



Mechanisms of Larval Growth Control in  
*Caenorhabditis elegans*

Joy Nyaanga

PhD Dissertation  
Northwestern University

This dissertation is submitted in partial fulfillment of the requirements  
for the degree Doctor of Philosophy: Field of Biological Sciences

Repository: <https://github.com/jnyaanga>

May 2022

# Contents

<b>Abstract</b>	<b>iv</b>
<b>Acknowledgements</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Phenotypic variation . . . . .	2
1.2 Quantitative trait loci mapping . . . . .	4
1.3 <i>Caenorhabditis elegans</i> . . . . .	5
<b>2 A role for cuticle stretch in developmental growth control</b>	<b>13</b>
2.1 Preface . . . . .	13
2.2 Abstract . . . . .	14
2.3 Introduction . . . . .	14
2.4 Materials and methods . . . . .	17
2.5 Results . . . . .	23
2.6 Discussion . . . . .	38
2.7 Contributions . . . . .	41
2.8 Supplement . . . . .	42
<b>3 Linkage mapping reveals loci that underlie differences in <i>C. elegans</i> body growth</b>	<b>52</b>
3.1 Preface . . . . .	52
3.2 Abstract . . . . .	53
3.3 Introduction . . . . .	53

<i>CONTENTS</i>	iii
3.4 Materials and methods . . . . .	53
3.5 Results . . . . .	53
3.6 Discussion . . . . .	53
3.7 Future directions . . . . .	53
3.8 Contributions . . . . .	53
<b>4 easyXpress: An R package to analyze and visualize <i>C. elegans</i> microscopy data generated using CellProfiler</b>	<b>54</b>
4.1 Preface . . . . .	54
4.2 Abstract . . . . .	55
4.3 Introduction . . . . .	55
4.4 Materials and methods . . . . .	57
4.5 Results . . . . .	59
4.6 Conclusions . . . . .	67
4.7 Contributions . . . . .	67
4.8 Supplement . . . . .	68
<b>5 Discussion</b>	<b>71</b>
<b>6 References</b>	<b>72</b>
<b>Appendix of R code</b>	<b>88</b>

# Abstract

Body size is one of the most discernible ways in which animal species vary. The blue whale (*Balaenoptera musculus*), the largest animal on earth, can reach up to 30 m in length and weigh up to 200 tonnes. At the other extreme, a species of frog called *Paedophryne amauensis* is among the smallest animals on earth with adults measuring just 8 mm in length, about the size of a pea. Understanding how organisms grow to their characteristic sizes is a fundamental biological question. Although a larger body size can increase an organism's competitive advantage, an increased body size also requires added time and nutrients to develop. As such, organisms must have mechanisms to both sense and adjust growth during development. Studies of single cells have revealed that growth regulation can be achieved using time or size sensing control methods. In multicellular organisms, however, regulatory mechanisms must not only control single cell growth but also integrate it across organs and tissues during development. In this dissertation, I investigate growth control in the roundworm nematode *Caenorhabditis elegans*. I first optimize a high-throughput phenotyping platform that facilitates quantitative assessment of thousands of individuals at high precision. Using this platform, I quantify changes in animal size and shape throughout development. I further explore the natural genetic variation underlying body size and shape. The results of this work lay the foundation for a mechanistic dissection of organismal growth control.

# Acknowledgements



# Chapter 1

## Introduction

A fundamental objective in biology is to understand what makes individuals, populations, and species different from each other. The observable characteristics of an organism, such as morphology, physiology, development, and behavior, make up its phenotype. Contrarily, genotype describes the inheritable genetic material that is responsible for determining a given trait [1]. The relationship between genotype and phenotype is often complex and multifaceted. Some traits, like color blindness can be linked to single gene mutations [2]. Most traits, however, are influenced by a combination of genetic and environmental factors. Investigating the variation in complex traits, such as animal growth and body size, provides deeper insights to phenotypic diversity and its causes.

This introduction starts by providing an overview on phenotypic heterogeneity, focusing on the genetic variation that can influence its emergence Section 1.1. I discuss how quantitative genetic mapping is used to study underlying variation in complex traits Section 1.2. Next, *Caenorhabditis elegans* is presented as an appropriate model system for the study of phenotypic variation Section 1.3. I discuss post-embryonic development and examine existing knowledge of how the processes of growth and body size are controlled in the nematode. Finally, I review principles of genetic mapping in *C. elegans*.

## 1.1 Phenotypic variation

The field of quantitative genetics classically postulates that variation among phenotypes arises from two major sources: genetics and environment [3,4]. Expressed as an equation, this can be written as,

$$V_P = V_G + V_E$$

where  $V_P$  represents variance in a phenotype,  $V_G$  is genetic variation, and  $V_E$  is variation in the environment. The genetic component of variation arises from genetic variants within a population [5]. Conversely, environmental variation arises when individuals of the same genotype experience changes in environmental conditions (i.e. food availability, temperature) and respond with multiple phenotypic differences [6]. Notably, genetic and environmental components of phenotypic variation are rarely independent; many traits exhibit a genetic component to plasticity. In this case, the observable change in a population's phenotype due to environmental differences is influenced by genetic variation. This form of variation indicates the presence of genotype-by-environment (GxE) interaction. Given this interaction term, total phenotypic variation is more appropriately expressed as,

$$V_P = V_G + V_E + V_{GxE}$$

where the contribution of genetic variance to plasticity is captured in the term,  $V_{GxE}$  [7].

Another important source of phenotype variation is stochastic variation. For a given environmental condition, random noise exists in the expression of the true phenotype. This uncontrolled variation manifests as differences among individuals of the same genotype raised in a common environment [8]. Accordingly, the total phenotypic variation of a population should now be represented as,

$$V_P = V_G + V_E + V_{GxE} + V_{Noise}$$

where  $V_{Noise}$  represents intra-individual variation in response to an intrinsic factor. Often, this noise is driven by stochastic fluctuations in the developmental process and called developmental noise [9]. The interaction of these various sources of variation ( $V_G$ ,  $V_E$ ,  $V_{GxE}$ ,  $V_{Noise}$ ), combined with the sheer magnitude of the number of phenotypic traits, makes the study of the phenotypic variation of a population strikingly complex.

### 1.1.1 From genotype to phenotype

In 1865, an Austrian monk by the name of Gregor Mendel published work investigating the connection between genotype and phenotype. For eight years he studied the pea plants that grew in the garden of his monastery, observing patterns in their color, texture, and height from one generation to the next [10]. Mendel's experiments laid the foundation for our understanding of genetic inheritance and the correlation between underlying genetic makeup and physically visible traits. More than 150 years later, scientists continue to study how genetic variation contributes to phenotypic differences.

Recent advances in sequencing technologies have accelerated the generation of vast amounts of genotype data [11–13] and have also enabled scientists to identify genetic perturbations causally linked to disease traits. However, discerning the relationship between genotype and phenotype remains a challenge. Although there is a strong correlation between the two, phenotype is not solely determined by the genotype of a few trait-associated genes but rather is the cumulative result of interactions among genetic and non-genetic factors over time [14]. This is particularly apparent in the context of disease. The clinical symptoms we observe in affected individuals are the result of interactions between their genetic background, mutations, and the environment over time. Disease phenotypes may follow Mendelian inheritance patterns and thus be predominantly driven by a single gene mutation, but oftentimes they involve multiple genes along with additional environmental factors [15]. As the complexity of

a trait increases, it becomes increasingly more difficult to resolve the genotype to phenotype relationship. However, the identification of genetic variants underlying phenotypes is key to understanding how genotype influences complex traits associated with growth, health, and disease.

## 1.2 Quantitative trait loci mapping

Quantitative trait loci (QTL) mapping is a method for identifying the genetic basis of complex traits. With this technique, regions of the genome that contain loci that affect a measurable trait can be identified. Mapping QTL allows scientists to study the underlying genetic architecture of complex traits. Many studies have mapped QTL that affect human disease and have uncovered new loci providing insights into the biology of disease [16–18].

There are two main approaches to identify QTL: linkage mapping and genome-wide association (GWA) mapping [19]. Both techniques rely on statistical methods to correlate genetic variation with phenotypic variation. Simply, these methods determine whether the genotype at a given position in the genome explains phenotypic differences among individuals. Linkage mapping uses populations of recombinants generated from a cross between phenotypically and genotypically diverse strains. In contrast, GWA mapping leverages the existing natural genetic diversity among genetically distinct individuals [19]. Though both methods can detect QTL, the power to do so differs depending on population size, allele frequency, and effect size [20].

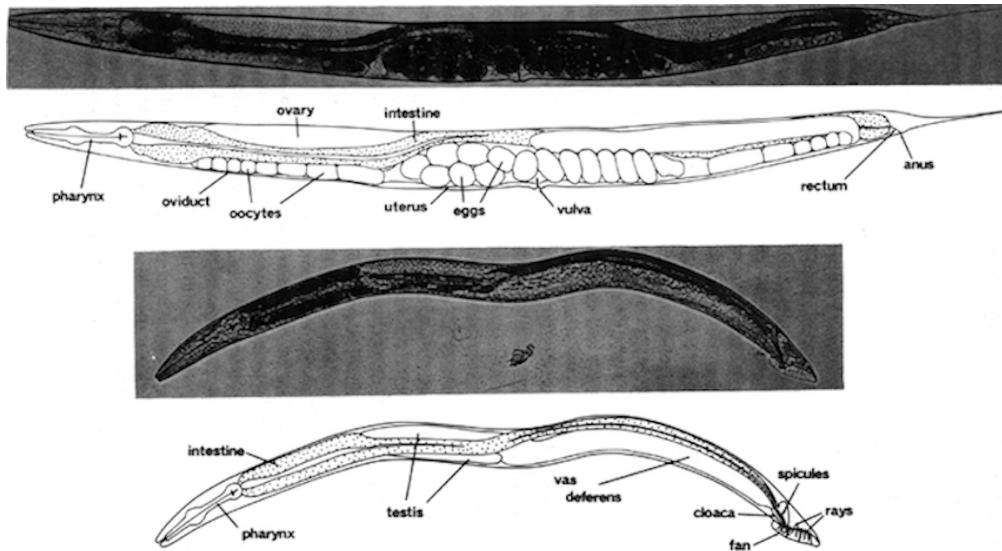
Oftentimes, QTL span large genomic regions containing many genes. To narrow the search space for potential genes contributing to phenotypic differences, the QTL region can be isolated in a clean genetic background with near-isogenic lines (NILs) [21,22]. The phenotypes of these NILs can then be used to refine the QTL interval by determining whether the introgressed genomic region recapitulates the QTL effect. Ultimately, in organisms with well-annotated genomes, causality can be tested with genome editing tools [23–25].

## 1.3 *Caenorhabditis elegans*

French biologist, Emile Maupas, first described isolating *Rhabditis elegans* (now *Caenorhabditis elegans*) near the city of Algiers (Algeria) in 1900 [26]. He conducted some of the first experiments on the nematode, elaborately detailing its anatomy and development. Several decades later, *C. elegans* was again isolated for study by Victor Nigon and Ellsworth Dougherty. Together they worked to refine culture conditions and experimental methods for propagating animals in the laboratory [27]. However, it wasn't until Sydney Brenner published work on the nematode in the 1970s that *C. elegans* became an established model organism [28]. Since then, research on the roundworm has expanded to explore diverse areas of modern biology and is actively studied in over a thousand laboratories worldwide [29].

### 1.3.1 As a model organism

In 1963, well-known bacterial geneticist, Sydney Brenner, decided to turn his research efforts from bacteria to an animal. He sought an organism that could serve as a genetic model to probe how mutations disrupt eukaryotic processes. After considering a number of organisms, he ultimately settled on *C. elegans* for several reasons [30]. This free-living bacterivore grows rapidly on agar plates or in liquid culture, reaching an adult length of approximately 1.5 mm in three days [31]. Its small size and transparent body make the nematode an ideal organism for microscopic observation (**Figure 1-1**). As hermaphrodites, *C. elegans* can produce up to 300 genetically identical progeny in a lifetime. These individuals can be maintained nearly indefinitely through cryopreservation, providing a method for long-term storage and safeguarding against the accumulation of mutations [31,32].



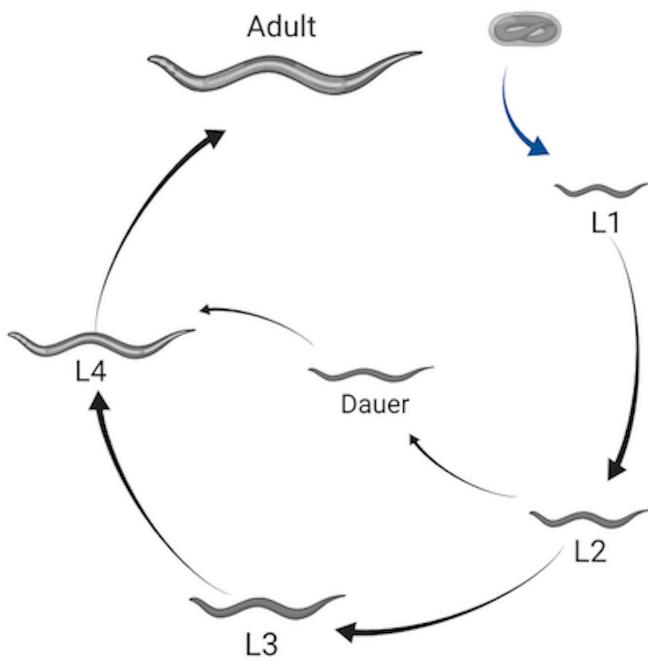
**Figure 1-1. Photomicrographs of *C. elegans*.** Major anatomical features of the *C. elegans* adult hermaphrodite (top) and male (bottom) under bright-field illumination. Bar, 20  $\mu\text{m}$ . (From [31]).

Since Brenner's first paper establishing *C. elegans* as a model organism, the animal has become a powerful system for genetic and molecular analysis. Notably, by the 1980s, scientists had documented the entire cell lineage of the nematode, from embryo to adult [33–36]. A decade later, *C. elegans* became the first multicellular organism to have its genome sequenced [37]. This 100 Mb genome, though relatively small, comprises six chromosomes and approximately 20,000 protein-coding genes that are homologous to protein-coding genes found in other organisms, including humans [38,39].

### 1.3.2 Post-embryonic development

*C. elegans* post-embryonic development consists of four larval stages (L1 - L4) that are punctuated by distinct developmental events called molts where animals shed an exoskeleton called the cuticle (**Figure 1-2**). Each larval stage begins with a period of active growth, and is followed by a period of quiescence (lethargus) where feeding and movement is temporarily

halted [40]. During this time, a new cuticle is synthesized under the old. A molt is completed when animals shed the old cuticle (ecdysis) and transition to the next stage [41].



**Figure 1-2. Illustration of the *C. elegans* life cycle.** Embryo's hatch to the first larval stage and proceed through four larval stages (L1 - L4) before becoming a mature adult. Transitions between larval stages are marked by molt events where the animal's outer collagen-rich cuticle is shed. An alternate developmental stage, dauer, occurs in response to unfavorable environmental conditions.

The rate of larval development in *C. elegans* is strongly dependent on environmental conditions. Dietary restriction has been shown to extend developmental progression up to ten-fold [42]. In extreme conditions, such as absence of food or high population density, *C. elegans* are able to halt development entirely by entering dauer, a long-lived morphologically distinct stage optimized for survival [43,44]. The decision to enter dauer is one of many checkpoints regulating developmental progression. Such checkpoints often occur early in the molting process and have been identified in every larval stage [45,46]. The well-coordinated timing of development is essential to the completion of stage-specific processes including cell-

differentiation and tissue morphogenesis [47,48], and underscores the importance of growth regulation during *C. elegans* larval development.

### 1.3.3 Growth regulation

Organismal growth is regulated on a genetic level, as changes in gene expression patterns and signaling dictate much of development. However, environmental conditions (*e.g.* nutrient availability and temperature) also have strong impacts on growth. Therefore, it is important to account for how the processes of growth and development are controlled genetically, and how they can be influenced environmentally.

#### 1.3.3.1 Genetic pathways regulate body size

The control of *C. elegans* developmental progression has been studied extensively on the genetic level. The very first body size mutants were isolated in genetic screens by Sydney Brenner [32]. Subsequent screens for small (Sma) and long (Lon) mutants revealed a major regulator of growth and body size: DBL-1 [49–52]. DBL-1 is a ligand in the transforming growth factor- $\beta$  (TGF- $\beta$ ) superfamily, and homologous to members of the mammalian bone morphogenetic protein (BMP) family [53]. Secreted by neurons and body wall muscle, DBL-1 is necessary for body size regulation, and other developmental and homeostatic processes [49]. Mutants deficient in dbl-1 have a marked reduction (40%) in both length and width compared to wild type throughout post-embryonic development [49,50]. In contrast, increased DBL-1 signaling causes animals to be longer (25%) than wild type [49,54].

#### 1.3.3.2 Environmental factors modulate body size

As previously mentioned, nutrient restriction is known to decrease *C. elegans* growth rate or, when extreme, induce complete developmental arrest [42,44,55]. Body size is also affected by the quality of food. As bacterivores, *C. elegans* obtain their nutrients primarily from bacteria.

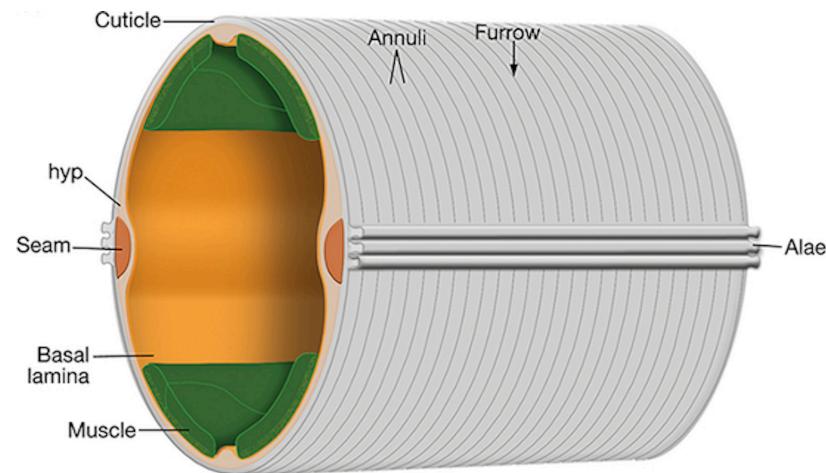
Historically, a strain of *Escherichia coli*, OP50, was chosen as the standard laboratory diet [32]. While most experimental studies use OP50 as the primary food source, the number of bacterial diets used to propagate *C. elegans* has expanded in recent years. Although these diets support development, research has identified several phenotypic effects that are altered. For example, when fed *Comamonas* DA1877 or *E. coli* HB101, animals develop faster and grow larger than those fed *E. coli* OP50 [56–58].

Aside from diet, other environmental stimuli can also regulate *C. elegans* body size. Standard wild-type animals used in most experiments grow larger when cultured at lower temperatures [59]. Additionally, research has shown that the *C. elegans* nervous system plays a key role in the regulation of body size by environmental stimuli. When raised in isolation, *C. elegans* exhibit decreased growth rates and reach smaller adult sizes compared to animals raised in groups [60]. This result, however, is reversed with mechanosensory stimulation during development, demonstrating a role for mechanosensory neurons in body size regulation. Similarly, a number of mutations affecting ciliated sensory neurons through which *C. elegans* perceive environmental stimuli are shown to be required for animals to grow to the proper size [61,62]. This suggests that perception of environmental stimuli is crucial for the regulation of body size in *C. elegans*.

#### 1.3.3.3 Cuticle structure influences body size

The *C. elegans* cuticle is a complex, multi-layered structure primarily composed of collagens. The major surface structures of the cuticle include the circumferential ridges (annuli) and troughs (furrows) [63] (**Figure 1-3**). As animals progress through their life-cycle, the structure and thickness of the cuticle changes but its role in the maintenance of body morphology and integrity remains. To date, 21 cuticle collagen mutants have been identified that cause a range of body morphology defects [63]. Several of these mutants (*dpy-2, 3, 7, 8, 10*) lack annuli and exhibit a disproportionate reduction in body size [64]. These *dpy* animals are noticeably shorter in length and wider than wild type, clearly demonstrating

the importance of the physical structure of the cuticle on growth.



**Figure 1-3. Illustration of the *C. elegans* cuticle.** The hypodermis (hyp) comprises hypodermal cells which secrete the cuticle (grey). Circumferential rings called annuli run the length of the cuticle. The valleys between annuli are furrows, which form immediately above circumferential bundles of actin (not shown) present in hypodermal cells during embryogenesis and molting. Seam cells (orange) that run along the left and right sides. (From [65]).

### 1.3.4 Natural genetic variation

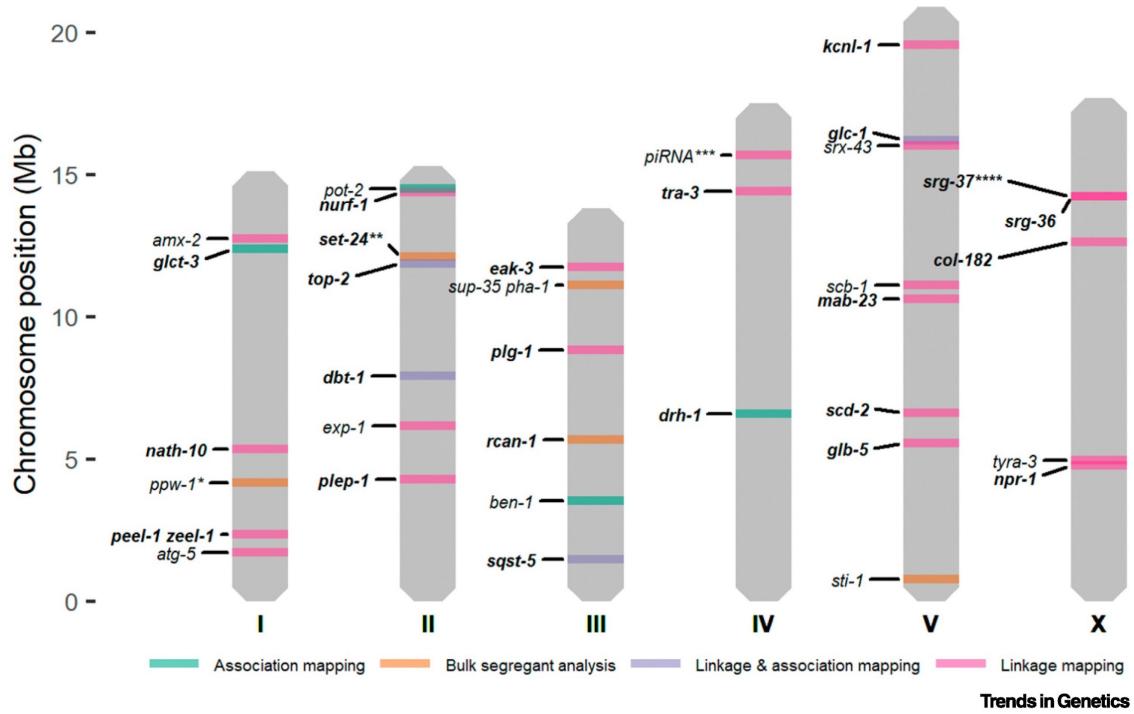
*C. elegans* present a scalable and tractable animal model to connect phenotypic differences to genetic variants. *C. elegans* are easily propagated as clonal cultures, eliminating genetic variance as a contributor to phenotypic variance in experimental studies. Additionally, hundreds of wild strains have been collected worldwide, providing access to extensive genetic diversity [66,67]. Two heavily studied strains are the laboratory-adapted strain, N2, and a wild isolate, CB4856. N2, derived from an isolate found in Bristol, UK in 1951, is recognized as the canonical laboratory wild-type strain with its genome serving as the reference [68]. Across the globe on the Hawaiian island of Maui, CB4856 was isolated in 1972 [69]. Compared against the N2 reference, the CB4856 genome is highly divergent [70,71]. This

extensive genetic diversity likely underlies the many phenotypic differences between the two strains, including social behavior, life-history traits (body size, fecundity, lifespan), and gene expression [7,72–76].

