

UNIVERSITY OF CALGARY

Phenotypic variability in *C. elegans* natural isolates reveals plasticity of gene essentiality for
mat-1 and *cgh-1*

by

Victoria Rodrigues Alves Barbosa

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN BIOCHEMISTRY AND MOLECULAR BIOLOGY

CALGARY, ALBERTA

MARCH, 2024

© Victoria Rodrigues Alves Barbosa 2024

Abstract

Essential genes are deemed crucial for survival and/or reproductive success, hence, their loss of function leads to lethality. However, gene essentiality is not static, but dependent on multiple factors, including the genetic background. The genetic background affects gene essentiality through genetic modifiers, second-site variants capable of interacting with the primary variant and altering phenotype. Genetic modifiers can act as suppressors, alleviating the phenotype, or enhancers, exacerbating the phenotype. The plasticity (modification) of gene essentiality caused by genetic modifiers has been shown in multiple species. In *Caenorhabditis elegans*, recent evidence shows phenotypic variability for knock-down of essential genes in two natural isolates: N2 and CB4856. N2 represents the laboratory-cultivated strain, while CB4856 is one of the strains with most diverse genome in comparison to N2. Much has been explored for these two genetic backgrounds, but little is known about the plasticity of gene essentiality in other *C. elegans* wild isolates. Thus, I here explore the effect of the genetic background on the plasticity of two genes, known as essential in the N2 background, but potentially dispensable in CB4856: the Metaphase-to-Anaphase Transition Defect gene (*mat-1*), and the Conserved Germline Helicase (*cgh-1*). These are involved in cell division and posttranscriptional regulation, respectively. Here, I further investigate the plasticity of *mat-1* and *cgh-1* in N2 and CB4856 backgrounds, and also include four other wild isolate backgrounds from diverse geographical locations (GXW1, KR314, JU1400, and AB1).

Using hatch rate and propagation assays, I explored the phenotype of *mat-1* and *cgh-1* knockdowns across the six natural isolates using temperature-sensitive alleles and uncovered phenotypic variability for the embryonic lethal phenotype, suggestively due to influence of each genetic background. Next, bioinformatics and genomics tools were utilized for identification of

candidate genetic modifiers for *mat-1* and *cgh-1* and led to prioritization of eight extragenic variants. Out of these, none showed notable modifying activity when studied in isolation. Importantly, this work shows the challenges associated with identifying true genetic modifiers in genomes with substantial variation. Undeniably, understanding more about a gene and its phenotype under different conditions and genetic backgrounds may be fundamental for elucidating fixed and plastic genetic interactions.

Preface

This thesis is the original, unpublished, independent work by the author, Victoria Rodrigues Alves Barbosa.

Acknowledgements

Pursuing my Master of Science degree in an unfamiliar, foreign (and cold) country marked one of the hardest things I have done in life so far. Here's to everyone who alleviated it:

Foremost, I express my deepest appreciation to my loving parents, André Alves Barbosa and Neiva Luzia Rodrigues, for going above and beyond so I could live this dream. Their consistent and unconditional support has been the cornerstone of my journey. Additionally, I extend my heartfelt thanks to my dear friends—Paloma Fontes, Morgana Genn, and Felipe Kajimoto—whose presence in Calgary proved to be the most invaluable gifts. Your companionship was fundamental for getting me through the toughest moments, and for celebrating the best ones.

I wish to acknowledge the outstanding colleagues in the laboratory for their constant presence throughout this academic voyage. The collaborative and welcoming environment fostered by them greatly influenced my learning experience. My knowledge previous and post this degree is *significantly different* because of you ($p\text{-value} < 0.05$). Special gratitude is extended to Dr. Francesca Jean, for guiding me through my first steps, and to Dr. Xiao Li and Amanda Stuart, whose continuous support in the wet lab was instrumental over the past years.

I extend my appreciation to Afiya Chida for the exemplary partnership in our shared project and the significant collaboration that contributed to the success of this work. Additionally, Dr. Tatiana Maroilley deserves my gratitude for her unwavering support in research and for affording me the opportunity to mentor students under her guidance.

My sincere thanks are directed to the friends I made in the coffee club—Dr. Danielle Blackwell, Dr. Aleksandra Tsoлова, Dr. Shuhiba Mohammad, and Alexa Krala— for always sharing a laugh on the good days, and a warm cup of coffee on the challenging ones. I also express

my gratitude to Shreya Tomar, for being a friend in and out of the lab. The support given by all of you was essential for *suppressing* the challenges that research can bring.

To the current members of the lab—Sue Ferris, Rumika Mascarenhas, Catherine Diao, Kyle Nickel, Tahsin Rahit, Dr. Ivette Menendez, and Vladimir Avramovic—I am indebted for the invaluable insights provided throughout my research journey and for collaborating on my thesis defense preparation. This achievement would not have been possible without your collective contribution.

Lastly, but certainly not least, I extend my heartfelt thanks to Dr. Tarailo-Graovac for the warm welcome into the lab and for the unwavering support and guidance provided. I am forever grateful for the opportunity to pursue my master’s degree in the Tarailo-Graovac Lab, an experience that has enriched both my professional and personal perspectives.

As I conclude this academic chapter, I reflect with gratitude on the past and anticipate the future with excitement. Each of you had a *high impact* in shaping my career, and there is a *high probability score* that I will be missing you :)

Contents

Abstract.....	ii
Preface.....	iv
Acknowledgements.....	v
Contents	vii
List of Tables	xi
List of Figures	xii
List of Abbreviations	xiv
Chapter 1: Introduction.....	1
1.1 Gene Essentiality	1
1.1.1 Plasticity of Essentiality.....	5
1.1.2 Genetic Modifiers	8
1.1.2.1 Genetic Nature of Modifiers	11
1.1.2.2 Genetic Modifier Screens	12
1.1.2.2.1 EMS Mutagenesis Screen	13
1.1.2.2.2 RNA interference (RNAi).....	15
1.2 Contemporary Techniques for Studying Gene Essentiality.....	16
1.3 <i>C. elegans</i> as a Model for Plasticity of Essentiality.....	17
1.3.1 The Metaphase-to-Anaphase Transition Defect Gene – <i>mat-1</i>	20
1.3.2 The Conserved Germline Helicase Gene – <i>cgh-1</i>	23
1.3.3 Diverse Genetic Backgrounds for Investigation of Plasticity of Essentiality	24
1.3.3.1 The Reference Strain, N2.....	25
1.3.3.2 The Hawaiian Strain, CB4856	26
1.3.3.3 The Chinese Strain, GXW1	26
1.3.3.4 The Australian Strain, AB1.....	27

1.3.3.5 The Spanish Strain, JU1400.....	27
1.3.3.6 The Canadian Strain, KR314	27
1.4 Hypothesis and Aims	28
Chapter 2: Materials and Methods	29
2.1 <i>C. elegans</i> Culture and Maintenance	29
2.2 CRISPR/Cas9 Design and Microinjection.....	29
2.3 Hatch Rate Assays	32
2.4 Titration/Propagation Assay	32
2.5 RNAi Treatment.....	33
2.6 DNA Extraction and WGS Long-Reads Preparation.....	33
2.7 Structural Variants Validation	34
2.8 DNA Extraction and WGS Short-Reads Preparation	35
2.9 srWGS Analysis and Search for CRISPR Off-Target Activity	35
2.10 Variant Curation Using Machine Learning.....	36
2.11 RNA Extraction and Sequencing	37
2.12 RNA-Seq Data Analysis	38
2.13 Grantham Scores	39
Chapter 3: <i>mat-1</i> and <i>cgh-1</i> Display Phenotypic Variability In Different Natural Isolate	
Backgrounds	40
3.1 RNA interference for <i>mat-1</i> and <i>cgh-1</i>	40
3.2 The Use of Temperature-Sensitive Alleles for Phenotype Investigation.....	41
3.2.1 Knock-in of Temperature-Sensitive Alleles in the Six Natural Isolates.....	42
3.3 Analyzing the Phenotype of <i>mat-1</i> (<i>ye121</i>) and <i>cgh-1</i> (<i>tn691</i>) across different genetic	
backgrounds	43
3.3.1 Hatch Rate Assays	43

3.3.1.1 Hatch rates at restrictive temperatures reveal phenotypic variability for <i>mat-1</i> (<i>ye121</i>)	44
3.3.1.2 Hatch rates at restrictive temperatures reveal phenotypic variability for <i>cgh-1</i> (<i>tn691</i>).....	47
3.3.2 Long-Term Propagation Assays	49
3.3.2.1 Propagation of <i>mat-1</i> (<i>ye121</i>)	50
3.3.2.2 Propagation of <i>cgh-1</i> (<i>tn691</i>).....	51
Chapter 4: Genomics Analyses Identify Candidate Modifier Variants	53
4.1 Whole Genome Sequencing Analysis for CRISPR/Cas9 Off-Target Effects	53
4.2 WGS Investigation of Genetic Modifiers for <i>mat-1</i> and <i>cgh-1</i>	56
4.2.1 Prioritization of SNVs as Candidate Modifiers for <i>mat-1</i> and <i>cgh-1</i>	57
4.2.1.1 Initial Bioinformatics Analysis	57
4.2.1.2 Machine Learning Application for Prioritization of Candidate Modifiers	59
4.3 SVs as Potential Candidate Modifiers for <i>mat-1</i> and <i>cgh-1</i>	60
4.4 Total mRNA-sequencing and Transcriptomics Analysis	62
4.4.1 RNA-Seq Analysis Reveals No Differential <i>mat-1</i> and <i>cgh-1</i> Expression Across All Six Natural Isolates	64
4.4.2 Natural Isolates Differ in Universal Gene Expression.....	65
Chapter 5: Investigation of Candidate Genetic Modifiers for <i>mat-1</i> and <i>cgh-1</i>	69
5.1 Knock-in of Candidate Modifiers With CRISPR/Cas9	69
5.2 Candidate modifiers for <i>mat-1</i>	70
5.2.1 The SUPpressor of spindle CHECKpoint defect gene, <i>such-1</i>	70
5.2.2 The yeast BUB homolog gene, <i>bub-3</i>	72
5.2.3 <i>Y73B6A.1</i>	74
5.3 Candidate modifiers for <i>cgh-1</i>	75
5.3.1 The HunchBack Like (fly gap gene related) gene, <i>hbl-1</i>	75

5.3.2 The NHL (ring finger b-box coiled coil) domain containing gene, <i>nhl-2</i>	77
5.3.3 The DEAD boX helicase homolog gene, <i>ddx-15</i>	79
5.3.4 The abnormal Germ Line Proliferation gene, <i>glp-1</i>	80
5.4 Propagation assays for candidate modifier strains.....	82
Chapter 6: Discussion	84
6.1 Phenotypic Variability of <i>mat-1</i> and <i>cgh-1</i> is Affected by Different Genetic Backgrounds	84
6.2 Natural Genetic Variation and its Influence on Gene Expression	88
6.3 Identifying Genetic Modifiers – An Ongoing Challenge	90
6.4 The Use of WGS + Reverse Genetics.....	93
6.5 Future Directions	95
6.6 Final Considerations – The Relevance of Investigating Genetic Modifiers of Gene Essentiality	96
References.....	98

