soc. B Biol. Sci.* 281(1778):20133055
- Gray JM, Karow DS, Lu H, Chang AJ, Chang JS, *et al.* 2004. Oxygen sensation and social feeding mediated by a *C. elegans* guanylate cyclase homologue. *Nature.* 430(6997):317–22
- Greene JS, Brown M, Dobosiewicz M, Ishida IG, Macosko EZ, *et al.* 2016a. Balancing selection shapes density-dependent foraging behaviour. *Nature.* 539(7628):254–58
- Greene JS, Dobosiewicz M, Butcher RA, McGrath PT, Bargmann CI. 2016b. Regulatory changes in two chemoreceptor genes contribute to a *Caenorhabditis elegans* QTL for foraging behavior. *Elife.* 5:1–19
- Grishok A, Mello C. 2002. RNAi (nematodes: *Caenorhabditis elegans*). In *Homology Effects*, Vol. 46, pp. 339–60. Elsevier



- Guabiraba R, Ryffel B. 2014. Dengue virus infection: Current concepts in immune mechanisms and lessons from murine models. *Immunology*. 141(2):143–56
- Guex N, Peitsch MC, Schwede T. 2009. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis*. 30(S1):S162–73
- Guo Y, Xun Z, Coffman SR, Chen F. 2017. The Shift of the Intestinal Microbiome in the Innate Immunity-Deficient Mutant *rde-1* Strain of *C. elegans* upon Orsay Virus Infection. *Front. Microbiol.* 8:1–11
- Guo YR, Hryc CF, Jakana J, Jiang H, Wang D, *et al.* 2014. Crystal structure of a nematode-infecting virus. *Proc. Natl. Acad. Sci.* 111(35):12781–86
- Hahnel SR, Zdraljevic S, Rodriguez BC, Zhao Y, McGrath PT, Andersen EC. 2018. Extreme allelic heterogeneity at a *Caenorhabditis elegans* beta-tubulin locus explains natural resistance to benzimidazoles. *PLoS Pathog.* 14(10):1–26
- Hamilton WD. 1980. Sex versus Non-Sex versus Parasite. *Oikos*. 35(2):282–90
- Han S, Lee S-J, Kim KE, Lee HS, Oh N, *et al.* 2016. Amelioration of sepsis by TIE2 activation–induced vascular protection. *Sci. Transl. Med.* 8(335):335ra55–335ra55
- Hancock DB, Gaddis NC, Levy JL, Bierut LJ, Kral AH, Johnson EO. 2015. Associations of common variants in the BST2 region with HIV-1 acquisition in African American and European American people who inject drugs. *Aids*. 29(7):767–77
- Heim MH, Bochud P-Y, George J. 2016. Host – hepatitis C viral interactions: The role of genetics. *J. Hepatol.* 65(1):S22–32
- Helsen J, Frickel J, Jelier R, Verstrepen KJ. 2019. Network hubs affect evolvability. *PLoS Biol.* 17(1):1–5
- Hodge KT, Viaene NM, Gams W. 1997. Two *Harposporium* species with *Hirsutella* synanamorphs. *Mycol. Res.* 101(11):1377–82
- Hodgkin J, Kuwabara PE, Corneliussen B. 2000. A novel bacterial pathogen, *Microbacterium nematophilum*, induces morphological change in the nematode *C. elegans*. *Curr. Biol.* 10(24):1615–18
- Hoffmann FS, Schmidt A, Dittmann Chevillotte M, Wisskirchen C, Hellmuth J, *et al.* 2015. Polymorphisms in melanoma differentiation-associated gene 5 link protein function to clearance of hepatitis C virus. *Hepatology*. 61(2):460–70
- Horby P, Sudoyo H, Viprakasit V, Fox A, Thai PQ, *et al.* 2010. What is the evidence of a role for host genetics in susceptibility to influenza A/H5N1? *Epidemiol. Infect.* 138(11):1550–58
- Hou Y, Zhao J, Martin W, Kallianpur A, Chung MK, *et al.* 2020. New insights into genetic susceptibility of COVID-19: An ACE2 and TMPRSS2 polymorphism analysis. *BMC Med.* 18(1):1–8
- Huang W, Massouras A, Inoue Y, Peiffer J, Ràmia M, *et al.* 2014. Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. *Genome Res.* 24(7):1193–1208
- Huang Y, Kammenga JE. 2020. Genetic Variation in *Caenorhabditis elegans* Responses to Pathogenic Microbiota. *Microorganisms*. 8(4):618
- Ingersoll MA. 2017. Sex differences shape the response to infectious diseases. *PLoS Pathog.* 13(12):6–11
- Iraqi FA, Mahajne M, Salaymah Y, Sandovski H, Tayem H, *et al.* 2012. The genome architecture of the collaborative cross mouse genetic reference population. *Genetics*. 190(2):389–401
- Jaenike J. 1978. An hypothesis to account for maintainance of sex within populations. *Evol. Theory*. 3:191–194

- Jansson HB. 1994. Adhesion of conidia of *Drechmeria coniospora* to *Caenorhabditis elegans* wild type and mutants. *J. Nematol.* 26(4):430–35
- Jiang H, Chen K, Sandoval LE, Leung C, Wang D. 2017. An Evolutionarily Conserved Pathway Essential for Orsay Virus Infection of *Caenorhabditis elegans*. *MBio.* 8(5):1–16
- Jiang H, Franz CJ, Wu G, Renshaw H, Zhao G, *et al.* 2014. Orsay virus utilizes ribosomal frameshifting to express a novel protein that is incorporated into virions. *Virology.* 450–451:213–21
- Johnson TE, Friedman DB. 1988. A Mutation in the *age-1* Gene in *Caenorhabditis elegans* Lengthens Life and Reduces Hermaphrodite Fertility. *Genetics.* 118:75–86
- Jones JT, Haegeman A, Danchin EGJ, Gaur HS, Helder J, *et al.* 2013. Top 10 plant-parasitic nematodes in molecular plant pathology. *Mol. Plant Pathol.* 14(9):946–61
- Jovic K, Grilli J, Sterken MG, Snoek BL, Riksen JAG, *et al.* 2019. Transcriptome resilience predicts thermotolerance in *Caenorhabditis elegans*. *BMC Biol.* 17(1):1–12
- Jovic K, Sterken MG, Grilli J, Bevers RPJ, Rodriguez M, *et al.* 2017. Temporal dynamics of gene expression in heat-stressed *Caenorhabditis elegans*. *PLoS One.* 12(12):1–16
- Jurado P, Kodama E, Tanizawa Y, Mori I. 2010. Distinct thermal migration behaviors in response to different thermal gradients in *Caenorhabditis elegans*. *Genes, Brain Behav.* 9(1):120–27
- Kaletsky R, Moore RS, Vrla GD, Parsons LR, Gitai Z, Murphy CT. 2020. *C. elegans* interprets bacterial non-coding RNAs to learn pathogenic avoidance. *Nature.* 586(7829):445–51
- Kammenga JE. 2017. The background puzzle: how identical mutations in the same gene lead to different disease symptoms. *FEBS J.* 284(20):3362–73
- Kasimatis KR, Moerdyk-Schauwecker MJ, Phillips PC. 2018. Auxin-mediated sterility induction system for longevity and mating studies in *Caenorhabditis elegans*. *G3 Genes|Genomes|Genetics.* 8(8):2655–62
- Kavaliers M, Choleris E. 2018. The role of social cognition in parasite and pathogen avoidance. *Philos. Trans. R. Soc. B Biol. Sci.* 373(1751):20170206
- Kemp C, Mueller S, Goto A, Barbier V, Paro S, *et al.* 2013. Broad RNA Interference–Mediated Antiviral Immunity and Virus-Specific Inducible Responses in *Drosophila*. *J. Immunol.* 190(2):650–58
- Kenney SJ, Anderson GL, Williams PL, Millner PD, Beuchat LR. 2006. Migration of *Caenorhabditis elegans* to manure and manure compost and potential for transport of *Salmonella* newport to fruits and vegetables. *Int. J. Food Microbiol.* 106(1):61–68
- Kenyon C, Chang J, Gensch E, Rudner A, Tabtiang R. 1993. A *C. elegans* mutant that lives twice as long as wild type. *Nature.* 366(6454):461–64
- Kiesecker JM, Skelly DK, Beard KH, Preisser E. 1999. Behavioral reduction of infection risk. *Proc. Natl. Acad. Sci.* 96(16):9165–68
- Kim C, Kim J, Kim S, Cook DE, Evans KS, *et al.* 2019. Long-read sequencing reveals intra-species tolerance of substantial structural variations and new subtelomere formation in *C. elegans*. *Genome Res.* 29(6):1023–35
- King EG, Macdonald SJ, Long AD. 2012a. Properties and power of the *Drosophila* synthetic population resource for the routine dissection of complex traits. *Genetics.* 191(3):935–49
- King EG, Merkes CM, Mcneil CL, Hoofer SR, Sen S, *et al.* 2012b. Genetic dissection of a model complex trait using the *Drosophila* Synthetic Population Resource. *Genome Res.* 22:1558–66