### 1.3.5 Linkage mapping

N2 and CB4856 can be studied to better understand how genetic variation contributes to phenotypic differences. Crossing N2 and CB4856 for many generations produces individuals with unique variants derived from each parental background. Performed at a large scale, these panels of recombinant inbred lines are a powerful tool for identifying genomic regions that are correlated with phenotypic variation.

Linkage mapping is a popular genetic mapping approach used to identify functional variants that contribute to phenotypic diversity. In *C. elegans*, linkage mapping leverages recombinant line panels, correlating genotype and phenotype to identify QTL. The Andersen Lab has developed an R package to facilitate such analyses [77]. By investigating the natural genetic variation underlying phenotypic differences, researchers are able to uncover genetic pathways involved in biological processes. Several studies have taken this approach to study the genetic underpinnings of complex traits like body size, social behavior, and drug response (**Figure 1-4**) [76,78–83].



**Figure 1-4. Overview of causal genes identified through QTL mapping approaches.** The colors represent the mapping technique(s) that were used for QTL mapping: bulk-segregant analysis (BSA) (orange); linkage mapping (pink); genome-wide association (GWA) mapping (green); linkage and GWA mapping (purple). (From [20])

# **Chapter 2**

## **A role for cuticle stretch in developmental growth control**

### **2.1 Preface**

In the spring quarter of 2019, I chose to complete my final first-year rotation in the Andersen Lab. At the time, I was coming off a rotation with Dr. Niall Mangan who was involved in a relatively new collaboration with Dr. Erik Andersen as part of the NSF-Simons Center for Quantitative Biology. The general goal of the project was to study the interplay between organismal growth and environmental variables such as diet and temperature. Needing a little extra experimental manpower, Erik recruited me to the team. Little did I know this project would become the heart of my thesis. What ensued was a series of daunting, exhausting, and incredibly rewarding 50+-hour-long time course experiments. The following chapter highlights the results of just one of these experiments, submitted as a first-author manuscript to Cells and Development in January 2022 [84].

## 2.2 Abstract

Growth control establishes organism size, requiring mechanisms to sense and adjust growth during development. Studies of single cells revealed that size homeostasis uses distinct control methods. In multicellular organisms, mechanisms that regulate single cell growth must integrate control across organs and tissues during development to generate adult size and shape. We leveraged the roundworm *Caenorhabditis elegans* as a scalable and tractable model to collect precise growth measurements of thousands of individuals, measure feeding behavior, and quantify changes in animal size and shape during a densely sampled developmental time course. As animals transitioned from one developmental stage to the next, we observed changes in body aspect ratio while body volume remained constant. Then, we modeled a physical mechanism by which constraints on cuticle stretch could cause changes in *C. elegans* body shape. The model-predicted shape changes are consistent with those observed in the data. Theoretically, cuticle stretch could be sensed by the animal to initiate larval-stage transitions, providing a means for physical constraints to influence developmental timing and growth rate in *C. elegans*.

## 2.3 Introduction

Growth is a complex process fundamental to development. Individual cells and whole animals must reach an appropriate size to remain competitive in their environment. A larger body size conveys many selective advantages to an organism, including increased success in predation, defense against predation, mating, or intraspecific as well as interspecific competition. Offsetting these advantages, larger organisms require more food resources to grow, take longer to develop, and produce fewer offspring [85]. Therefore, it is critical for multicellular organisms to effectively coordinate the growth of both individual cells and the whole body. Additionally, growth at both of these scales must be coupled with developmental progression to ensure the proper timing of irreversible developmental events.

In recent years, efforts have focused on understanding how organisms control growth to achieve size homeostasis [86–88]. Many of these studies are motivated by the decades-long debate about whether growth is linear or exponential; two separate models each having unique implications for size regulation. In a linear model with constant growth rate, smaller organisms must grow proportionally more than larger organisms to maintain size homeostasis. In this paradigm, organism size can be controlled simply by specifying growth duration. Subsequently, this method of growth control was named the “Timer” model [89,90]. Instead of regulating growth duration, organisms can monitor size and adjust duration of growth to reach an optimal size, often named the “Sizer” model [91,92].

In an exponential model, growth rate is proportional to size. Here, a time-based control mechanism alone would fail to maintain size homeostasis because larger organisms would grow proportionally more during a specified period of time. This difference in growth requires a size-based control mechanism to ensure that growth is halted once a maximum size is reached. Although “Timer” and “Sizer” are the most often proposed size-control models, other models have been suggested. The “Adder” model proposes that a fixed volume is added to a cell or organism during growth [93,94], whereas the “Folder” model specifies that an organism increases in volume by a fixed proportion in order to control growth [95]. It is not trivial to determine which model most accurately describes growth of individual cells or whole organisms because quantitative measurements of growth must be collected at high precision and throughput under tightly controlled experimental conditions. In unicellular organisms, the development of high-throughput experimental techniques in combination with theoretical models have advanced the understanding of size control [96–100]. Progress has been slower for multicellular organisms because cell growth within tissues and tissue growth within organisms often progress at different rates, suggesting that they are likely not regulated in the same ways [42,101,102].

The nematode *Caenorhabditis elegans* presents both a scalable and tractable multicellular animal model to study growth control. With an adult body length of approximately 1 mm, hundreds of thousands of individuals are easily cultured in controlled laboratory conditions

[30]. Moreover, *C. elegans* post-embryonic development is marked by several molts that provide clear developmental milestones [63]. Each molt is initiated by a period of quiescence (lethargus) and terminated once the animal successfully sheds its collagen-rich outer cuticle (ecdysis) [40]. Four molts separate the *C. elegans* life cycle into five distinct stages: four larval stages (L1-L4) and adult. The timing of these molts determines the completion of stage-specific development [47,48] and underscores the importance of growth regulation during *C. elegans* larval development.

A full description of an organism’s development includes the assessment of how growth and body size are regulated. Initial studies of *C. elegans* development described whole-organism growth as a sigmoidal curve characterized by continuous larval growth in length that reaches saturation in adulthood [103]. These early studies hypothesized that molt events had little effect on continuous growth as the *C. elegans* cuticle allowed for stretch during larval stages. Later work determined that larval progression was not continuous but rather piecewise in nature [104]. This study showed that *C. elegans* volumetric growth rate increased from stage to stage such that L1 animals had the slowest rate of growth and L4 animals had the fastest. This finding suggests that *C. elegans* have a mechanism for regulating growth rate, potentially at each molt. Next, researchers using single-animal imaging strategies observed that animals did not advance to the next developmental stage until a critical volume was reached [42]. This finding suggests that *C. elegans* growth follows a “Sizer” model with each molt decision controlled by a volume threshold and further implies that individual cells are able to communicate information about body size to precisely regulate growth. Most recently, live imaging and characterization of body volume heterogeneity revealed that with respect to the start of a larval stage, *C. elegans* volume fold change within a stage is nearly invariant thereby preventing rapid divergence in volume between fast- and slow-growing animals [95]. A mechanism that maintains a constant volume fold change within each larval stage relies on the coupling between growth rate and developmental timing. Notably, such coupling is consistent with recent observations of temporal scaling in *C. elegans* development where despite inter-individual variability in the absolute duration of a larval stage, relative

timing of a stage is similar [105,106].

To understand *C. elegans* growth control at the whole-organism level, we used a combination of quantitative measurements and mathematical modeling. We performed a high-resolution longitudinal study of *C. elegans* larval progression and captured high-precision details about animal length, width, volume, and feeding dynamics. By investigating *C. elegans* feeding and growth in tandem for thousands of individual animals, we found decreases in feeding behavior associated with each larval transition that were also correlated in time with changes in growth rate. We used our large-scale measurements of body shape to further analyze the periods of time surrounding each larval transition. At each molt, we observed simultaneous increases in length, decreases in width, and maintenance of volume. Based on the physical characteristics of the cuticle, we propose a “Stretcher” mechanism whereby physical constraints on cuticle stretch influence body shape. We find the Stretcher model-predicted shape changes are consistent with observed data. Animals may be able to detect when the cuticle reaches its maximum capacity for stretch providing a signal to initiate larval-stage transitions.

## 2.4 Materials and methods

### 2.4.1 Worm culture

The canonical laboratory strain N2 was obtained from the *C. elegans* Natural Diversity Resource [107]. Animals were cultured at 20°C on 6 cm plates of modified nematode growth media (NGMA), which contained 1% agar and 0.7% agarose seeded with *E. coli* OP50 bacteria [76].

### 2.4.2 Bacterial food

*E. coli* HB101 bacteria were prepared from cultures grown for 15 hours in Superbroth and then pelleted by centrifugation. HB101 bacteria were diluted to OD100 in K medium (51

mM NaCl, 32 mM KCl, 3 mM  $CaCl_2$ , and 3 mM MgSO<sub>4</sub> in distilled water) and stored at  $-80^{\circ}C$ . Bacteria were thawed and fed to animals at a concentration sufficient to sustain population growth from hatching to adulthood (OD20).

### 2.4.3 Growth of the animals

Populations of animals were propagated on NGMA plates for two generations without starvation. In the third generation, gravid adults were bleach-synchronized [108]. Embryos were resuspended in K medium, aliquoted into a 500 mL flask at a concentration of one embryo per  $\mu L$ , and allowed to hatch overnight. The following day, arrested L1s were fed HB101 bacteria at a final concentration of OD20 in a final flask volume of 100 mL K medium and HB101 food. Animals were grown for three days at  $20^{\circ}C$  with constant shaking. Following these three days, adult animals were bleach-synchronized once more and embryos were aliquoted to seven replicate 500 mL flasks at a concentration of one embryo per  $\mu L$  in 100 mL. The following morning, six flasks were fed HB101 bacterial food at a final concentration of OD20 in a final flask volume of 100 mL K medium and HB101 food. Two additional flasks were included to control for L1 animal size and possible clumping of bacterial food: one flask contained L1 larvae but did not have food added and one flask contained no larvae but the same concentration of HB101 bacteria as the six flasks containing L1 larvae. All replicate flasks were kept in an incubator at  $20^{\circ}C$  with shaking for the duration of the experiment. A small temperature gradient of  $1.25^{\circ}C$  was recorded in the shaking incubator with the highest temperature readings on the right side and lowest temperature readings on the left side. This slight variation in temperature contributed to variation in developmental rate among replicates based on position within the incubator (replicates were placed in numerical order with replicate 1 positioned on the far right side of the incubator).

#### 2.4.4 High-throughput measurements of body size and fluorescence

Flasks were sampled each hour beginning one hour after feeding and continuing for 72 consecutive hours. At each hour, 500  $\mu\text{L}$  was removed from each flask and transferred to a well of a deep 96-well plate. Each flask was sampled at each time point. Fluorescent polychromatic beads (Polysciences, 19507-5) with a 0.5  $\mu\text{m}$  particle size were added to each well at a final concentration of  $3.64 \times 10^8$  beads/mL and incubated at 20°C for 10 minutes with shaking. Following the bead incubation, 30  $\mu\text{L}$  from each well of the deep 96-well plate was aliquoted to a 96-well microtiter plate. The process was repeated 11 times to 11 separate wells of the same microtiter plate with pipetting to mix the well contents from the deep 96-well plate. Animals were then treated with sodium azide at a final concentration of 50 mM to paralyze and prevent defecation of the ingested beads. The 96-well plate was imaged with an ImageXpress Nano (Molecular Devices, SanJose, CA) using both 2x (Nikon MRD00025) and 10x (Nikon MRH00101) objectives. The ImageXpress Nano acquires brightfield images using a 4.7 megaPixel CMOS camera. Images are stored in 16-bit TIFF format. Finally, animals were scored using a large-particle flow cytometer (COPAS BIOSORT, Union Biometrica, Holliston MA). The COPAS BIOSORT sheath flow rate was kept at a constant  $10.3 \pm 0.1$  mL per minute to reduce variability in length measurements.

#### 2.4.5 Image processing

Manual measurements of animal size were performed using the free Java image-processing program ImageJ [109]. Well images for the six replicate flasks, excluding controls, were loaded into ImageJ software. Length was measured from head to tail, and width was measured at the widest point of the animal. Five animals were measured per well across thirty total wells for each hour. Measurements were repeated for all 72 time points in the assay. Body length and width were used to estimate cross-sectional area ( $length * width$ ). This metric was used to describe animal area for the extent of the text. Volume was calculated from body length

and width by approximating the animal as a cylinder. Pixels were converted to  $\mu\text{m}$  using a conversion factor of  $3.2937 \text{ pixels}/\mu\text{m}$ .

### 2.4.6 Data processing

The COPAS BIOSORT was used to collect measurements of animal length (TOF), optical extinction (EXT), and fluorescence for every animal in each well. These traits measure properties of nematode development and, as such, increase as animals progress to adulthood [110]. Optical extinction measurements correspond to the amount of light absorbed over the full length of an animal as it passes through the instrument. An additional measurement (norm.EXT) can be calculated by normalizing optical extinction by length. The raw data collected were imported and processed using the *easysorter* R package [77].

The COPAS BIOSORT data were analyzed further using Gaussian finite mixture modeling as implemented in the *mclust* R package [111]. These probabilistic models assume that data are generated from a mixture of multivariate normal distributions and, therefore, can be used to classify unstructured data into meaningful groups. Specifically, the *mclust* package fits a mixture model to data and selects the optimal number of clusters using an expectation-maximization algorithm and Bayes Information Criteria. For model-based clustering, log transformed animal length (logTOF) and log transformed optical extinction (logEXT) were used as inputs for the *Mclust* function. Data from each hour of the experiment was analyzed by replicate and clusters that did not appear to include majority animal objects were identified and removed as described previously [112]. This processing removed non-animal objects such as bacterial clumps, shed cuticles, and next generation larval animals from the time-course data.

We used a numpy polyfit regression of well-median data from the COPAS BIOSORT and image measurements to convert TOF and norm.EXT data to microns. Only the unit-corrected BIOSORT data were used for further analysis.

### 2.4.7 Molt analysis

Fluorescence data obtained from the COPAS BIOSORT was used as a proxy for feeding behavior to distinguish animals in a molt from growing animals. First, fluorescence was normalized by EXT to account for the ability of larger animals to consume more food and beads. Next, an analysis of variance statistical model was fit to the fluorescence data normalized by EXT to determine the amount of variance contributed by replicate and well (Table S1). A local kernel regression smoothing method was then applied to the residuals of the variance analysis using the *lokern* R package [113]. Residuals were used to address only the differences over time and ignore minor variation among replicates and wells. The local minima of the regression function were found by solving for where the first derivative of this function equaled zero. The time associated with each local minimum was used to represent the timing of each molt. Molts occurred at 14, 25, 36, and 48 hours.

To identify periods of time that contained a majority of growing animals, the inflection points of the regression function were calculated by solving for where the second derivative of the function equaled zero. Time points between inflection points that did not contain a local fluorescence minimum were considered as growth periods. These hours were 1-13, 17-22, 27-32, and 39-45 corresponding to L1, L2, L3, and L4 growth periods.

Each molt is initiated when animals enter lethargus: a behavioral state where animals cease active feeding. To classify individual animals as in a molt or growing, we set a quiescence threshold using fluorescence measurements at each local minimum. The fluorescence measurement at each local minimum was as follows: 0.07, 0.06, 0.06, 0.06. The average of these measurements (0.06) was used as the fluorescence threshold signifying quiescent behavior. Any individual animals that fell below this threshold fluorescence value were designated as in a molt and animals above this threshold value were classified as growing.

### 2.4.8 Comparison of model fits

To determine the volume growth model, we fit linear, exponential, and cubic functions to data designated as growth periods for each larval stage. Both linear and nonlinear functions were fitted using least-squares regression. Akaike's information criterion (AIC) [114] and Bayesian information criterion (BIC) [115] were goodness of fit criteria used to evaluate candidate models. To assess the strength of evidence for each candidate model, we identified the model with the smallest AIC/BIC value and assessed the difference between this value and the AIC/BIC of the other two models. The magnitude of the difference was used to determine the level of support for each candidate model as previously described [116,117]. All model fits and analysis were performed using the *stats* R package.

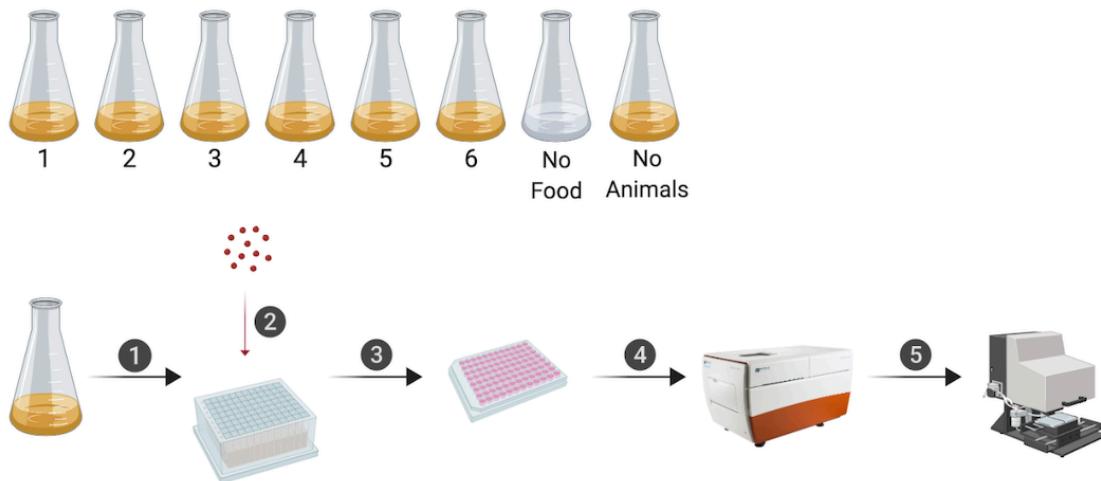
### 2.4.9 Stretcher model analysis

To analyze shape dynamics, length and width data from growth time periods were extracted from the full COPAS BIOSORT population data and analyzed from each replicate separately to avoid issues with replicate variability. For replicate 2, the hours defining growth periods were 1-13, 16.37-22.39, and 26.93-32.96; corresponding to L1, L2, and L3. Hours defining larval stages were rounded as data was collected at exact hour increments. The L4 stage was excluded from the analysis because of the high variability within the population. We applied a local kernel regression, *lokern* R package [113], to smooth the population dynamics of length and width. To calculate mean and standard deviation, the smoothed population measurements were bootstrapped using 2,000 samples with replacement. To determine cuticle properties throughout larval stages, we calculated the mean ratio of derivatives of regression width and length.

## 2.5 Results

### 2.5.1 Quantitative measurements of *C. elegans* growth

We have optimized a quantitative growth assay that reliably measures small changes in *C. elegans* body size throughout development (Fig. 2-1). Our method provides both high-throughput and high-precision assessment of developmental growth. In brief, populations of 100,000 animals were cultured in flasks. We cultured six replicate populations of *C. elegans* for a total of 600,000 synchronized and growing animals. Every hour after feeding, a sample of the population from each flask (~300 animals/flask) was collected to measure animal length, width, and feeding rate. Feeding rate, examined using fluorescent microspheres, and body size were measured using the COPAS BIOSORT (Union Biometrica). Then, the ImageXpress system (Molecular Devices) was used to collect images of sampled animals. This platform enabled further analysis of life stage and body size, contributing added precision to our measurements.

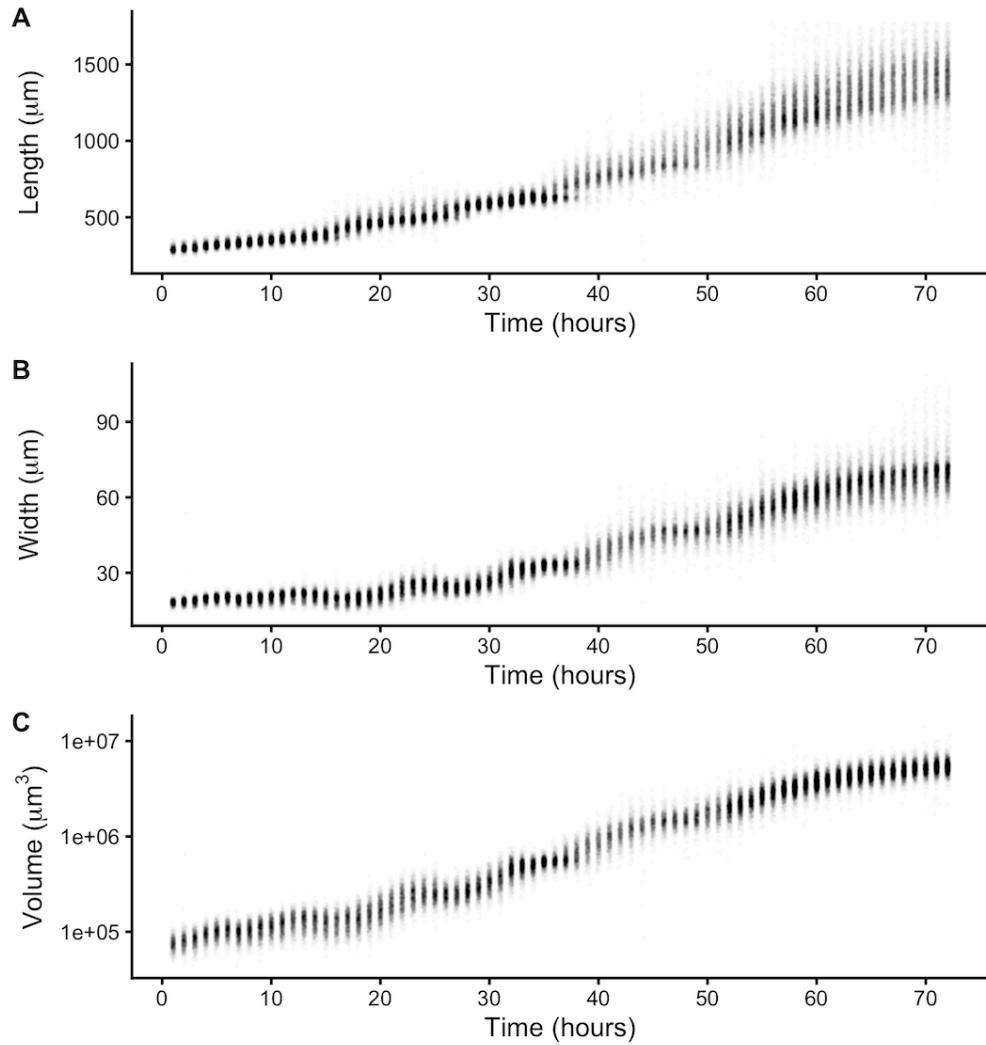


**Fig. 2-1. An overview of the quantitative growth assay.** Synchronized animals were cultured in flasks where six flasks contained replicate populations of nematodes, one flask had a population of unfed animals, and one flask only contained bacterial food. At each hour of the experiment, all eight flasks were sampled. In step 1, animals were transferred

from each flask to a single well of a 96-well microtiter plate. In step 2, fluorescent beads were added to each well. Following a 10-minute incubation period, animals from each well of the deep-well plate were transferred to several wells of a 96-well microtiter plate for step 3. In step 4, animals in each well of the microtiter plate were imaged. In step 5, the same animals were measured using the COPAS BIOSORT. This process was repeated every hour after feeding for 72 consecutive hours (see Methods). Schematic of the experimental workflow was created with BioRender.com.