List of Tables

Table 1. PCR-based genotyping strategy for natural isolates.....	29
Table 2. CRISPR/Cas9 targets and components for homology directed repair.....	30
Table 3. PCR genotyping strategy and primer sequences for CRISPR-created strains.....	31
Table 4. Primers utilized for SV validation in different <i>C. elegans</i> natural isolate strains.	34
Table 5. mRNA-sequencing samples concentration and ratios.	38
Table 6. Strains carrying temperature-sensitive alleles for <i>mat-1</i> and <i>cgh-1</i> created by Dr. Xiao Li using CRISPR/Cas9.	42
Table 7. Propagation of wild-type strains at restrictive temperature.....	50
Table 8. Propagation of <i>mat-1(ye121)</i> strains at restrictive temperature.....	51
Table 9. Propagation of <i>cgh-1(tn691)</i> strains at restrictive temperature.	52
Table 10. Variants of interest, related to <i>mat-1</i> and <i>cgh-1</i> , identified in initial WGS and bioinformatics analysis.	58
Table 11. Candidate modifiers for <i>mat-1</i> and <i>cgh-1</i> identified using machine learning approach.	59
Table 12. Candidate modifier prioritized in this study and their presence or absence in each of the six natural isolates' backgrounds AB1, CB4856, GXW1, JU1400, KR314 and N2.....	60
Table 13. <i>p</i> -values for <i>alg-2</i> expression from AB1, CB4856, KR314, GXW1 and JU1400 in comparison with N2 control.....	62
Table 14. <i>p</i> -values for <i>mat-1</i> and <i>cgh-1</i> expression from AB1, CB4856, KR314, GXW1 and JU1400 in comparison with N2 control.....	66
Table 15. <i>p</i> -values for candidate modifier genes investigated in this study as potential modifiers of <i>mat-1</i> and <i>cgh-1</i>	68
Table 16. Propagation of <i>mat-1(ye121)</i> candidate modifier strains at restrictive temperature.....	82
Table 17. Propagation of <i>cgh-1(tn691)</i> candidate modifier strains at restrictive temperature.. ...	83

List of Figures

Figure 1. Schematic representation of how gene essentiality might be modified through the effect of the genetic background.	8
Figure 2. A. gene essentiality of <i>mat-1</i> and <i>cgh-1</i> in different <i>C. elegans</i> genetic backgrounds. B. Original location from the six natural isolates used in this study.	25
Figure 3. RNAi results for comparison of knockdown of <i>mat-1</i> (B) or <i>cgh-1</i> (C)... ..	41
Figure 4. Hatch rates for WT isolates under restrictive conditions (23°C, 24°C, 25°C, and 25.3°C, respectively), showing little impairment of survival.	44
Figure 5. Hatch rates for <i>mat-1(ye121)</i> strains depict phenotypic variability for embryonic lethality..	46
Figure 6. Hatch rates for <i>cgh-1(tn691)</i> strains depict phenotypic variability for embryonic lethality.	48
Figure 7. Propagation assay. After final observation, plates are classified into three categories according to the level of propagation present.	49
Figure 8. <i>soap-1</i> (V:7207103 A > G) information. (A) depicts IGV screenshot confirming presence of this variant exclusively in the N2; <i>cgh-1(tn691)</i> background. (B) depicts summarized interactions of <i>soap-1</i> and its indirect linkage to <i>cgh-1</i> . (C) STRING <i>soap-1</i> protein interactions search (https://string-db.org/ Accessed on September 5 th , 2023).	55
Figure 9. Investigation of <i>soap-1</i> as a potential interaction with <i>cgh-1</i> and influence on the <i>cgh-1(tn691)</i> phenotype.	56
Figure 10. Genetic modifiers identification workflow.	57
Figure 11. IGV visualization of gene expression of <i>alg-2</i> in JU1400, AB1, CB4856, KR314, GXW1 and N2 WT strains (left) and <i>cgh-1(tn691)</i> strains (right). Peaks (gray) represent coverage within exon.	62
Figure 12. IGV visualization of gene expression of <i>cgh-1</i> in JU1400, AB1, CB4856, KR314, GXW1 and N2 WT strains (left) and <i>cgh-1(tn691)</i> strains (right).	64
Figure 13. IGV visualization of gene expression of <i>mat-1</i> in JU1400, AB1, CB4856, KR314, GXW1 and N2 WT strains (left) and <i>mat-1(ye121)</i> strains (right).	65
Figure 14. MD plots of each natural isolate against N2 control strain depicting differential expression for both over- (red) or under-expression (blue).	67

Figure 15. AB1 strains and IGV visualization of <i>hbl-1</i> gene expression, showing reduced expression in the WT background.	68
Figure 16. <i>C. elegans</i> target region for CRISPR/Cas9 microinjection highlighted in orange (rachis)..	70
Figure 17. Hatch rate assays for investigation of <i>such-1</i> as a modifier for <i>mat-1</i>	71
Figure 18. Hatch rate assays for investigation of <i>bub-3</i> as a modifier for <i>mat-1</i>	73
Figure 19. Hatch rate assays for investigation of <i>Y73B6A.1</i> as a modifier for <i>mat-1</i>	75
Figure 20. Hatch rate assays for investigation of <i>hbl-1</i> as a modifier for <i>cgh-1</i>	77
Figure 21. Hatch rate assays for investigation of <i>nhl-2</i> as a modifier for <i>cgh-1</i>	78
Figure 22. Hatch rate assays for investigation of <i>ddx-15</i> as a modifier for <i>cgh-1</i>	80
Figure 23. Hatch rate assays for investigation of <i>glp-1</i> as a modifier for <i>cgh-1</i>	82

List of Abbreviations

A-P	Anterior - Posterior
APC/C	Anaphase Promoting Complex/Cyclosome
bps	Base pairs
BH	Benjamini and Hochberg
CeNDR	Caenorhabditis elegans Natural Diversity Resource
CGC	Caenorhabditis Genetics Centre
<i>cgh-1</i>	Conserved Germline Helicase
CGR	Complex Genomic Rearrangement
CHGI	The Centre for Health Genomics and Informatics
CNVs	Copy Number Variants
CPC	Chromosomal Passenger Complex
CRISPR	Clustered Regularly Interspaced Palindromic Repeats
CVL	Candidate Variant List
DEG	Database of Essential Genes
DTC	Distal Tip Cell
dsRNA	Double-stranded RNA
EMS	Ethyl Methanesulfonate
eQTL	Expression Quantitative Trait Loci
ExAC	Exome Aggregation Consortium
GOI	Gene of Interest
gRNA	Guide RNA
HVR	Hypervariable regions
IGV	Integrative Genome Viewer
IPTG	Isopropyl- β -D-thiogalactopyranoside
L4	Larval Stage 4
logCPM	log-2-counts-per-million

lrWGS	Long Reads Whole Genome Sequencing
<i>mat-1</i>	Metaphase-to-Anaphase Transition Defect
MCC	Mitotic Checkpoint Complex
MGI	Mouse Genome Informatics
miRNA	Micro RNA
ML	Machine Learning
MMP	Million Mutation Project
mRNA	Messenger RNA
NGM	Nematode Growth Media
PCR	Polymerase Chain Reaction
QTL	Quantitative Traits Locus
RNAi	RNA-mediated interference
SAC	Spindle Assembly Checkpoint
SILAC	Stable Isotope Labeling by Amino Acids In Cell Culture
SMRT	Single-molecule real-time
SNP	Single Nucleotide Polymorphism
SNV	Single Nucleotide Variant
srWGS	Short Reads Whole Genome Sequencing
ssODN	Single Stranded Donor Oligonucleotide
SV	Structural Variant
T _M	Temperature of Melting
UC	Ulcerative Colitis
UV	Ultraviolet
VCF	Variant Call Format
WGS	Whole Genome Sequencing
WT	Wild Type

Chapter 1: Introduction

1.1 Gene Essentiality

The essential genes in a genome are those in which a loss-of-function leads to a lethal phenotype or an impairment of reproductive success (Jordan *et al.* 2002). Essential genes can be sub-classified into four categories, from genes that are only essential under specific conditions, to the ones that are indispensable for minimal cellular life (Zhang and Ren 2015). Meanwhile, “non-essential” genes are the ones that may still encode a crucial function in the organism but are dispensable for survival or reproduction. The collection of the minimal number of genes essential for an organism’s survival and fitness is known as “essentialome”. For instance, in vertebrates such as Zebrafish (*Danio rerio*), the number of essential genes responsible for embryonic development is approximately 1,400 (Amsterdam *et al.* 2004). Others, such as mice (*Mus musculus*), carry an essentialome of roughly 1,302 genes, identified using the Mouse Genome Informatics (MGI) resource (Kabir *et al.* 2017). In general, essential genes encode proteins related to fundamental biological processes important for the existence of the cell itself; therefore, essentialomes are enriched in genes required for DNA, RNA and protein synthesis (Kim *et al.* 2010; Zhan and Boutros 2015).

The gene essentiality concept was first proposed by Gluecksohn-Waelsch (1963) and prokaryotes were the first organisms in which studies with essential genes and lethal variants were conducted. Since then, researchers tried to identify the minimal number of genes indispensable for regular cellular function and survival in microorganisms such as bacteria and yeast. However, the methods selected and score thresholds may diverge from study to study (Cacheiro *et al.* 2020). In microorganisms, 80% of *Mycoplasma genitalium* genes, (Martínez-Carranza *et al.* 2018) and 20% of *Saccharomyces cerevisiae* genome are found to be indispensable for survival (Zhang and Ren

2015). In *Haemophilus influenzae*, ~22% of its ~1,600 genes are considered essential (Gawronski *et al.* 2009), while only ~7% are essential in the *Escherichia coli* (*E. coli*) genome (Baba *et al.* 2006; Mori *et al.* 2015). In *E. coli*, however, the set of genes common to all strains within this species (called core genome) represents at most half the number of genes in any of these strains. This highlights how bacteria from the same species can differ widely in their gene content, and consequently in their phenotype (Rousset *et al.* 2021). Bacteria such as *Bacillus subtilis*, for example, clearly show evidence of variable gene essentiality. In this species, *FtsZ* is a tubulin-like gene essential for cell division in the naturally prevalent (wild type - WT) bacteria. However, a single amino acid change in *yqiD* gene can modify this essentiality, turning *FtsZ* into a non-essential gene in L-form *B. subtilis* (Leaver *et al.* 2009).

S. cerevisiae was also one of the first organisms where the concept of gene essentiality was studied (Winzeler *et al.* 1999). Under laboratory conditions, it has been reported that over 80% of its genes are essential in this species (Papp *et al.* 2004). Zhang & Ren (2015) classified yeast gene essentiality in 4 distinct groups: conditional essential (i.e., only essential under specific conditions), essential (i.e., required for survival under optimal conditions), redundant essential (i.e., synthetic lethal/redundant pathways) and absolute essential (extremely important for minimal cellular life). As an example, the double knock-out of *CEP1* and *CEP3* leads to a lethal phenotype in yeast (synthetic lethality) (Baker *et al.* 1998).

Studies in *Drosophila melanogaster* also provided relevant insights to this matter. Initial identification of fruit fly's essential genes for adult viability was done using *P* element transposon mutagenesis screens by the Berkeley *Drosophila* Genome Project. Later on, a study from 2004 revealed that only 2% of *D. melanogaster* genes are essential (Boutros *et al.* 2004). The method used, RNA interference (RNAi), is far less effective in comparison to complete knockouts and

could lead to merely hypomorphic phenotypes, which might explain the low percentage of essential genes identified (Boutros *et al.* 2004). Nevertheless, this study represents another example of a genome in which only a small portion of genes may be indispensable for survival. More specifically, ~2,000 out of the ~16,000 genes in the fruit fly genome are suggestively essential (St. Pierre *et al.* 2014), and about 30% of the 195 young genes (genes arisen 3 to 35 million years ago) are considered essential in this species. Even though essential genes are often portrayed as conserved and ancient, newly-arisen genes can also perform relatively indispensable organismal functions (Chen *et al.* 2010).