- Kleemann GA, Basolo AL. 2007. Facultative decrease in mating resistance in hermaphroditic *Caenorhabditis elegans* with self-sperm depletion. *Anim. Behav.* 74(5):1339–47
- Klein SL, Flanagan KL. 2016. Sex differences in immune responses. *Nat. Rev. Immunol.* 16(10):626–38
- Kondrashov AS. 1984. Deleterious mutations as an evolutionary factor: I. The advantage of recombination. *Genet. Res.* 44(2):199–217
- Kondrashov AS. 1985. Deleterious mutations as an evolutionary factor. II. Facultative apomixis and selfing. *Genetics.* 111(3):635–53
- Koubkova-Yu TC-T, Chao J-C, Leu J-Y. 2018. Heterologous Hsp90 promotes phenotypic diversity through network evolution. *PLOS Biol.* 16(11):e2006450
- Kruijer W, Boer MP, Malosetti M, Flood PJ, Engel B, *et al.* 2014. Marker-based estimation of heritability in immortal populations. *Genetics.* 199(2):379–98
- Kulesa A, Krzywinski M, Blainey P, Altman N. 2015. Points of Significance: Sampling distributions and the bootstrap. *Nat. Methods.* 12(6):477–78
- Kuznetsova A, Brockhoff PB, Christensen RHB. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *J. Stat. Softw.* 82(13):1–26
- Langmead B, Salzberg S. 2013. Fast gapped-read alignment with Bowtie2. *Nat. Methods.* 9(4):357–59
- Laplana M, Caruz A, Pineda JA, Puig T, Fibla J. 2013. Association of BST-2 gene variants with HIV disease progression underscores the role of BST-2 in HIV type 1 infection. *J. Infect. Dis.* 207(3):411–19
- Larkin MA, Blackshields G, Brown NP, Chenna R, Mcgettigan PA, *et al.* 2007. Clustal W and Clustal X version 2.0. *Bioinformatics.* 23(21):2947–48
- Lässig C, Hopfner KP. 2017. Discrimination of cytosolic self and non-self RNA by RIG-I-like receptors. *J. Biol. Chem.* 292(22):9000–9009
- Lazear HM, Govero J, Smith AM, Platt DJ, Fernandez E, *et al.* 2016. A Mouse Model of Zika Virus Pathogenesis. *Cell Host Microbe.* 19(5):720–30
- Le Pen J, Jiang H, Di Domenico T, Kneuss E, Kosalka J, *et al.* 2018. Terminal uridylyltransferases target RNA viruses as part of the innate immune system. *Nat. Struct. Mol. Biol.* 25(9):778–86
- LeBoeuf B, Correa P, Jee C, García LR. 2014. *Caenorhabditis elegans* male sensory-motor neurons and dopaminergic support cells couple ejaculation and post-ejaculatory behaviors. *Elife.* 3:1–32
- Lee D, Yang H, Kim J, Brady S, Zdraljevic S, *et al.* 2017. The genetic basis of natural variation in a phoretic behavior. *Nat. Commun.* 8(1):273
- Lee D, Zdraljevic S, Cook DE, Frézal L, Hsu J-C, *et al.* 2019. Selection and gene flow shape niche-associated variation in pheromone response. *Nat. Ecol. Evol.* 3(10):1455–63
- Lee D, Zdraljevic S, Stevens L, Wang Y, Tanny RE, *et al.* 2020. Balancing selection maintains ancient genetic diversity in *C. elegans*. *bioRxiv*
- Lee H, Choi MK, Lee D, Kim HS, Hwang H, *et al.* 2012. Nictation, a dispersal behavior of the nematode *Caenorhabditis elegans*, is regulated by IL2 neurons. *Nat. Neurosci.* 15(1):107–12
- Lee J-H, Muhsin M, Atienza GA, Kwak D-Y, Kim S-M, *et al.* 2010. Single Nucleotide Polymorphisms in a Gene for Translation Initiation Factor (eIF4G) of Rice (*Oryza sativa*) Associated with Resistance to Rice tungro spherical virus. *Mol. Plant-Microbe Interact.* 23(1):29–38

- Lee RYN, Howe KL, Harris TW, Arnaboldi V, Cain S, *et al.* 2018. WormBase 2017: Molting into a new stage. *Nucleic Acids Res.* 46(D1):D869–74
- Lehtonen J, Jennions MD, Kokko H. 2012. The many costs of sex. *Trends Ecol. Evol.* 27(3):172–78
- Leyva-Díaz E, Stefanakis N, Carrera I, Glenwinkel L, Wang G, *et al.* 2017. Silencing of repetitive DNA is controlled by a member of an unusual *Caenorhabditis elegans* gene family. *Genetics.* 207(2):529–45
- Li Y, Álvarez OA, Gutteling EW, Tijsterman M, Fu J, *et al.* 2006. Mapping determinants of gene expression plasticity by genetical genomics in *C. elegans*. *PLoS Genet.* 2(12):2155–61
- Li Y, Breitling R, Snoek LB, Van Der Velde KJ, Swertz MA, *et al.* 2010. Global genetic robustness of the alternative splicing machinery in *Caenorhabditis elegans*. *Genetics.* 186(1):405–10
- Li YI, van de Geijn B, Raj A, Knowles DA, Petti AA, *et al.* 2016. RNA splicing is a primary link between genetic variation and disease. *Science.* 352(6285):600–604
- Lim JK, Lisco A, McDermott DH, Huynh L, Ward JM, *et al.* 2009. Genetic variation in OAS1 is a risk factor for initial infection with West Nile virus in man. *PLoS Pathog.* 5(2):e1000321
- Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, *et al.* 1996. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell.* 86(3):367–77
- Liu X, Li YI, Pritchard JK. 2019. Trans Effects on Gene Expression Can Drive Omnigenic Inheritance. *Cell.* 177(4):1022–1034.e6
- Loh P-R, Bhatia G, Gusev A, Finucane HK, Bulik-Sullivan BK, *et al.* 2015. Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. *Nat. Genet.* 47(12):1385–92
- Lopes PC, Sucena É, Santos ME, Magalhães S. 2008. Rapid Experimental Evolution of Pesticide Resistance in *C. elegans* Entails No Costs and Affects the Mating System. *PLoS One.* 3(11):e3741
- Luallen RJ, Reinke AW, Tong L, Botts MR, Félix MA, Troemel ER. 2016. Discovery of a Natural Microsporidian Pathogen with a Broad Tissue Tropism in *Caenorhabditis elegans*. *PLoS Pathog.* 12(6):1–28
- Lynch ZR, Penley MJ, Morran LT. 2018. Turnover in local parasite populations temporarily favors host outcrossing over self-fertilization during experimental evolution. *Ecol. Evol.* 8(13):6652–62
- Mackay TFC, Richards S, Stone EA, Barbadilla A, Ayroles JF, *et al.* 2012. The *Drosophila melanogaster* Genetic Reference Panel. *Nature.* 482(7384):173–78
- Magwire MM, Fabian DK, Schweyen H, Cao C, Longdon B, *et al.* 2012. Genome-Wide Association Studies Reveal a Simple Genetic Basis of Resistance to Naturally Coevolving Viruses in *Drosophila melanogaster*. *PLoS Genet.* 8(11):e1003057
- Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, *et al.* 2009. Finding the missing heritability of complex diseases. *Nature.* 461(7265):747–53
- Mariani V, Kiefer F, Schmidt T, Haas J, Schwede T. 2011. Template Based Assessment Assessment of template based protein structure predictions in CASP9. *Proteins Struct. Funct. Bioinforma.* 79(S10):37–58
- Marroquin LD, Elyassnia D, Griffiths JS, Feitelson JS, Aroian R V. 2000. Bacillus thuringiensis (Bt) toxin susceptibility and isolation of resistance mutants in the nematode *Caenorhabditis elegans*. *Genetics.* 155(4):1693–99
- Mashimo T, Lucas M, Simon-Chazottes D, Frenkiel M-P, Montagutelli X, *et al.* 2002. A nonsense mutation in the gene encoding 2'-5'-oligoadenylate synthetase/L1 isoform is associated with West Nile virus susceptibility in laboratory mice. *Proc. Natl. Acad. Sci.* 99(17):11311–16

- Masri L, Schulte RD, Timmermeyer N, Thanisch S, Crummenerl LL, *et al.* 2013. Sex differences in host defence interfere with parasite-mediated selection for outcrossing during host-parasite coevolution. *Ecol. Lett.* 16(4):461–68
- Maynard Smith J. 1978. *The Evolution of Sex*. Cambridge University Press
- McLaren PJ, Coulonges C, Bartha I, Lenz TL, Deutsch AJ, *et al.* 2015. Polymorphisms of large effect explain the majority of the host genetic contribution to variation of HIV-1 virus load. *Proc. Natl. Acad. Sci.* 112(47):14658–63
- Mestas J, Hughes CCW. 2004. Of Mice and Not Men: Differences between Mouse and Human Immunology. *J. Immunol.* 172(5):2731–38
- Mitani S. 2009. Nematode, an experimental animal in the National BioResource Project. *Exp. Anim.* 58(4):351–56
- Mitchell-Olds T, Willis JH, Goldstein DB. 2007. Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nat. Rev. Genet.* 8(11):845–56
- Morran LT, Cappy BJ, Anderson JL, Phillips PC. 2009. Sexual Partners for the Stressed: Facultative Outcrossing in the Self-Fertilizing Nematode *C. elegans*. *Evolution.* 63(6):1473–82
- Morran LT, Schmidt OG, Gelarden IA, Parrish RC, Lively CM. 2011. Running with the Red Queen: Host-parasite coevolution selects for biparental sex. *Science.* 333(6039):216–18
- Mussabekova A, Daeffler L, Imler JL. 2017. Innate and intrinsic antiviral immunity in *Drosophila*. *Cell. Mol. Life Sci.* 74(11):1–16
- Nadeau JH, Singer JB, Matin A, Lander ES. 2000. Analysing complex genetic traits with chromosome substitution strains. *Nat. Genet.* 24(3):221–25
- Ndungo E, Herbert AS, Raaben M, Obernosterer G, Biswas R, *et al.* 2016. A Single Residue in Ebola Virus Receptor NPC1 Influences Cellular Host Range in Reptiles. *mSphere.* 1(2):1–15
- Nedelko T, Kollmus H, Klawonn F, Spijker S, Lu L, *et al.* 2012. Distinct gene loci control the host response to influenza H1N1 virus infection in a time-dependent manner. *BMC Genomics.* 13(1):411
- Neil SJD, Zang T, Bieniasz PD. 2008. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature.* 451(7177):425–30
- Ng M, Ndungo E, Kaczmarek ME, Herbert AS, Binger T, *et al.* 2015. Filovirus receptor NPC1 contributes to species-specific patterns of ebolavirus susceptibility in bats. *Elife.* 4:e11785
- Nguyen A, David JK, Maden SK, Wood MA, Weeder BR, *et al.* 2020. Human Leukocyte Antigen Susceptibility Map for Severe Acute Respiratory Syndrome Coronavirus 2. *J. Virol.* 94(13):1–12
- Noble LM, Chang AS, McNelis D, Kramer M, Yen M, *et al.* 2015. Natural Variation in *plep-1* Causes Male-Male Copulatory Behavior in *C. Elegans*. *Curr. Biol.* 25(20):2730–37
- Noble LM, Chelo I, Guzella T, Afonso B, Riccardi DD, *et al.* 2017. Polygenicity and Epistasis Underlie Fitness-Proximal Traits in the *Caenorhabditis elegans* Multiparental Experimental Evolution (CeMEE) Panel. *Genetics.* 207(4):1663–85
- Obbard DJ, Jiggins FM, Halligan DL, Little TJ. 2006. Natural Selection Drives Extremely Rapid Evolution in Antiviral RNAi Genes. *Curr. Biol.* 16:580–85
- Osman GA, Fasseas MK, Koneru SL, Essmann CL, Kyrou K, *et al.* 2018. Natural Infection of *C. elegans* by an Oomycete Reveals a New Pathogen-Specific Immune Response. *Curr. Biol.* 28(4):640–648.e5