The raw data from the quantitative growth assay provides measurements of body size and feeding behavior, which are traits related to animal growth. Two measurements of body size were collected from raw data taken from the COPAS BIOSORT: time of flight (TOF) and optical extinction (EXT) (Figure S1). Time of flight is a measurement of body length, and optical extinction corresponds to optical density, a measurement influenced by body length, thickness, and composition [110,118]. We investigated whether optical extinction could be correlated to a different measure of body size using the collection of manual size measurements from images (see Methods). We calculated the median length, width, area, and volume of animals in a subset of imaged wells from each hour of the experiment. We then compared these values to well median statistics from the processed COPAS BIOSORT data. We found a strong correlation between manual measurements of animal length from the image analysis and TOF measurements from the COPAS BIOSORT (Figure S2-2). We also observed an equally strong correlation between manual measurements of animal area and EXT as well as animal width and EXT normalized by body length (norm.EXT). We then approximated animal volume using measurements from the COPAS BIOSORT by using a cylindrical approximation for *C. elegans* shape (see Methods). This result expanded the number of body size parameters that we were able to assess using the COPAS BIOSORT data, allowing us to investigate growth dynamics in length, width, and volume (**Fig. 2-1 A-C**). To disentangle nematode objects from non-animal objects (bacteria clumps, detritus, shed cuticles), we employed model-based clustering to remove unwanted objects and better examine growth of animals (Figure S2-3). Lastly, we converted COPAS BIOSORT

measurements into microns (see Methods).



**Fig 2-2. Quantitative measurements of animal size.** COPAS BIOSORT data of animal length (A), width (B), and volume (C) after the removal of non-animal objects using model-based clustering methods (see Methods).

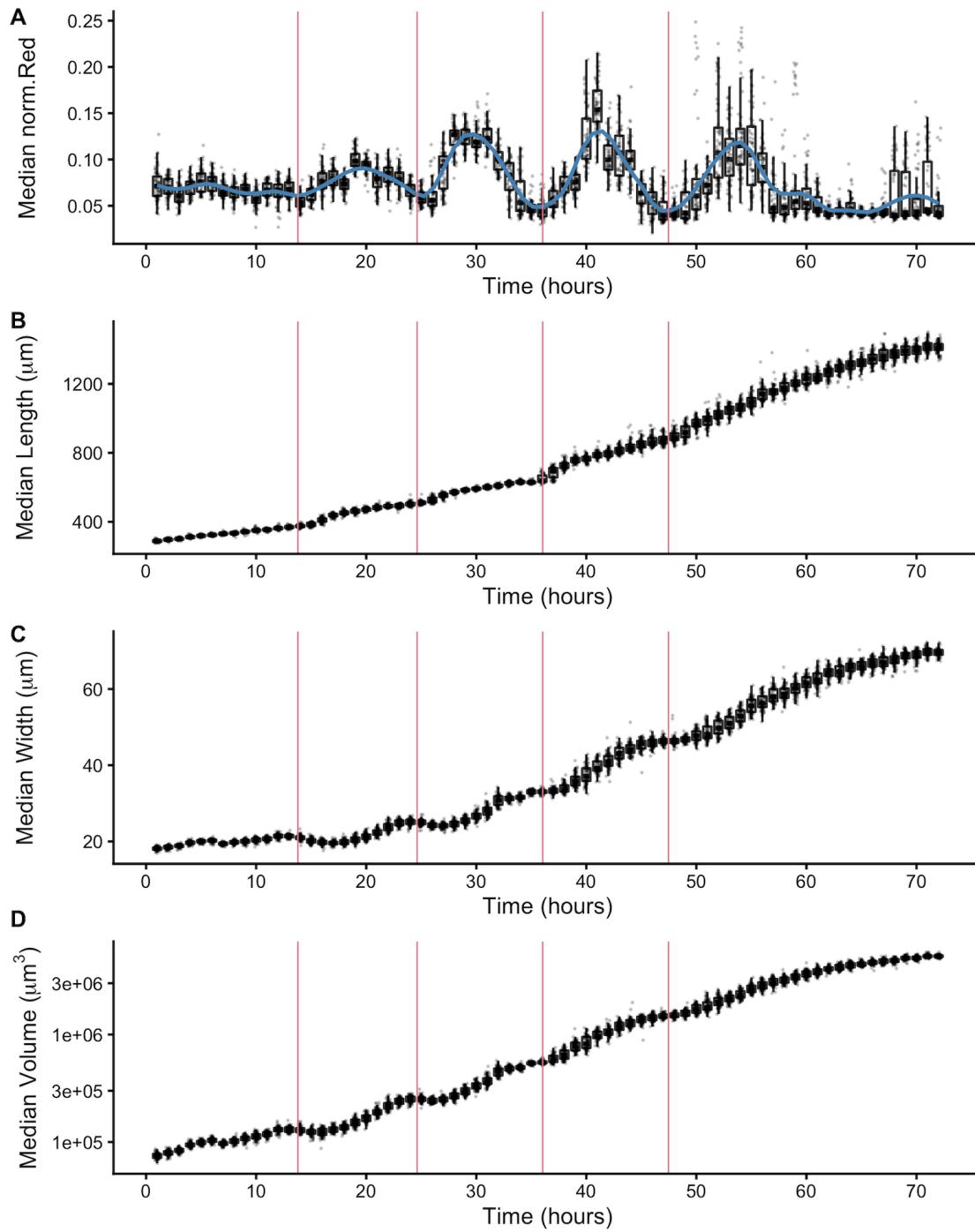
We report body length, width, and volume of animals at each hour of development from L1 to adult (Figure S2-1 and **Fig. 2-2**). Historically, growth of *C. elegans* has been shown as a sigmoidal curve where exponential growth during larval stages reaches a maximum rate in adulthood [103]. More recently, researchers have identified that growth curves are discontinuous during development punctuated by larval transitions [42,104]. Using our quantitative

growth assay, we captured these small-scale discontinuities in larval growth rate as well as an apparent growth maximum during early adulthood. We noticed that all size variables (length, width, and volume) displayed these dynamics. Objects identified as animals appear to grow in size. However, in particular time windows during development, growth dynamics visibly shift, producing discontinuities in animal growth rate. With these data, we were able to further investigate *C. elegans* growth and size control.

### 2.5.2 Fluorescence provides a quantitative measurement of animal feeding behavior and developmental progression

In addition to body size and shape, the raw data from the quantitative growth assay described above measured fluorescence of each animal object. To readily assess the thousands of measurements acquired at each hour, we generated summary statistics of median well measurements (Table S2-1). With these summarized data, we investigated the relationship between feeding behavior and developmental stage. It is well established that temporary suspensions of *C. elegans* feeding occur during each molt [43,103]. As such, active feeding is frequently used to distinguish growing animals from individuals in a molt. We quantified feeding behavior by exposing animals to fluorescent beads the approximate size of bacteria and measuring fluorescence of animals [119]. Because larger animals are able to consume more food and therefore contain more ingested food, we normalized fluorescence by animal area to account for increases in body size (Figure S2-4). The resulting fluorescence data showed a dynamic pattern (**Fig. 2-3A**). At approximately 15 hours, fluorescence steadily increased to a peak before decreasing back to initial levels at approximately hour 27. This pattern, repeated three additional times, established clear time windows of local minimal fluorescence. These local minima represent periods of time where a large proportion of the population had reduced or ceased feeding and therefore suggests time windows where a majority of animals were likely not feeding because they were in a molt. We used a local kernel regression method to estimate a smooth curve and calculate the derivative to identify the

time associated with each local minimum (see Methods). We then assessed images collected from the growth assay and demonstrated that periods of decreased feeding are concurrent with the presence of shed cuticles, supporting that animals are undergoing a molt during these periods of time (Figure S2-5). When we overlaid the timing of each local minimum on the population size data, we were able to outline the start and end of each larval stage (**Fig. 2-3 B-D**). Notably, local minima occurred approximately every ten hours, consistent with well established observations of molt timing [103]. Furthermore, we observed a clear relationship between changes in feeding behavior and growth dynamics where decreases in feeding occurred simultaneously with discontinuous growth in length, width, and volume.



**Fig 2-3. Fluorescence dynamics outline larval stages.** (A) Median normalized red fluorescence (y-axis) over time (x-axis) is shown. The blue line represents the kernel regression fit to the data. The red vertical lines correspond to the local minima of the regression

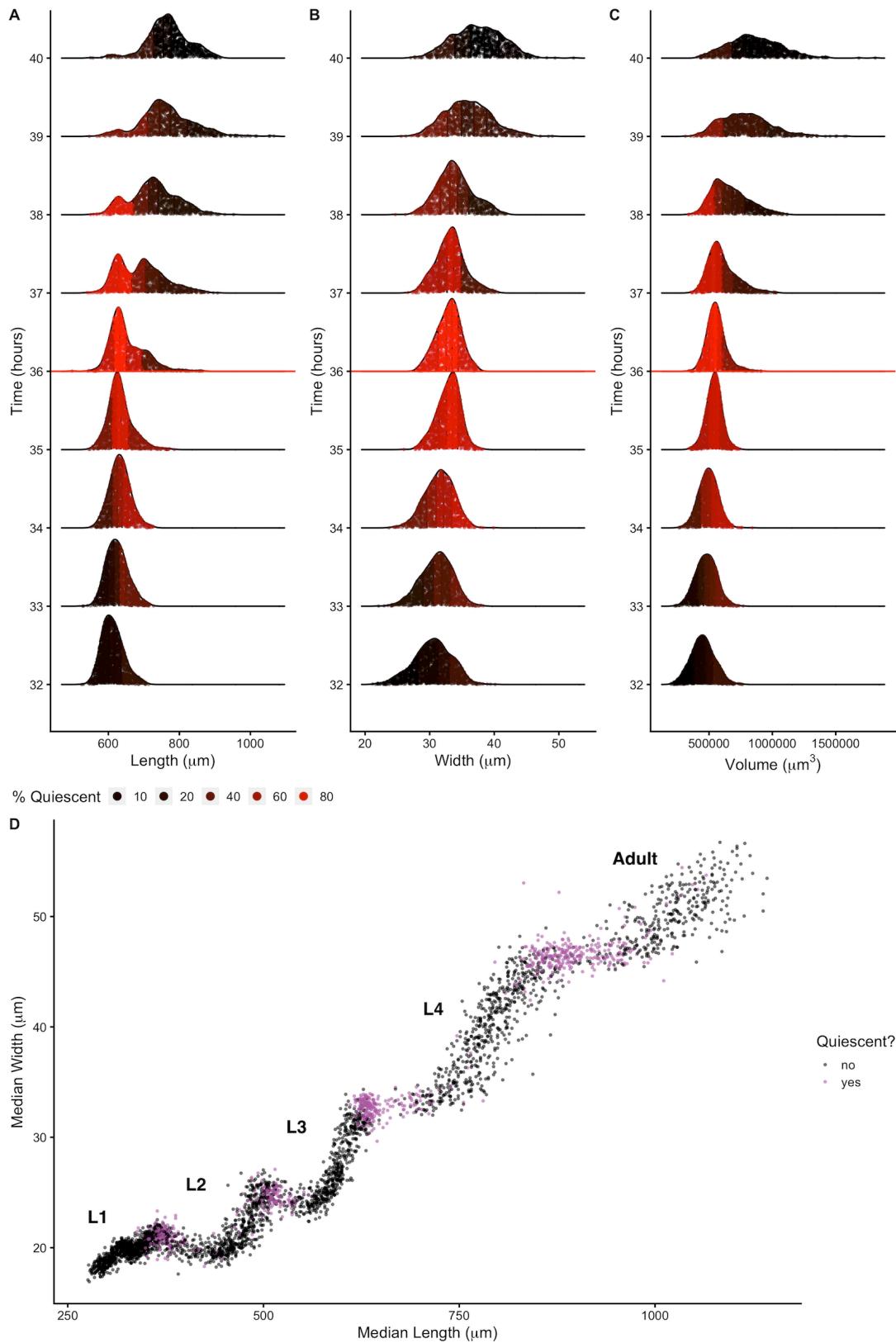
and represent the transition between larval stages. Median length (B), median width (C), and median log volume (D) are shown with larval-stage transitions as well. Upper and lower bounds of the box plots correspond to the first and third quartiles. The upper and lower whiskers extend to 1.5 times the value of the interquartile range.

### 2.5.3 Changes in *C. elegans* body shape occur at larval-stage transitions

Adult body size is ultimately determined by the coordination of developmental progression and rate of growth. To understand how *C. elegans* achieve final size, we must first examine how *C. elegans* grow. Quantitative studies of *C. elegans* growth frequently assess changes in length or volume over time; however, to fully characterize changes associated with growth, it is also important to consider the dynamics of width. Two general models were proposed for *C. elegans* growth in volume: linear and exponential [42,103,104]. Notably, these volume growth models require different dynamics in length and width. To achieve linear volume growth, length and width must increase at precise sublinear rates that together result in a linear increase in volume. If animal length and width increased at a constant linear rate, then volume would increase at a cubic rate. Alternatively, if both length and width grew exponentially, then volume would fit an exponential model. We sought to identify which model best described *C. elegans* growth behavior but were unable to consistently distinguish between linear, exponential, and cubic models using statistical information criterion because of the similarity in the shapes of the growth curves (Figure S2-6 and Table S2-2). This result is not surprising because computational simulations have shown that increases in experimental noise, above 2% added noise, limit the correct identification of growth models [120].

Growth has important implications for how animals regulate size. Size homeostasis requires that growth rate and developmental rate are coordinated. *C. elegans* reach a similar volume at each larval transition despite significant variation in individual growth rates [42]. Because

individuals in a population maintain similar sizes despite differences in growth rate, a control mechanism to regulate developmental progression must exist. Early work proposed a size-based growth control model in *C. elegans* [42], although recent work suggests that size homeostasis is achieved through a folder mechanism where growth rate and development are coupled [95]. To assess changes in body size and shape during a larval transition, we examined the dynamics of animal length, width, and volume in the hours before, during, and after each molt. We find that for each shape variable, larger animals enter molt first (**Fig. 2-4**). We also observe differences in the distributions of lengths during a larval transition compared to widths and volumes. Measurements of animal width and volume remain unimodal throughout a molt, but length does not. As larger animals begin to exit the molt, an increase in body length occurs that leads to the appearance of bimodality of lengths across the population. Importantly, volume remains constant while length increases and width decreases, indicating a change in body geometry not size. Notably, the length increase occurs simultaneously with a decrease in width across the population (**Fig. 2-4D**). These changes in the physical dimensions at each larval transition suggests that body shape, in addition to size, is involved in the control of *C. elegans* growth.



**Fig. 2-4. Changes in body shape occur during larval-stage transitions.** Population density curves of length (A), width (B) and volume (C) for the hours surrounding the L3 - L4 larval transition (red horizontal line at 36 hours corresponds to the molt). Each distribution was divided into five quantiles. The percentage of quiescent animals present within each quantile was calculated (see Methods), and each quantile was colored to reflect this percentage. In all shape variables, quantiles that contain the largest animals displayed an increase in quiescence earlier than quantiles that contain the smallest animals. These dynamics were consistent across all larval-stage transitions (Figure S2-7). (D) Median width vs. median length for experimental hours 1 - 55. Red indicates measurements that fall above the quiescence threshold (see Methods). Simultaneous changes in length and width occur during periods of increased quiescence.

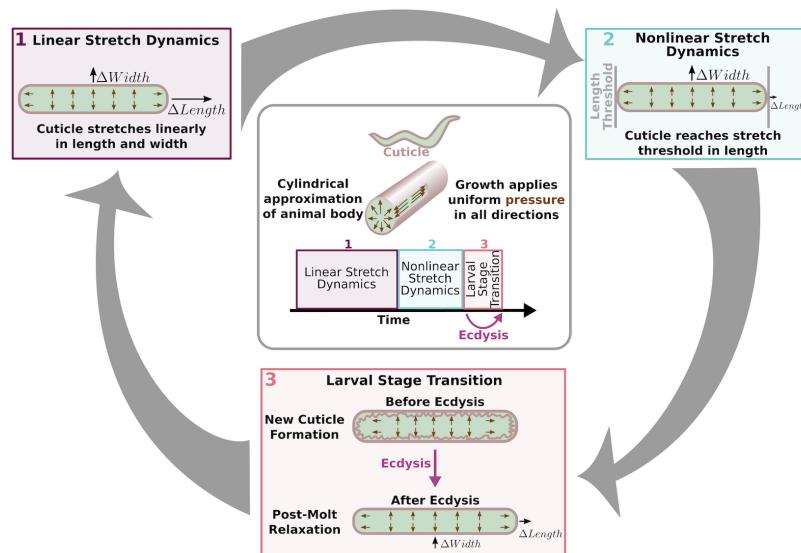
### 2.5.4 Modeling *C. elegans* cuticle stretch dynamics

#### 2.5.4.1 Sensing of cuticle stretch as a trigger for larval-stage transitions

Previous studies theorized that the internal mechanism for sensing body size and triggering molts in *C. elegans* is driven, in part, by the properties of the collagen-rich cuticle [42,95]. Many cuticle collagen mutations cause morphological defects in nematode shape some of which cause animals to be shorter but do not impact animal width, implying that the cuticle affects length and width independently [32]. The *C. elegans* cuticle does not grow through the addition of new material, but rather stretches to accommodate increases in animal body size. Cuticle stretch is likely limited by the material properties of the cuticle. The *C. elegans* cuticle is primarily made of cross-linked collagens organized into lateral ridges and circumferential bands [63]. Commonly found in many biological systems, collagen-based materials are fairly flexible under low stress conditions. However, as stress increases, collagen fibrils may become elongated and orient in the load bearing direction leading to a decrease in elasticity [121]. Previous work using atomic force microscopy revealed a high level of biomechanical stiffness at the circumferential bands [122], leading others to speculate that

mechanical strain on these structures is likely adjusted as internal body pressure changes [123]. Additionally, in *nekl-3(sv3)* molting mutants, the cuticle is not properly removed from the middle part of their body, leaving the free head and tail to grow normally while the encased middle is constricted by the old cuticle to pre-molt dimensions [124]. Given this body restriction, we speculate that the old cuticle stretches beyond its tolerance, becomes stiff, and constricts the center of the nematode relative to the growing head and tail size. We hypothesize that *C. elegans* may be able to sense when the cuticle becomes restrictive. If so, the reduction of elasticity or “stretchiness” of the cuticle can act as one signal, among many, used to determine when to initiate a molt.

To understand the impact of cuticle elasticity on body shape, we developed a “Stretcher” model. We propose that the nematode passes through three distinct regimes related to cuticle stretch: linear stretch dynamics, non-linear stretch dynamics, and larval stage transition (**Fig. 2-5**). These regimes arise naturally from the following biologically supported assumptions. The cuticle structure is anisotropic, possibly leading to distinct properties in the length and width directions [125,126]. We approximated the cuticle as a hollow cylinder of negligible thickness filled by the body of the nematode. Growth was modelled as internal pressure evenly applied to the cuticle in all directions. An anisotropic, elastic cuticle responds differently to pressure from growth during linear stretch, nonlinear stretch, and post-molt relaxation, leading to differences in shape during each stage of development.



**Fig. 2-5. Cuticle stretch dynamics guide larval-stage transitions.** The Stretcher model describes each larval stage as a cycle. Nematodes are modeled as a cylindrical object with a thin cuticle epidermis. (Box 1) Linear Stretch Dynamics: uniform growth pressure stretches the cuticle linearly in both length and width. (Box 2) Nonlinear Stretch Dynamics: the cuticle has reached a stretch threshold in length, and under uniform growth pressure the length stretches less (sub-linear) and width stretches linearly. (Box 3) Larval Stage Transition: a new cuticle is formed and the old cuticle is shed (ecdysis), removing constraints in length. The nematode body “relaxes” in length, causing an increase in length, a decrease in width, and constant volume.