In the nematode *Caenorhabditis elegans* (*C. elegans*), as in other species, the road to discovery of its essentialome has been continuous. An early ethyl methanesulfonate (EMS) mutagenesis screen identified at least 2,850 essential genes in this species, a small portion of its genome which contains 20,000 genes (Johnsen and Baillie 1991; Hodgkin 2001). The development of RNAi in this species facilitated gene manipulation and gene knockdown, resulting in initial confirmation of over 1,000 essential genes (Fire *et al.* 1998; Kamath *et al.* 2003), which rapidly increased to almost 4,000 genes linked to essential phenotypes (WormBase WS275). However, this collection continues to change as more work towards gene essentiality remains to be elucidated. A considerable piece of essentialome is conserved between species (Zhan and Boutros 2015). A comparative study showed that over 50% of essential genes in worms are orthologous to genes also considered essential in *S. cerevisiae*, highlighting the conservation of essential processes in eukaryotes (Fraser 2015).

Accordingly, many studies have been performed regarding the human essentialome. For instance, about ~10% of human genes were reported to be essential in cultured human cells (Blomen *et al.* 2015). Although this might not fully correspond to gene essentiality *in vivo*, the

study of this concept is relevant for scrutinizing Mendelian diseases, since these usually involve severe impairment of human fitness due to disruption of crucial processes in the human organism (Bartha *et al.* 2018). In 2006, it was estimated that 7.9 million children worldwide are born with a serious birth defect of genetic or partially genetic origin, and that 3.3 million of them under the age of 5 died each year (Zarocostas 2006).

Systematic genetic screens in humans are often impractical, and mammalian cells have shown to be less amenable to RNAi screens than invertebrate models. This challenge was overcome with the advent of CRISPR/Cas9 technology, which allowed for precise gene disruption and further experimenting in human cell lines (Fraser 2015). Investigation of the mice essential genes has also been impactful for understanding human essentialome and gene-disease associations: a study conducted with 400 human cancer cell lines revealed that 35% of orthologous genes lethal in mice were also indispensable for cell fitness in these human cell populations. Additionally, a combination study on human cancer cell lines and yeast model revealed differences in essential genes specific to each cell line and cancer type due to developmental origin, oncogenic drivers, paralogous gene expression pattern, and chromosomal structure, which may be a crucial factor for potential treatments (Wang *et al.* 2015). Overall, essential genes in humans were found to be very similar to essential genes in other species such as yeasts and worms, being mostly involved in core cellular functions such as transcription, translation, and metabolism. They also are highly expressed and suffer from increased selective pressure, noted by both the lower variation rates within species and cross-species conservation (Fraser 2015).

Gene essentiality has had an increasing interest as a topic of research for the past decades. Its massive investigation led to the assortment of the Database of Essential Genes (DEG), which carries essentialome information for both prokaryote and eukaryote organisms (Zhang *et al.* 2004;

Luo *et al.* 2021). Much has been discovered towards similarities and divergencies between species, however, the variation of the essentialome within species still lacks elucidation. Additionally, identification of genes essential for cell or organism viability not only collaborates to the understanding of human genetic disorders, but also allows novel drug development approaches to target parasites, pathogens, and cancer cells, through impairment of genes specifically crucial for survival of these (Doyle *et al.* 2010; Shi *et al.* 2015; Zhang *et al.* 2018).

1.1.1 Plasticity of Essentiality

Although at first gene essentiality was viewed as immutable within a species, much has been considered around the concept of this static essentiality. Nowadays, it is believed that the essentiality of genes is a variable condition and depends on both environmental and genetic contexts (Figure 1), as well as subjected to evolutionary change. This flexibility in essential genes is what we call “plasticity of essentiality”. Regarding environmental influence, the auxotrophy phenomenon is a clear example of this. In this concept, genes encoding basic molecules for metabolism are essential only if the same molecule is absent from the growth medium or natural environment (Rancati *et al.* 2017). As another example, 37–68% genes considered dispensable for survival in a rich media were found to be essential for fitness in other growth conditions in yeast (Papp *et al.*, 2004). Additionally, deletions in non-essential genes, might be beneficial for survival in a determined environment, while detrimental in others (Qian *et al.* 2012). Nevertheless, gene essentiality may also be influenced by evolution, as probed in bacteria and yeast (Baba *et al.*, 2006a; Kim *et al.*, 2010). During evolution, genes or functions could arise separately, be lost, or replaced by others, and some pathways and fundamental processes might become redundant (Koo *et al.* 2017).

Proofs of plasticity of essentiality in human cells have also been observed when comparing multiple cell lines. Three recent studies have, in conjunction, disrupted every human gene in a variety of cell lines, and determined a range of ~2,000 essential genes (Blomen *et al.* 2015; Wang *et al.* 2015; Hart *et al.* 2015). However, as seen in model organisms, essentiality in human cells also seems to act in a context-dependent manner. A small set of these genes were only essential in specific cell lines, which might shed light into cancer cells and their unique essential genes set (Wang *et al.* 2015; Hart *et al.* 2015; Fraser 2015).

Gene essentiality is also influenced by the genetic background - defined as “the genotype of all other related genes that may interact with the gene of interest, and therefore potentially influence the specific phenotype” (Yoshiki and Moriwaki 2006). The genetic context might influence essentiality in multiple ways, for example, a gene that becomes essential for survival only if another gene performing the same or similar function is also lost, leading to synthetic lethality (Nijman 2011), as seen in the iron transporter genes *AFT1*, *FET3* and *FET4* in yeast. Deletion of *FET3* dysregulates the function of high-affinity iron transport complex, while loss of *AFT1* also affects the mobilization of iron stores inside the cell. *FET4*, involved in low-affinity Fe(II) transporter of the plasma membrane, displayed synthetic lethality in association to deletion of either of the other two genes (Berthelet *et al.* 2010; Chowdhury *et al.* 2015). It is also possible that an essential gene becomes “dispensable” when differences between individuals of the same species within the genetic background are considered.

In two strains of *S. cerevisiae*, for example, 44 genes are uniquely essential in the Sigma1278b strain, whereas 13 others are essential in the S288c strain, and other 894 were unconditionally essential in both backgrounds (Dowell *et al.* 2010). Additionally, comparison of essential genes in yeast under three distinct conditions uncovered two sets of genes – core essential

genes (i.e., indispensable in all tested conditions), and context-dependent genes, in which the effect of the genetic background in gene essentiality can be observed (Bosch-Guiteras and van Leeuwen 2022). Variation within species exemplifies how the assessment of phenotypic variability requires not only the use of lab-conditioned strains, but also incorporation of wild isolates (and consequently their natural variation in the genetic background) (Gasch *et al.* 2016).

Evidence of conditional essentiality and the influence of genetic background in *C. elegans* have been observed, in a comparative study of ~1,500 genes knockdowns between reference strain N2 and Hawaiian strain CB4856. The phenotypic variability observed for multiple essential genes in their dataset was suggestively linked to differences in the genetic backgrounds of these two wild isolates. Natural variations in the genomic background of different natural isolates have been the topic of multiple studies and may strongly influence phenotypic variability. A study conducted in 2014 identified several single nucleotide variants (SNVs) and insertion type variants in the CB4856 strain when compared to N2 (Vergara *et al.* 2014), and evidence of variable phenotype among these strains has been reported using RNAi knockdown (Vu *et al.* 2015). Moreover, the investigation of gene expression of 207 *C. elegans* wild isolates traced back the transcriptomics variability to corresponding genetic differences. Using genome-wide association mapping, Zhang *et al.* (2022) linked phenotypic variation of complex traits to the differential gene expression among these strains, which was then explained by natural genetic variation among them (Zhang *et al.* 2022). Another study between reference strain N2 and Hawaiian strain (CB4856) has previously indicated the presence of extremely high divergent regions in CB4856 genome (Thompson *et al.* 2015). It also seems like genes related to substantial functions such as sensory perception and pathogen response lie on those regions and might enable this species to survive in perturbing environmental conditions (Lee *et al.* 2021). Additionally, such delicate regions might represent

evolution's way to maintain genetic diversity in a species that is 99% self-crossing (Barrière and Félix 2005b).

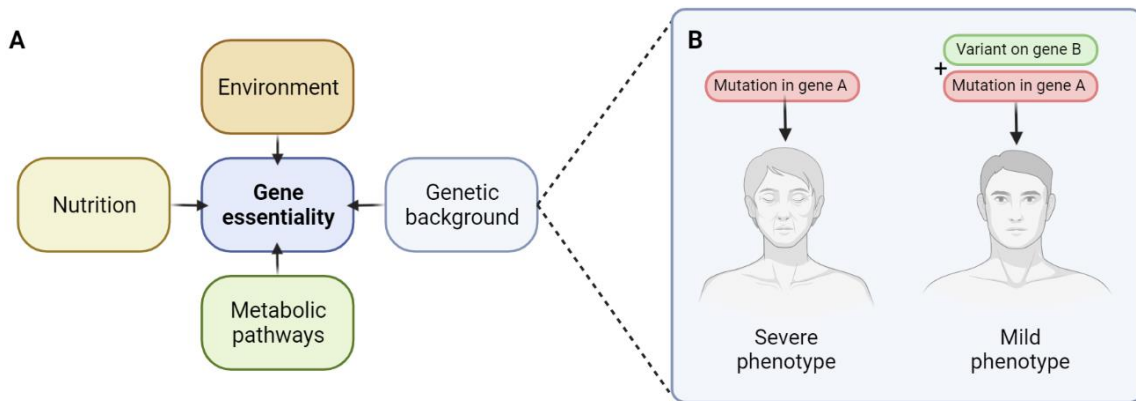


Figure 1. Schematic representation of how gene essentiality might be modified through the effect of the genetic background. Image designed using BioRender [<https://www.biorender.com/>].

1.1.2 Genetic Modifiers

The effect of the genetic background in a certain phenotype is known as genetic modification and is due to genetic interactions. Genetic modifiers, second-site variants that influence the phenotype of a primary variant, can lead to the suppression or exacerbation of that phenotype (Figure 1). Although modifying variants do not always have phenotypes of their own, they can have significant effects on phenotypic outcome of other variants in the genome. Their influence can result from direct interaction with the primary variant, acting within similar biological processes or even compensation through alternative pathways, in a subtle or significant manner (Prior et al., 2009). The first comprehensive genetic network for any model organism conducted by Costanzo et al. (2016) in yeast shows that genetic interactions are more likely to occur among genes with related function. They analyzed thousands of genetic interactions and classified them into two main groups: negative genetic interactions (enhancers), including synthetic lethality, in which combined variants lead to a lethal phenotype, and positive interactions

(suppressors) in which genetic suppression alleviate the phenotype of primary variant. Overall, ~500,000 negative genetic interactions and ~350,000 positive ones were catalogued in this study.

Modifiers can alter penetrance, expressivity, pleiotropy, and dominance. Although these represent distinct biological concepts, their effects can often be found to overlap. The effect of modifiers on penetrance can be observed when individuals carrying a disease-causing variant are asymptomatic, and the proportion of affected individuals with the expected phenotype is not 100% (Riordan and Nadeau 2017). Incomplete penetrance can be observed in both dominant and recessive conditions, and its effect on genotype–phenotype correlations can be difficult to elucidate, even when identical environment and genetic background is considered (Dickinson *et al.* 2016; Kingdom and Wright 2022). On the other hand, modifiers can also contribute to variable expressivity among affected individuals, when the severity of a genetic trait is not consistent between them, even in individuals of the same family (Riordan and Nadeau 2017). Pleiotropy is influenced when a novel phenotype is observed only in the presence of a genetic modifier, and lastly, partial dominance can be observed in a genetic background containing dominant modifiers. This way, a single copy of a target allele is sufficient to induce the phenotype in those backgrounds, but not in others (Riordan and Nadeau 2017).