- Ovsyannikova IG, Ryan JE, Vierkant RA, Pankratz VS, Jacobson RM, Poland GA. 2005. Immunologic significance of HLA class I genes in measles virus-specific IFN- $\gamma$  and IL-4 cytokine immune responses. *Immunogenetics*. 57(11):828–36
- Paciência FMD, Rushmore J, Chuma IS, Lipende IF, Caillaud D, *et al.* 2019. Mating avoidance in female olive baboons (*Papio anubis*) infected by *Treponema pallidum*. *Sci. Adv.* 5(12):1–8
- Palopoli MF, Rockman M V., TinMaung A, Ramsay C, Curwen S, *et al.* 2008. Molecular basis of the copulatory plug polymorphism in *Caenorhabditis elegans*. *Nature*. 454(7207):1019–22
- Panek J, Gang SS, Reddy KC, Luallen RJ, Fulzele A, *et al.* 2020. A cullin-RING ubiquitin ligase promotes thermotolerance as part of the intracellular pathogen response in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* 117(14):7950–60
- Park J-H, Gail MH, Weinberg CR, Carroll RJ, Chung CC, *et al.* 2011. Distribution of allele frequencies and effect sizes and their interrelationships for common genetic susceptibility variants. *Proc. Natl. Acad. Sci.* 108(44):18026–31
- Peitsch MC. 1997. SWISS-MODEL and the Swiss-PdbViewer : An environment for comparative protein modeling. *Electrophoresis*. 18:2714–23
- Perelygin AA, Scherbik S V, Zhulin IB, Stockman BM, Li Y, Brinton MA. 2002. Positional cloning of the murine flavivirus resistance gene. *Proc. Natl. Acad. Sci.* 99(14):9322–27
- Perez MF, Francesconi M, Hidalgo-Carcedo C, Lehner B. 2017. Maternal age generates phenotypic variation in *Caenorhabditis elegans*. *Nature*. 552(7683):106–9
- Petersen C, Dirksen P, Prah S, Strathmann E, Schulenburg H. 2014. The prevalence of *Caenorhabditis elegans* across 1.5 years in selected North German locations: the importance of substrate type, abiotic parameters, and *Caenorhabditis* competitors. *BMC Ecol.* 14(1):4
- Petersen C, Dirksen P, Schulenburg H. 2015. Why we need more ecology for genetic models such as *C. elegans*. *Trends Genet.* 31(3):120–27
- Pfeifer B, Wittelsbürger U, Ramos-Onsins SE, Lercher MJ. 2014. PopGenome: An efficient swiss army knife for population genomic analyses in R. *Mol. Biol. Evol.* 31(7):1929–36
- Piasecka B, Duffy D, Urrutia A, Quach H, Patin E, *et al.* 2018. Distinctive roles of age, sex, and genetics in shaping transcriptional variation of human immune responses to microbial challenges. *Proc. Natl. Acad. Sci.* 115(3):E488–97
- Pillay D. 2009. *Principles of Virology*. The Lancet Infectious Diseases
- Pisarski K. 2019. The global burden of disease of zoonotic parasitic diseases: Top 5 contenders for priority consideration. *Trop. Med. Infect. Dis.* 4(1):44
- Pittman KJ, Glover LC, Wang L, Ko DC. 2016. The Legacy of Past Pandemics: Common Human Mutations That Protect against Infectious Disease. *PLoS Pathog.* 12(7):e1005680
- Pollard TD. 2010. A guide to simple and informative binding assays. *Mol. Biol. Cell.* 21(23):4061–67
- Pungalija C, Srinivasan J, Fox BW, Malik RU, Ludewig AH, *et al.* 2009. A shortcut to identifying small molecule signals that regulate behavior and development in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* 106(19):7708–13
- Raj A, Tyagi S. 2010. *Detection of Individual Endogenous RNA Transcripts in Situ Using Multiple Singly Labeled Probes*. Elsevier Inc.
- Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S. 2008. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods*. 5(10):877–79



- Rasmussen AL, Okumura A, Ferris MT, Green R, Feldmann F, *et al.* 2014. Host genetic diversity enables Ebola hemorrhagic fever pathogenesis and resistance. *Science*. 346(6212):987–91
- Reddy KC, Dror T, Sowa JN, Panek J, Chen K, *et al.* 2017. An Intracellular Pathogen Response Pathway Promotes Proteostasis in *C. elegans*. *Curr. Biol.* 27(22):3544–3553.e5
- Reddy KC, Dror T, Underwood RS, Osman GA, Desjardins CA, *et al.* 2019. Antagonistic paralogs control a switch between growth and pathogen resistance in *C. elegans*. *Plos Pathog.* 15(1):e1007528
- Remmert M, Biegert A, Hauser A. 2012. HHblits : lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat. Methods*. 9(2):
- Richaud A, Zhang G, Lee D, Lee J, Félix M-A. 2018. The Local Coexistence Pattern of Selfing Genotypes in *Caenorhabditis elegans* Natural Metapopulations. *Genetics*. 208(2):807–21
- Rockman M V. 2012. The QTN program and the alleles that matter for evolution: All that's gold does not glitter. *Evolution*. 66(1):1–17
- Rockman M V., Kruglyak L. 2009. Recombinational landscape and population genomics of *Caenorhabditis elegans*. *PLoS Genet.* 5(3):e1000419
- Rockman M V., Skrovanek SS, Kruglyak L. 2010. Selection at Linked Sites Shapes Heritable Phenotypic Variation in *C. elegans*. *Science*. 330(6002):372–76
- Rodriguez M, Snoek LB, Riksen JAG, Bevers RP, Kammenga JE. 2012. Genetic variation for stress-response hormesis in *C. elegans* lifespan. *Exp. Gerontol.* 47(8):581–87
- Rogers C, Persson A, Cheung B, de Bono M. 2006. Behavioral Motifs and Neural Pathways Coordinating O<sub>2</sub> Responses and Aggregation in *C. elegans*. *Curr. Biol.* 16(7):649–59
- Sabin LR, Zhou R, Gruber JJ, Lukinova N, Bambina S, *et al.* 2009. Ars2 regulates both miRNA- and siRNA-dependent silencing and suppresses RNA virus infection in *Drosophila*. *Cell*. 138(2):340–51
- Salek-Ardakani S, Arrand JR, Mackett M. 2002. Epstein–Barr Virus Encoded Interleukin-10 Inhibits HLA-Class I, ICAM-1, and B7 Expression on Human Monocytes: Implications for Immune Evasion by EBV. *Virology*. 304(2):342–51
- Šali A, Blundell TL. 1993. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* 234:779–815
- Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, *et al.* 1996. Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature*. 382(6593):722–25
- Samuel BS, Rowedder H, Braendle C, Félix M-A, Ruvkun G. 2016. *Caenorhabditis elegans* responses to bacteria from its natural habitats. *Proc. Natl. Acad. Sci.* 113(27):E3941–49
- Sandoval LE, Jiang H, Wang D. 2019. The Dietary Restriction-Like Gene *drl-1*, Which Encodes a Putative Serine/Threonine Kinase, Is Essential for Orsay Virus Infection in *Caenorhabditis elegans*. *J. Virol.* 93(3):e01400-18
- Sarkies P, Ashe A, Le Pen J, McKie MA, Miska EA. 2013. Competition between virus-derived and endogenous small RNAs regulates gene expression in *Caenorhabditis elegans*. *Genome Res.* 23(8):1258–70
- Satterthwaite FE. 1941. Synthesis of variance. *Psychometrika*. 6(5):309–16
- Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, *et al.* 2012. Fiji: An open-source platform for biological-image analysis. *Nat. Methods*. 9(7):676–82
- Schlee M. 2013. Master sensors of pathogenic RNA - RIG-I like receptors. *Immunobiology*. 218(11):1322–35

- Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods.* 9(7):671–75
- Schulenburg H, Félix M. 2017. The Natural Biotic Environment of *Caenorhabditis elegans*. *Genetics.* 206(1):55–86
- Schulte RD, Hasert B, Makus C, Michiels NK, Schulenburg H. 2012. Increased responsiveness in feeding behaviour of *Caenorhabditis elegans* after experimental coevolution with its microparasite *Bacillus thuringiensis*. *Biol. Lett.* 8(2):234–36
- Scully EP, Haverfield J, Ursin RL, Tannenbaum C, Klein SL. 2020. Considering how biological sex impacts immune responses and COVID-19 outcomes. *Nat. Rev. Immunol.* 20(7):442–47
- Shi H, Kichaev G, Pasaniuc B. 2016. Contrasting the Genetic Architecture of 30 Complex Traits from Summary Association Data. *Am. J. Hum. Genet.* 99(1):139–53
- Shin D-L, Hatesuer B, Bergmann S, Nedelko T, Schughart K. 2015. Protection from Severe Influenza Virus Infections in Mice Carrying the *Mx1* Influenza Virus Resistance Gene Strongly Depends on Genetic Background. *J. Virol.* 89(19):9998–10009
- Shultz LD, Brehm MA, Garcia JV, Greiner DL. 2012. Humanized mice for immune system investigation: progress, promise and challenges. *Nat. Rev. Immunol.* 12(11):786–98
- Sivasundar A, Hey J. 2005. Sampling from natural populations with RNAi reveals high outcrossing and population structure in *Caenorhabditis elegans*. *Curr. Biol.* 15(17):1598–1602
- Slowinski SP, Morran LT, Parrish RC, Cui ER, Bhattacharya A, et al. 2016. Coevolutionary interactions with parasites constrain the spread of self-fertilization into outcrossing host populations. *Evolution.* 70(11):2632–39
- Smyth GK, Speed T. 2003. Normalization of cDNA microarray data. *Methods.* 31(4):265–73
- Snoek BL, Sterken MG, Bevers RPJ, Volkers RJM, van't Hof A, et al. 2017. Contribution of trans regulatory eQTL to cryptic genetic variation in *C. elegans*. *BMC Genomics.* 18(1):1–15
- Snoek BL, Sterken MG, Hartanto M, van Zuilichem AJ, Kammenga JE, et al. 2020. WormQTL2: an interactive platform for systems genetics in *Caenorhabditis elegans*. *Database.* 2020:1–17
- Snoek BL, Volkers RJM, Nijveen H, Petersen C, Dirksen P, et al. 2019. A multi-parent recombinant inbred line population of *C. elegans* allows identification of novel QTLs for complex life history traits. *BMC Biol.* 17(1):1–17
- Snoek LB, Sterken MG, Volkers RJM, Klatter M, Bosman KJ, et al. 2015. A rapid and massive gene expression shift marking adolescent transition in *C. elegans*. *Sci. Rep.* 4(1):3912
- Sowa JN, Jiang H, Somasundaram L, Tecle E, Xu G, et al. 2019. The *Caenorhabditis elegans* RIG-I Homolog DRH-1 Mediates the Intracellular Pathogen Response upon Viral Infection. *J. Virol.* 94(2):e01173–19
- Speed D, Hemani G, Johnson MR, Balding DJ. 2012. Improved heritability estimation from genome-wide SNPs. *Am. J. Hum. Genet.* 91(6):1011–21
- Srinivasan J, Kaplan F, Ajredini R, Zachariah C, Alborn HT, et al. 2008. A blend of small molecules regulates both mating and development in *Caenorhabditis elegans*. *Nature.* 454(7208):1115–18
- Stein L, Sternberg PW, Durbin R, Thierry-Mieg J, Spieth J. 2002. WormBase: network access to the genome and biology of *Caenorhabditis elegans*. *Nucleic Acids Res.* 29(1):82–86
- Sterken MG, Snoek LB, Bosman KJ, Daamen J, Riksen JAG, et al. 2014. A heritable antiviral RNAi response limits orsay virus infection in *Caenorhabditis elegans* N2. *PLoS One.* 9(2):e89760