In the linear stretch regime (Fig. 2-5), the cuticle would be linearly elastic in both the length and width directions, stretching proportionally to the pressure exerted on the cuticle. Previous work found evidence for a linearly elastic cuticle [127,128] in animals expanded in a negative external pressure environment or after positive force was applied to the cuticle. Gilpin *et al.* have found evidence of linear elasticity in the nematode body. We conjecture that this linear elasticity is caused by the constraints applied by the cuticle [127,128]. A linearly elastic cuticle will have  $\Delta L$  stretch in the length direction and  $W$  stretch in the width direction, each related to growth-applied pressure  $\Delta p$  by

$$\Delta L = a_L \Delta p$$

$$\Delta W = a_W \Delta p.$$

The “stretch coefficients” in length,  $a_L$ , and width,  $a_W$ , measure the stiffness of the cuticle. Smaller values correspond to a stiffer material, which is less able to stretch in response to pressure. The stretch coefficients are constant in the linearly elastic regime and are determined by geometric constants and material properties. The ratio of the change in length (Eq. 1) and width (Eq. 2) produces a pressure-independent relationship that depends only on the ratio of the geometric and material properties, which can be verified using measurements of length and width (Fig. 2-3). During the linearly elastic regime, the ratio of growth in width to growth in length is constant throughout a larval stage where the cuticle properties are fixed as in

$$\Delta W / \Delta L = a_W / a_L = \text{constant}.$$

In the non-linear stretch regime (**Fig. 2-5**), growth continues to apply pressure to the cuticle uniformly in all directions. As observed in *nekl-3(sv3)* mutants, the cuticle can restrict body growth [124]. Once outside of the linearly elastic regime, the cuticle would hardly stretch, even under large forces. We hypothesize that this shift from linear to nonlinear regimes can provide a mechanism for size-sensing and cues the larval-stage transition (**Fig. 2-5**). In principle, this transition could occur in either the width or length directions. For simplicity we illustrate a transition from linear to non-linear stretch in the length direction while linear stretch in the width direction is maintained. In the nonlinear regime, the stretch in the length direction in response to pressure becomes

$$\Delta L \approx \tilde{a}_L(p) \Delta p.$$

The nonlinear “stretch coefficient,”  $\tilde{a}_L(p)$ , is no longer constant and decreases with increasing pressure. It is smaller than  $a_L$  because the cuticle has become less elastic than in the linear regime. If the length-direction enters the nonlinear regime and has reduced stretch response, while width has the same constant stretch response then, we expect the  $\Delta W/\Delta L$  ratio to increase

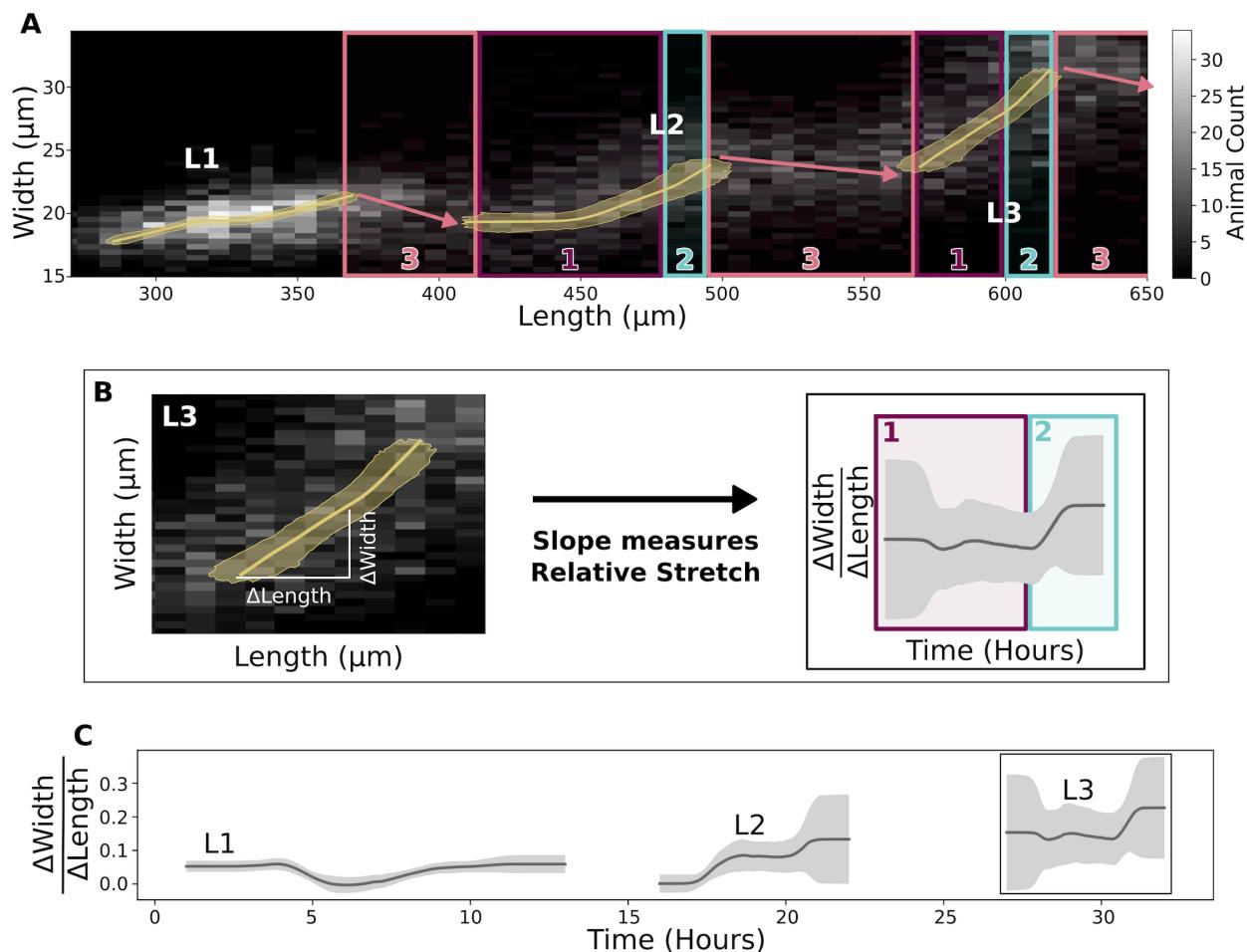
$$\Delta W/\Delta L|_{non-linear} = a_W/\tilde{a}_L(p) > a_W/a_L = \Delta W/\Delta L|_{linear}.$$

During the larval-stage transition (Fig. 2-5), a new, larger cuticle is formed beneath the old cuticle that is shed during ecdysis. Because the old cuticle constrained growth in length, we predict a rapid increase in the length direction when the constraint is removed. Animal volume is conserved as growth does not occur during this process. Therefore, the relaxation in length is accompanied by a corresponding decrease in width.

#### 2.5.4.2 Discontinuities in animal growth rate driven by limits of cuticle stretch

To test whether the shape dynamics predicted by the Stretcher model are supported by data, we analyzed the relationship between measured nematode length and width over developmental time. Although the shape relaxation for an individual animal is expected to happen at a much shorter time scale than data collection (seconds to minutes), the measured shape change within the population occurs on a larger time scale (several hours). As such, we can observe the sudden shape change at transitions by assessing changes in length and width of the population (Fig. 2-6). Doing so, we detected all three regimes predicted by the Stretcher model in the COPAS BIOSORT data: linear stretch, non-linear stretch, and relaxation (Fig. 2-6A). In all larval stages, we observed an approximately constant instantaneous  $\Delta W/\Delta L$  ratio, consistent with a linear stretch regime (Fig. 2-6B and Fig. 2-6C). We also detected a large slope decrease during the L1 stage, which could correspond in time to the metabolic decision for entry into the dauer stage [129] or divisions of the seam cells [33]. Following the linear stretch regime, we noted a simultaneous increase in length and

decrease in width at the transition between larval stages, consistent with a length threshold in the Stretcher model (**Fig. 2-5D**, **Fig. 2-6A**, and Figure S2-8). Although transitions in the slope ( $\Delta W/\Delta L$  ratio) are difficult to detect at all larval stages because of noise amplification and population effects (Figure S2-9), near the end of L2 and L3 stages we observed a sharp slope increase, consistent with a non-linear stretch regime in length prior to lethargus. As a whole, these results suggest that the material properties of the cuticle influence body shape at larval-stage transitions, leading us to speculate about its role in molt initiation. The ability to sense when a critical cuticle stiffness is reached would allow animals to monitor fold changes in body growth within each larval stage, serving as a connection between growth rate and developmental timing.



**Fig. 2-6.** Experimental data are consistent with a length threshold in cuticle

**stretch.** (A) A grayscale histogram of the width (y-axis) vs length (x-axis) of all sampled animals in replicate 2. The range of all bootstrap regressions is in gold. (B) Demonstration of calculating the ratio of width-to-length stretch as the local slope using L3. Left panel is a repetition of L3 data from **Fig. 2-6A**. Right panel is a repetition of results from **Fig. 2-6C**. (C) Within a larval stage, the ratio of width to length stretch varies over time. The standard deviation captures population variation (grey).

## 2.6 Discussion

Using an integrated image-based and flow-based phenotyping strategy to precisely evaluate feeding, growth, and molt dynamics at high replication, we detected oscillations in feeding behavior consistent with larval progressions and used these dynamics to define larval stages. We observed changes in body shape at each larval-stage transition that can be driven by differences in physical cuticle properties along length or width (anisotropy). These results suggest a mechanism by which animals sense their size and control molt timing by detecting the physical stretch of the cuticle, demonstrating how physical constraints can influence developmental timing and growth rate.

### 2.6.1 A potential role for cuticle stretch in the timing of larval-stage transitions

Measurement of both animal length and width allowed us to observe changes in body shape as well as body size. We propose a mechanism in which a stretch threshold along the body length axis acts as a trigger to larval-stage transitions. Previously, a folder mechanism for *C. elegans* growth has been suggested [95]. Mechanical stretch sensing could provide organisms a way to couple the rate of growth and development to maintain a constant volume fold change within a larval stage and achieve size homeostasis within a population. In this way, smaller animals would reach a stretch limit at a smaller size as the cuticle would only stretch a

percentage of its original size before reaching a threshold. For cuticle stretch to trigger larval-stage transitions, animals must either have the ability to measure the amount the cuticle has stretched or the stiffness of the cuticle. Across biological systems, cells can respond to the stiffness of their environment using mechanosensitive components [130,131], but few examples in tissues or whole-organisms are known. In *C. elegans*, it has been demonstrated that hemidesmosomes, which connect the cuticle and the epidermis, are mechanosensitive during embryogenesis [132,133]. Additionally, dense bodies, which connect the epidermis and muscles, are hypothesized to be mechanosensitive as well [48,134–136]. Changes in cuticle composition, and presumably stiffness, have been shown to also affect well known growth controlling pathways such as the BMP signaling pathway [137]. These possible mechanosensitive components could monitor the stiffness of the cuticle and be part of the signaling pathways that regulate the initiation of molts. Further experiments are required to explicitly test whether these components control larval-stage transitions. It is important to note, however, that evidence for a stretch-based mechanism for growth control does not preclude the possibility of a developmental timer. It is likely that physical constraint-based events are part of a larger regulatory system coordinating developmental decisions.

Our analysis of width-to-length ratio variation over larval stages provides a first approximation of the timing of cuticle stretch properties. The sudden increase we observed in the width-to-length ratio suggests that the elasticity in the length direction decreases substantially before a molt and may be detectable by the animal. These observations of shape change also indicate that *C. elegans* do not preempt the shape change by molting before they hit the “stretch limit,” supporting the role of a physical constraint mechanism. Interestingly, when observing the L4 to adult transition, others have detected anisotropic constriction on the transverse (width) axis followed by gradual relaxation driven by rearrangements in cortical actin networks [138]. Single-worm, high frequency measurements targeting hours surrounding the sudden width-to-length ratio increase, are needed to better resolve cuticle shape dynamics. Although analysis of the larval growth dynamics for *C. elegans* body shape mutants (*dpy*, *lon*, *sma*) may provide insight into width-to-length ratio variation, measure-

ments of whole animal length and width only provide a total stiffness estimate, leaving us unable to distinguish the contributions of cuticle stiffness from other tissues. For instance, previous work has shown that bodies of *dpy-10(e128)* are twice as stiff as wild type [139] while *dpy-5(e61)* animals are softer than wild type animals [127]. Authors have speculated that this difference in stiffness is due to an increased internal glycerol concentration in *dpy-10(e128)* animals that is absent in *dpy-5(e61)* [139], indicating that body shape alone does not predict body stiffness. To assess cuticle properties independent of other nematode tissues and organs, future experiments probing the stiffness of free cuticles are necessary.

Additionally, within the L1 stage, the relative stretch measured in width and length did not follow the pattern observed in other larval stages. We observed a mid-stage dip in the width-to-length ratio that is otherwise approximately constant throughout the L1 stage. Experiments exploring the structural properties of cuticles at all larval stages can help to determine where the L1 shape changes originate.

### 2.6.2 Development comprises complex interactions of growth regulation across diverse scales

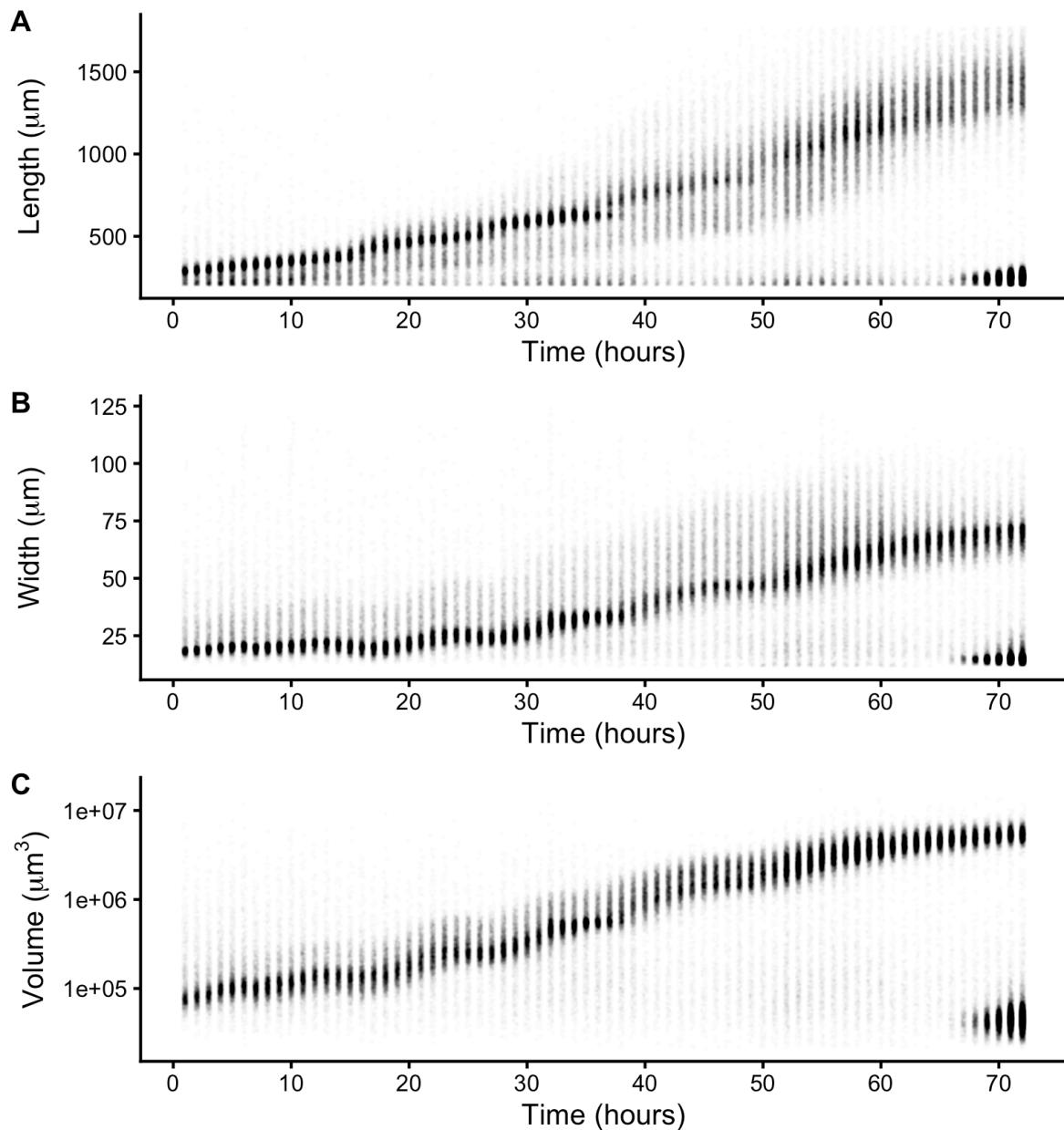
Our results demonstrate that *C. elegans* may use physical constraints on animal size, in tandem with other regulatory mechanisms, to control growth rate and determine developmental transitions. This type of regulation could be applicable to organisms with stiff cuticles or other physical barriers to growth, like many species of Ecdysozoa. The control of whole-organism growth requires cells, tissues, and organs to orchestrate final size and cell number. In *C. elegans*, cell number is precisely defined and invariant from animal to animal [140], so the final adult size of an individual must come from cell size as opposed to number. Future studies should focus on how whole-organism size is determined by the integration of cell, tissue, and organ size. By incorporating these different developmental scales, the Stretcher model can be refined to completely describe how physical constraints on parts of the organism impact the whole. *C. elegans* gives investigators a method to investigate animal-to-animal

variation in developmental trajectories across each of these scales.

## 2.7 Contributions

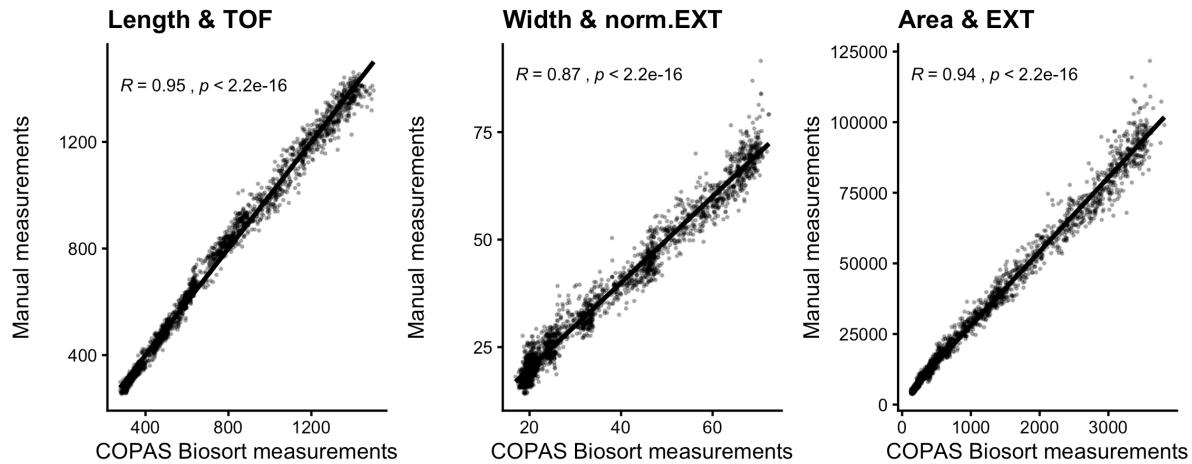
This work was supported by the NSF-Simons Center for Quantitative Biology at Northwestern University. Dr. Erik Andersen, Dr. Gaotian Zhang, and Nicole Roberto assisted with the high-throughput sampling. Dr. Sasha Shirman and Christina Goss built and tested the mathematical models. Hannah N. Ahmed, Elliot J. Andersen, Isabella R. Miller, Justine K. Rozenich, Iris L. Swarthout, and Jordan A. Vaughn helped collect manual size measurements of animals from images. Drs. Erik Andersen and Niall Mangan supervised the project.

## 2.8 Supplement

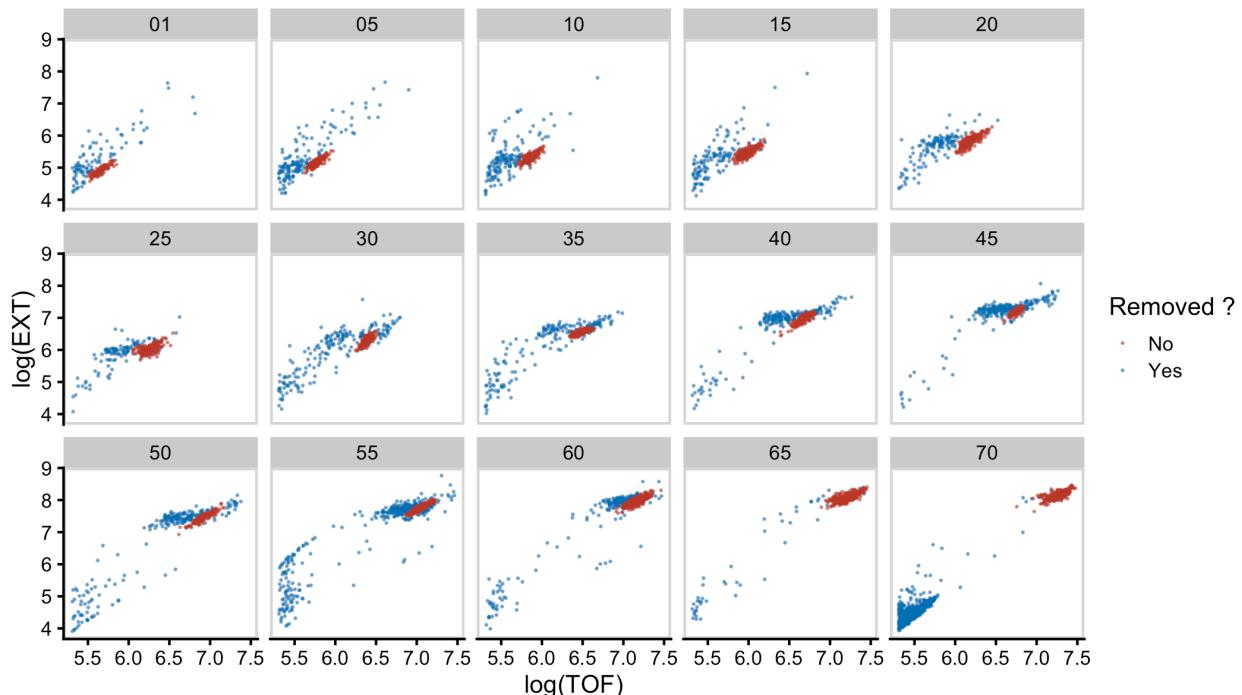


**Figure S2-1. Raw measurements of animal size.** Raw COPAS BIOSORT data of animal length (A), width (B), and volume (C) are shown here. After 60 hours, animals have developed to the adult stage. Smaller objects observed after 65 hours were the next generation of newly hatched L1 larvae laid by the animals that developed during the time

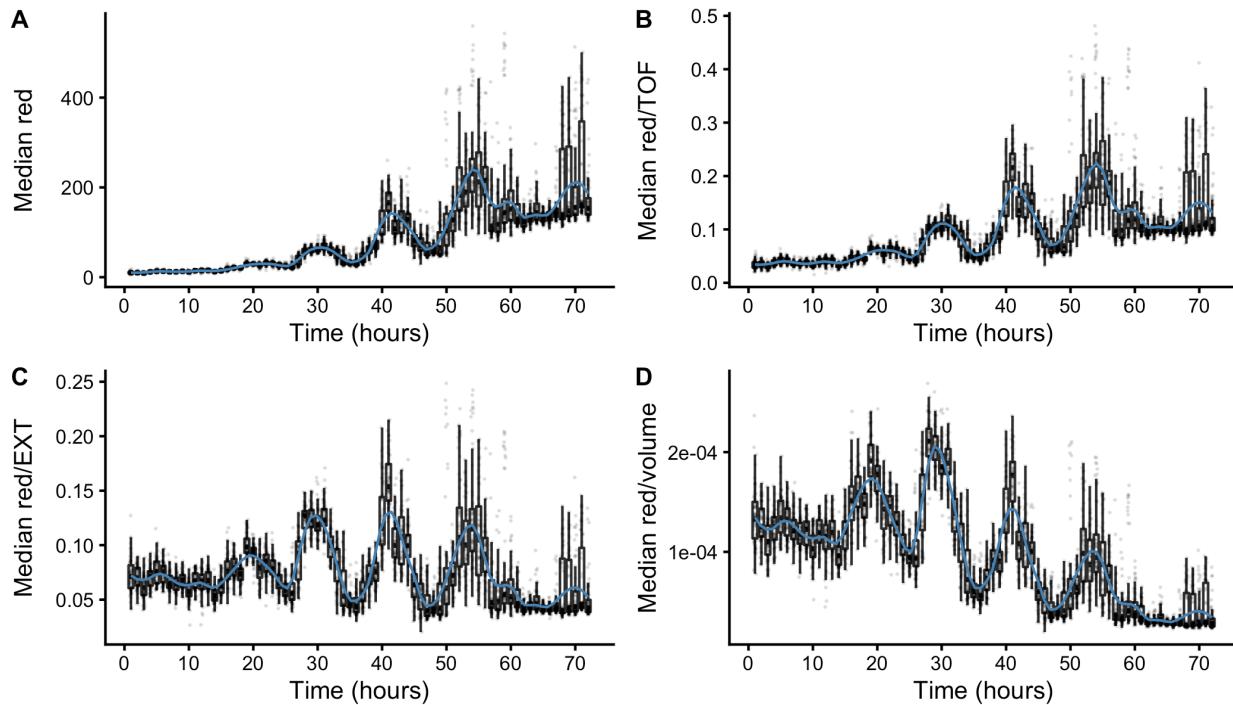
course.