The simplest type of genetic modification is the one where there is one modifier to one variant, identified by examining a single second-site variant and its influence on the primary phenotype (enhancement or suppression) (Zhou *et al.* 2003). However, this does not completely depict the reality of genetic modifiers, and phenotypic variability is often caused by a complex modifier network (Hou *et al.* 2019). In that case, it should be considered if modifiers work individually with additive effects on the primary phenotype, or if they have synergistic or antagonistic effects that could “balance out” to provide the final phenotype (Riordan & Nadeau,

2017). For instance, either duplication of *cyb-3*, or its combination with a missense variant in *dhc-1* compensates for impairment of *mdf-1* (Tarailo-Graovac *et al.* 2010, 2014).

Throughout the last decades, genetic modifiers have been identified in a variety of model organisms. In mice, for example, *Nrd1d1* was discovered as a suppressor of another nuclear hormone receptor, *Nr2e3*, in which variants cause eyesight impairment due to photoreceptor degeneration (Meyer and Anderson 2017). Additionally, mouse models of hearing loss showed a *Atp2b2^{dfw-2J}-Cdh23^{753A/G}* interaction in age-related hearing loss which also translated to observations in human patients (Johnson *et al.* 2006).

The identification of genetic modifiers and the discovery of genetic interactions have always overlapped and helped with deciphering complex pathways. In *C. elegans*, examples of genetic modifiers as positive genetic interactions include *sup-11* as a suppressor of the *unc-93(e1500)* and *such-1* as a suppressor of *mdf-1(gk2)* (Greenwald and Horvitz 1982; Tarailo *et al.* 2007a). Suppression of the *unc-93(e1500)* by *sup-11* was identified by missense mutations that induce a novel activity to SUP-11 (neomorph), suggesting that altered *sup-11* is a modified product of a gene that is a member of an *unc-93* gene family (Greenwald and Horvitz 1982). On the other hand, suppression of *mdf-1(gk2)* was performed by *such-1(h1960)*, a point mutation that induces a G27D alteration in the amino acid sequence in the semiconserved region, and results in delayed mitosis and consequent suppression of *mdf-1(gk2)* (Tarailo *et al.* 2007a). A more extensive study conducted in 2006 systematically tested ~65,000 pairs of genes in *C. elegans* to their ability to interact with each other, leading to the identification of 350 genetic interactions relevant to signaling pathways that are mutated in human diseases (Lehner *et al.* 2006).

In humans, comparative expression, genome-wide association studies or family-based analysis have helped to drive attention to the effects of genetic background in inherited conditions

(Chen *et al.* 2016; Rahit and Tarailo-Graovac 2020). The Resilience Project, for example, is a great achievement in that matter, in which the search for healthy carriers of variant related to severe genetic disorders have been conducted (Chen *et al.* 2016). The conducted analysis of patients' genomes allowed the identification of 13 individuals with variants for 8 severe Mendelian disorders who did not present any clinical manifestation (Chen *et al.* 2016). In addition, investigation of the Exome Aggregation Consortium (ExAC) dataset analysed over 60,000 exomes from unrelated individuals and pointed out several variants implicated in Mendelian disorders with suggestive phenotypic variability, ranging from mild to severe (Tarailo-Graovac *et al.* 2017).

Primary observations of variable age of onset and phenotype of inherited diseases dates back to the 1940s (Haldane 1941). However, discovery of modifiers for human genetic diseases is difficult due to low population frequency of rare trait, and lack of pedigree data (Chen *et al.* 2016). In contrast, model organisms such as nematodes are genetically tractable and offer a plethora of experimental tools to start addressing the modifiers relevant to human rare diseases (Hou *et al.* 2018). Therefore, the identification of genetic modifiers using model organisms could clarify why patients with the same genetic disease have variable clinical presentations and potentially lead to improved prognosis and novel therapeutics.

1.1.2.1 Genetic Nature of Modifiers

Despite the challenges in identification of genetic modifiers, several SNVs have been characterized as the source of genetic modification in multiple organisms and human disorders (Prior *et al.* 2009; Paaby *et al.* 2015; Tamura *et al.* 2017). Discovery of variants may be facilitated with the use of Whole Genome Sequencing (WGS) thorough bioinformatics analysis (Jean *et al.* 2021). However, SNVs are not uniquely responsible for genetic modification, although being more commonly observed in literature. In fact, structural variants (SVs), which includes copy number

variants (CNVs), may also contribute to phenotypic variability. A deletion of *Thra* locus in mice was able to compensate for *Thrb* deficiency and restore hearing loss in these animals (Ng *et al.* 2001). In another study, SVs were identified as modifiers for the variegating variant *brown*^(Dominant) (*bw*^(D)) of *Drosophila melanogaster* (Talbert *et al.* 1994). SVs such as duplications were shown to act as modifiers of essential genes in addition to SNVs, as discovered in a yeast study (Van Leeuwen *et al.* 2020), and were also found to modify human genetic disorders, such as Hirschsprung disease (Jiang *et al.* 2011). Their extensive analysis identified 188/380 strains carrying aneuploidy or duplications with potential suppressor effect. As another example, a duplication of *cyb-3* gene in *C. elegans* is capable of suppressing MDF-1/MAD-1 Spindle Assembly Checkpoint (SAC) sterile phenotype, and in addition to a missense variant in *dhc-1* gene, was capable to make up for MDF-1 impairment (Tarailo-Graovac *et al.* 2010, 2014). Duplication of *cyb-3* seems to result in a less error-prone spindle formation that lowers the requirement for a functional SAC, while its association with reduction of DHC-1 through missense allele *dhc-1(dot168)* also compensates for *mdf-1* loss, without causing a constant anaphase onset delay (Tarailo-Graovac *et al.* 2010, 2014).

1.1.2.2 Genetic Modifier Screens

Several approaches for studying genetic modifiers and genetic interaction networks have been developed, each carrying its own advantages and challenges. For instance, the use of mutagenesis screens and backcrossing strategies can filter out unrelated variants in an isogenic background. Mutagenesis screens in *C. elegans* have been conducted for decades and are often applied to study gene functions, genetic interactions and genetic diseases due to both cost and efficiency (Brenner 1974; Sin *et al.* 2014). Additionally, RNAi has been developed and widely applied to investigate genetic interactions in *C. elegans*, and provide a targeted impairment of a

gene, which cannot be reached with mutagenesis-screens, that instead, induce random mutations. Both these techniques have been extensively used for genetic interaction studies in *C. elegans*.

1.1.2.2.1 EMS Mutagenesis Screen

The most common method for genome perturbation in *C. elegans* is treatment with EMS that induces SNVs in the background of interest (Brenner 1974). In this method, the parental strain (P0), usually in the larval stage 4 (L4) or older, is exposed to the chemical, and the F1 generation, (heterozygous for the induced variants) is allowed to self to produce F2 homozygous animals, that are then screened for the phenotype of interest (Jorgensen and Mango 2002). EMS mutagen typically induces guanine/cytosine and adenine/thymine modifications causing missense, splice site, or nonsense variants throughout the genome, and it is widely used due to its relatively low toxicity and good efficiency (Jorgensen and Mango 2002; Flibotte *et al.* 2010; Kutscher and Shaham 2014). It also induces larger variants: roughly 13% of EMS effects are deletions or other chromosomal rearrangements (Kutscher and Shaham 2014).

Mutagenesis screens can be used not only for identifying new gene functions, but also to screen different variants responsible for the same phenotype and unravel genetic interactions (epistasis – genetic modifiers) (Brenner 1974; Sulston and Horvitz 1981; Stuart 2022). Typical EMS mutagenesis induces ~400 variants per genome in *C. elegans*, in both coding and non-coding regions (Thompson *et al.* 2013). In the search for genetic modifiers, not all the variants modify the primary phenotype, but the phenotypic readout can be used to determine which genomes contain the modifying variant (showing either enhancement or suppression of the phenotype). Further, the specific variants responsible for such modifications are typically identified through extensive rounds of backcrossing to the parental strain under detrimental/selection conditions. Backcrossing involves repeatedly crossing the mutagenized worms back with the parental strain, which allows

for the selection of variants related to the primary phenotype: each round of crosses reduces the number of variants from the mutagenized strain in half, and populations carrying the potential modifier can be selected using specific conditions, until a final strain with a minimal number of variants is achieved (Zuryn *et al.* 2010).

Mutagenesis studies have been useful for deciphering multiple gene functions in *C. elegans*, and also unraveling genetic modifiers for human diseases (Sin *et al.* 2014). The first EMS screen identified 619 mutants in *C. elegans* (Brenner 1974). Although effective, when we consider the initial pool of variants induced by mutagenesis, these genetic screens become often slow and laborious and may be delayed if fitness of the studied strains is affected (Sega 1984). Additionally, although identification of a suppressor is quite efficient, there is no guarantee that other suppressors or variants of interest might not be missed using this method. The inability of observing the entirety of a genetic network might preclude our understanding of complex traits in the genome. This influenced the development of new techniques for phenotypic studies and genetic modifiers identification.

A significant improvement in identification of modifiers in mutagenesis screens was the combination with next-generation sequencing, which allowed for faster mapping of the EMS-induced variants (Doitsidou *et al.* 2010; Lehrbach *et al.* 2017). The use of WGS facilitates the prioritization of candidate modifiers, which can then be tested using approaches such as CRISPR/Cas9 (Jean *et al.* 2021). This strategy was successfully applied in the identification of intragenic and extragenic variants for *zyg-1*, a gene involved in centriole function (Jean *et al.* 2021; Stuart 2022). Application of WGS and bioinformatics analyses was able to replace the backcrossing steps, and selected candidate modifiers were experimentally validated using CRISPR/Cas9. Additionally, this strategy identified that ~20% of mutagenized strains had more

than one modifier, exemplifying the power of WGS in facilitating the identification of true interaction networks (Stuart 2022).

1.1.2.2 RNA interference (RNAi)

Reverse genetics techniques such as RNAi have been widely applied for the analysis of loss-of-function phenotypes. The effects of double-stranded RNA for genetic interference were in fact discovered in *C. elegans* and revolutionized genetic studies in worms (Fire *et al.* 1998). RNAi acts by introduction of double-stranded RNA (dsRNA) which targets a specific mRNA and induces its degradation in the organism. *C. elegans* has a comprehensive RNAi library that is commercially available and carries over 16,000 bacterial clones, targeting almost its entire genome (Kamath *et al.* 2003). An advantage of using RNAi in *C. elegans* is that it has a systemic effect in this species, meaning that introduction of dsRNA in one tissue is capable of silencing RNA throughout the whole organism, which is credited to a process called transitive RNAi (Timmons *et al.* 2003). Another advantage is its easy introduction in *C. elegans*: RNAi can be delivered by microinjection, soaking worms in dsRNA solution, or by simply feeding worms with dsRNA-producing bacteria (Fire *et al.* 1998; Tabara *et al.* 1998; Timmons *et al.* 2001). It is also amenable to studies throughout different developmental stages, and its genome perturbation is targeted and specific to a gene of interest, which is not possible with EMS screens (Maine 2008; Sin *et al.* 2014).