- Sterken MG, Snoek LB, Kammenga JE, Andersen EC. 2015. The laboratory domestication of *Caenorhabditis elegans*. *Trends Genet.* 31(5):224–31
- Stern S, Kirst C, Bargmann CI. 2017. Neuromodulatory Control of Long-Term Behavioral Patterns and Individuality across Development. *Cell.* 171(7):1649–1662.e10
- Sulston JE, Horvitz HR. 1977. Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev. Biol.* 56(1):110–56
- Sulston JE, Schierenberg E, White JG, Thomson JN. 1983. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* 100(1):64–119
- Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, *et al.* 1999. The *rde-1* gene, RNA interference, and transposon silencing in *C. elegans*. *Cell.* 99(2):123–32
- Tajima F. 1989. Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics.* 123:585–95
- Tan MW, Mahajan-Miklos S, Ausubel FM. 1999. Killing of *Caenorhabditis elegans* by *Pseudomonas aeruginosa* used to model mammalian bacterial pathogenesis. *Proc. Natl. Acad. Sci.* 96(2):715–20
- Tanguy M, Véron L, Stempor P, Ahringer J, Sarkies P, Miska EA. 2017. An alternative STAT signaling pathway acts in viral immunity in *Caenorhabditis elegans*. *MBio.* 8(5):1–16
- Teotonio H, Estes S, Phillips PC, Baer CF. 2017. Experimental Evolution with *Caenorhabditis* Nematodes. *Genetics.* 206:691–716
- Teotónio H, Manoel D, Phillips PC. 2006. Genetic variation for outcrossing among *Caenorhabditis elegans* isolates. *Evolution.* 60(6):1300–1305
- Tepper RG, Ashraf J, Kaletsky R, Kleemann G, Murphy CT, Bussemaker HJ. 2013. PQM-1 complements DAF-16 as a key transcriptional regulator of DAF-2-mediated development and longevity. *Cell.* 154(3):676–90
- Thach DC, Kleeberger SR, Tucker PC, Griffin DE. 2001. Genetic Control of Neuroadapted Sindbis Virus Replication in Female Mice Maps to Chromosome 2 and Associates with Paralysis and Mortality. *J. Virol.* 75(18):8674–80
- Thangavel RR, Bouvier NM. 2014. Animal models for influenza virus pathogenesis, transmission, and immunology. *J. Immunol. Methods.* 410:60–79
- The *C. elegans* Sequencing Consortium. 1998. Genome Sequence of the Nematode *C. elegans*: A Platform for Investigating Biology. *Science.* 282(5396):2012–18
- Thomas CG, Li R, Smith HE, Woodruff GC, Oliver B, Haag ES. 2012. Simplification and desexualization of gene expression in self-fertile nematodes. *Curr. Biol.* 22(22):2167–72
- Thomas DL, Thio CL, Martin MP, Qi Y, Ge D, *et al.* 2009. Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature.* 461(7265):798–801
- Thomas JH. 2006. Adaptive evolution in two large families of ubiquitin-ligase adapters in nematodes and plants. *Genome Res.* 16(8):1017–30
- Thompson O, Edgley M, Strasbourger P, Flibotte S, Ewing B, *et al.* 2013. The million mutation project: A new approach to genetics in *Caenorhabditis elegans*. *Genome Res.* 23(10):1749–62
- Thompson OA, Snoek LB, Nijveen H, Sterken MG, Volkers RJMM, *et al.* 2015. Remarkably Divergent Regions Punctuate the Genome Assembly of the *Caenorhabditis elegans* Hawaiian Strain CB4856. *Genetics.* 200(3):975–89

- Timmons L, Luna H, Martinez J, Moore Z, Nagarajan V, *et al.* 2014. Systematic comparison of bacterial feeding strains for increased yield of *Caenorhabditis elegans* males by RNA interference-induced non-disjunction. *FEBS Lett.* 588(18):3347–51
- Todesco M, Owens GL, Bercovich N, Légaré J-S, Soudi S, *et al.* 2020. Massive haplotypes underlie ecotypic differentiation in sunflowers. *Nature.* 584(7822):602–7
- Trapnell C, Furlan SN, Waterston RH, Huynh C, Cao J, *et al.* 2017. Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science.* 357(6352):661–67
- Travers THE. 1999. How the worm was won. *Trends Genet.* 15(2):51–58
- Traversa D. 2012. Pet roundworms and hookworms: A continuing need for global worming. *Parasites and Vectors.* 5(1):91
- Troemel ER, Félix MA, Whiteman NK, Barrière A, Ausubel FM. 2008. Microsporidia are natural intracellular parasites of the nematode *Caenorhabditis elegans*. *PLoS Biol.* 6(12):
- Turan K, Mibayashi M, Sugiyama K, Saito S, Numajiri A, Nagata K. 2004. Nuclear MxA proteins form a complex with influenza virus NP and inhibit the transcription of the engineered influenza virus genome. *Nucleic Acids Res.* 32(2):643–52
- Tybur JM, Gangestad SW. 2011. Mate preferences and infectious disease: Theoretical considerations and evidence in humans. *Philos. Trans. R. Soc. B Biol. Sci.* 366(1583):3375–88
- Van Den Berg MCW, Woerlee JZ, Ma H, May RC. 2006. Sex-dependent resistance to the pathogenic fungus *Cryptococcus neoformans*. *Genetics.* 173(2):677–83
- van den Hoogen J, Geisen S, Routh D, Ferris H, Traunsperger W, *et al.* 2019. Soil nematode abundance and functional group composition at a global scale. *Nature.* 572(7768):194–98
- van der Sijde MR, Ng A, Fu J. 2014. Systems genetics: From GWAS to disease pathways. *Biochim. Biophys. Acta - Mol. Basis Dis.* 1842(10):1903–9
- van Manen D, van 't Wout AB, Schuitemaker H. 2012. Genome-wide association studies on HIV susceptibility, pathogenesis and pharmacogenomics. *Retrovirology.* 9(70):70
- Van Valen L. 1973. A new evolutionary theory. *Evol. Theory.* 1:1–30
- Vasseur E, Patin E, Laval G, Pajon S, Fornarino S, *et al.* 2011. The selective footprints of viral pressures at the human RIG-I-like receptor family. *Hum. Mol. Genet.* 20(22):14–16
- Victor Garcia J. 2016. Humanized mice for HIV and AIDS research. *Curr. Opin. Virol.* 19:56–64
- Viñuela A, Snoek LB, Riksen JAG, Kammenga JE. 2010. Genome-wide gene expression regulation as a function of genotype and age in *C. elegans*. *Genome Res.* 20(7):929–37
- Viñuela A, Snoek LB, Riksen JAG, Kammenga JE. 2012. Aging Uncouples Heritability and Expression-QTL in *Caenorhabditis elegans*. *G3 Genes[Genomes]Genetics.* 2(5):597–605
- Visscher PM, Medland SE, Ferreira MAR, Morley KI, Zhu G, *et al.* 2006. Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLoS Genet.* 2(3):0316–25
- Volkers RJ, Snoek L, Hubar CJ van H, Coopman R, Chen W, *et al.* 2013. Gene-environment and protein-degradation signatures characterize genomic and phenotypic diversity in wild *Caenorhabditis elegans* populations. *BMC Biol.* 11(1):93
- Von Reuss SH, Bose N, Srinivasan J, Yim JJ, Judkins JC, *et al.* 2012. Comparative Metabolomics Reveals Biogenesis of Ascarosides, a Modular Library of Small-Molecule Signals in *C. elegans*. *J. Am. Chem. Soc.* 134(3):1817–24