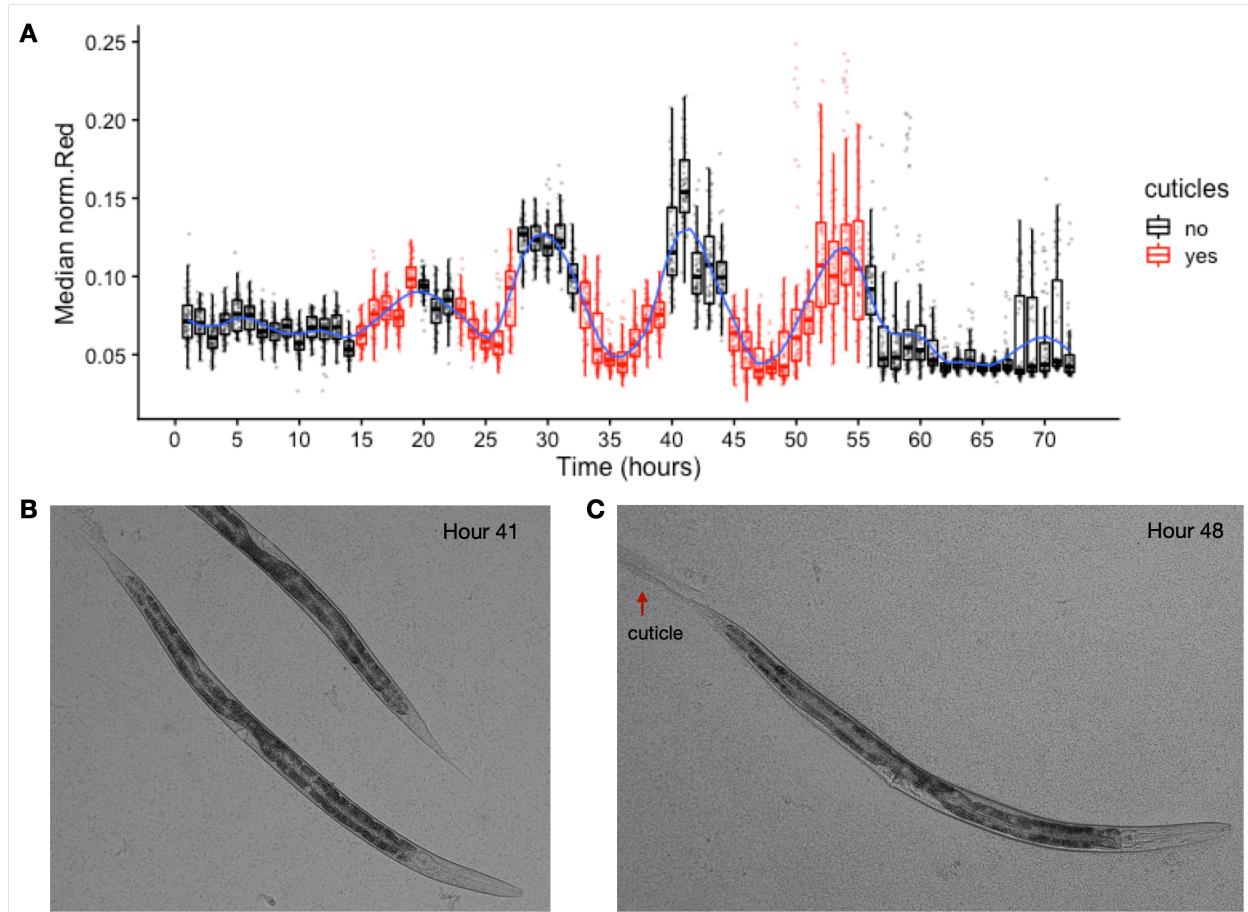
**Figure S2-2. Correlation analysis of body size measurements.** Manual measurements of animal length, width, and estimated area were compared to COPAS BIOSORT measurements of TOF, norm.EXT, and EXT. Kendall correlation value is shown in each plot.



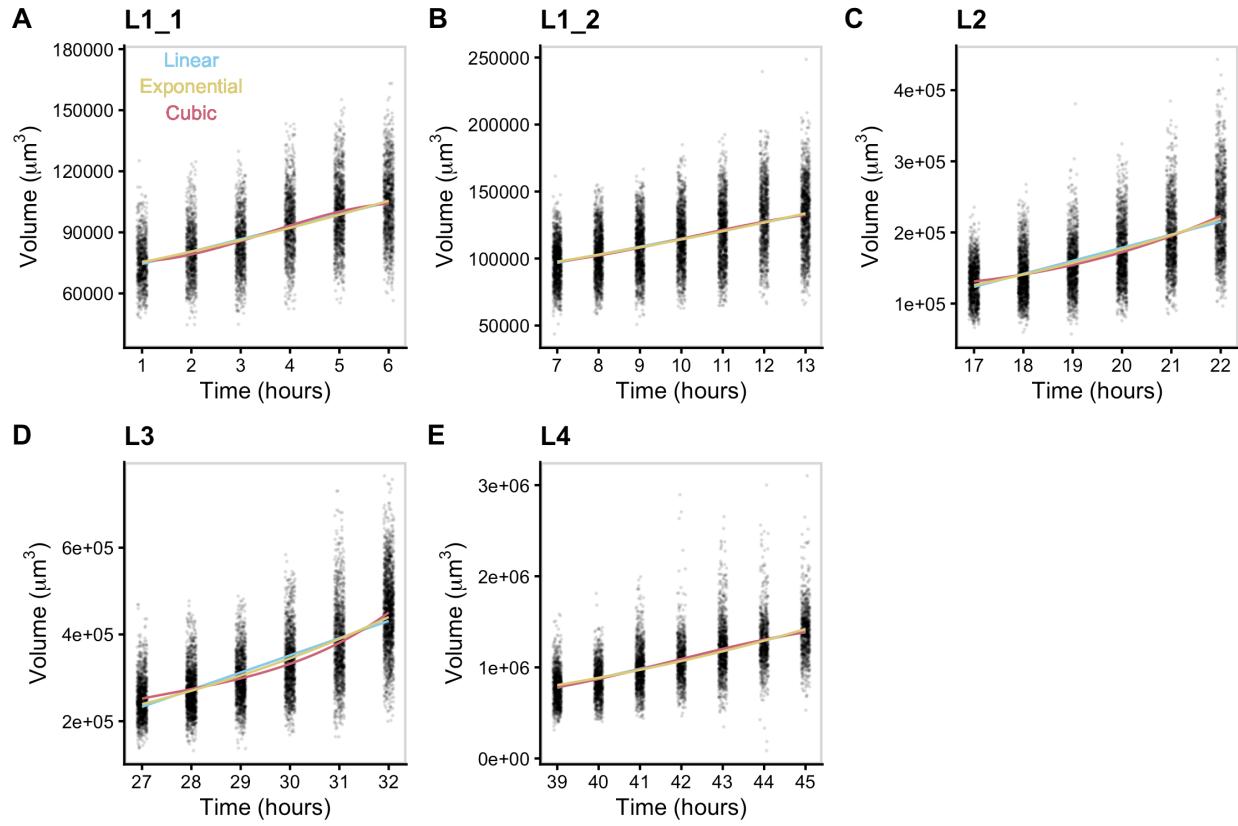
**Figure S2-3.** Mixture modeling of COPAS BIOSORT data was used to prune data. Mixture models of Gaussian distributions were fit to log transformed animal length (x-axis) and log transformed optical extinction (y-axis). Data from each hour of the experiment was analyzed and processed to remove clusters that did not include animal objects. All replicates were pruned independently; a subset of data from replicate 2 is shown here. Panels indicate experimental hours from which data were taken.



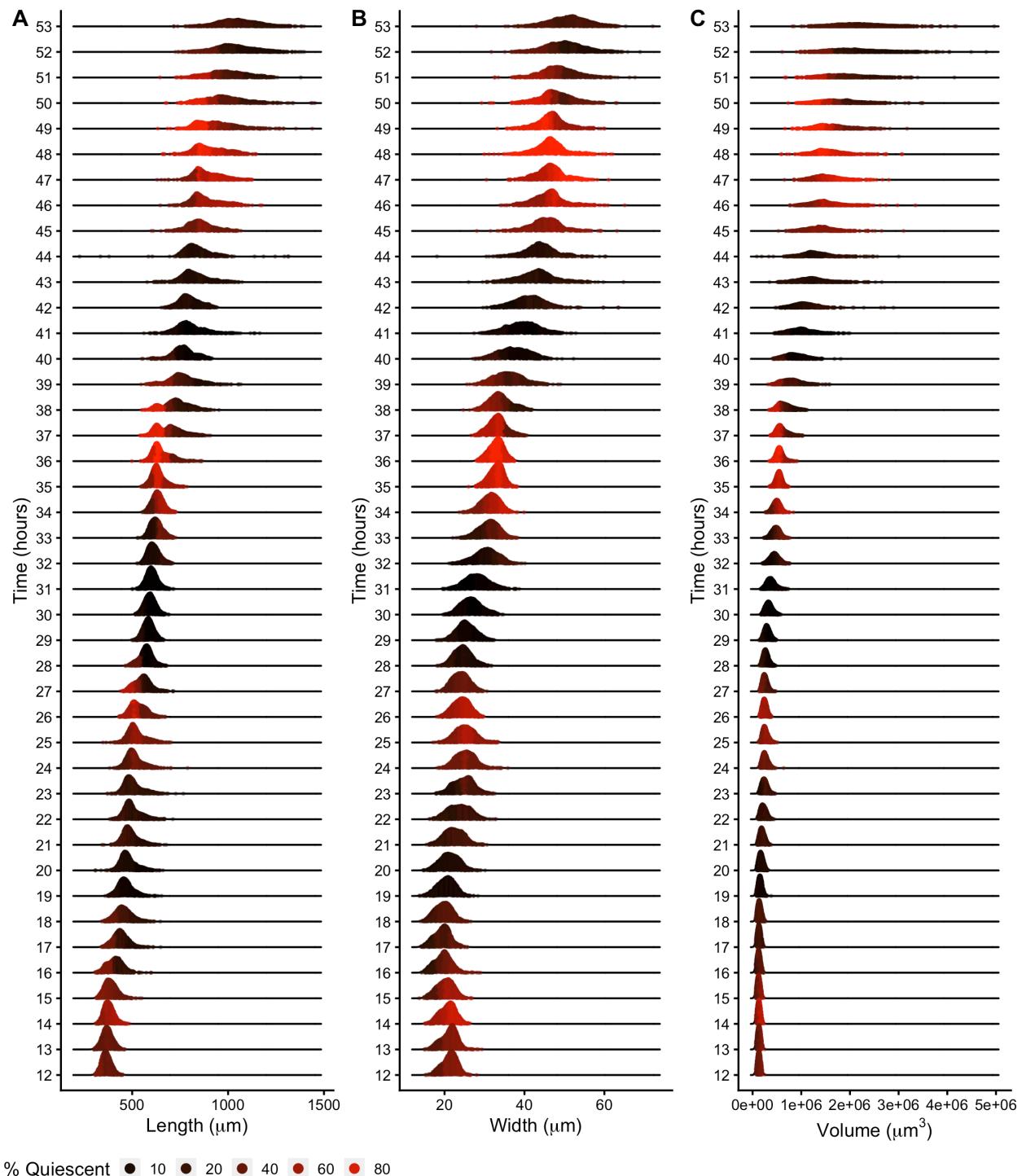
**Figure S2-4.** Fluorescence measurements normalized by body size. Red fluorescence beads were fed to animals during experimentation and fluorescence data was collected by the COPAS BIOSORT. Fluctuations in fluorescence indicate fluctuations in feeding behavior. Fluorescence data was normalized by body size measurements to account for increases in body size. Dividing fluorescence by area was most successful in normalizing fluorescence dynamics to account for changes in animal size over time.



**Figure S2-5. Cuticles identified during periods of decreased feeding.** Images of wells collected during the experiment were examined for evidence of shed cuticles. (A) Experimental hours where cuticles were identified from images overlap with hours where population feeding behavior is low. Cuticles shed from the L4-Adult molt persisted longer than previous larval stage cuticle debri. (B) Example image of animal without visible cuticle during a period of elevated feeding. (C) Example image of an animal with visible cuticle indicating completion of molt during a period of decreased feeding.

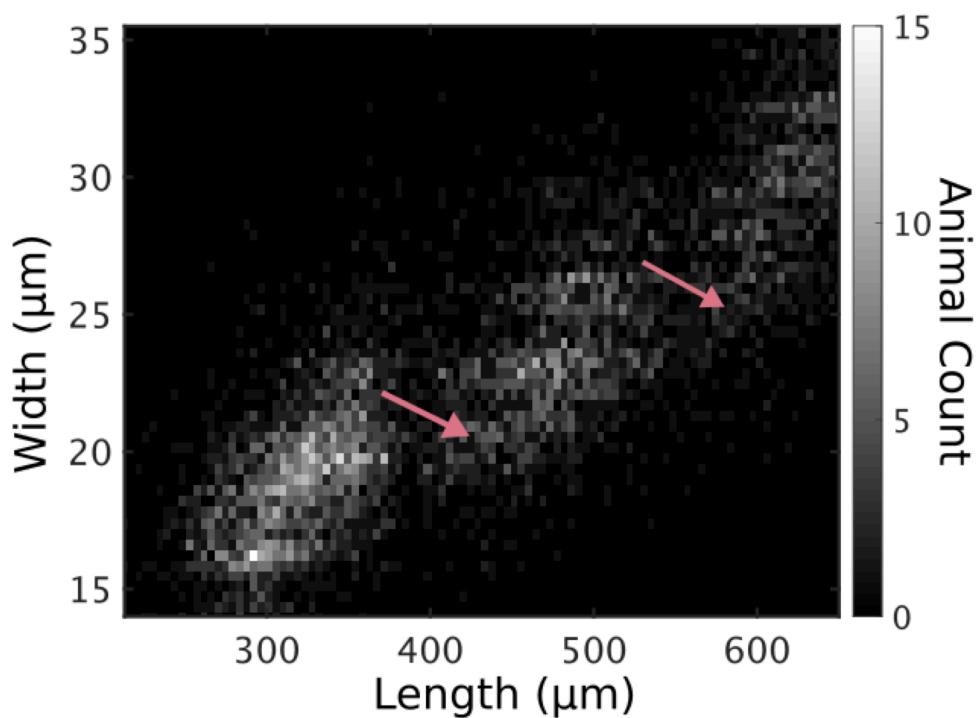


**Figure S2-6. Volume growth data fit with linear, exponential, and cubic models.**  
 Volume data of individuals in time points defined as growth periods are analyzed for each stage. L1 stage was further separated into two periods to account for the volume dip that occurs mid-stage.

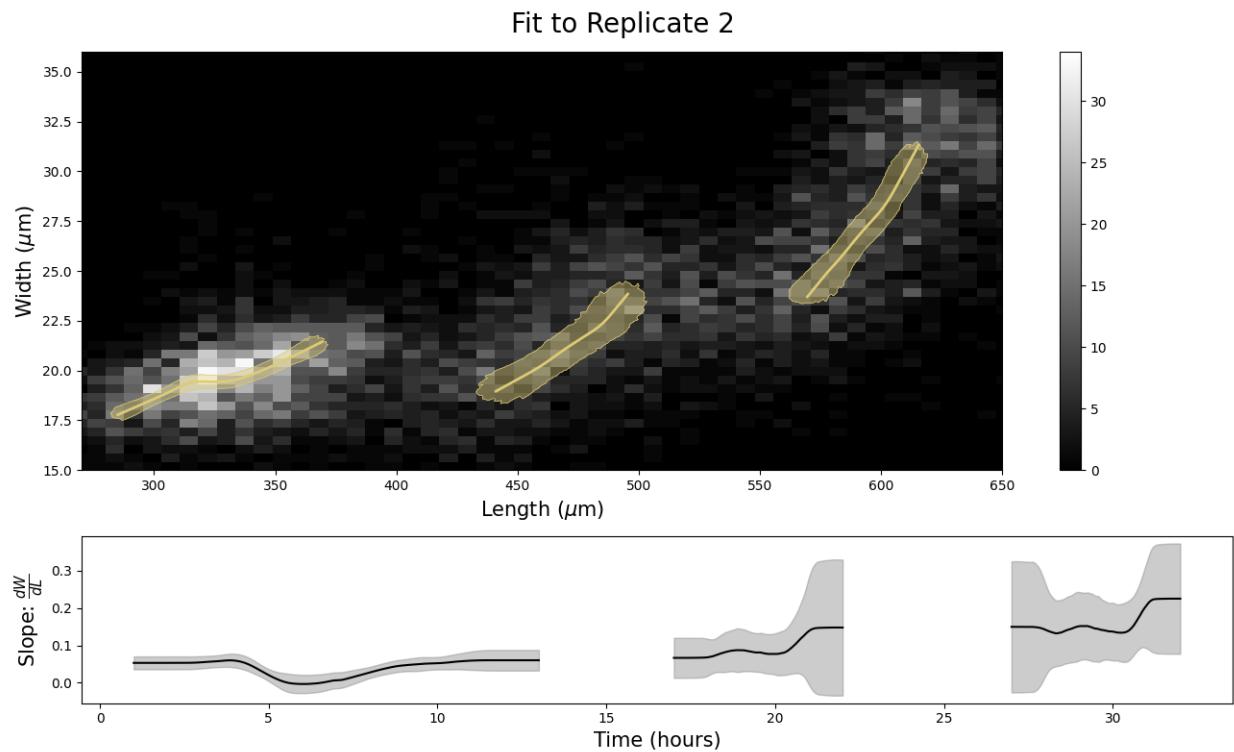


**Figure S2-7. Density plots of population size dynamics across all larval transitions.** Density curves of length (A), width (B) and volume (C). Curves are divided into five

quantiles and colored by the percentage of quiescent animals present within that quantile. Molt s are estimated to occur at experimental hours 14, 25, 36, and 48 (see Methods).



**Figure S2-8. Animals in all replicates, measured from images.** Animal length and width over *C. elegans* development captured from image data. Higher noise levels in these measurements preclude accurate regressions to individual larval stages. Length jumps and width dips are still apparent. Compare with Fig 6.



**Figure S2-9. Stretcher model analysis of replicate 2 COPAS BIOSORT data for different stage thresholding.** Compare to Figure 2-5. Larval hours were defined by taking the ceiling of the lower boundary and the floor of the upper boundary. This rounding method for larval stage definition demonstrates the sensitivity of the analysis to edge effects. The unexpected step in the L2 larval stage (Fig 5) was significantly reduced with this rounding method.

**Response = Norm.Red**

<i>Terms</i>	<i>Df</i>	<i>Sum Sq</i>	<i>Mean Sq</i>	<i>F value</i>	<i>Pr(&gt;F)</i>	<i>% Var Explained</i>
hour	1	439.46	439.46	217762.01	0	54.34
replicate	6	165.82	27.64	13694.48	0	20.51
well	10	0.32	0.03	15.62	0	0.04
Residuals	100619	203.06	0	NA	NA	25.11

**Response = Length**

<i>Terms</i>	<i>Df</i>	<i>Sum Sq</i>	<i>Mean Sq</i>	<i>F value</i>	<i>Pr(&gt;F)</i>	<i>% Var Explained</i>
hour	1	86190107879	86190107879	8223506	0	98.4
replicate	6	349438944	58239824	5557	0	0.4
well	10	834970	83497	8	0	0
Residuals	100619	1054582098	10481	NA	NA	1.2

**Response = Width**

<i>Terms</i>	<i>Df</i>	<i>Sum Sq</i>	<i>Mean Sq</i>	<i>F value</i>	<i>Pr(&gt;F)</i>	<i>% Var Explained</i>
hour	1	209090438.47	209090438.47	6950760.3	0	97.99
replicate	6	1266985.87	211164.31	7019.7	0	0.59
well	10	1495.32	149.53	4.97	0	0
Residuals	100619	3026786.99	30.08	NA	NA	1.42

**Response = Volume**

<i>Terms</i>	<i>Df</i>	<i>Sum Sq</i>	<i>Mean Sq</i>	<i>F value</i>	<i>Pr(&gt;F)</i>	<i>% Var Explained</i>
hour	1	621712474191714688	621712474191714688	843342	0	84.08
replicate	6	43356968850687072	7226161475114512	9802	0	5.86
well	10	152458715405862	15245871540586	21	0	0.02
Residuals	100619	74176459811167296	737201321929	NA	NA	10.03

**Table S2-1. Results of analysis of variance models fit to COPAS BIOSORT data.** Analysis of variance tests were used to quantify the amount of variance in our data contributed by the sampling technique. The sampling technique involved unbiased sampling

of animals from six replicate populations and subsequent distribution into multiple wells of a microtiter plate for analysis. We quantified the amount of variance contributed by replicate and well. We find that the variance explained by well is nearly negligible whereas replicate contributes minor variance in some measurements. Given this information, we deem the generated summary statistics an appropriate representation of the population.

Stage	$\Delta AIC$			$\Delta BIC$			Best model by AIC	Best model by BIC
	Linear	Exponential	Cubic	Linear	Exponential	Cubic		
L1_1	17	21	0	4	9	0	Cubic	Likely Cubic
L1_2	2	4	0	0	2	12	Can't distinguish	Can't distinguish
L2	142	43	0	128	28	0	Cubic	Cubic
L3	374	145	0	360	131	0	Cubic	Cubic
L4	4	44	0	0	40	10	Likely Cubic	Linear

**Table S2-2. Model fit criteria used to assess candidate growth models.** To determine the level of support for each model, the candidate model with the smallest raw AIC/BIC was identified and compared to other AIC/BIC values. If the delta value was greater than 6, the model with the smallest AIC/BIC value was denoted as the best model. If the delta value was less than 6 but greater than 2, the model with the smallest AIC/BIC value was determined to likely be the best model. If the delta value was less than 2, we are unable to distinguish the model of best fit.

# Chapter 3

## Linkage mapping reveals loci that underlie differences in *C. elegans* body growth

### 3.1 Preface

The Andersen lab is particularly adept at connecting phenotypic variation in a population to unique genetic variants. As a follow-up to the work described in the previous chapter, I was interested in leveraging this expertise to explore how natural genetic variation influenced growth in *C. elegans*. This meant running another 50-hour-long time course experiment (on my golden birthday no less!), digging into a treasure trove of “throw-away” data collected in 2014, and working with a few new strains other than our good pal N2. The following chapter is based off of my work on this project, submitted as a first-author manuscript to G3 in March 2022.

## 3.2 Abstract

## 3.3 Introduction

## 3.4 Materials and methods

## 3.5 Results

## 3.6 Discussion

## 3.7 Future directions

## 3.8 Contributions

# Chapter 4

## easyXpress: An R package to analyze and visualize *C. elegans* microscopy data generated using CellProfiler

### 4.1 Preface

What I feel to be the most rewarding project of my doctoral work forms the basis of this chapter. I joined the Andersen lab with an introductory level of coding experience in Python but had little knowledge of R. Making quick study of the new programming language, I have since fallen head over heels in love with it. Now, I use R to not only analyze data but also generate reports, build websites, and even write this thesis (thank you, bookdown! [141]). This chapter describes the R package I helped develop for handling high-throughput *C. elegans* microscopy data which was published in PLOS ONE in 2021 [142]. I am beyond ecstatic to contribute to the open-source and #rstats worlds and hope to continue to do so in my future.

## 4.2 Abstract

High-throughput imaging techniques have become widespread in many fields of biology. These powerful platforms generate large quantities of data that can be difficult to process and visualize efficiently using existing tools. We developed easyXpress to process and review *C. elegans* high-throughput microscopy data in the R environment. The package provides a logical workflow for the reading, analysis, and visualization of data generated using CellProfiler’s WormToolbox. We equipped easyXpress with powerful functions to customize the filtering of noise in data, specifically by identifying and removing objects that deviate from expected animal measurements. This flexibility in data filtering allows users to optimize their analysis pipeline to match their needs. In addition, easyXpress includes tools for generating detailed visualizations, allowing the user to interactively compare summary statistics across wells and plates with ease. Researchers studying *C. elegans* benefit from this streamlined and extensible package as it is complementary to CellProfiler and leverages the R environment to rapidly process and analyze large high-throughput imaging datasets.