During the past decades, RNAi has been widely used for identification of new gene functions in *C. elegans* and genetic interactions in the worm. A study utilizing RNAi for testing over 65,000 gene pairs in *C. elegans* was able to identify ~350 genetic interactions in genes associated with pathways involved in human disease (Lehner *et al.* 2006). In addition, RNAi in *C. elegans* models of neurodegenerative disorders led to the identification of genetic modifiers for Parkinson's disease, Alzheimer's disease and polyglutamine aggregation in Huntington's disease

(Morley *et al.* 2002; Kraemer *et al.* 2006; Van Ham *et al.* 2008).

On the other hand, RNAi also presents some disadvantages and can lead to variable results (Kamath and Ahringer 2003). Firstly, it is only capable of reducing gene activity, but does not provide a full knock-out. A gene-phenotype association might be biased when using RNAi, as less-effective knockdowns may lead to false negative interpretations. Additionally, the nature of its effect (knockdown) does not suffice to investigate the whole spectrum of variants and gene perturbations, such as missense SNVs, and SVs. Moreover, RNAi interferes in the messenger RNA (mRNA) level, influencing the expression of a protein. However, if a protein is highly stable and already produced in the cell, RNAi may fail in impairing its function and also influence result interpretation. Another obstacle for results interpretation is the natural variation in the response to RNAi, which was observed in a study comparing differential gene expression of multiple *C. elegans* wild isolates after germline gene knockdown (Bell *et al.* 2023). Finally, other aspects, such as freshness of bacteria (feeding dsRNA method) and quality of reagents (antibiotics, Isopropyl- β -D-thiogalactopyranoside - IPTG, etc) may also impact the given results (Sin *et al.* 2014).

1.2 Contemporary Techniques for Studying Gene Essentiality

The advancement of genetic techniques allowed for the optimized identification of genotype-phenotype associations in multicellular organisms and more complex systems. Further development of gene editing tools, for example CRISPR/Cas9, enabled faster and more cost-effective methods for reverse genetics studies and the identification of essential components and genetic modifiers (Jean *et al.* 2021). CRISPR enables a precise gene editing with minimal off-target effects and can be applied to a number of species (Luo *et al.* 2014). Throughout the past decade, CRISPR/Cas9 has been successfully established in multiple species: bacteria, yeast,

zebrafish, mice, worms, and human cells (Cong *et al.* 2013; Cho *et al.* 2013; Hwang *et al.* 2013; Jiang *et al.* 2013; Dicarlo *et al.* 2013; Wang *et al.* 2013; Tzur *et al.* 2013). It represents the greatest revolution in gene editing to date and allows for precise study of disease-causing variants in model organisms (Kim *et al.* 2022). For instance, CRISPR/Cas9 was used to model retinitis pigmentosa in *C. elegans*, allowing investigation of disease mechanism, novel drug targets, and genetic modifiers (Kukhtar Kukhtar 2021). It was also applied for large-scale loss-of-function knockout screens, study of synthetic lethality and virus-host interactions in worms (Liu and Li 2019).

Additionally, next-generation sequencing also plays an extremely helpful role for the study of variations in the genetic background and identification of genetic modifiers (Jean *et al.* 2021). Currently, genetic modifier screens can be aided by modern bioinformatics analysis that facilitate the prioritization of candidate variants, which can then be molecularly confirmed as modifiers using CRISPR/Cas9. The application of WGS in modifier search optimizes high-throughput analysis while reducing research time and labor. This was observed in studies in *C. elegans* including genes associated with human conditions (Jean *et al.* 2021). Altogether, these modern methodologies are helping scientists to unveil some mechanisms, including genetic modifiers, responsible for such phenotypic variability.

1.3 *C. elegans* as a Model for Plasticity of Essentiality

C. elegans is a free-living worm belonging to the phylum *Nematoda* that is approximately 1 millimetre long (adults) and is well understood in both cellular and molecular levels. It was the first multicellular organism to have the genome completely sequenced and it has an immutable number of somatic cells (959) that were thoroughly mapped in the hermaphrodite worm (The *C. elegans* Sequencing Consortium*, 1998; Sulston & Horvitz, 1977). Additionally, approximately ~52% of human genes have orthologues in *C. elegans*, allowing modelling of human conditions

for genetic research with this model organism (Kuwabara and O'Neil 2001; Harris *et al.* 2004; Kaletta and Hengartner 2006; Kim *et al.* 2018). *C. elegans* is found globally in the wild, offers easy maintenance, and it is amenable to genetic manipulation, which makes it an outstanding model for genomic studies (Brenner 1974). Nowadays, *C. elegans* strains can be readily obtained from the Caenorhabditis Genetics Center (CGC), that offers hundreds of strains and genotypes for worms' scientific research (<https://cgc.umn.edu/>).

C. elegans genome consists of 5 pairs of autosomes (often referred to as LG I, II, III, IV, V) plus 1 pair of sex chromosomes (XX = hermaphrodites) (Coulson *et al.* 1991). Males can also be present in the population, but only in rare cases of chromosomal non-disjunction (<0.2%), in which they would carry only one copy of the X chromosome (XO) (Kaletta and Hengartner 2006). Males and hermaphrodites have different anatomical features that are very easy to distinguish, such as their size, tail, and specific internal structures. The species is self-fertilizing, meaning that hermaphrodites produce their own oocytes and sperm. Hermaphrodites have a limited capacity to produce sperm and usually generate around 300 eggs but are also able to mate with males and produce viable progeny, laying up to 1000 eggs instead. In fact, when males mate with hermaphrodites, male sperm is preferentially used, and the offspring has a ratio of 50% males (Corsi *et al.* 2015; Albert Hubbard and Schedl 2019).

A strain isolated in Bristol, England in 1951 has been widely used for *C. elegans* experiments: the N2 wild isolate was initially kept and propagated in the laboratory for almost two decades, which collaborated to its “domestication”. Several variants arose and remained in the N2 background, which is partially associated with the agar plate environment (Frézal and Félix 2015). It was only many years later that its use as a model organism became widespread, and at that point, it had been already modified by domestication (Frézal and Félix 2015). It is currently used as a

wild-type control for most of the studies on this species and it has not developed many variations in its genome in comparison to other wild isolates due to its strict laboratory development over the past five decades.

Although *C. elegans* is considered to present low molecular diversity within a population due to its self-reproduction (Barrière and Félix 2005a), there is great genetic variability found in the genome of strains isolated in different parts of the world. Therefore, it is a great ally for understanding phenotypic variability, as each strain presents diverse genomic features (Crombie *et al.* 2019). A great resource for this purpose is CeNDR, the *Caenorhabditis elegans* Natural Diversity Resource (<https://www.elegansvariation.org/>). This platform gathers information from over 1,500 wild strains, which contributes to a better understanding of natural diversity and underlying phenotypic variability (Cook *et al.* 2017). CeNDR provides extensive data on SNVs across diverse natural isolates. However, it lacks information on larger SVs that could also play a role in genetic modification.

Most of the discoveries using *C. elegans* were made using the reference strain N2. However, much more can be discovered when analysing natural variations in the genetic background of other wild isolates (Kamkina *et al.* 2016; Andersen and Rockman 2022). The use of isogenic lab strains for most interaction studies, albeit useful, can hide how more natural complex interactions may arise, highlighting the importance of studies in natural isolates. For example, a study performed by Vu *et al.* (2015) examined phenotypic variability of RNAi knockdown in both N2 and CB4856 *C. elegans* backgrounds for over 1400 genes and observed extensive phenotypic variability in these strains. About 20% of these genes presented different phenotypes in the backgrounds of N2 and CB4856. CB4856 has distinct levels of genetic variants in its background and represents one of the most variable genomes, and may be closer to the origin

of *C. elegans* in comparison with N2 (Crombie *et al.* 2019). Additionally, these two wild isolates also differ in behavioral responses, including the responses to simultaneous changes in environmental O₂ and CO₂ (McGrath *et al.* 2009).

Two genes shown by Vu *et al.* (2015) to display phenotypic variability in N2 and CB4856 were selected to be studied further as part of my thesis: the Metaphase-To-Anaphase Transition Defect gene (*mat-1*) and the Conserved Germline Helicase gene (*cgh-1*). Both *mat-1* and *cgh-1* are known for being involved in essential biological processes in *C. elegans* and the loss of either of these genes results in lethality in the N2 background, yet reduced level of lethality was observed in the CB4856 background. In Vu *et al.* (2015), by using RNAi to knock-down these genes, the two natural isolates showed variable phenotypes: while N2 background showed lower embryonic survival, since the eggs failed to hatch, most CB4856 eggs were able to hatch, suggesting the suppression of the lethal phenotype to some degree in this strain. In addition to the RNAi screens, Vu *et al.* (2015) also analyzed the embryonic lethal phenotype of *cgh-1* with the use of the temperature-sensitive allele *tn691*: hypomorphic temperature-sensitive variants in this gene allow propagation at a permissive temperature (15°C), while causing embryonic arrest in the N2, but not the CB4856 background, at a restrictive temperature of 22°C. The *tn691* allele is a point mutation that induces protein instability due to structure alteration when placed under higher temperatures. The phenotypic variability in the lethal phenotype proven by the aforementioned work not only suggests the presence of genetic modifiers, but further shows that gene essentiality can vary with genomic context, i.e., the genetic background can in fact influence the phenotype.

1.3.1 The Metaphase-to-Anaphase Transition Defect Gene – mat-1

mat-1 is the *C. elegans* ortholog of human *CDC27* and is involved in an essential process: the cell cycle and asymmetric cell division. It is a component of the Anaphase-Promoting

Complex/Cyclosome (APC/C) (Golden *et al.* 2000) and its impairment leads to mitotic arrest (Juo and Kaplan 2004). The APC, also known as cyclosome, is a multi-subunit E3 ubiquitin ligase that promotes mitosis exit and acts in two important steps during that process (Peters 2002; Wade Harper *et al.* 2002). First, APC drives cells from metaphase to anaphase by poly-ubiquitinating securin, the inhibitory binding protein of separase (Cohen-Fix *et al.* 1996; Funabiki *et al.* 1996). Free-form separase can then cleave cohesion proteins present between the sister chromatids, allowing the microtubules to pull them towards opposing poles (Uhlmann *et al.* 1999, 2000). It is responsible for separating paired homologs in meiosis I, and sister chromatids in meiosis II. Second, APC/C is responsible for ubiquitinating M-phase cyclins, also forcing mitotic exit (Golden *et al.* 2000; Peter *et al.* 2001; Taieb *et al.* 2001). In *C. elegans*, exit from meiosis II is complex and involves the direct M to S transition and the conversion of fertilized oocyte into a zygote, which also seems affected in APC mutants (Shakes *et al.* 2003).

Shakes *et al.* (2003) studied seven alleles of *mat-1* mutants that displayed temperature sensitivity and conditional phenotype at semi-permissive temperatures and confirmed meiotic arrest associated with *mat-1* defects, in addition to defective embryo eggshells and asymmetries. This was fundamental for the association of *mat-1* as the ortholog of *CDC27* and its role as part of APC/C. Impairment of *mat-1* resulted in defects in both meiosis and mitosis, including altered pronuclear migration, defective polarization and altered timing and orientation of embryo's mitotic cell divisions. Additionally, the lack of proper cell division in *mat-1(ax144)* and *mat-1(ax212)* temperature-sensitive alleles was shown to cause male tail abnormalities, such as missing or complete absence of rays (Shakes *et al.* 2003). The multicellular embryonic lethality in *mat-1* mutants is also observed in RNAi experiments, in which the mRNA expression is knocked down. The treated worms initially produce viable offspring, but once the RNAi takes effect, these animals

start to produce multicellular dead embryos in the first 20-24 hours, and later generate only one-cell arrested embryos (Shakes *et al.* 2003). Phenotype of *mat-1* RNAi reflected the absence of APC/C function itself, and it replicated the phenotype of knockdown of seven different APC/C subunits (Davis *et al.* 2002; Shakes *et al.* 2003).