- Vu V, Verster AJ, Schertzberg M, Chuluunbaatar T, Spensley M, *et al.* 2015. Natural Variation in Gene Expression Modulates the Severity of Mutant Phenotypes. *Cell*. 162(2):391–402
- Wang L, Pittman KJ, Barker JR, Salinas RE, Stanaway IB, *et al.* 2018. An Atlas of Genetic Variation Linking Pathogen-Induced Cellular Traits to Human Disease. *Cell Host Microbe*. 24(2):308–323.e6
- Wang YA, Snoek BL, Sterken MG, Riksen JAG, Stastna JJ, *et al.* 2019. Genetic background modifies phenotypic and transcriptional responses in a *C. elegans* model of  $\alpha$ -synuclein toxicity. *BMC Genomics*. 20(1):1–12
- Webster CL, Waldron FM, Robertson S, Crowson D, Ferrari G, *et al.* 2015. The discovery, distribution, and evolution of viruses associated with *Drosophila melanogaster*. *PLoS Biol*. 13(7):e1002210
- Wegewitz V, Schulenburg H, Streit A. 2008. Experimental insight into the proximate causes of male persistence variation among two strains of the androdioecious *Caenorhabditis elegans* (Nematoda). *BMC Ecol*. 8:1–12
- Wei Z, Liu Y, Xu H, Tang K, Wu H, *et al.* 2015. Genome-Wide Association Studies of HIV-1 Host Control in Ethnically Diverse Chinese Populations. *Sci. Rep.* 5(1):10879
- Welton AR, Chesler EJ, Sturkie C, Anne U, Hirsch GN, *et al.* 2005. Identification of Quantitative Trait Loci for Susceptibility to Mouse Adenovirus Type 1. *J. Virol.* 79(17):11517–22
- Wilfert L, Jiggins FM. 2010. Disease association mapping in *Drosophila* can be replicated in the wild. *Biol. Lett.* 6(5):666–68
- Wilke CO, Sawyer SL. 2016. At the mercy of viruses. *Elife*. 5:e12469
- Woodruff GC, Knauss CM, Mangel TK, Haag ES. 2014. Mating damages the cuticle of *C. elegans* hermaphrodites. *PLoS One*. 9(8):1–5
- Xiao H, Killip MJ, Staeheli P, Randall RE, Jackson D. 2013. The Human Interferon-Induced MxA Protein Inhibits Early Stages of Influenza A Virus Infection by Retaining the Incoming Viral Genome in the Cytoplasm. *J. Virol.* 87(23):13053–58
- Yin D, Haag ES. 2019. Evolution of sex ratio through gene loss. *Proc. Natl. Acad. Sci.* 116(26):12919–24
- Yun S, Song B, Frank JC, Julander JG, Olsen AL, *et al.* 2018. Functional Genomics and Immunologic Tools : The Impact of Viral and Host Genetic Variations on the Outcome of Zika Virus Infection. *Viruses*. 10(422):1–28
- Zahurak M, Parmigiani G, Yu W, Scharpf RB, Berman D, *et al.* 2007. Pre-processing Agilent microarray data. *BMC Bioinformatics*. 8:1–13
- Zdravljec S, Fox BW, Strand C, Panda O, Tenjo FJ, *et al.* 2019. Natural variation in *C. elegans* arsenic toxicity is explained by differences in branched chain amino acid metabolism. *Elife*. 8:1–28
- Zhang G, Sachse M, Prevost M-C, Luallen RJ, Troemel ER, Felix M-A. 2016. A Large Collection of Novel Nematode-Infecting Microsporidia and Their Diverse Interactions with *Caenorhabditis elegans* and Other Related Nematodes. *PLoS Pathog*. 12(12):
- Zhang S-Y, Jouanguy E, Ugolini S, Smahi A, Elain G, *et al.* 2007. TLR3 Deficiency in Patients with Herpes Simplex Encephalitis. *Science*. 317(5844):1522–27
- Zhang Y, Lu H, Bargmann CI. 2005. Pathogenic bacteria induce aversive olfactory learning in *Caenorhabditis elegans*. *Nature*. 438(7065):179–84
- Zheng N, Schulman BA, Song L, Miller JJ, Jeffrey PD, *et al.* 2002. Structure of the Cul1–Rbx1–Skp1–F boxSkp2 SCF ubiquitin ligase complex. *Nature*. 416(6882):703–9
- Zigáčková D, Vaňáčová Š. 2018. The role of 3' end uridylation in RNA metabolism and cellular physiology. *Philos. Trans. R. Soc. B Biol. Sci.* 373(1762):20180171

- Zignego AL, Wojcik GL, Cacoub P, Visentini M, Casato M, *et al.* 2014. Genome-wide association study of hepatitis C virus- and cryoglobulin-related vasculitis. *Genes Immun.* 15(7):500–505
- Zúñiga J, Buendía-Roldán I, Zhao Y, Jiménez L, Torres D, *et al.* 2012. Genetic variants associated with severe pneumonia in A/H1N1 influenza infection. *Eur. Respir. J.* 39(3):604–10

## English summary

The nematode *Caenorhabditis elegans* belongs to the world's most powerful genetic model organisms. Studying the genome of this nematode is facilitated by the androdiecious (male-hermaphrodite) mode of reproduction. A single hermaphrodite can start a population that will contain hundreds of genetically identical individuals after only a couple of days. Males are not necessary for reproduction but are used to recombine genomes, for instance to introduce a mutation into a population. Although the nematode has provided a wealth of genetic knowledge, a large part of the *C. elegans* genome does not have a known function yet. Many of the genes without an assigned role will likely have a function in natural populations, for example by providing protection against the natural pathogens that are ubiquitous in nature. Viruses strongly shape the genome of their host and in this thesis, the interaction between the positive-strand RNA virus, the Orsay virus (OrV), and *C. elegans* was studied.

Chapter 2 reviews quantitative genetic studies investigating antiviral defense and discusses the practical tools to perform these studies in model organisms. Natural genetic variation in the genome of the host leads to different viral susceptibilities for individuals of the same species. Understanding the consequences of host genetic variation is expected to lead to better treatments and personalized medicine for human patients. But studying viral infections in humans comes with ethical and practical challenges. Therefore, quantitative genetic studies in model organisms can help to better understand the mechanisms by which host genetic variation defines viral susceptibility. To study host-virus interactions in three genetic model organisms, mouse (*Mus musculus*), fruit fly (*Drosophila melanogaster*) and nematode (*C. elegans*), different tools have been developed over the years. These include two- or multiparent Recombinant Inbred Lines (RILs) and Introgression Lines (ILs) that are genetic mosaics of their parental strains. Additionally, Genome Wide Association Studies (GWAS) link genetic variation in populations of genetically distinct individuals to viral susceptibility. The use of these tools has led to identification of genetic variants that contribute to viral susceptibility in evolutionary conserved pathways and has improved the understanding of human-infecting viruses such as West Nile, Influenza and Ebola.

Chapter 3 investigates natural genetic variation in antiviral defense in *C. elegans* populations. Because viruses represent such a strong selective pressure, their natural presence also leads to genetic changes in the genome of their host. Traces of pathogenic selection were identified in *C. elegans* by analyzing the genetic composition of antiviral genes of the Intracellular Pathogen Response (IPR) in strains that were collected worldwide. This led to the discovery that natural strains carry a limited set of *pals*-gene variants, a class of genes fundamental within the IPR, that are maintained in *C. elegans* populations by balancing selection. Only three haplotypes were found worldwide for the IPR regulators *pals-22* and *pals-25* and two strains with distinct *pals-22/pals-25* haplotypes had a different IPR activity. The *pals-22/pals-25* haplotype of the



standard reference strain, Bristol N2, was most common in currently samples strains. The N2 strain had low basal IPR expression that strongly increased upon OrV infection. Contrary, the Hawaiian strain CB4856, had a constantly active IPR that did not change after OrV exposure. The strain CB4856 had lower viral susceptibility than the N2 strain after short term exposure, which might relate to IPR activity.

Chapter 4 studies the genetic loci that determine viral susceptibility in N2 and CB4856. To uncover genetic variants that underlie the different viral susceptibility of N2 and CB4856, a quantitative trait locus (QTL) mapping approach was used. Thereto, a panel of N2xCB4856 RILs was infected and statistical associations between viral susceptibility and genetic background linked to chromosome IV. Using ILs from both genetic backgrounds, a small region containing 34 polymorphic genes was found to affect the viral susceptibility of N2 and CB4856. One of the genes located in this region is the IPR gene *cul-6*. This gene contains a single nucleotide polymorphism between N2 and CB4856 at a conserved site near the binding domain with another *cul-6* complex member. After infection of CRISPR-Cas9 allele swap strains it became clear that *cul-6* contributes to the viral susceptibility. Nevertheless, having a CB4856 *cul-6* allele only did not confer resistance. Together with the finding that multiple genetic loci contribute to differences between N2 and CB4856 this shows that viral susceptibility is a complex trait with a polygenic basis.

Chapter 5 investigated viral infection in mixed-sex populations of *C. elegans*. Sex is another factor that is genetically determined, and sexual differences can underlie different viral susceptibilities. Both sexes (males and hermaphrodites) of three genetically distinct strains were exposed to the OrV. Males of the reference strain N2 were more resistant to the OrV than hermaphrodites. This could result from higher IPR activity in males under standard conditions, possibly protecting them from infection. Viral presence can also change population and mating dynamics. Indeed, male frequencies increased in the isogenic populations of the genotypes CB4856 and JU1580. Moreover, males rather mated with healthy than infected hermaphrodites. Together, this shows that viral infection can result in flexible outcrossing in *C. elegans* populations.

The research presented in this thesis displays how viral presence can have shaped the genome of *C. elegans* and how genetic variation determines viral susceptibility to date. Although these observations have been made in the laboratory, they were placed in the context of field observations to provide a better ecological context for this model organism. The findings made here suggest an evolutionary advantage for individuals with an active IPR haplotype under pathogenic conditions. Moreover, these results indicate an ecological advantage of having males, even in the frequently isogenic populations that are found in nature. Together, this thesis invites for investigation of host-virus interactions in a more natural set-up to fully incorporate lab and field studies.

## Nederlandse samenvatting

De nematode *Caenorhabditis elegans* is een aaltje dat wereldwijd wordt gebruikt als modelorganisme voor genetische studies. Deze studies proberen het functioneren van onze genen te begrijpen. Zo wordt er bijvoorbeeld onderzocht wat de invloed is van natuurlijke variaties in de genen op allerlei eigenschappen, zoals bijvoorbeeld de vatbaarheid voor ziekten. *C. elegans* is hiervoor een uitermate geschikt model, omdat het een bijzondere manier van voorplanten heeft. De zichzelf bevruchtende *C. elegans* hermafrodieten hebben geen mannetje nodig om zich voort te planten. Daardoor is slechts één hermafrodit nodig om een populatie te verkrijgen van honderden genetisch identieke individuen. De veel zeldzamere mannetjes kunnen wel worden ingezet voor het recombineren van genomen of het introduceren van mutaties in de populatie. Dit is waardevol voor onderzoekers, want met behulp van genetische recombinatie kan het effect van natuurlijke voorkomende genetische variaties achterhaald worden. Door middel van mutaties kan een gen worden uitgezet en het effect van dat gen in het functioneren van het aaltje worden bestudeerd. Hoewel *C. elegans* één van de best bestudeerde organismen op aarde is, weten we van behoorlijk wat genen nog niet waarvoor ze dienen. Veel van deze genen hebben waarschijnlijk een functie in de natuur, en zullen daarom niet opvallen in de steriele en onveranderlijke omgeving waar lab studies plaatsvinden. Een deel van deze genen zou bijvoorbeeld kunnen beschermen tegen virale infectie. In deze thesis is daarom de interactie tussen een natuurlijk voorkomend virus, het Orsay virus (OrV) en *C. elegans* bestudeerd.