## 4.3 Introduction

Developments in high-throughput imaging techniques have led to a rapid increase in these data. Researchers are able to move away from the laborious manual collection of images that typically limits large-scale analyses [143]. Furthermore, these advances have enabled scientists to collect data of intact cells, tissues, and whole-organisms with increased temporal and spatial resolution [144]. However, typical users require software methods for efficient handling, analysis, and visualization to make the most of these extensive image datasets.

*C. elegans* is a globally distributed, free-living roundworm nematode that is amenable to many types of experimental biology. The *C. elegans* cell lineage is completely characterized [33], and the *C. elegans* connectome is completely mapped [145], making these animals an exemplary model for developmental biology and neurobiology. The species can also be rapidly

reared in large, genetically diverse populations in laboratory settings, providing unparalleled statistical power for experimental biology compared to any other metazoan [146]. Furthermore, metabolic and developmental pathways in *C. elegans* are conserved in humans [39]. High-throughput imaging technologies can improve *C. elegans* studies by increasing experimental efficiency, scalability, and quality. Existing systems for automated image acquisition, such as the Molecular Devices ImageXpress platforms generate images of nematodes that can be analyzed with software like CellProfiler’s WormToolbox [147] to extract nematode phenotype information. This software uses probabilistic nematode models trained on user-selected animals to automate the segmentation of nematodes from the background of images in high-throughput. As a result, CellProfiler’s WormToolbox is able to measure hundreds of phenotypes related to animal shape, intensity, and texture. Implementing this software for large-scale imaging experiments can generate large quantities of data that requires additional analysis software for reliable and reproducible handling, processing, and visualization. CellProfiler Analyst was developed to offer tools for the analysis of image-based datasets, but this software is not integrated with modern statistical environments. We sought to design a resource that facilitates the exploration of CellProfiler data in the R environment [148], where this limitation can be eliminated. The R language provides extensive open-source statistical and data visualization tools that are well supported by the user community. In leveraging R, we are able to create a flexible tool that can be rapidly integrated with other statistical R packages to suit project-specific analysis needs.

We developed easyXpress, a software package for the R statistical programming language, to assist in the processing, analysis, and visualization of *C. elegans* data generated using CellProfiler. easyXpress provides tools for quality control, summarization, and visualization of image-based *C. elegans* phenotype data. Built to be complementary to CellProfiler, this package provides a streamlined workflow for the rapid quantitative analysis of high-throughput imaging datasets.

## 4.4 Materials and methods

### 4.4.1 Preparation of animals for imaging

Bleach-synchronized animals were fed *E. coli* HB101 bacteria suspended and allowed to develop at 20°C with continuous shaking. Animals in 96-well microtiter plates were titered to approximately 30 animals per well. Prior to imaging, animals were treated with sodium azide (50 mM in 1X M9) for 10 minutes to paralyze and straighten their bodies.

### 4.4.2 Imaging

Animals in microtiter plates were imaged at 2X magnification with an ImageXpress Nano (Molecular Devices, San Jose, CA). The ImageXpress Nano acquires brightfield images with a 4.7 megaPixel CMOS camera and are stored in 16-bit TIFF format. The images were processed using CellProfiler software (for details see <https://github.com/AndersenLab/CellProfiler>).

### 4.4.3 Paraquat dose response

A 1.5 M solution of paraquat (Methyl viologen dichloride, Sigma, 856177-1G) was prepared in sterile water, aliquoted, and frozen at -20°C until used. Experimental animals were grown at 20°C and fed OP50 bacteria spotted on modified nematode growth medium, containing 1% agar and 0.7% agarose to prevent animals from burrowing. After three generations of passaging, animals were bleach-synchronized and embryos were transferred to the wells of 96-well microplates. Each well contained approximately 30 embryos in 50  $\mu$ L of K medium [149]. Microplates were incubated overnight at 20°C with continuous shaking. The following day, arrested L1 animals were fed HB101 bacteria suspended in K medium. At the time of feeding, the animals were also exposed to paraquat at one of six concentrations (0, 7.81, 31.25, 125, 500, 2000  $\mu$ M) by serial dilution of a freshly thawed aliquot of 1.5 M paraquat

solution. The final volume in each well after dosing and feeding was  $75 \mu\text{L}$ . The animals were then grown for 48 hours at  $20^\circ\text{C}$  with continuous shaking, afterwards the microplates were imaged to assess the effects of paraquat exposure on nematode development.

#### 4.4.4 Naming conventions

Several functions in the easyXpress package require specific naming conventions to work properly. For full details regarding essential file naming and directory structure see the package repository (<https://github.com/AndersenLab/easyXpress>). Importantly, when using the Metadata module in CellProfiler to extract information describing your images, specific column names are suggested.

Suggested naming conventions for CellProfiler metadata.						
Image_FileName_RawBF	Image_PathName_RawBF	Metadata_Date	Metadata_Experiment	Metadata_Plate	Metadata_Magnification	Metadata_Well
20191119-growth-p05-m2X_C03.TIF	/CellProfiler/example/raw_images	20191119	growth	p05	m2x	C03
20191119-growth-p06-m2X_C09.TIF	/CellProfiler/example/raw_images	20191119	growth	p06	m2x	C09
20191119-growth-p09-m2X_C06.TIF	/CellProfiler/example/raw_images	20191119	growth	p09	m2x	C06

Showing 1 to 3 of 3 entries

Previous  Next

The naming of “Metadata\_Plate” and “Metadata\_Well” are essential to the *setflags()*, *viewPlate()*, *viewWell()*, and *viewDose()* functions. Additionally, “Image\_fileName\_RawBF” and “Image\_PathName\_RawBF” are necessary for the proper function of *viewDose()*.

#### 4.4.5 Data availability

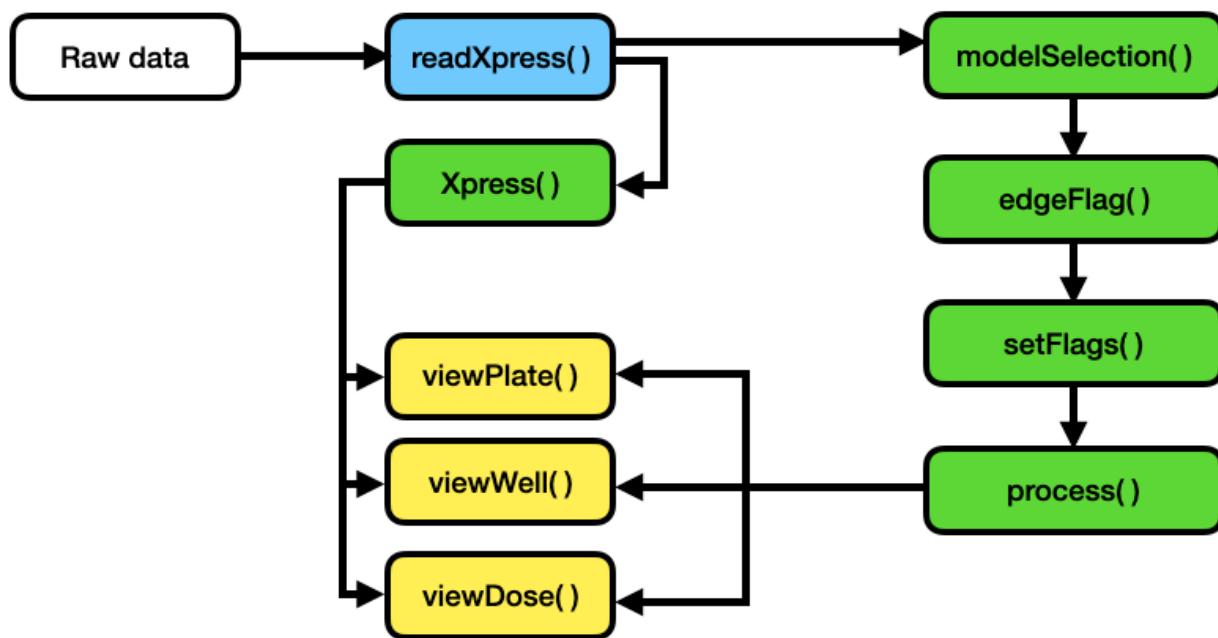
The entirety of the easyXpress package is written in the R language and is free to install across any system supporting R, including Linux, MacOSX, and Windows. The complete

source code, example data, extensive documentation, and installation details are available on GitHub. A tutorial on the usage of easyXpress and the available functions, can be found at <https://rpubs.com/jnyaanga/765641>. This package is open-source; for updates and to submit comments, visit <https://github.com/AndersenLab/easyXpress>.

## 4.5 Results

### 4.5.1 Design and implementation

The easyXpress package is designed to be simple and accessible to users familiar with the R environment. The easyXpress package comprises nine functions for reading, processing, and visualizing large high-throughput image-based datasets acquired from microplate-based assays processed with CellProfiler (**Figure 4-1**). Because our software is built to handle CellProfiler data as input, we suggest users review the overview and applications of CellProfiler as a prerequisite description of data generation [147]. Below, we describe the workflow for users to analyze their image data with easyXpress.



**Figure 4-1. easyXpress workflow.** The suggested workflow for using the easyXpress package starts with raw data generated from CellProfiler. For more information on implementing CellProfiler to generate data, see <https://github.com/AndersenLab/CellProfiler>. Reading steps are shown in blue, processing steps are shown in green, and visualization steps are shown in yellow.

#### 4.5.2 Data import and model assignment

To read in CellProfiler data files, we provide `readXpress()`. Measurements calculated by CellProfiler can be exported in a comma-separated value (csv) file and accessed using `readXpress()`. For large-scale, high-throughput experiments, users can employ a computing cluster for increased analysis speed (<https://github.com/AndersenLab/CellProfiler>). In this case, CellProfiler data stored in .RData format is accessed using `readXpress()`. Additionally, the function can optionally import a design file created by the user containing experimental treatments and conditions. This design file is joined to the CellProfiler data and output as a single dataframe.

CellProfiler’s WormToolbox detects and measures the phenotypes of individual animals based on user-calibrated models of variability in body size and shape [147]. To effectively detect animals in a mixed-stage population, multiple worm models must be used. However, using multiple worm models creates a one-to-many relationship between real animals and their measured phenotype (S4-1 Fig). We have included the function `modelSelection()` to annotate this information for downstream analysis. In instances where multiple worm model objects are assigned to a single primary object, `modelSelection()` will identify the best fitting model. Models are first ranked by frequency in the dataset such that the smallest model is classified as the most frequently occurring and the largest model is the least frequently occurring. In our experience, the most frequently occurring model in the dataset has the smallest size because it is often repeatedly assigned to a single primary object. Conversely, the least frequently occurring model in the dataset has the largest size as it is able to define the entire

size of an animal, and is thus assigned to a primary object only once. The largest ranked model occurring within a single primary object is then selected as the best fitting model for that primary object. If necessary, *modelSelection()* will also specify whether the selected model object was repeatedly assigned to the same primary object and flag this event as a cluster. This problem occurs in instances where a model object is repeatedly assigned to a single primary object. If the largest model object is found to occur repeatedly in a single primary object, this model will be selected and a cluster flag will be added (S4-1C Fig). The *modelSelection()* step is essential to resolve cases where multiple instances of a selected model object are assigned to a single primary object, thus contributing to inaccurate phenotype measurements.

#### 4.5.3 Data pruning and summarization

Once the data are read into the R statistical environment, it is crucial to optimize data quality before in-depth analysis. Uneven well illumination can hinder the performance of CellProfiler’s object identification and phenotype extraction. Despite correcting for uneven illumination within a well, discerning foreground objects from background can be especially challenging near the periphery of the well and can add noise to nematode phenotype data (S4-2 Fig). The function *edgeFlag()* was written to identify and flag animals located near the edge of circular wells using the centroid coordinates of the selected model object. By default, the function sets the radius of even illumination from the image center to 825 pixels, but this parameter can be adjusted by the user to serve project-specific analyses.

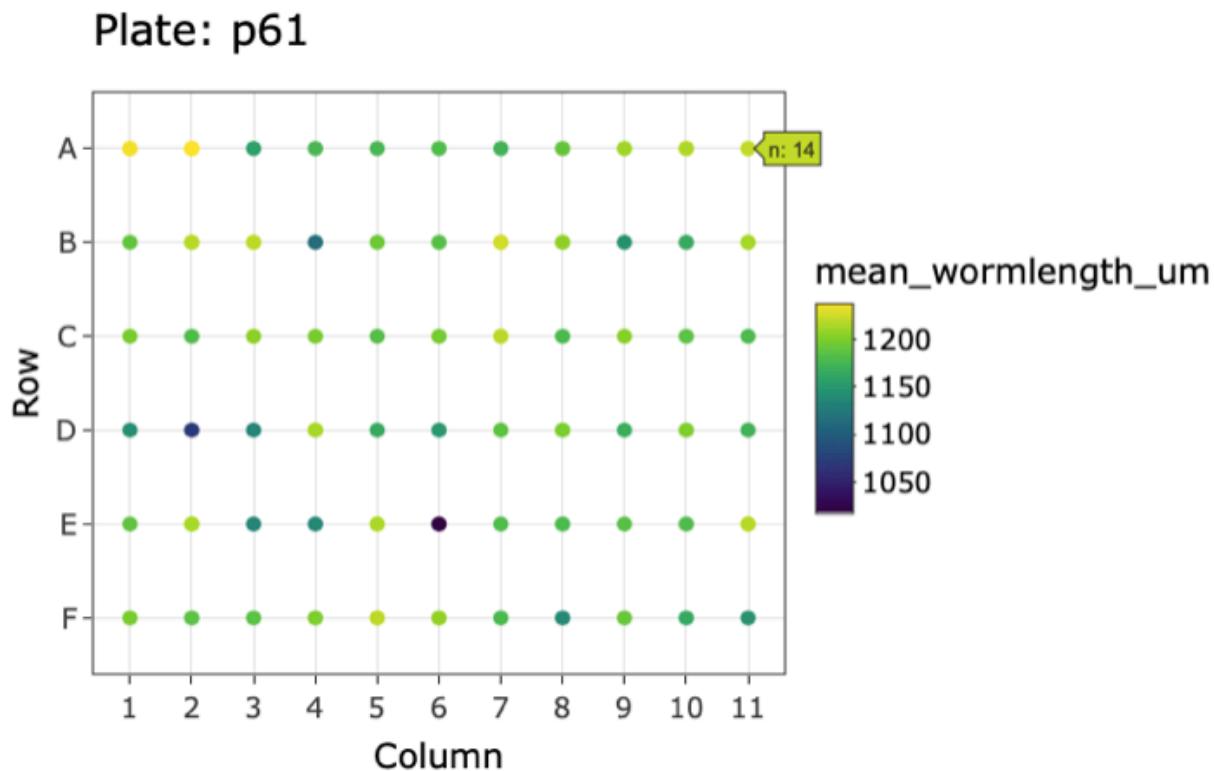
We also developed *setFlags()* in conjunction with *edgeFlag()* to further address data points that deviate from the expected animal measurements. The function *setFlags()* takes the output of *edgeFlags()* and detects outlier measurements among all measurements within a well using Tukey’s fences [150]. By default, outlier calculations are performed by excluding data identified by *modelSelection()* as part of a cluster as well as data in close proximity to the well edge. However, *setFlags()* is customizable, allowing the user to specify which

filters to include. `edgeFlag()` and `setFlags()` were designed to allow for analysis-specific optimization when handling various experimental datasets. This flexibility in data filtering makes easyXpress extensible to many unique projects.

Once data are adequately flagged, the function `process()` organizes the data into a list containing four elements: raw data, processed data, and summaries for both datasets. The raw data element is the CellProfiler data following `modelSelection()` and flag annotation. The processed data are generated by default after subsequent removal of all cluster, edge, and outlier flags. If a user includes data annotated as clusters or edge cases in `setFlags()`, cluster and edge cases will be retained in the processed data output. Finally, it is often useful to summarize data by well to interpret patterns specific to experimental variables. Alternatively, measurements may be summarized by other experimental factors according to the individual experimenter's plate design. `process()` aids in the summarization of both the raw and processed data elements. This function comprehensively calculates the means, variances, quantiles, minimum, and maximum values of animal length for any experimental unit (e.g. well). We have also included the wrapper function `Xpress()` to accelerate the import and processing of CellProfiler data. `Xpress()` will perform the above functions with all default settings, but a user can alter input arguments to better suit project specific needs.

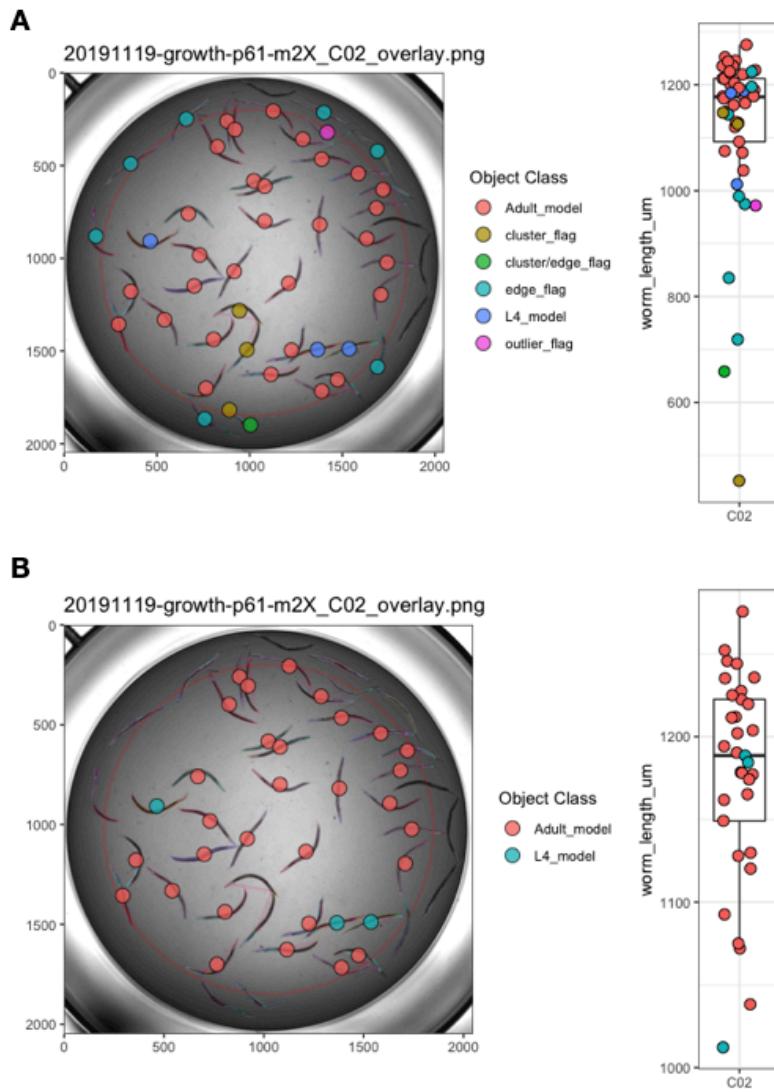
#### 4.5.4 Visualization

The easyXpress package provides several plotting functions to allow users to explore the data through detailed and elegant visualizations. After data summarization, it is often useful to inspect the values of the summary statistics in order to recognize patterns or identify potential outlier data. We provide `viewPlate()` to assist with the visualization of mean animal length within each well across a microtiter plate (**Figure 4-2**). This function accepts either raw or processed data to generate an interactive plot that allows users to scan across a plate to determine the number of objects present within individual wells.



**Figure 4-2.** Example plot generated by *viewPlate()*. Well-wise plot of mean animal length ( $\mu\text{m}$ ) from the summarized processed data. Interactive feature enables the assessment of the number of animals per well.

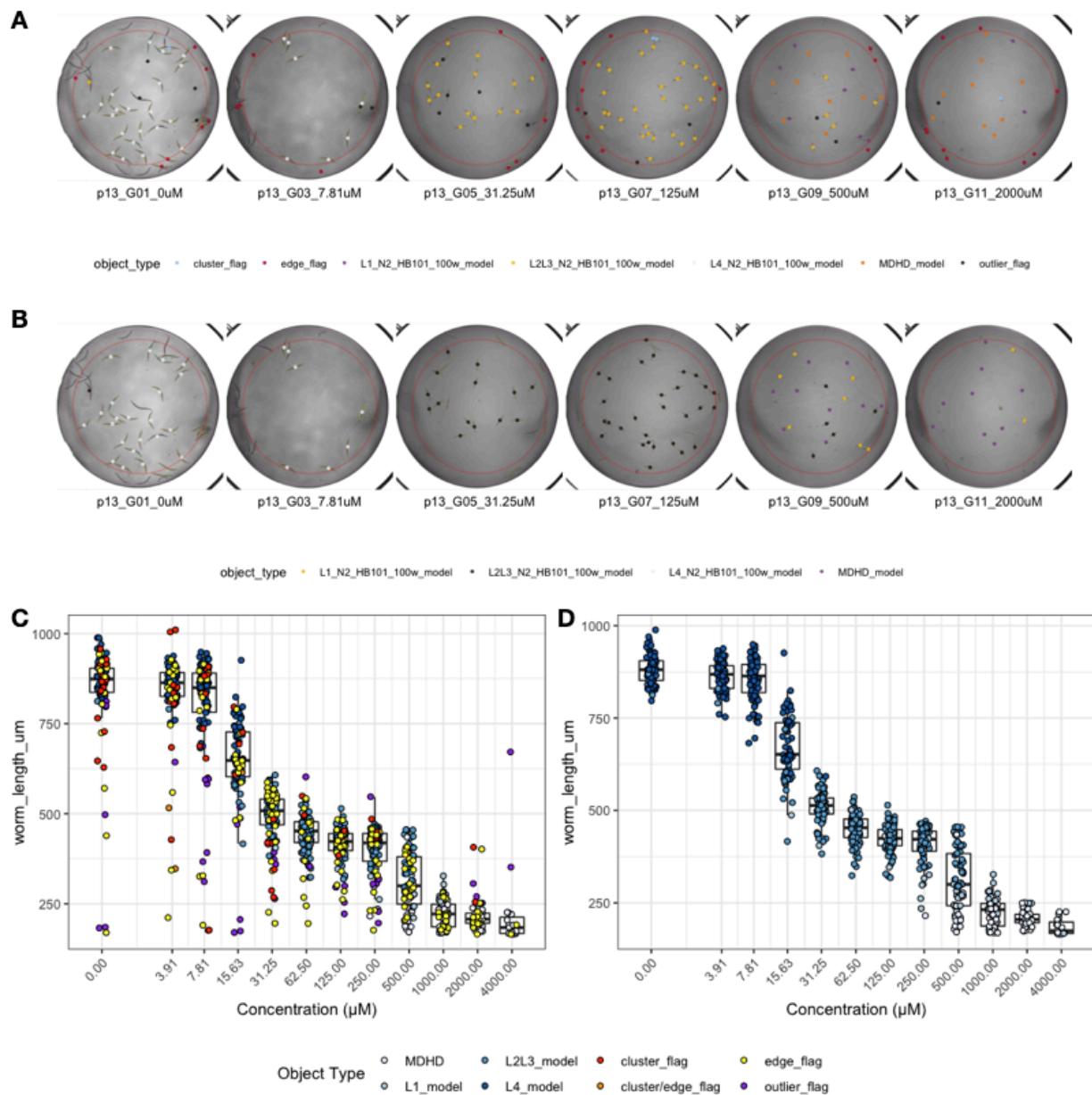
To complement the top-level data visualization provided by *viewPlate()*, we have included *viewWell()* to allow users to deeply explore data within individual wells. This function generates a plot of the well image following CellProfiler analysis with all objects annotated with their assigned class (**Figure 4-3**). Additionally, *viewWell()* can optionally generate a boxplot of the length values for each object. This plotting function is especially useful because it enables rapid qualitative assessment of object classification performance. By overlaying the model object classifications on the well image, users can quickly determine whether CellProfiler classified objects as expected or whether errors in model selection or data flags occurred.