Developmental roles for APC/C during embryonic axis formation and anterior/posterior (A-P) polarity during oocyte fertilization have also been identified in *C. elegans* (Rappleye *et al.* 2002). A-P polarization is important to promote asymmetric first mitotic cell division in the one-cell embryo, resulting in formation of sister blastomeres (Shakes *et al.* 2003). APC activity in the A-P formation, induced by CDC-20, promotes restriction of PAR-3 in the anterior pole, allowing PAR-2 to act in posterior end. APC impairment leads to defective association between paternal pronucleus and centrosome with the cortex, which is deemed essential for the restriction of PAR-3 to the anterior pole and successful formation of A-P axis. The role of APC in such association is influenced by separin, and lack of this protein also resulted in defects corresponding to A-P polarity loss (Rappleye *et al.* 2002). *mat-1* mutants show disrupted cytoplasm segregation towards the sperm pronucleus. Since fertilization occurs opposite the oocyte nucleus, these mutant embryos arrest with the oocyte chromatin in meiosis at one end of the embryo and condensed sperm chromatin at the other end (Wallenfang and Seydoux 2000). Oocytes lacking *mat-1* activity can still be fertilized, but their division is arrested in metaphase of meiosis I and they are blocked from further growth (Golden *et al.* 2000). This embryonic arrest is apparently variable: while some mutants can complete meiosis I and still show an abnormal chromosome separation, others skip meiosis II completely and fail to produce normal eggshells.

For analyzing the *mat-1* phenotype, I make use of the temperature-sensitive allele *ye121*, a missense variant at the position I:5,125,258. This allele impacts protein folding, and high

temperatures can destabilize an intermediate in the folding pathway. The *ye121* variant is immediately adjacent to highly conserved regions, and it was isolated in a screen for osmotically sensitive mutants in a study of defects in axis formation (Rappleye *et al.* 2002). In the uterus of *mat1(ye121)* animals, there is a significant increase in the number of embryos that undergo meiotic arrest. At restrictive temperatures (25°C), these mutants can only generate embryos that are arrested at the one-cell stage, and hatch rate is close to 0% in the N2 background. Under permissive temperatures (15°C), these worms have mild complications and carry only 1-2 embryos arrested in meiosis, while still being able to produce viable eggs and having hatch rates close to 100% (Shakes *et al.* 2003).

1.3.2 The Conserved Germline Helicase Gene – cgh-1

cgh-1 is the *C. elegans* ortholog of the human gene *DDX6*, and it is expressed specifically in the germline and early embryo. It enables RNA helicase activity, another essential process, and mRNA binding, and it is relevant for both oocyte and sperm function. Also, it is highly conserved, with about 70% homology to its yeast, human, and *Drosophila* orthologs within the helicase region (Navarro *et al.* 2001). During normal oocyte formation in *C. elegans*, up to half of the developing cells undergo apoptosis, to maintain oocyte homeostasis. This apoptotic process is different from any other in somatic cells and maintains the cytoplasm to be used by the survivor cells (Gumienny *et al.* 1999). However, loss of function of *cgh-1* seems to induce the apoptosis of all these developing oocytes instead and results in non-functional sperm (Navarro *et al.* 2001). With the inhibition of apoptosis, *cgh-1*(RNAi) oocytes can still develop and proliferate normally, however, they cannot be fertilized (Navarro *et al.* 2001). Although *cgh-1*(RNAi) still produce a normal number of germline cells and gonad chromosome morphologies appear normal, these oocytes usually die at high rates (Navarro *et al.* 2001).

Additionally, CGH-1 is localized in the P-granules, germline-specific mRNA compartments, and its inactivation in the germline leads to transformation of P-granules into a rigid matrix (Boag *et al.* 2005; Hubstenberger *et al.* 2013). Its function in the germline relates to sequestering and stabilization of maternally transcribed mRNAs, which is potentially associated with other RNA-binding proteins (Boag *et al.* 2005). Besides, CGH-1 also plays a role in promoting miRISC-target interactions (Hammell *et al.* 2009) and regulation of the endoplasmic reticulum in the germline (Langerak *et al.* 2019). Knock-out of *cgh-1* results in sterility during adulthood, due to abnormal production of oocytes and sperm (Navarro *et al.* 2001; Boag *et al.* 2005). Impairment of *cgh-1* function also strongly enhances the delayed larval development in *mir-48*; *mir-84* mutants, all being suppressed by *hbl-1* RNAi (Hammell *et al.* 2009).

RNAi experiments targeting *cgh-1* conducted by Vu *et al.* (2015) in the N2 reference strain showed embryonic arrest and lower hatching rates in comparison with the CB4856 genetic background. Here, we make use of temperature-sensitive allele to decipher the phenotype of *cgh-1*. The *tn691* temperature-sensitive allele is a missense variant present in *cgh-1* in the position III:7,497,545. It presents a hypomorphic phenotype in the N2 background: under restrictive temperatures, these worms result in embryonic arrest and eggs fail to hatch properly, while hatching rates can be considered normal when observed under permissive conditions (15°C) (Vu *et al.* 2015).

1.3.3 Diverse Genetic Backgrounds for Investigation of Plasticity of Essentiality

Considering that the majority of *C. elegans* essential genes have only been probed in the N2 strain, which is not representative of the broad diversity of *C. elegans* strains, we here expand the study performed by Vu *et al.* (2015) and the investigation of phenotypic variability of *mat-1* and *cgh-1* with the use of six natural isolates (Figure 2). The use of six divergent genetic

backgrounds provides wider scope of natural variations to help with investigation of the plasticity of essentiality, in this case, for *mat-1* and *cgh-1*, potentially explained by the influence of genetic modifiers in these phenotypes. The natural isolates present in this study were chosen based on geographical location, aiming on including the broadest genetic backgrounds as possible: N2 (England), CB4856 (Hawaii), GXW1 (China), AB1 (Australia), JU1400 (Spain), and KR314 (Canada). Hereby, I hope to enlighten the influence of each of these genetic backgrounds in the “essential” function of our two target genes, *mat-1* and *cgh-1*.

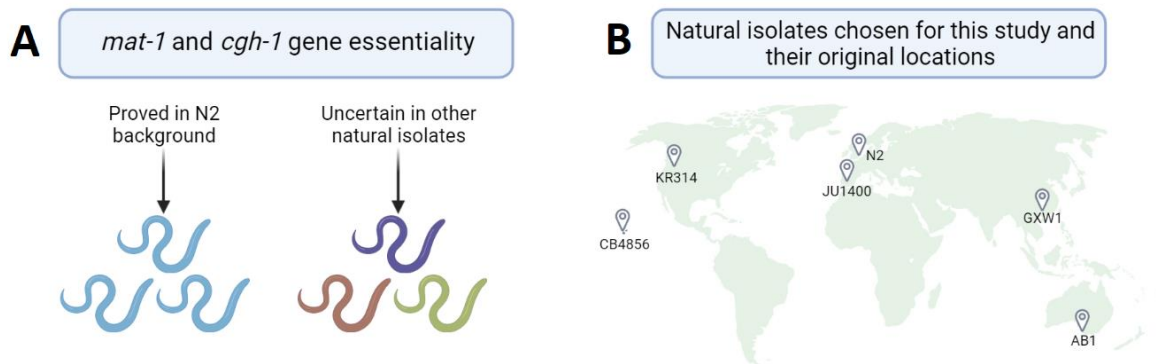


Figure 2. A. gene essentiality of *mat-1* and *cgh-1* in different *C. elegans* genetic backgrounds. B. Original location of the six natural isolates used in this study. We focused on a broad variety of strains to prioritize the diversity of their genetic backgrounds. Image designed using BioRender [<https://www.biorender.com/>].

1.3.3.1 The Reference Strain, N2

The N2 strain is broadly used as a wild-type control for most studies in *C. elegans*. Its strict laboratory development over the past five decades resulted in a stable strain that has not developed many variations in its genome in comparison with other wild isolates. *C. elegans* N2 was the first multicellular organism with a fully-sequenced genome, and this data still remains as the reference genome for this species (The *C. elegans* sequencing consortium, 1998).

1.3.3.2 The Hawaiian Strain, CB4856

The *C. elegans* strain CB4856 (originally called HA8) was isolated from a pineapple tree in Hawaii in 1972 and contains one of the most diverse genomes in comparison with the reference N2 (Stewart *et al.* 2005). CB4856 and N2 differ by one polymorphism per 840 bp (Wicks *et al.* 2001; Swan *et al.* 2002). Based on this high genetic diversity, it is hypothesized that Hawaii may actually be the origin of the *C. elegans* species (Crombie *et al.* 2019). This higher genetic diversity can be explained by remaining ancestral diversity present since the transition of the species to self crossing, abundant supply of habitats – i.e. allowing a more stable population - and stable temperatures throughout the year (Frézal and Félix 2015; Richaud *et al.* 2018).

Additionally, gene flow to other global isolates of *C. elegans* suggests emigration of alleles from specific Hawaiian populations to other regions of the world (Crombie *et al.* 2019). This could be influenced by human activity and recent production and export of sugarcane and tropical fruits initiated in the nineteenth century (Bartholomew *et al.* 2012). In comparison to other genetic backgrounds selected for this project, CB4856 carries 12,285 unique SNVs and 341 unique SVs (Chida 2023). Some interesting phenotypes associated with this wild isolate include its cryophilic behaviour. CB4856 tends to migrate towards the cold area of a temperature gradient, and tendency was higher than N2 and CB4658 strains (Jurado *et al.* 2010). Its survival rates are also higher when compared with N2, AB1 and KR314 wild isolates during cold tolerance experiments (Okahata *et al.* 2016).

1.3.3.3 The Chinese Strain, GXW1

The GXW1 wild isolate was sampled from soil under a kiwi fruit tree in Wuhan City, China in 2010. Its original isolation is credited to A Cutter & E Dolgin, and it represents one of the few Asian wild isolates identified so far. This strain was whole-genome sequenced as part of the

Million Mutation Project (MMP) (<https://cgc.umn.edu/strain/GXW1>). GXW1 srWGS analysis identified 2,139 unique SNVs and 97 unique SVs, when aligned to N2 reference genome and in comparison to the other five natural isolates used in this study (CB4856, N2, AB1, KR314 and JU1400) (Chida 2023).

1.3.3.4 The Australian Strain, AB1

The AB1 strain was isolated in Adelaide, Australia, AB1. Its phylogenetic study suggests its genome is the result of recombination of two diverse genotypes (Barrière and Félix 2005a), and it differs even from other strains that are geographically close, such as AB4 (Koch *et al.* 2000). This is the only wild isolate from south hemisphere included in this study and differs from N2 at a molecular level of 0.4×10^{-3} SNVs per nucleotide. When compared to the other strains in this project, AB1 carries 1,821 unique SNVs and 62 unique SVs (Chida 2023).

1.3.3.5 The Spanish Strain, JU1400

JU1400 was isolated from rotting orange fruits in Sevilla, Spain, in 2008. Its isolation is attributed to Marie-Anne Felix. Its genome contains 6,709 unique SNVs and 144 unique SVs in comparison with the other strains in our study (Chida 2023).