In hoofdstuk 2 worden studies bediscussieerd die het effect van natuurlijk voorkomende genetische variaties op de vatbaarheid voor virussen bestudeerden. Het uitgangspunt is het verkrijgen van een duidelijker beeld waarom de ene persoon soms wel ziek wordt (bijvoorbeeld van de griep) en de ander niet. Het begrijpen van deze genetische aanleg voor virale vatbaarheid leidt naar verwachting tot betere behandelingen en medicijnen voor patiënten. Virale ziekten kunnen vaak echter niet op een moleculair niveau in mensen worden bestudeerd, en er kunnen al helemaal geen gerichte studies worden uitgevoerd waarin mensen bijvoorbeeld expres ziek zouden worden gemaakt. Daarom gebruiken we zogenaamde modelorganismen om virale infecties te begrijpen. Dit hoofdstuk richt zich op drie modelorganismen: muis, fruitvlieg en nematode. In het hoofdstuk worden studies beschreven die zich richtten op de interactie tussen natuurlijke variaties in het genoom van de gastheer en virus in deze organismen. Daarbij worden de genetische hulpmiddelen die beschikbaar zijn om onderzoekers te helpen efficiënt en gericht onderzoek uit te voeren samengevat. Het gebruik van deze middelen heeft ervoor gezorgd dat we nu genetische variaties kennen die bijdragen aan de vatbaarheid voor humane virussen, zoals het West Nijl virus (veroorzaker Westnijkooorst), Influenza virus (veroorzaker griep) en Ebolavirus.

In hoofdstuk 3 wordt de natuurlijke variatie in antivirale genen in *C. elegans* onderzocht. Omdat virussen in de natuur een sterke selectiedruk geven (enkel de sterkste individuen overleven en kunnen hun genen doorgeven aan een volgende generatie), zorgen virussen na verloop van tijd voor genetische veranderingen in het genoom van hun gastheer. Om te kijken of virussen ook sporen hebben nagelaten in *C. elegans* werd de genoomdata van 330 wilde *C. elegans* lijnen vergeleken. Daaruit bleek dat de veel zogenaamde *pals*-genen die zorgen voor de antivirale Intracellulaire Pathogeen Reactie (IPR) slechts weinig variabel waren. En van de genen die relatief variabel waren, komen er wereldwijd maar een paar combinaties van de genen voor. De IPR regulatoren *pals-22* en *pals-25* sturen de IPR aan en wereldwijd vonden we slechts drie combinaties van genetische variaties. De *pals-22/pals-25* combinatie van de standaard gebruikte Engelse *C. elegans* lijn N2 kwam het meest voor. De N2 lijn laat een sterke IPR zien na infectie. De Hawaïaanse lijn CB4856 daarentegen had geen sterke IPR na infectie, maar onder normale omstandigheden waren de IPR genen al veel actiever in deze lijn. We vonden ook dat CB4856 een lagere virale vatbaarheid had dan N2 na korte blootstelling aan het Orsay virus, en dat zou daarom gerelateerd kunnen zijn aan de IPR activiteit vóór infectie.

In hoofdstuk 4 wordt er gezocht naar genen die het verschil in virale vatbaarheid tussen N2 en CB4856 kunnen verklaren. Hiervoor wordt een zogenaamde QTL analyse uitgevoerd. Hierbij wordt voor verschillende nakomelingen van een kruising tussen N2 en CB4856 de vatbaarheid voor virus gemeten. Die verschillende nakomelingen verschillen net zo van elkaar als broers en zussen doen. Van iedere van deze zogenaamde RIL lijnen kennen we ook alle genetische variaties. Vervolgens hebben we deze genetische variaties gelinkt aan de vatbaarheid voor virus van de RIL lijnen en zo twee genetische locaties op chromosoom IV ontdekt die waarschijnlijk bedragen aan het verschil in vatbaarheid tussen N2 en CB4856. Van één van deze locaties konden we experimenteel bevestigen dat deze een rol speelt. Dat deden we door IL lijnen te infecteren. In tegenstelling tot de RIL lijnen, die een mix van genetische fragmenten van de ouders zijn, bevatten IL lijnen slecht een klein genetisch fragment van één ouder in het genoom van de andere ouder. Hiermee kun je heel specifiek aantonen dat dit fragment een effect heeft als de vatbaarheid van de IL anders is dan die van de ouder. Op de locatie die wij vonden liggen 34 interessante genen, waaronder één IPR gen: het gen *cul-6*. Tussen N2 en CB4856 zit er een codeverschil in het gen dat resulteert in een ander eiwitproduct en mogelijk het functioneren van het eiwit in de cel verandert. Het effect van deze natuurlijk variatie hebben we getest door twee lijnen te gebruiken die het verschil in dit specifieke gen uitwisselen tussen N2 en CB4856. Deze lijnen zijn gemaakt met CRISPR, waardoor zulke specifieke uitwisselingen mogelijk zijn. Infecties in de CRISPR lijnen bevestigden dat *cul-6* bijdraagt aan virale vatbaarheid. De experimenten lieten ook zien dat je wel ziek kan worden door de verkeerde genetische variatie, maar niet automatisch beter wordt van de goede. Dat komt omdat meerdere genen een rol spelen in het bepalen van vatbaarheid voor een virus.

In hoofdstuk 5 is de virale vatbaarheid van mannelijke en hermafrodiete *C. elegans* nematoden onderzocht in drie verschillende *C. elegans* lijnen (N2, CB4856 en JU1580). De mannetjes van N2 bleken minder vatbaar voor het Orsay virus dan de vrouwtjes. De N2 mannetjes hadden ook een hogere IPR activiteit zonder dat er virus aanwezig is en wellicht biedt dit bescherming tijdens een infectie. Daarnaast hebben we het effect van een infectie op de dynamiek in de populatie en het paren van de nematoden bekeken. Normaal gesproken bevatten *C. elegans* populaties slechts weinig mannetjes, maar we vonden dat dit aantal tijdens een virusinfectie kan toenemen. We zagen ook dat mannetjes een voorkeur hadden om te paren met gezonde in plaats van geïnfecteerde hermafrodieten. Mogelijk verklaart deze partnerkeuze de toename van mannetjes: paren met gezonde hermafrodieten leidt immers tot een hoger aantal mannelijke nakomelingen. Dit toont aan dat *C. elegans* nematoden zich flexibel kunnen voorplanten: door middel van zelfbevruchting onder gunstige omstandigheden en door middel van paringen wanneer de populatie geïnfecteerd is met virus.

Het onderzoek in deze thesis toont aan hoe virale infecties het genoom van *C. elegans* kunnen hebben veranderd en hoe natuurlijke genetische variaties de vatbaarheid voor virus beïnvloeden. De observaties zijn allemaal in het laboratorium gedaan, maar wel vergeleken met veldobservaties om een beter ecologisch inzicht te krijgen in dit modelorganisme. De bevindingen suggereren dat individuen in virus geïnfecteerde populaties profijt kunnen hebben van een hoge IPR activiteit. Verder laten ze zien dat mannetjes een voordeel zouden kunnen opleveren voor *C. elegans* als soort, zelfs wanneer de populatie genetisch weinig variabel is (wat vaak wordt waargenomen in de natuur). Uiteindelijk zouden deze eigenschappen in een meer natuurlijke proefopstelling bestudeerd kunnen worden om lab- en veldstudies in de toekomst beter te verbinden.



## Acknowledgments

Here I would like to thank those that supported me during the last five years when I worked on my PhD thesis. I enjoyed working on my PhD and that is for a large part because you have made this time fun and satisfying.

Als eerste wil ik mijn begeleiders, Jan, Gorben en Mark bedanken. Jan, bedankt voor je vertrouwen in mijn werk en de kritische blik waarmee je mijn werk (met name het schrijven) hebt verbeterd. Jij zorgt voor samenhang binnen de *C. elegans* groep en ook in het bijzonder de studenten (waartoe ik ook eens behoorde) door iedereen te betrekken bij discussies over wetenschap, maar ook de rafelranden daarvan. Gorben, jij bedankt voor je virologische kijk op mijn gastheer-virus (of is het een virus-gastheer) project. Je hielp me om me ook thuis te voelen binnen de vakgroep Virologie en je wist altijd een positieve draai te geven aan mijn werk. Daarnaast waardeer (en deel) ik jouw enthousiasme voor allerlei niet-wetenschappelijke zaken, zoals auto's, tussen Kunst en Kitsch en aquaria (hoewel het sequencen van een blokje visvoer ook laat zien hoe je dan toch weer je werk hieraan kan verbinden). Mark, jij was vanaf het moment dat ik als MSc student startte bij Nematologie mijn begeleider (hoewel je dat op papier pas later werd voor mijn PhD). Dankzij jou heb ik leren programmeren (iets wat ik erg leuk blijkt te vinden) en je stond altijd klaar om mij te helpen (ook om 6 uur 's ochtends, 11 uur 's avonds of 3 dagen nadat onze vakanties begonnen waren..). We delen het plezier in het in de natuur verzamelen van nematoden, voor mij altijd een van de favoriete dagen van het jaar. Daarnaast werd in jouw proefschrift de basis gelegd voor mijn thesis en zorgde dat ervoor dat ik een goede start had voor mijn project.

Tijdens mijn project heb ik heel wat uren in het lab doorgebracht. Joost, mede dankzij jou was dit een prettige plek om te werken. Je hebt heel veel ervaring met het werken met *C. elegans* en het was altijd fijn jou te kunnen vragen om advies (of advies te krijgen als je zag dat ik iets niet handig aan het aanpakken was). Ook heb je me regelmatig geholpen met het uitvoeren van experimenten, van mannetjes tellen tot het uitvoeren van microarrays (in de inmiddels goede oude tijd). Verder heb je ook mijn thesis studenten geholpen op het lab, waardoor ze snel aan de slag konden en er altijd iemand in de buurt was als ze praktische vragen hadden. Ook bedankt dat je de camera set-up lang genoeg hebt gedoogd, zodat ik er toch een experiment mee uit kon voeren (na jarenlang gezegd te hebben hoe leuk ik die set-up vond en dat ik hem toch echt wel binnen een paar maanden een keer zou willen testen). Daarnaast wordt het lab draaiende gehouden door de andere technicians: Sven, Casper, Hein, Debbie en Rikus. Ik wil jullie bedanken voor de prettige, veilige en ook erg gezellige werkomgeving. Rikus, extra bedankt voor de keren dat je mij bijstond bij het aanzetten van de Sorval (een enorme centrifuge), zodat ik daadwerkelijk op de startknop durfde te drukken.