**Figure 4-3. Example plots generated by *viewWell()*.** The function *viewWell()* facilitates the exploration of data within an individual well. Well images displaying easyXpress raw (A) and processed (B) data are annotated with the location of each model object centroid (circles) and are colored by object class in the legend (left). Animals are outlined in different colors to indicate the model object(s) identified for each primary object (see S4-1 Fig). The length of each object is displayed as a boxplot (right). Well edge circumference defined by the function *edgeFlag()* is shown in red.

Lastly, we have developed the function *viewDose()* to allow for the visualization of dose

response data. *C. elegans* are often used to study conserved responses to various compounds [80,82,110,151,152]. *viewDose()* allows a user to visually examine the effect of a compound on animal size and shape over a range of concentrations (**Figure 4-4**). By specifying the strain and compound of interest, a plot of representative wells will be generated that includes labels for each identified object.

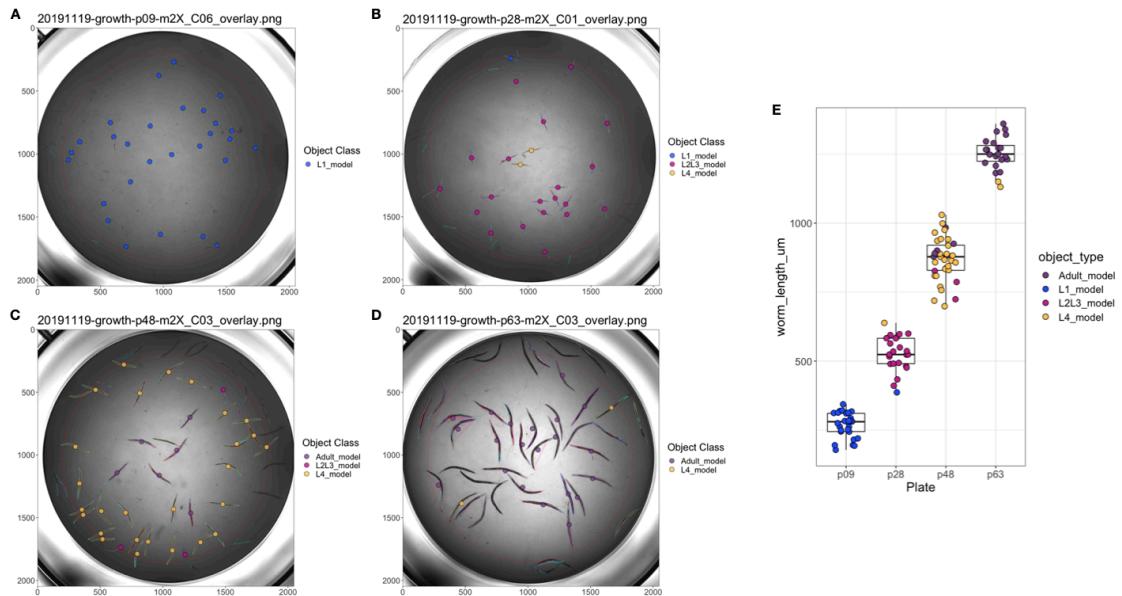


**Figure 4-4.** Example plots generated by *viewDose()*. The function *viewDose()* plots

representative raw (A) or processed (B) well images with objects annotated by model class for each dose of a selected drug and strain. The length measurements of raw (C) and processed (D) are also shown.

#### 4.5.5 Application to *C. elegans* growth data

We evaluated easyXpress using data collected from a *C. elegans* growth experiment [84]. Animals were imaged throughout the entire life cycle, beginning at the first larval (L1) stage and continuing until adulthood. Images were then processed with CellProfiler's WormToolbox and analyzed using easyXpress. During the implementation of easyXpress, four unique worm models representing *C. elegans* life stages were calibrated and applied: L1, L2/L3, L4, and Adult. These worm models do not designate stage assignments but rather represent the approximate sizes of animals that fall within the respective age groups (S4-1 Fig). The function *modelSelection()* assigned the appropriate model object to animals at each life stage, *edgeFlag()* and *setFlags()* identified outlier data points, and *viewWell()* provided clear visualizations of both the processed (**Figure 4-5**) and raw (S4-3 Fig) data.



**Figure 4-5. easyXpress applied to *C. elegans* growth data.** A subset of well images

acquired during *C. elegans* development displaying easyXpress processed data are shown here. Images taken at (A) 9 hours indicating the L1 stage, (B) 28 hours indicating the L2/L3 stage, (C) 46 hours indicating the L4 stage, and (D) 63 hours indicating the adult stage were analyzed with CellProfiler using four worm models. The easyXpress workflow was then used to process and visualize the data. The length of each object identified after processing is shown in (E).

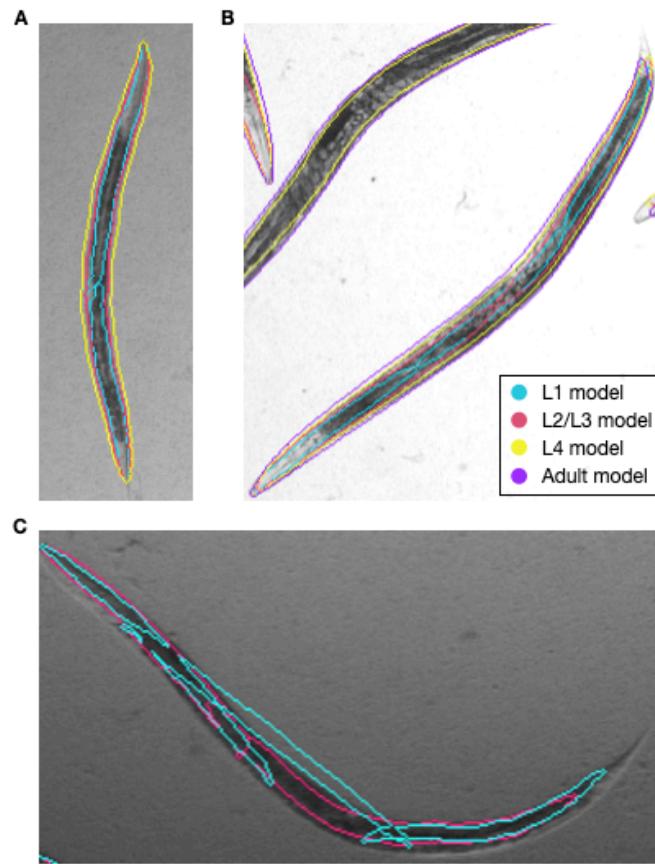
## 4.6 Conclusions

The easyXpress package presents an organized workflow for managing *C. elegans* phenotype data generated using CellProfiler. This package provides tools for the reading, processing, and visualization of these data in a simple and efficient way. By leveraging existing R infrastructure, easyXpress enables reproducible analysis, integration with other statistical R packages, and extensibility to many research projects using an open-source analysis pipeline.

## 4.7 Contributions

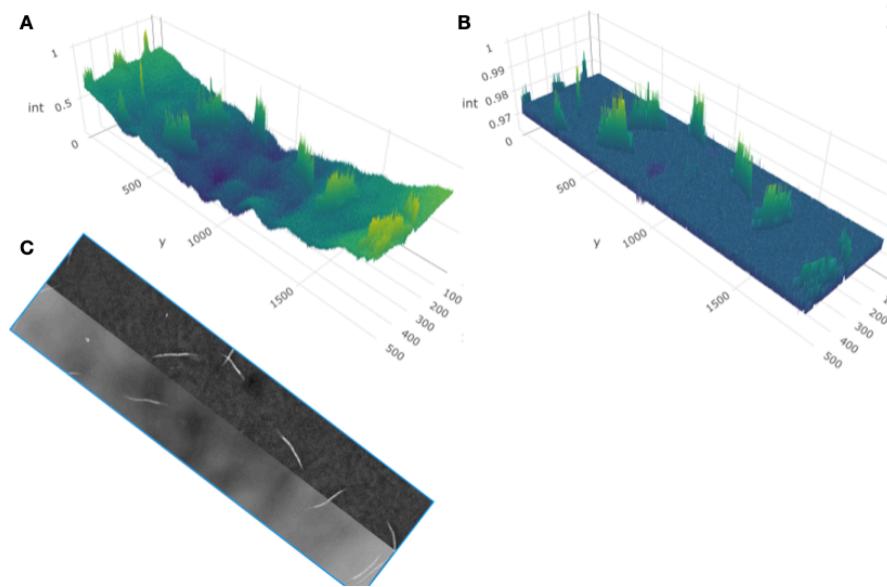
This work was supported by the NSF-Simons Center for Quantitative Biology at Northwestern University. This package was built by Joy Nyaanga and Drs. Timothy Crombie and Sam Widmayer. Erik Andersen supervised the construction and development of the package.

## 4.8 Supplement

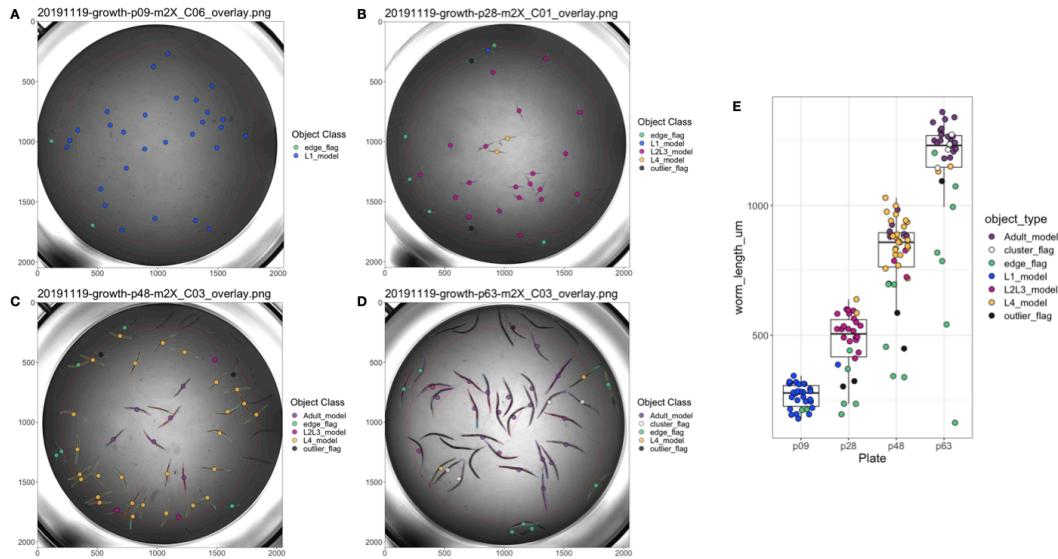


**S1 Fig. Multiple model objects assigned to a single primary object.** When running CellProfiler’s WormToolbox with multiple worm models, multiple model objects can be assigned to a single primary object (real animal). Different colors are used to outline each worm model object. Here four unique models were used: L1, L2/L3, L4, and Adult. These worm models represent the approximate sizes of animals at each life stage. For example, some mutant or diverse wild genetic backgrounds might have differently sized adult animals as compared to the laboratory-adapted N2 strain. We have included this “soft matching” to account for small differences in the sizes of developmental stages across different genetic backgrounds, laboratories, and environmental conditions. (A) An animal detected by CellProfiler as a primary object has been assigned three unique worm models: two L1 model objects, one L2/L3 model object, and one L4 model object. *modelSelection()* classifies this

animal as an L4 model object. (B) An animal detected as a primary object has been assigned four unique worm models: three L1 model objects, two L2/L3 model objects, one L4 model object, and one Adult model object. Here, *modelSelection()* identifies the Adult model as the best fitting model object. (C) An animal detected as a primary object has been assigned two unique worm models: three L1 model objects, and two L2/L3 model objects. In this case, *modelSelection()* classifies this animal as an L2/L3 model object and adds a cluster flag annotation to indicate the repeated assignment of the selected model object to the primary object.



**S4-2 Fig. Uneven illumination along well edge hinders CellProfiler's ability to segment animals from background.** (A) Left is raw intensity values across well. (B) Right is with background correction. Intensities of object illumination are displayed on each z-axis. Objects near the edge of the well ( $y < 500$  and  $y > 1500$ ) have similar raw detected intensities (int) to more medial objects ( $y \sim 1000$ ) in (A) but lower corrected intensities in (B) because of uneven background correction. Raw and background-corrected image segments are displayed in (C). Notice animals on the edges of the well do not stand out from the background as much as animals in the center of the well and therefore are more challenging to discern.



**S4-3 Fig.** Raw data from *C. elegans* growth experiment displayed by the function `viewWell()`. Similar to Fig 5, well images taken at (A) 9 hours indicating the L1 stage, (B) 28 hours indicating the L2/L3 stage, (C) 46 hours indicating the L4 stage, and (D) 63 hours were analyzed. Here, the raw data results are displayed. The length of each identified object identified is shown in (E).

# Chapter 5

## Discussion

# Chapter 6

## References

1. Johannsen W. The genotype conception of heredity. *Am Nat.* 1911;45: 129–159. doi:10.1086/279202
2. Nathans J, Piantanida TP, Eddy RL, Shows TB, Hogness DS. Molecular genetics of inherited variation in human color vision. *Science.* 1986;232: 203–210. doi:10.1126/science.3485310
3. Lynch M, Walsh B, Others. Genetics and analysis of quantitative traits. Sinauer Sunderland, MA; 1998.
4. Falconer DS, Others. Introduction to quantitative genetics. Introduction to quantitative genetics. 1960.
5. Griffiths AJF, Miller JH, Suzuki DT, Lewontin RC, Gelbart WM. Genetic variation. W. H. Freeman; 2000.
6. Davis MB, Shaw RG, Etterson JR. Evolutionary responses to changing climate. *Ecology.* 2005;86: 1704–1714. doi:10.1890/03-0788
7. Gutteling EW, Riksen JAG, Bakker J, Kammenga JE. Mapping phenotypic plasticity and genotype-environment interactions affecting life-history traits in *caenorhabditis elegans*. *Heredity.* 2007;98: 28–37. doi:10.1038/sj.hdy.6800894

8. Félix M-A, Barkoulas M. Pervasive robustness in biological systems. *Nat Rev Genet.* 2015;16: 483–496. doi:10.1038/nrg3949
9. Kiskowski M, Glimm T, Moreno N, Gamble T, Chiari Y. Isolating and quantifying the role of developmental noise in generating phenotypic variation. *PLoS Comput Biol.* 2019;15: e1006943. doi:10.1371/journal.pcbi.1006943
10. Mendel G. Experiments in plant hybridization. Oliver & Boyd; 1965.
11. 1000 Genomes Project Consortium, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, et al. A map of human genome variation from population-scale sequencing. *Nature.* 2010;467: 1061–1073. doi:10.1038/nature09534
12. Telenti A, Pierce LCT, Biggs WH, Iulio J di, Wong EHM, Fabani MM, et al. Deep sequencing of 10,000 human genomes. *Proc Natl Acad Sci U S A.* 2016;113: 11901–11906. doi:10.1073/pnas.1613365113
13. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 2016;536: 285–291. doi:10.1038/nature19057
14. Cooper DN, Krawczak M, Polychronakos C, Tyler-Smith C, Kehrer-Sawatzki H. Where genotype is not predictive of phenotype: Towards an understanding of the molecular basis of reduced penetrance in human inherited disease. *Hum Genet.* 2013;132: 1077–1130. doi:10.1007/s00439-013-1331-2
15. Antonarakis SE, Beckmann JS. Mendelian disorders deserve more attention. *Nat Rev Genet.* 2006;7: 277–282. doi:10.1038/nrg1826
16. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. *Science.* 2008;322: 881–888. doi:10.1126/science.1156409
17. Donnelly P. Progress and challenges in genome-wide association studies in humans. *Nature.* 2008;456: 728–731. doi:10.1038/nature07631

18. Frazer KA, Murray SS, Schork NJ, Topol EJ. Human genetic variation and its contribution to complex traits. *Nat Rev Genet.* 2009;10: 241–251. doi:10.1038/nrg2554
19. Mackay TFC, Stone EA, Ayroles JF. The genetics of quantitative traits: Challenges and prospects. *Nat Rev Genet.* 2009;10: 565–577. doi:10.1038/nrg2612
20. Evans KS, Wijk MH van, McGrath PT, Andersen EC, Sterken MG. From QTL to gene: *C. Elegans* facilitates discoveries of the genetic mechanisms underlying natural variation. *Trends Genet.* 2021;37: 933–947. doi:10.1016/j.tig.2021.06.005
21. Doroszuk A, Snoek LB, Fradin E, Riksen J, Kammenga J. A genome-wide library of CB4856/N2 introgression lines of *caenorhabditis elegans*. *Nucleic Acids Res.* 2009;37: e110. doi:10.1093/nar/gkp528
22. Wong GT. Speed congenics: Applications for transgenic and knock-out mouse strains. *Neuropeptides.* 2002;36: 230–236. doi:10.1054/npep.2002.0905
23. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science.* 2007;315: 1709–1712. doi:10.1126/science.1138140
24. Paix A, Wang Y, Smith HE, Lee C-Y, Calidas D, Lu T, et al. Scalable and versatile genome editing using linear DNAs with microhomology to Cas9 sites in *caenorhabditis elegans*. *Genetics.* 2014;198: 1347–1356. doi:10.1534/genetics.114.170423
25. Doudna JA, Charpentier E. The new frontier of genome engineering with CRISPR-Cas9. *Science.* 2014;346: 1258096. doi:10.1126/science.1258096
26. Maupas E. Modes et formes de reproduction des nematodes. 1900.
27. Nigon V. Le déterminisme du sexe chez un nématode libre hermaphrodite, *rhabditis elegans* maupas. *C R Seances Soc Biol Fil.* 1943;137: 40–41.
28. Corsi AK, Wightman B, Chalfie M. A transparent window into biology: A primer on *caenorhabditis elegans*. *Genetics.* 2015;200: 387–407. doi:10.1534/genetics.115.176099

29. All registered worm labs. <http://wbg.wormbook.org/the-worm-lab-project/all-labs/>; 2016.
30. Wood WB. The nematode *caenorhabditis elegans*. Cold Spring Harbor Laboratory; 1988.
31. Riddle DL, Blumenthal T, Meyer BJ, Priess JR, editors. *C. Elegans II*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2011.
32. Brenner S. The genetics of *caenorhabditis elegans*. *Genetics*. 1974;77: 71–94.
33. Sulston JE, Horvitz HR. Post-embryonic cell lineages of the nematode, *caenorhabditis elegans*. *Dev Biol*. 1977;56: 110–156. doi:10.1016/0012-1606(77)90158-0
34. Sulston JE, Schierenberg E, White JG, Thomson JN. The embryonic cell lineage of the nematode *caenorhabditis elegans*. *Dev Biol*. 1983;100: 64–119. doi:10.1016/0012-1606(83)90201-4
35. Deppe U, Schierenberg E, Cole T, Krieg C, Schmitt D, Yoder B, et al. Cell lineages of the embryo of the nematode *caenorhabditis elegans*. *Proceedings of the National Academy of Sciences*. 1978;75: 376–380.
36. Kimble J, Hirsh D. The postembryonic cell lineages of the hermaphrodite and male gonads in *caenorhabditis elegans*. *Dev Biol*. 1979;70: 396–417. doi:10.1016/0012-1606(79)90035-6
37. *C. elegans* Sequencing Consortium. Genome sequence of the nematode *c. elegans*: A platform for investigating biology. *Science*. 1998;282: 2012–2018. doi:10.1126/science.282.5396.2012
38. Kaletta T, Hengartner MO. Finding function in novel targets: *C. Elegans* as a model organism. *Nat Rev Drug Discov*. 2006;5: 387–398. doi:10.1038/nrd2031

39. Kim W, Underwood RS, Greenwald I, Shaye DD. OrthoList 2: A new comparative genomic analysis of human and *caenorhabditis elegans* genes. *Genetics*. 2018;210: 445–461. doi:10.1534/genetics.118.301307
40. Singh RN, Sulston JE. Some observations on moulting in *caenorhabditis elegans*. *Nematologica*. 1978;24: 63–71. doi:10.1163/187529278X00074
41. Monsalve GC, Frand AR. Toward a unified model of developmental timing: A “moult-ing” approach. *Worm*. 2012.
42. Uppaluri S, Brangwynne CP. A size threshold governs *caenorhabditis elegans* developmental progression. *Proc Biol Sci*. 2015;282: 20151283. doi:10.1098/rspb.2015.1283
43. Cassada RC, Russell RL. The dauerlarva, a post-embryonic developmental variant of the nematode *caenorhabditis elegans*. *Dev Biol*. 1975;46: 326–342. doi:10.1016/0012-1606(75)90109-8
44. Hu PJ. Dauer. *WormBook*; 2018.
45. Schindler AJ, Baugh LR, Sherwood DR. Identification of late larval stage developmental checkpoints in *caenorhabditis elegans* regulated by insulin/IGF and steroid hormone signaling pathways. *PLoS Genet*. 2014;10: e1004426. doi:10.1371/journal.pgen.1004426
46. Baugh LR, Sternberg PW. DAF-16/FOXO regulates transcription of cki-1/Cip/Kip and repression of lin-4 during c. Elegans L1 arrest. *Curr Biol*. 2006;16: 780–785. doi:10.1016/j.cub.2006.03.021
47. Monsalve GC, Van Buskirk C, Frand AR. LIN-42/PERIOD controls cyclical and developmental progression of c. Elegans molts. *Curr Biol*. 2011;21: 2033–2045. doi:10.1016/j.cub.2011.10.054
48. Zaidel-Bar R, Miller S, Kaminsky R, Broday L. Molting-specific downregulation of c. Elegans body-wall muscle attachment sites: The role of RNF-5 E3 ligase. *Biochem Biophys Res Commun*. 2010;395: 509–514. doi:10.1016/j.bbrc.2010.04.049