1.3.3.6 The Canadian Strain, KR314

The Canadian wild isolate KR314 was originally harvested in Vancouver (Kitsilano) from rotting avocados in 1987. Its isolation is attributed to Ann Rose. It is known for producing males with abnormal, swollen tails, that still carry all other male-specific structures (e.g. fan, rays, and spicules). This characteristic is associated with a natural variant in the gene *mab-23(e2518)*, suggestively resulting in male mating inability. This represents the capacity of this wild isolate surviving in its natural habitat reproducing entirely by self-fertilization (Hodgkin and Doniach 1997). Its genome contains 4,139 unique SNVs and 97 unique SVs when compared to the other

isolates from my project (Chida 2023).

1.4 Hypothesis and Aims

Hereby, I aim to study the concept of plasticity of essentiality using two essential genes in *C. elegans* important for two unrelated processes: *mat-1* and *cgh-1*. My research goal is to test their essentiality in genetic backgrounds not yet explored and investigate the consistency of such essentiality under the influence of natural genetic variation. Currently, these genes have been extensively studied in the N2 laboratory cultivated background but not in many others natural isolate *C. elegans* backgrounds (with the exception of CB4856). It is my hypothesis that the requirement for these essential processes may be different in diverse genetic backgrounds presumably due to the presence of genetic modifiers. This hypothesis will be addressed via three main aims. Firstly, my objective is to study the phenotype of the *mat-1(ye121)* and *cgh-1(tn691)* in six different genetic backgrounds (Chapter 3). Secondly, I will make use of multi-omics dataset to thoroughly investigate each genetic background in an attempt to identify the potential variants responsible for any given phenotypic variability observed in the first part of this project (Chapter 4). Finally, I will apply CRISPR/Cas9 methods to probe the modifying effect of selected variants identified in the genomes of the natural isolates as potential candidate modifiers (Chapter 5). Altogether, the assessment of the phenotype of *mat-1* and *cgh-1* and its modification may collaborate to the interpretation of plasticity of essentiality in *C. elegans*, as well as assist with tailoring the identification of genetic modifiers.

Chapter 2: Materials and Methods

2.1 *C. elegans* Culture and Maintenance

C. elegans strains were maintained following standard laboratorial conditions and kept at 15°C unless noted otherwise (Brenner 1974). Nematode Growth Media (NGM) spread with a lawn of OP50 *Escherichia coli* was used for routine worm growth. For maintaining worm populations, around five young adults were placed onto fresh plates every seven days. Strains carrying temperature-sensitive alleles were periodically unfrozen fresh from stocks every three months for minimizing the appearance of spontaneous *de novo* variants. All strains were genotyped frequently to assure the quality of produced data using a Polymerase Chain Reaction (PCR) strategy. The markers for genotyping were based on SVs unique to each background and identified after thorough investigation by Afiya Chida (Chida 2023). Genotyping information and PCR details can be found in Table 1.

Table 1. PCR-based genotyping strategy for natural isolates. For primer sequences, please refer to Table 4.

Strain	SV	Position	Forward	Reverse	Size	T _M
AB1	Inversion	V:16025076-16122556	oMTG1178	oMTG1179	693bp	62
CB4856	Deletion	IV:10207379-10208603	oMTG1154	oMTG1155	981bp	62
GXW1	Tandem Duplication	I:6123787-6128960	oMTG1291	oMTG1292	866bp	62
JU1400	Tandem Duplication	II:8243775-8249999	oMTG1287	oMTG1288	451bp	62
KR314	Tandem Duplication	IV:17044133-17049698	oMTG1168	oMTG1169	853bp	62
N2	Deletion	V:1645712-1647498	oMTG1331	oMTG1332	432bp	62

2.2 CRISPR/Cas9 Design and Microinjection

For CRISPR design, guide RNAs (gRNAs) were selected based on proximity with the target site and predicted efficiency given by the UCSC genome browser [genome.ucsc.edu]. Selected sequences were ordered from IDT as Alt-R CRISPR-Cas9 gRNA. For repair template, single stranded oligo donor nucleotides (ssODNs) were designed using Horizon Discoveries tool

[<https://horizondiscovery.com/en/products/tools/Edit-R-HDR-Donor-Designeroligo>] and ordered as a 4 nM Ultramer DNA Oligo from IDT. For CRISPR/Cas9 transfection, a well-established microinjection approach was used. L4 larvae were plated a day before and let for growth at 20°C until young adulthood was reached. Young adults were injected targeting the rachis, which contains premature germline cells. The system would take effect and show phenotype in the F1 generation, in a heterozygote setup. CRISPR injection mix contained 0.5 uL of Cas9 (Integrated DNA Technologies; ID: 1081058), 1.25 uL of each 10 uM pre-annealed gRNA and tracrRNA, and 1 uL of each ssODN, all eluted into duplex buffer and RNase-free water. CRISPR constructs and targets used here can be found on Table 2.

Table 2. CRISPR/Cas9 targets and components for homology directed repair. Underlined sequence refers to gRNA site. Blue nucleotides represent synonymous variants induced to avoid CRISPR/Cas9 secondary activity. Red represents variant of interest (mutant).

Candidate Variant	Position	gRNA (3'-5')	Repair template (3'-5')
<i>such-1</i>	III:11517531	ACTGTGTATTCAGGATATGA	AGACTACGAACTACACGCTTTTTTGGTGTCAATGTC CAAGTTTTT <u>ACTGTGTATCCAAGACATGATGGATGG</u> GAAATTCCTCAAgtaggcagccaa
<i>such-1</i> (Hawaiian)	III:11517714	CAAAAAGAAGAGGAATCGG	GCTGCGTGAACCGAATTCTGCTGGACACGACGCTCC TTTTCT <u>ACTGATTCCTCTCTTTTTTGCCTGAAATTAT</u> TATCAT
<i>hbl-1</i>	X:5824753	GTTGTTTCAAACGCTATCAA	TGCATTCTCCATTTCGACTGCTTTTCAGGTGACGGCAT GTCTCC <u>CTTAATAGCGTTTGAAACA</u> ACTCCTAATGCG GAAGCCAA
<i>nhl-2</i>	III:4897421	CTTCGTTAACTTGTGATTTA	TCGAAGAGGCGCGAATCCGTCAAGAGTCTAACATTG ATGACATGTACAGAAAATCACAAGTTAACGAAGCAC GAGTACATGATGCCATCAAC
<i>bub-3</i>	II:13728035	GTTTGTGGTGTTCGTCGATCG	GGGCGGTGCAGTTCCTTTCCGACTGGCGAAGCGTTTGT <u>CCTGTCTTCTATCGAGGGAAGAGTCGCTGTAGAGTAT</u> GTAGATCAG
<i>ddx-15</i>	III:5591654	GCGGACGCGGAGACAGAAGA	TTATTATTATATTGAACTGGGGAAGCAACAATTTTAC TTACCTTGTCTGTCTCCGCGTCCGCTGCCATCTAAAT CCAAT
<i>glp-1</i> (Hawaiian)	III:9097961	TTCAATTCGTTGTAATCACC	GTGCGAACGTCTCAGATTGAAAGTCAAAGTTCAATGC <u>ATTATAGTCACCGGATATCCATTTGCGAGAGTTTGTG</u> GATA

A co-CRISPR technique was used to help with the identification of the worms that were successfully transfected, using *dpy-10(cn64)* as the secondary target. This affects the organismal morphology and creates a visible phenotype to screen for potential heterozygotes for the gene of interest. This additional variant is dominant and presents as “dumpy” phenotype in homozygous worms and a “roller” phenotype in heterozygous worms. Successful CRISPR worms were identified in the F1 generation having the roller phenotype. Rollers were then plated individually and genotyped for the variant of interest with PCR and Sanger sequencing. When a heterozygous worm for the target gene was identified, its F2 generation is also genotyped to isolate the ones homozygous for the variant of interest and wild type for the *dpy-10* phenotype.

Genotyping PCR was performed using NEB OneTaq (New England BioLabs, ID: M0480). Sanger sequencing was performed by the CHGI (The Centre for Health Genomics and Informatics) at the University of Calgary, using the forward flanking primer for each variant. F2 generation was also genotyped to isolate the ones homozygous for the variant of interest and wild type for the *Dpy-10* phenotype. Finally, the new, confirmed strain was assigned a strain name and frozen to ensure maintenance of stocks for future experiments. Strains in the Tarailo-Graovac lab are given a strain name beginning with “MTG” (i.e., MTGxxx) and alleles isolated in the lab are denoted with *mod* (i.e., *bub-3(mod495)*). Details for genotyping can be found on Table 3.

Table 3. PCR genotyping strategy and primer sequences for CRISPR-created strains.

Strain	Gene	Primer Name	Primer Sequence (3'-5')	Size (bps)	T _M
MTG667	<i>such-1</i>	oMTG1045	GAAGAGGAATCGGTGGAAAACG	493	58.5
		oMTG1046	TGTCCAGTAGAAGCCGAAGACG		
MTG649	<i>nhl-2</i>	oMTG1049	CGGCTTGTAAGCAATGCCTGG	707	60
		oMTG1050	CACCTCCTGGTTGAGAACCCTCC		
MTG687	<i>such-1 (Hawaiian)</i>	oMTG1251	TGGCTGCCTACTTGAAGAATTTC	856	59.5
		oMTG1273	TCACAGGTGAACGTGGACGATATG		
MTG685	<i>hbl-1</i>	oMTG1252	ATGCTGTCTCTTCCACCTCTACAAC	472	60

		oMTG1253	CACTGAAATGGCTTGCTCCCTATG		
MTG728	<i>bub-3</i>	oMTG1528	GTTACAACAAAGTAGCGTTCGGAGG	516	60
		oMTG1529	TCAGCTCAGTTCATCAGTATCCTTC		
MTG729	<i>ddx-15</i>	oMTG1531	TCCGACAAGAGTGATGCACTGG	868	59
		oMTG1532	GGCGAGTCTGAGAGCCATTC		
MTG730	<i>glp-1 (Hawaiian)</i>	oMTG1541	GGAAGTGTGTCGCTGGTGTTATC	571	60
		oMTG1542	GCCTTCAAGCATCGAATTGCCTC		
MTG758	<i>Y73B6A.1</i>	oMTG1580	TGAGAAGTCTTACGCGGAAATGG	899	60
		oMTG1581	AATGGATTCTTGTGCTGCTGGTG		

2.3 Hatch Rate Assays

Experiments were conducted at either permissive or restrictive temperatures. For each hatching rate trial, five L4 worms per strain were plated onto individual seeded NGM plates. L4s were selected for being the last stage before adulthood. This guarantees the observance of the entire fertile stage of each worm and its entire brood. Every 24 hours, the parental worm was moved into a fresh plate and the number of F1 eggs and larvae was assessed. This was followed until the end of reproduction stage (usually after 4-5 days) and three trials per strain at each temperature were conducted. Hatch rate was calculated for each worm (n=5) as the ratio of the total number of hatched embryos to the total number of progeny (eggs and larvae). Experiments were conducted in a range between 15°C and 25.3°C, depicting the phenotype under both permissive and restrictive conditions. To ratify that the phenotype observed under restrictive conditions was uniquely due to presence of temperature-sensitive alleles, experiments with the wild type natural isolate strains were performed under the same restrictive conditions (23°C, 24°C, 25°C, and 25.3°C).