De vakgroep Nematologie is een hele fijne plek om te werken. Jaap, Geert, Aska, Hans, Arjen en Jan, bedankt voor het organiseren van de verschillende onderzoeken binnen Nema, en jullie inzet voor samenwerking tussen de groepen. Deze instelling heeft al tot veel successen geleid, en geeft iedereen een kans zichzelf te ontwikkelen.

Lisette, Manouk en Christel, jullie hulp zorgt ervoor dat de vakgroep soepel blijft draaien. Manouk, nogmaals bedankt voor het regelen van de betaling van een half-betaalde factuur nadat de financiële afdeling de rekening slechts half had geprint en betaald (nadat het papier in de printer op was..). Omdat we naast elkaar zaten weet ik hoeveel telefoontjes en e-mails het regelen hiervan hebben gekost en heeft me wederom laten inzien hoe blij we met jullie mogen zijn. Lisette, jij bedankt voor de structuur die je aanbrengt in de groep en het beantwoorden van al mijn vragen over het invullen van formulieren (niet mijn sterkste kant).

Liesbeth, Jet en Jessica, bedankt voor jullie inzet in het onderwijs. Nematologie verzorgt ontzettend veel onderwijs en jullie zorgen ervoor dat studies, vakken en thesen soepel verlopen. Ik ben blij dat ik vanaf komend jaar dit team mag versterken!

Ruud, Kim en Koen, bedankt voor jullie hulp met het werken aan eiwitten en ook voor de hulp die jullie hebben geboden aan mijn studenten om zich deze technieken eigen te maken.

Basten, bedankt voor je vele inzichten in statistiek en wetenschap. Ook nu je al een tijdje in Utrecht werkt sta je altijd klaar voor advies of een vrijdagmiddagborrel.

Sanja, Stefan, André, Martijn, Casper, José, Lotte and Erik, you work on a range of interesting scientific projects and provide inspiration for what a post-PhD career could look like.

I think few jobs involve experiencing as many failures as performing a PhD. Experiencing this together has always been better than suffering alone. Therefore, I like to thank my fellow PhDs for all the awesome and efficient moments during the last years. Kim, Koen, Ava and Yiru, we have been on a lab trip together to Canterbury. Although severe weather conditions (5cm of snow) shut down the university and cancelled our flight, once we arrived, we had a great symposium in a pub (sympubium) and a lot of fun together. Ava, we started our PhD on the same day (and handed in the reading version of the thesis five days apart) and I always felt happy that we could share our PhD process together. Kim, you are attentive and makes sure everyone feels well (not at least by providing all sorts of tea). Yiru, you are always happy and kind and I keep good memories of our 'outside work' hours in Barcelona. Koen, you bring nice chaos and energy to the group. Koen and Ava, we will see you again at the end of the hike. Paula, thank you for your jokes, honesty and entertaining during the coffee breaks. Jaap-Jan thank you for your humorous comments on life and running every task force Nema has ever known. Sonja, Amalia, Katharina and Octavina you were already somewhat further in your PhD progress and it was nice to learn from your experiences. Katharina, you have always impressed me with your professional presentations and you could help me out with some



data challenges. Octavina, you defended your PhD just before I hope to defend mine and I was happy we could share our final steps and discussions about future plans. Yiru, Yuqing and Qi, it was great fun celebrating Chinese New Year with you (twice) with hotpot and karaoke (I am afraid there is evidence). Yuqing, you followed up the work with the Orsay virus (after performing your MSc thesis on the topic first) and I am glad we worked on this topic together, especially since you such a sweet person. It always cheers me up to run into you or when you found me having 'maybe one question' (sadly this does not happen often anymore). Marijke, you already came to Nema as a BSc student and I am glad you decided to come back for your PhD to amuse us up with your humor. Iqbal and Myrna, both of you link the *C. elegans* group with different subgroups which provides all sorts of new insights. Myrna, I am also grateful that you proved that the CRISPR system works well for other mutations other than those that create dumpy or roller worms and that you were able to really integrate technique in our lab (together with the essential injection skills of Joost of course). Sara, I like the drive you have for your work (and your pizza). Nina, you are one of the gentler personalities in Nema, and at the same time you show great strength when it comes to improving things. Joris, you naturally blended into the group and I hope that we will organize a board game evening soon. Matthijs, you organized the PhD meetings that helped keeping in touch during the covid crises and you always have a plethora of interesting stories to tell. Vera, Alejandro, Jorge and Alex, I am sure you will provide novel ideas and new spirit to the group of PhDs.

My PhD project was shared between the Laboratory of Nematology and the Laboratory of Virology. Although I spent most of my time in the first, I have always felt welcome and part of the latter. I enjoyed the Monday morning seminars and joined during what I called 'my yearly virologists integration day' (the DAVS) where I could also get to know you better personally. I would like to especially thank the people that helped me directly and the PhDs of 'my cohort': Gorben, Corinne, Monique, Marleen, Melanie, Giel, Sandra, Simone, Irene, Bob, Han, Tessy, Haidong, Mandy, John, Judith and Linda, thank you! Corinne, I would like to thank you specifically for introducing the microscope to quite some of my students and helping me when I could not get the settings to work (let's forget that one time when we also had to ask Jan for help though). There will always be some inevitable discrepancy between scientists looking from a host or virus perspective (as witnessed by an anonymous reviewer stating: "It appears that the review has been written by a non-virologist."), yet I am glad to have been part of both worlds.

Over the last years I have had the pleasure to supervise BSc and MSc thesis students. I would like to thank Maarten, Yuqing, Marijke, Mitra, Niels, Tatiana, Mels, Emma, Marlieke, Jie, Sanne, Josh, Sophie, Marèl, Jenny, Nika, Steven and Victoria for the great time I had supervising you. Many of you have contributed also to the data in this thesis, but for sure each of you has provided me with new insights. I am glad we keep in touch every now and then and enjoy hearing how you are doing.

Ook via Utrecht Universiteit heb ik hulp gekregen bij mijn PhD project. Juliane en Amir, bedankt voor jullie hulp met de Biosorter, waardoor het voor mij mogelijk was enorme hoeveelheden wormen te sorteren. Jullie hulp was onmisbaar om het apparaat goed te laten lopen. Mike, jij bedankt voor je uitleg over CRISPR in *C. elegans* en het verstrekken van de plasmiden die daarvoor nodig zijn.

De eerste keer dat ik met CRISPR in aanraking kwam, was het enkel nog bekend als bacterieel immuunsysteem. Edze, bedankt voor de eerste echte kennismaking met wetenschap tijdens mijn BSc thesis. Ook Stineke wil ik bedanken voor de goede samenwerking, waardoor het uiteindelijk gelukt is dit werk te publiceren. De ervaringen die ik dankzij jullie heb opgedaan hebben me erg geholpen tijdens mijn PhD.

The courses and many of the seminars I have followed have been organized by the graduate school PE&RC. Claudius, Lennart, Jacqueline, Anneloes, Naomi and Sabine, your work is essential for us PhDs. I am glad we got to know each other better during my years in the PE&RC PhD Council (PPC). I would also like to thank my fellow PPC members for the teamwork and interesting meetings. A special thanks for Jannike, Alejandro, Ying, Rima, Rens, Zhilei, Judith, Jeroen, Martijn and Milou with whom I have organized the PE&RC day in 2017 and 2018. Thijs, also thanks for your help in that perspective and willingness to drive to Amsterdam so early that we arrived even before the opening of the building. Thijs, Naomi, Tessa and Tessy, thank you for chairing and organizing the meetings during these years.

Stefan en Silke, sinds de eerste dag van onze studie zijn we bevriend (onze whatsappgroep heet nog steeds naar het nummer dat ons groepje toen kreeg) en aangezien we allemaal een PhD zijn gaan doen hebben we de laatste jaren daar vaak over gesproken tijdens het etentjes in Wageningen, Heesum, Den Haag, Amsterdam en Utrecht.

Evine, bedankt voor het vertrouwen om voor jouw paarden te mogen zorgen. Het is nog steeds de beste manier om even helemaal je hoofd leeg te maken. Straks kunnen we samen een rit of wandeling maken.

Gabrielle en Suus, we zien elkaar helaas weinig, maar als we elkaar zien is het altijd als vanouds. Jullie hebben zelfs het lab gezien en weten hoe *C. elegans* eruitziet op een petrischaal. Het is fijn met jullie te spreken over alles dat niet met werk te maken heeft.

Ellen, Tessa, Renee, Jeanne, Joanne, Femke en Gill, bedankt voor alle leuke afspraken en mooie avonturen! Onze gezamenlijke reizen naar Limburg, Zweden en Thailand waren uiteraard mooie hoogtepunten de afgelopen jaren, maar het fijnste is de vertrouwde sfeer binnen een groep waarin iedereen elkaar inmiddels alweer tien jaar kent.

Linda, Andries, Niels, Bart, Daan, Siti, Gatske en Henk, jullie zijn een fantastische schoonfamilie: genoeg gekke gewoontes om om te kunnen lachen en ik heb me altijd zeer thuis gevoeld bij jullie. Jullie stonden altijd klaar als wij weer eens een projectje hadden bedacht: van verhuizen (ik weet nog die dag met windkracht 8 toen we een flinke kast kochten en op de dakdragers van de auto vervoerden), het opbouwen van de winkel en het maken van de konijnenren (zo groot dat er al door de burens werd gespeculeerd wat er toch gebouwd werd daar). Alles was een stuk lastiger geworden zonder jullie hulp.

Ook wil ik pap, mam, Tessa, René, Lennart, opa Frits, oma Ria en oma Tijger, mijn ooms, tantes, neven en nichten, bedanken voor de interesse in mijn project de afgelopen jaren. Het is altijd fijn thuis te komen en ik zie ernaar uit als jullie hier langskomen. Opa Frits en oma Ria, het moge duidelijk zijn waar de liefde voor natuur en scheikunde hun oorsprong vinden. Oma Tijger, het is altijd fijn om bezoek te komen en nieuws over de familie uit te wisselen. Pap en mam, bedankt voor al jullie wijze lessen en de onvoorwaardelijke steun en praktische hulp (de andere helft van onze verhuizing, de kippenren en het rijdend houden van onze fietsen). Tessa en René, altijd leuk om met jullie op stap te zijn en samen de natuur in te trekken. Lennart, ik vind het leuk dat je ook een PhD bent gaan doen, om daar binnen de familie met iemand over te sparren. Ik hoop dat we binnenkort allemaal weer eens een bordspelletje van je kunnen verliezen.