49. Suzuki Y, Yandell MD, Roy PJ, Krishna S, Savage-Dunn C, Ross RM, et al. A BMP homolog acts as a dose-dependent regulator of body size and male tail patterning in *caenorhabditis elegans*. *Development*. 1999;126: 241–250.
50. Morita K, Chow KL, Ueno N. Regulation of body length and male tail ray pattern formation of *caenorhabditis elegans* by a member of TGF-beta family. *Development*. 1999;126: 1337–1347.
51. Savage-Dunn C, Maduzia LL, Zimmerman CM, Roberts AF, Cohen S, Tokarz R, et al. Genetic screen for small body size mutants in *c. elegans* reveals many TGFbeta pathway components. *Genesis*. 2003;35: 239–247. doi:10.1002/gene.10184
52. Gumienny TL, Savage-Dunn C. TGF- $\beta$  signaling in *c. elegans*. WormBook; 2018.
53. Wu MY, Hill CS. Tgf-beta superfamily signaling in embryonic development and homeostasis. *Dev Cell*. 2009;16: 329–343. doi:10.1016/j.devcel.2009.02.012
54. Gumienny TL, MacNeil LT, Wang H, Bono M de, Wrana JL, Padgett RW. Glypican LON-2 is a conserved negative regulator of BMP-like signaling in *caenorhabditis elegans*. *Curr Biol*. 2007;17: 159–164. doi:10.1016/j.cub.2006.11.065
55. Baugh LR. To grow or not to grow: Nutritional control of development during *caenorhabditis elegans* L1 arrest. *Genetics*. 2013;194: 539–555. doi:10.1534/genetics.113.150847
56. MacNeil LT, Watson E, Arda HE, Zhu LJ, Walhout AJM. Diet-induced developmental acceleration independent of TOR and insulin in *c. elegans*. *Cell*. 2013;153: 240–252. doi:10.1016/j.cell.2013.02.049
57. So S, Miyahara K, Ohshima Y. Control of body size in *c. elegans* dependent on food and insulin/IGF-1 signal. *Genes Cells*. 2011;16: 639–651. doi:10.1111/j.1365-2443.2011.01514.x
58. Stuhr NL, Curran SP. Bacterial diets differentially alter lifespan and healthspan trajectories in *c. elegans*. *Commun Biol*. 2020;3: 653. doi:10.1038/s42003-020-01379-1

59. Kammenga JE, Doroszuk A, Riksen JAG, Hazendonk E, Spiridon L, Petrescu A-J, et al. A *caenorhabditis elegans* wild type defies the Temperature–Size rule owing to a single nucleotide polymorphism in *tra-3*. *PLoS Genet.* 2007;3: e34. doi:10.1371/journal.pgen.0030034
60. Rose JK, Sangha S, Rai S, Norman KR, Rankin CH. Decreased sensory stimulation reduces behavioral responding, retards development, and alters neuronal connectivity in *caenorhabditis elegans*. *J Neurosci.* 2005;25: 7159–7168. doi:10.1523/JNEUROSCI.1833-05.2005
61. Fujiwara M, Ishihara T, Katsura I. A novel WD40 protein, CHE-2, acts cell-autonomously in the formation of c. *Elegans* sensory cilia. *Development.* 1999;126: 4839–4848.
62. Fujiwara M, Sengupta P, McIntire SL. Regulation of body size and behavioral state of c. *Elegans* by sensory perception and the EGL-4 cGMP-dependent protein kinase. *Neuron.* 2002;36: 1091–1102. doi:10.1016/s0896-6273(02)01093-0
63. Page AP, Johnstone IL. The cuticle. *WormBook*; 2007.
64. McMahon L, Muriel JM, Roberts B, Quinn M, Johnstone IL. Two sets of interacting collagens form functionally distinct substructures within a *caenorhabditis elegans* extracellular matrix. *MBoC.* 2003;14: 1366–1378. doi:10.1091/mbc.e02-08-0479
65. Lints R, Hall DH. WormAtlas hermaphrodite handbook - the cuticle. *WormAtlas.* 2003. doi:10.3908/wormatlas.1.12
66. Andersen EC, Gerke JP, Shapiro JA, Crissman JR, Ghosh R, Bloom JS, et al. Chromosome-scale selective sweeps shape *caenorhabditis elegans* genomic diversity. *Nat Genet.* 2012;44: 285–290. doi:10.1038/ng.1050
67. Cook DE, Zdraljevic S, Roberts JP, Andersen EC. CeNDR, the *caenorhabditis elegans* natural diversity resource. *Nucleic Acids Res.* 2017;45: D650–D657. doi:10.1093/nar/gkw893

68. Nicholas WL, Dougherty EC, Hansen EL. AXENIC CULTIVATION OF CAENORHARDITIS BRIGGSAE (NEMATODA: RHABDITIDAE) WITH CHEMICALLY UNDEFINED SUPPLEMENTS; COMPARATIVE STUDIES WITH RELATED NEMATODES\*. Ann N Y Acad Sci. 2006;77: 218–236. doi:10.1111/j.1749-6632.1959.tb36902.x
69. Hodgkin J, Doniach T. Natural variation and copulatory plug formation in *caenorhabditis elegans*. Genetics. 1997;146: 149–164.
70. Thompson OA, Snoek LB, Nijveen H, Sterken MG, Volkers RJM, Brenchley R, et al. Remarkably divergent regions punctuate the genome assembly of the *caenorhabditis elegans* hawaiian strain CB4856. Genetics. 2015;200: 975–989. doi:10.1534/genetics.115.175950
71. Kim C, Kim J, Kim S, Cook DE, Evans KS, Andersen EC, et al. Long-read sequencing reveals intra-species tolerance of substantial structural variations and new subtelomere formation in *c. elegans*. Genome Res. 2019;29: 1023–1035. doi:10.1101/gr.246082.118
72. Capra EJ, Skrovanek SM, Kruglyak L. Comparative developmental expression profiling of two *c. Elegans* isolates. PLoS One. 2008;3: e4055. doi:10.1371/journal.pone.0004055
73. Rockman MV, Skrovanek SS, Kruglyak L. Selection at linked sites shapes heritable phenotypic variation in *c. elegans*. Science. 2010;330: 372–376. doi:10.1126/science.1194208
74. Viñuela A, Snoek LB, Riksen JAG, Kammenga JE. Aging uncouples heritability and Expression-QTL in *caenorhabditis elegans*. G3. 2012;2: 597–605. doi:10.1534/g3.112.002212
75. Bono M de, Bargmann CI. Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *c. elegans*. Cell. 1998;94: 679–689. doi:10.1016/s0092-8674(00)81609-8

76. Andersen EC, Bloom JS, Gerke JP, Kruglyak L. A variant in the neuropeptide receptor npr-1 is a major determinant of *caenorhabditis elegans* growth and physiology. *PLoS Genet.* 2014;10: e1004156. doi:10.1371/journal.pgen.1004156
77. Shimko TC, Andersen EC. COPASutils: An R package for reading, processing, and visualizing data from COPAS large-particle flow cytometers. *PLoS One.* 2014;9: e111090. doi:10.1371/journal.pone.0111090
78. Bendesky A, Bargmann CI. Genetic contributions to behavioural diversity at the gene–environment interface. *Nat Rev Genet.* 2011;12: 809–820. doi:10.1038/nrg3065
79. Brady SC, Zdraljevic S, Bisaga KW, Tanny RE, Cook DE, Lee D, et al. A novel gene underlies Bleomycin-Response variation in *caenorhabditis elegans*. *Genetics.* 2019;212: 1453–1468. doi:10.1534/genetics.119.302286
80. Evans KS, Andersen EC. The gene scb-1 underlies variation in *caenorhabditis elegans* chemotherapeutic responses. *G3 Genes|Genomes|Genetics.* 2020;10: 2353–2364. doi:10.1534/g3.120.401310
81. Hahnel SR, Zdraljevic S, Rodriguez BC, Zhao Y, McGrath PT, Andersen EC. Extreme allelic heterogeneity at a *caenorhabditis elegans* beta-tubulin locus explains natural resistance to benzimidazoles. *PLoS Pathog.* 2018;14: e1007226. doi:10.1371/journal.ppat.1007226
82. Zdraljevic S, Fox BW, Strand C, Panda O, Tenjo FJ, Brady SC, et al. Natural variation in *c. Elegans* arsenic toxicity is explained by differences in branched chain amino acid metabolism. 2019. doi:10.7554/eLife.40260
83. Zdraljevic S, Strand C, Seidel HS, Cook DE, Doench JG, Andersen EC. Natural variation in a single amino acid substitution underlies physiological responses to topoisomerase II poisons. *PLoS Genet.* 2017;13: e1006891. doi:10.1371/journal.pgen.1006891

84. Nyaanga J, Goss C, Zhang G, Ahmed HN, Andersen EJ, Miller IR, et al. Highly scaled measurements of *C. elegans* development suggest that physical constraints guide growth trajectories and animal shape. *bioRxiv*. 2021. p. 2021.04.01.438121. doi:10.1101/2021.04.01.438121
85. Hone DWE, Benton MJ. The evolution of large size: How does Cope's rule work? *Trends Ecol Evol*. 2005;20: 4–6. doi:10.1016/j.tree.2004.10.012
86. Björklund M. Cell size homeostasis: Metabolic control of growth and cell division. *Biochim Biophys Acta Mol Cell Res*. 2019;1866: 409–417. doi:10.1016/j.bbamcr.2018.10.002
87. Willis L, Huang KC. Sizing up the bacterial cell cycle. *Nat Rev Microbiol*. 2017;15: 606–620. doi:10.1038/nrmicro.2017.79
88. Turner JJ, Ewald JC, Skotheim JM. Cell size control in yeast. *Curr Biol*. 2012;22: R350–9. doi:10.1016/j.cub.2012.02.041
89. Donnan L, John PC. Cell cycle control by timer and sizer in chlamydomonas. *Nature*. 1983;304: 630–633. doi:10.1038/304630a0
90. Wang P, Hayden S, Masui Y. Transition of the blastomere cell cycle from cell size-independent to size-dependent control at the midblastula stage in *Xenopus laevis*. *J Exp Zool*. 2000;287: 128–144. doi:10.1002/1097-010x(20000701)287:2<128::aid-jez3>3.0.co;2-g
91. Tzur A, Kafri R, LeBleu VS, Lahav G, Kirschner MW. Cell growth and size homeostasis in proliferating animal cells. *Science*. 2009;325: 167–171. doi:10.1126/science.1174294
92. Sveiczer A, Novak B, Mitchison JM. The size control of fission yeast revisited. *J Cell Sci*. 1996;109 ( Pt 12): 2947–2957.
93. Campos M, Surovtsev IV, Kato S, Paintdakhi A, Beltran B, Ebmeier SE, et al. A constant size extension drives bacterial cell size homeostasis. *Cell*. 2014;159: 1433–1446. doi:10.1016/j.cell.2014.11.022

94. Taheri-Araghi S, Bradde S, Sauls JT, Hill NS, Levin PA, Paulsson J, et al. Cell-size control and homeostasis in bacteria. *Curr Biol.* 2015;25: 385–391. doi:10.1016/j.cub.2014.12.009
95. Towbin BD, Grosshans H. A folder mechanism ensures size uniformity among *C. Elegans* individuals by coupling growth and development. *bioRxiv*. 2021.
96. Soifer I, Robert L, Amir A. Single-Cell analysis of growth in budding yeast and bacteria reveals a common size regulation strategy. *Curr Biol.* 2016;26: 356–361. doi:10.1016/j.cub.2015.11.067
97. Osella M, Nugent E, Cosentino Lagomarsino M. Concerted control of *Escherichia coli* cell division. *Proc Natl Acad Sci U S A.* 2014;111: 3431–3435. doi:10.1073/pnas.1313715111
98. Jorgensen P, Tyers M. How cells coordinate growth and division. *Curr Biol.* 2004;14: R1014–27. doi:10.1016/j.cub.2004.11.027
99. Wang P, Robert L, Pelletier J, Dang WL, Taddei F, Wright A, et al. Robust growth of *Escherichia coli*. *Curr Biol.* 2010;20: 1099–1103. doi:10.1016/j.cub.2010.04.045
100. Cadart C, Monnier S, Grilli J, Sáez PJ, Srivastava N, Attia R, et al. Size control in mammalian cells involves modulation of both growth rate and cell cycle duration. *Nat Commun.* 2018;9: 3275. doi:10.1038/s41467-018-05393-0
101. Moss-Taylor L, Upadhyay A, Pan X, Kim M-J, O'Connor MB. Body size and Tissue-Scaling is regulated by Motoneuron-Derived activin $\beta$  in *Drosophila melanogaster*. *Genetics.* 2019;213: 1447–1464. doi:10.1534/genetics.119.302394
102. Spence AJ. Scaling in biology. *Curr Biol.* 2009;19: R57–61. doi:10.1016/j.cub.2008.10.042
103. Byerly L, Cassada RC, Russell RL. The life cycle of the nematode *Caenorhabditis elegans*. I. Wild-type growth and reproduction. *Dev Biol.* 1976;51: 23–33. doi:10.1016/0012-1606(76)90119-6

104. Knight CG, Patel MN, Azevedo RBR, Leroi AM. A novel mode of ecdysozoan growth in *caenorhabditis elegans*. *Evol Dev.* 2002;4: 16–27. doi:10.1046/j.1525-142x.2002.01058.x
105. Faerberg DF, Gurarie V, Ruvinsky I. Inferring temporal organization of postembryonic development from high-content behavioral tracking. *Dev Biol.* 2021;475: 54–64. doi:10.1016/j.ydbio.2021.02.007
106. Filina O, Haagmans R, Zon JS van. Temporal scaling in *c. Elegans* larval development. *bioRxiv.* 2020. p. 2020.09.21.306423. doi:10.1101/2020.09.21.306423
107. Cook DE, Zdraljevic S, Tanny RE, Seo B, Riccardi DD, Noble LM, et al. The genetic basis of natural variation in *caenorhabditis elegans* telomere length. *Genetics.* 2016;204: 371–383. doi:10.1534/genetics.116.191148
108. Stiernagle T. Maintenance of *c. elegans*. WormBook; 2006.
109. Abràmoff MD, Magalhães PJ, Ram SJ. Image processing with ImageJ. *Biophotonics international.* 2004;11: 36–42.
110. Andersen EC, Shimko TC, Crissman JR, Ghosh R, Bloom JS, Seidel HS, et al. A powerful new quantitative genetics platform, combining *caenorhabditis elegans* High-Throughput fitness assays with a large collection of recombinant strains. *G3.* 2015;5: 911–920. doi:10.1534/g3.115.017178
111. Scrucca L, Fop M, Murphy TB, Raftery AE. Mclust 5: Clustering, classification and density estimation using gaussian finite mixture models. *R J.* 2016;8: 289–317.
112. Smith MV, Boyd WA, Kissling GE, Rice JR, Snyder DW, Portier CJ, et al. A discrete time model for the analysis of medium-throughput *c. Elegans* growth data. *PLoS One.* 2009;4: e7018. doi:10.1371/journal.pone.0007018
113. Hermann E. Lokern: Kernel regression smoothing with local or global plug-in bandwidth. R package version 1.1-8. <https://CRAN.R-project.org/package=lokern>; 2016.

114. Sakamoto Y, Ishiguro M, Kitagawa G. Akaike information criterion statistics. Dordrecht, The Netherlands: D Reidel. 1986;81: 26853.
115. Schwarz G. Estimating the dimension of a model. *aos.* 1978;6: 461–464. doi:10.1214/aos/1176344136
116. Kass RE, Raftery AE. Bayes factors. *J Am Stat Assoc.* 1995;90: 773. doi:10.2307/2291091
117. Burnham KP, Anderson DR. Model selection and multimodel inference: A practical Information-Theoretic approach. Springer Science & Business Media; 2007.
118. Pulak R. Techniques for analysis, sorting, and dispensing of c. *Elegans* on the CO-PAS™ Flow-Sorting system. In: Strange K, editor. *C Elegans: Methods and applications*. Totowa, NJ: Humana Press; 2006. pp. 275–286. doi:10.1385/1-59745-151-7:275
119. Nika L, Gibson T, Konkus R, Karp X. Fluorescent beads are a versatile tool for staging *caenorhabditis elegans* in different life histories. *G3.* 2016;6: 1923–1933. doi:10.1534/g3.116.030163
120. Vuaridel-Thurre G, Vuaridel AR, Dhar N, McKinney JD. Computational analysis of the mutual constraints between Single-Cell growth and division control models. *Adv Biosyst.* 2020;4: e1900103. doi:10.1002/adbi.201900103
121. Holzapfel GA. Similarities between soft biological tissues and rubberlike materials. Constitutive Models for Rubber IV. 2017; 607–617. doi:10.1201/9781315140216-105
122. Essmann CL, Elmi M, Shaw M, Anand GM, Pawar VM, Srinivasan MA. In-vivo high resolution AFM topographic imaging of *caenorhabditis elegans* reveals previously unreported surface structures of cuticle mutants. *Nanomedicine.* 2017;13: 183–189. doi:10.1016/j.nano.2016.09.006
123. Dodd W, Tang L, Lone J-C, Wimberly K, Wu C-W, Consalvo C, et al. A damage sensor associated with the cuticle coordinates three core environmental stress responses in *caenorhabditis elegans*. *Genetics.* 2018;208: 1467–1482. doi:10.1534/genetics.118.300827

124. Yochem J, Lažetić V, Bell L, Chen L, Fay D. C. *Elegans* NIMA-related kinases NEKL-2 and NEKL-3 are required for the completion of molting. *Dev Biol.* 2015;398: 255–266. doi:10.1016/j.ydbio.2014.12.008
125. Petzold BC, Park S-J, Ponce P, Roozeboom C, Powell C, Goodman MB, et al. *Caenorhabditis elegans* body mechanics are regulated by body wall muscle tone. *Biophys J.* 2011;100: 1977–1985. doi:10.1016/j.bpj.2011.02.035
126. Cox GN, Staprants S, Edgar RS. The cuticle of *caenorhabditis elegans*. II. Stage-specific changes in ultrastructure and protein composition during postembryonic development. *Dev Biol.* 1981;86: 456–470. doi:10.1016/0012-1606(81)90204-9
127. Park S-J, Goodman MB, Pruitt BL. Analysis of nematode mechanics by piezoresistive displacement clamp. *Proc Natl Acad Sci U S A.* 2007;104: 17376–17381. doi:10.1073/pnas.0702138104
128. Gilpin W, Uppaluri S, Brangwynne CP. Worms under pressure: Bulk mechanical properties of *c. Elegans* are independent of the cuticle. *Biophysical Journal.* 2015. pp. 1887–1898. doi:10.1016/j.bpj.2015.03.020
129. Golden JW, Riddle DL. The *caenorhabditis elegans* dauer larva: Developmental effects of pheromone, food, and temperature. *Dev Biol.* 1984;102: 368–378. doi:10.1016/0012-1606(84)90201-x
130. Schiller HB, Fässler R. Mechanosensitivity and compositional dynamics of cell-matrix adhesions. *EMBO Rep.* 2013;14: 509–519. doi:10.1038/embor.2013.49
131. Wolfenson H, Bershadsky A, Henis YI, Geiger B. Actomyosin-generated tension controls the molecular kinetics of focal adhesions. *J Cell Sci.* 2011;124: 1425–1432. doi:10.1242/jcs.077388
132. Suman SK, Daday C, Ferraro T, Vuong-Breder T, Tak S, Quintin S, et al. The plakin domain of VAB-10/plectin acts as a hub in a mechanotransduction pathway to promote morphogenesis. *Development.* 2019;146. doi:10.1242/dev.183780

133. Zhang H, Landmann F, Zahreddine H, Rodriguez D, Koch M, Labouesse M. A tension-induced mechanotransduction pathway promotes epithelial morphogenesis. *Nature*. 2011;471: 99–103. doi:10.1038/nature09765
134. Moerman DG, Williams BD. Sarcomere assembly in *c. Elegans* muscle. *WormBook*. 2006; 1–16. doi:10.1895/wormbook.1.81.1
135. Costa M, Draper BW, Priess JR. The role of actin filaments in patterning the *caenorhabditis elegans* cuticle. *Dev Biol*. 1997;184: 373–384. doi:10.1006/dbio.1997.8530
136. Broday L, Hauser CA, Kolotuev I, Ronai Z. Muscle-epidermis interactions affect exoskeleton patterning in *caenorhabditis elegans*. *Dev Dyn*. 2007;236: 3129–3136. doi:10.1002/dvdy.21341
137. Madaan U, Faure L, Chowdhury A, Ahmed S, Ciccarelli EJ, Gumienny TL, et al. Feedback regulation of BMP signaling by cuticle collagens. *Mol Biol Cell*. 2020;31: 825–832. doi:10.1091/mbc.E19-07-0390
138. Katz SS, Maybrun C, Maul-Newby HM, Frand AR. Non-canonical apical constriction shapes emergent matrices in *c. elegans*. *bioRxiv*. 2018. p. 189951. doi:10.1101/189951
139. Fechner S, Loizeau F, Nekimken AL, Pruitt BL, Goodman MB. The bodies of *dpy-10(e128)* are twice as stiff as wild type. *MicroPubl Biol*. 2018;2018. doi:10.17912/ecsm-mp67
140. Horvitz HR, Sulston JE. Isolation and genetic characterization of cell-lineage mutants of the nematode *caenorhabditis elegans*. *Genetics*. 1980;96: 435–454.
141. Xie Y. *Bookdown: Authoring books and technical documents with R markdown*. CRC Press; 2016.
142. Nyaanga J, Crombie TA, Widmayer SJ, Andersen EC. *easyXpress: An R package to analyze and visualize high-throughput c. Elegans microscopy data generated using CellProfiler*. *PLoS One*. 2021;16: e0252000. doi:10.1371/journal.pone.0252000

143. Swedlow JR. Innovation in biological microscopy: Current status and future directions. *Bioessays*. 2012;34: 333–340. doi:10.1002/bies.201100168
144. Cassidy PJ, Radda GK. Molecular imaging perspectives. *J R Soc Interface*. 2005;2: 133–144. doi:10.1098/rsif.2005.0040
145. White JG, Southgate E, Thomson JN, Brenner S. The structure of the nervous system of the nematode *caenorhabditis elegans*. *Philos Trans R Soc Lond B Biol Sci*. 1986;314: 1–340. doi:10.1098/rstb.1986.0056
146. Teotónio H, Estes S, Phillips PC, Baer CF. Experimental evolution with nematodes. *Genetics*. 2017;206: 691–716. doi:10.1534/genetics.115.186288
147. Wählby C, Kamentsky L, Liu ZH, Riklin-Raviv T, Conery AL, O'Rourke EJ, et al. An image analysis toolbox for high-throughput *c. Elegans* assays. *Nat Methods*. 2012;9: 714–716. doi:10.1038/nmeth.1984
148. Team RC. R: A language and environment for statistical computing <http://www.R-project.org>. 2014.
149. Boyd WA, Smith MV, Freedman JH. *Caenorhabditis elegans* as a model in developmental toxicology. In: Harris C, Hansen JM, editors. *Developmental toxicology: Methods and protocols*. Totowa, NJ: Humana Press; 2012. pp. 15–24. doi:10.1007/978-1-61779-867-2\3
150. Tukey JW, Others. *Exploratory data analysis*. Reading, Mass.; 1977.
151. Dilks CM, Hahnel SR, Sheng Q, Long L, McGrath PT, Andersen EC. Quantitative benzimidazole resistance and fitness effects of parasitic nematode beta-tubulin alleles. *Int J Parasitol Drugs Drug Resist*. 2020;14: 28–36. doi:10.1016/j.ijpddr.2020.08.003
152. Wit J, Rodriguez BC, Andersen EC. Natural variation in *caenorhabditis elegans* responses to the anthelmintic emodepside. *Int J Parasitol Drugs Drug Resist*. 2021;16: 1–8. doi:10.1016/j.ijpddr.2021.04.001

## Appendix of R code

```
model <- lm(y ~ x1 + x2, data = df)
summary(model)
```