Sander, bedankt voor alle steun de afgelopen jaren. Jij hebt menig presentatie van mijn PhD gezien, de hele thesis nagelezen op spelfouten (die je altijd genadeloos spot gelukkig) en bent ook aardig wat weekenden mee gegaan naar het lab voor de gezelligheid (wormen groeien nou eenmaal ook dan). Gelukkig houden we elkaar ook vaak genoeg van het werk af en ik heb al zin in de avonturen die we nog gaan beleven samen. Want of het nou gaat om kamperen op een ver eiland of op een vulkaan (die in uitbarsting is), vogels spotten in het bos of op een van de kliffen op de Faeröer (niet te dicht bij de rand graag) of een ritje in onze auto (je weet nooit wanneer dat van een verplaatsing ineens een belevenis wordt), samen is het altijd goed!



## About the author

Lisa van Sluijs was born on the 3<sup>rd</sup> of February 1992 in Vlissingen, the Netherlands. In 2010 she started her study Molecular Life Sciences at Wageningen University. During the last year of her BSc, she became enthusiastic about investigating host-parasite interactions that take place at a microscopic scale. This resulted in performing a BSc thesis on the bacterial immune system CRISPR-Cas at the Laboratory of Microbiology under the supervision of prof. dr Edze Westra and prof. dr John van der Oost (this work was published in *FEMS Microbiology Letters*). She continued her study with a thesis at the Laboratory of Nematology on host-virus interactions in the nematode *Caenorhabditis elegans* supervised by dr Mark Sterken and prof. dr Jan Kammenga. During her internship she continued working with the model nematode and worked at the University of Manchester to investigate epigenetic inheritance of lifespan under the supervision of dr Gino Poulin.



In 2016 Lisa returned to Wageningen and started her PhD at the Laboratory Nematology and Laboratory of Virology. During her PhD she studied different natural aspects that determine the viral susceptibility of *C. elegans* to its natural pathogen the Orsay virus. Lisa presented her work at nine (inter)national conferences. She received three travel grants to visit international courses and conferences and her oral presentation at the Zoology conference in 2017 was awarded with the best presentation prize. During her PhD, Lisa was keen on sharing knowledge with students. She has been involved in teaching seven courses and supervised eighteen BSc and MSc thesis students. Moreover, Lisa enjoyed teaching science for 'groep 5/6' children at elementary schools in Ede. Other highlights of her PhD involved the yearly *C. elegans* sampling trips in the surroundings of Wageningen and looking for naturally occurring nematodes at her parents' place (when she took some binoculars there for reparations). Lisa also took part in the PE&RC PhD Council for four years and helped organizing the yearly PE&RC seminar day twice.

## Publication list

### Peer-reviewed publications

- van Sluijs L, Pijlman GP, Kammenga JE. 2017. Why do individuals differ in viral susceptibility? A story told by model organisms. *Viruses*. 9(10):1–13
- van Sluijs L\*, van Houte S\*, van der Oost J, Brouns SJ, Buckling A, Westra ER. 2019. Addiction systems antagonize bacterial adaptive immunity. *FEMS Microbiol. Lett.* 366(5):fnz047
- Wang YA, van Sluijs L, Nie Y, Sterken MG, Harvey SC, Kammenga JE. 2020. Genetic Variation in Complex Traits in Transgenic  $\alpha$ -Synuclein Strains of *Caenorhabditis elegans*. *Genes*. 11(7):778

### Preprints

- van Sluijs L, Bosman K, Pankok F, Blokhina T, Riksen JAG, Snoek LB, Pijlman GP, Kammenga JE, Sterken MG. 2019. Balancing selection shapes the Intracellular Pathogen Response in natural *Caenorhabditis elegans* populations. *bioRxiv*.
- Sterken MG\*, van Sluijs L\*, Wang YA, Ritmahan W, Gultom ML, Riksen JAG, Volkers RJM, Snoek LB, Pijlman GP, Kammenga JE. 2020. Punctuated loci on chromosome IV determine natural variation in Orsay virus susceptibility of *Caenorhabditis elegans* strains Bristol N2 and Hawaiian CB4856. *bioRxiv*.

\*These authors contributed equally

## PE&RC Training and Education Statement

With the training and education activities listed below the PhD candidate has complied with the requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)



### Review of literature (4.5 ECTS)

- Why do Individuals Differ in Viral Susceptibility? A Story Told by Model Organisms

### Writing of project proposal

- Identification of novel generic viral defence pathways in *Caenorhabditis elegans* by linking genetic variation to viral susceptibility (*written during MSc course Research Master Cluster*)

### Post-graduate courses (5.5 ECTS)

- The molecular basis of diseases: Can we infer phenotypes from protein variant analysis? (*including poster and oral presentation*)
- Principles of Ecological and Evolutionary Genomics (*including poster presentation*)
- Wageningen Evolution and Ecology Workshops
- Masterclass Virology
- R big data
- Linear models
- Generalized linear models
- Overcoming the challenges of designing efficient and specific CRISPR gRNAs (*webinar*)

### Laboratory training and working visits (2.1 ECTS)

- Working visit CRISPR/Cas9; Utrecht University
- Laboratory training and working visits BioSorter; Utrecht University

### Invited review of (unpublished) journal manuscripts (2 ECTS)

- BMC Biology: *C. elegans* transcriptome under non-gravitational conditions
- PlosOne: *C. elegans* vector ecology



### Competence strengthening / skills courses (4 ECTS)

- Competence assessment; WGS (2016)
- Brain training; WGS (2018)
- An introduction to (La)TeX; PE&RC (2019)
- Reviewing a scientific manuscript; WGS (2020)
- Adobe booklet course library; WUR Library (2020)
- Career orientation; WGS (2020)
- PhD workshop carousel; WGS (2016, 2017, 2019)
- Symposium Publish for Impact; WUR Library (2017)

### Scientific integrity / ethics in science activity (0.3 ECTS)

- CRISPR/Cas9 lectures and discussion: rewriting our genes; Wageningen Young Academy (2016)
- Lecture Darwins Engelen; Rad van Wageningen (2019)

### PE&RC Annual meetings, seminars and the PE&RC weekend (3 ECTS)

- PE&RC First years weekend (2016)
- PE&RC Day (2016-2020)
- PE&RC Last years weekend (2019)

### Discussion groups / local seminars / other scientific meetings (5.7 ECTS)

- CT Networks in ecology (2016)
- Presentation prof. A. Yonath: a noble event (2016)
- Dutch Annual Virology Symposium DAVS (2016- 2019)
- Guest lectures Nematology (2016-2020)
- Wageningen Evolution and Ecology Seminars (2016-2020)
- Guest lectures virology (2016, 2018)
- Rob Goldbach lectures (2016, 2018)
- Nederlands zaden symposium (2017)
- Frontiers in resilience symposium (2017)
- Sympubium; Canterbury Christ Church University (2018) (*including oral presentation*)
- CRISPR-Cas From evolution to revolution (2018)
- NWO Life congress (2019) (*including poster presentation*)
- Dutch worm meeting (2019) (*including oral presentation*)

### Societally relevant exposure (1 ECTS)

- Primary school lecture and practical 'Ageing and disease' (2019, 2020)

### International symposia, workshops and conferences (11.1 ECTS)

- Evolutionary Biology of *Caenorhabditis* & Other Nematodes; oral presentation; Cold Spring Harbor, USA (2016)
- Zoology: genotype to phenotype symposium; oral presentation; Wageningen, the Netherlands (2017)
- European *C. elegans* meeting; poster presentation; Barcelona, Spain (2018)
- Evolutionary Biology of *Caenorhabditis* & Other Nematodes; pitch and poster presentation; Hinxton, UK (2018)
- Evolutionary Biology of *Caenorhabditis* & Other Nematodes; oral presentation; virtual (2020)
- Worm/Virus meeting; oral presentation; virtual (2020)

### Lecturing / supervision of practicals / tutorials (13.6 ECTS)

- Research methodology for plant sciences (2016, 2017)
- Introduction environmental sciences (2016, 2017)
- Ecology I (2016, 2017)
- Research master cluster (2016, 2018)
- Introduction to molecular sciences and biotechnology (2016, 2019)
- Host-parasite interactions (2018, 2020)
- Ecophysiology (2020)

### Supervision of BSc / MSc students (41 ECTS)

- Natural genetic variation determines the susceptibility of *Caenorhabditis elegans* strains to Orsay virus infection
- Optimizing the CRISPR-Cas9 system to explore the function of *cul-6* natural variants in *Caenorhabditis elegans* - Orsay virus interaction
- Comparable susceptibility to Orsay virus infection in four genetically different wild strains of *Caenorhabditis elegans*
- Quantitative trait loci (QTL) associated with viral susceptibility in *Caenorhabditis elegans* are affected by the genetic background
- Discovery of antiviral defence genes of *C. elegans* against the Orsay virus through the use of knockout mutants
- Short-term exposure to Orsay virus infection of *C. elegans* does not result in the significant gene expression alterations based on microarray analysis
- The complex inheritance mechanism underlying resistance to the Orsay virus in *C. elegans* wildtype CB4856
- Design of a high-throughput fecundity assay and comparison experiment to elucidate a resistance-fitness trade-off in the CB4856 *Caenorhabditis elegans* strain upon Orsay virus infection
- Determining the effects of Orsay virus infection on Parkinson's Disease in *Caenorhabditis elegans*
- The effect of sex on immune response against Orsay viral infection in *Caenorhabditis elegans*
- Unravelling the cause for the differences in Orsay virus susceptibility between male and hermaphrodite *Caenorhabditis elegans* nematodes
- Studying the underlying mechanism for increased  $\alpha$ -synuclein accumulation in virally infected *Caenorhabditis elegans*
- Exploring the link between viral susceptibility and sex in *Caenorhabditis elegans*
- *C. elegans pals*-gene expression in hermaphrodites and males in different genetic backgrounds
- The effect of viral infection on alpha-synuclein accumulation in *C. elegans*
- Inter-sex transcriptional profiling of Orsay virus infected *C. elegans*
- The absence of QTLs in *C. elegans* that directly influence the gene expression of  $\alpha$ -synuclein in a Parkinson's disease model
- Sex-specific gene expression influencing viral susceptibility in *C. elegans*

The research described in this thesis was conducted at the laboratory of Nematology and the laboratory of Virology at Wageningen University (Wageningen, the Netherlands). The research in this thesis was financially supported by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) and by the NWO (Nederlandse Organisatie voor Wetenschappelijk Onderzoek) (824.15.006).