2.4 Titration/Propagation Assay

Titration assays were performed to test a gradient of temperatures and identifying the temperature at which one strain could be long-term propagated while other(s) cannot. For each assay, five L4 worms per strain were plated onto individual NGM plates. Plates were incubated

for two weeks, which was a sufficient time to discern ability of the strain to propagate next generations. After that period, plates were scored for propagation and classified into three categories: “No” propagation was considered if only presence of parental worm and unhatched eggs was observed. “Mild” propagation was considered if <10 F1 generation worms were present, and most progeny consisted of unhatched eggs. True propagation (“Yes”) was considered if there was the presence of several worms, multiple generations, and little or no unhatched eggs were observed. Considering the use of temperature-sensitive alleles in this project, propagation assays were performed at different temperatures in a range from 15°C to 25.3°C, to expand the phenotype inquiry.

2.5 RNAi Treatment

RNAi by feeding was performed using *Y110A7A.d (mat-1)* and *C07H6.5 (cgh-1)* clones provided by Dr. David Hansen’s lab and using clones carrying the empty L4440 plasmid as controls. For validation of clones, DNA was purified using the Sigma GenElute Plasmid Miniprep Kit (ID: PLN350) and confirmed through Sanger sequencing before the start of the experiment. RNAi bacteria was grown with selection media using ampicillin (100 mg/mL) and tetracycline (2.5mg/mL) until log phase was reached and spread onto NGM plates. NGM media contained 1 M IPTG and 100 mg/mL ampicillin and were kept away from ultraviolet (UV) light. Seeded RNAi plates were stored at 4°C for up to one week. L4 worms were used for the hatch rate assay conducted at 20°C. L4s were kept on plate for 72 hours for RNAi induction, then moved onto a fresh plate for another 24 hours. Parental worm was then removed from replica and progeny was scored for phenotype.

2.6 DNA Extraction and WGS Long-Reads Preparation

Genomic DNA for long-reads WGS was extracted using Wizard® HMW DNA Extraction

Kit (ID: A2920) following standard protocol. DNA was eluted in DNA Rehydration Solution provided in the kit and its concentration measured. Samples with a minimum A260/280 ratio of 1.8 were submitted for sequencing at Sick Kids in Toronto (PacBio - Single-molecule real-time (SMRT) developed Pacific BioSciences) [<http://tcag.ca/facilities/dnaSequencingSynthesis.html>]. Obtained data was analyzed by Afiya Chida. Using Integrative Genomes Viewer (IGV), the screening considered if the SV is affecting protein-coding regions of genes that have similar function to *mat-1* or *cgh-1* or that are known to interact with them. Variants that affect protein production potentially have a stronger modifier effect. PCR reactions were used to confirm the presence of the candidate SVs, and Sanger sequencing was applied for further confirmation, as described below.

2.7 Structural Variants Validation

Breakpoints of SVs and complex rearrangements were confirmed by PCR and Sanger Sequencing. Primers were designed with PrimerQuest IDT tool (<https://www.idtdna.com/PrimerQuest>) by Afiya Chida, an MSc graduate from our lab. Sanger sequencing was performed by the CHGI at the University of Calgary. Primer sequences and PCR conditions are listed on Table 4.

Table 4. Primers utilized for SV validation in different *C. elegans* natural isolate strains.

Strain Name	Location	Primer IDs	Primer Sequence (5'-3')	Size (bps)	T _M
CB4856	II:10006742-10016704	oMTG1148	TCGTTCTCTACTCGTCTCTCTC	897	62
		oMTG1149	GGCACAGATCTCGAACCTAAA		
CB4856	II:14034295-14034709	oMTG1150	CGAGTAGCAGCATTGGGTAGAG	938	62.5
		oMTG1151	GCGATACGTGTACTGTGTAACC		
CB4856	IV:10207379-10208603	oMTG1154	CAATATGGGTCCGCAGGTATTA	981	62
		oMTG1155	TGCTTTGTCAGCCGTACTION		
CB4856	V:15452708-15457000	oMTG1156	GCTTCCGCGCTACTTAGAAA	451	62
		oMTG1157	GGCGGGACTCTATCTAGTCATA		

CB4856	V:10380363-10384404	oMTG1158	CCCAATCATGCTCTCAATCCT	807	62
		oMTG1159	CTGGACTGAATGGATTGTCTCTC		
CB4856	I:13012535-13023893	oMTG1160	CCCGAATCCTGTGAAATGAAATG	375	62
		oMTG1161	CGGTAGCTGTTGTCATAGTTGT		
CB4856	V:15862955-15867242	oMTG1162	GAAACAAAGCTTCAGGCACAA	650	62
		oMTG1163	AACTGCGCAGAGAGAATCAA		
CB4856	V:16679981-16684986	oMTG1164	CCCTTGCGAACTGAACATTAAC	872	62
		oMTG1165	AGTTTCCTGCGAATGAGAGATAG		
CB4856	II:14938828-14939477	oMTG1166	AACACTCTCTCCCTATCCCTTC	662	62.5
		oMTG1167	CCCGGAAGTCCTATTCATTTC		
KR314	IV:17044133-17049698	oMTG1168	CATAGCCCTGGTACCCAATAAA	853	62
		oMTG1169	GAGTGGCGTGCGGATATAAA		
AB1	V:17172415-17175843	oMTG1176	CACCTCCACAACAAGCAAAC	447	62
		oMTG1177	GCTCATATCTAAGCAGCCGATAG		
AB1	V:16025076-16122556	oMTG1178	TCAGTCAATCCAAGACCATCAC	693	62
		oMTG1179	GCGAGTCGCAAAGCATTATC		
AB1	V:1085800-1088470	oMTG1180	GGAAACTGGCAATGCTGTTATAG	799	62
		oMTG1181	CCTGAAGATCCATCGGAACAA		
JU1400	III:13333471-13339461	oMTG1285	GTTCCAAAGCTCTGCAAGAAAT	732	62
		oMTG1286	CTGGTAGGCACGTCTATGTTT		
GXW1	II:9404532-9410508	oMTG1289	AAAGGGAGCGAAGCAGAAA	817	62
		oMTG1290	CCTGACCACAGTTCATTGGATA		
GXW1	I:6123787-6128960	oMTG1291	AGGAGAACTGGAACGAGAATG	866	62
		oMTG1292	GCGTGAAGTGAGAGTGTGAA		
GXW1	IV:1124616-1130844	oMTG1293	GGAGGAGGAGGTTATGGAAATG	501	62
		oMTG1294	CCAAGTGAAGTGTGCTCAGA		

2.8 DNA Extraction and WGS Short-Reads Preparation

Genomic DNA for short-reads WGS (srWGS) was extracted using Qiagen Blood and Tissue kit (ID: 13323) following standard procedure. DNA was suspended in 10 mM Tris-HCl (pH 8.0), and its concentration measured. Samples with a minimum A260/280 ratio of 1.8 were submitted for sequencing at Sick Kids in Toronto (Illumina NovaSeq 6000) [<http://tcag.ca/facilities/dnaSequencingSynthesis.html>].

2.9 srWGS Analysis and Search for CRISPR Off-Target Activity

WGS data was analyzed using an in-house developed pipeline (Maroilley *et al.* 2023).

Quality control was done with FASTQC v0.11.9 (S. 2010). Read sequences were trimmed with Trimmomatic v0.38 (Bolger *et al.* 2014a) and were aligned with BWA-MEM v0.7.17 (Li and Durbin 2009) using the Wormbase *C. elegans* genome version WS265 as the reference (Li and Durbin 2009; Bolger *et al.* 2014b; Maroilley *et al.* 2021). FreeBayes v1.3.1 (Garrison and Marth 2012), and SNPeff v4.3.1 (Cingolani *et al.* 2012) were used for variant calling and annotation, respectively, and Bcftools v1.10 (Danecek *et al.* 2021) was used to manipulate variant calls in the Variant Call Format (VCF).

The output Candidate Variant Lists (CVLs) were manually curated for confirmation of presence of the temperature-sensitive alleles *mat-1(ye121)* and *cgh-1(tn691)* and searched for potential CRISPR off-target effects. Final CVLs contained variants with high confidence that were prioritized based on its *in silico* predicted type (e.g., missense, splice-region variant, stop/start, gain/loss variant) and its gene function and correlation with genes of interest in this study.

2.10 Variant Curation Using Machine Learning

An in-house Machine Learning (ML) approach developed by a current PhD student at the Tarailo-Graovac lab, Tahsin Rahit, was used for prioritizing the variants. The ML, named ModSpy, was used to analyse genome data, identify variants unique to each background, and rank the variants according to a final “probability score” (y -score) of its modifying activity for the target gene. The parameters considered for variant prioritizations by the ML model include protein interaction data from STRING database, gene ontology, and phenotype ontology, which is based on experimentally validated genetic modifiers data used for model training.

After final CVL was obtained, I performed manual curation for further evaluation of the best candidates, focusing on the strains in which phenotypic experiments described in Chapter 3 indicated presence of modifiers. Initial CVL lists contained hundreds of variants for each

background. To minimize the list, I first filtered out any variants that were not truly unique to each background through comparison of CVLs and visual confirmation in IGV. Variants not supported by >90% of reads may end up absent in some CVLs and therefore be considered unique to a specific strain, which can be ruled out after the visualization step. Next, I initially searched for variants that could be acting as suppressors in the KR314, and enhancers in the CB4856 and AB1, based on phenotype presented for *mat-1(ye121)* or *cgh-1(tn691)* alleles (Chapter 3). Backgrounds containing each of the candidate modifiers studied here can be found in Table 11. Selection criteria included association with target gene (gene function and gene ontology listed on WormBase), impact of variant (Grantham score), and y-probability score given by the ML model.

2.11 RNA Extraction and Sequencing

Worm strains were cultured on NGM plates spread with OP50 *E. coli*, at permissive temperature (15°C). Strains were synchronized using a hypochlorite + NaOH solution in a 1:1 ratio. Plates were fed with additional OP50 culture until crowded with adult worms. Synchronized plates were collected into pellets using M9 solution and kept at -80°C for later RNA extraction. For this step, a Trizol-chloroform based RNA extraction protocol was applied. In-column DNase treatment was also performed using ZYMO Research RNA Clean & Concentrator kit. (ID: R1013). Purified RNA was eluted in DNase-RNase free water provided in the kit and concentration was measured using NanoDrop. For facilitating the extraction step, samples were split into three batches (see Table 5), and one N2 WT control was included in each batch. Samples were sent for total mRNA-sequencing at Sick Kids in Toronto (Illumina NovaSeq 6000) [<http://tcag.ca/facilities/dnaSequencingSynthesis.html>].

Table 5. mRNA-sequencing samples concentration and ratios. Samples were split into three batches for facilitating RNA extraction.

Batch#	Strain	MTG#	ng/ul	A260/280	A260/230
1	<i>N2; mat-1(ye121)</i>	MTG561	1046.4	2.20	2.53
	<i>AB1; mat-1(ye121)</i>	MTG562	671.0	2.17	2.47
	<i>CB4856; mat-1(ye121)</i>	MTG563	843.0	2.18	2.53
	<i>GXW1; mat-1(ye121)</i>	MTG564	973.7	2.19	2.49
	<i>JU1400; mat-1(ye121)</i>	MTG565	420.7	2.14	2.42
	<i>KR314; mat-1(ye121)</i>	MTG566	1009.9	2.21	2.53
	N2	MTG7-M	696.6	2.17	2.4
2	<i>N2; cgh-1(tn691)</i>	MTG558	825.2	2.18	2.52
	<i>AB1; cgh-1(tn691)</i>	MTG567	555.6	2.19	2.47
	<i>CB4856; cgh-1(tn691)</i>	MTG568	896.3	2.17	2.48
	<i>GXW1; cgh-1(tn691)</i>	MTG569	690.6	2.15	2.45
	<i>JU1400; cgh-1(tn691)</i>	MTG559	100.1	2.10	2.51
	<i>KR314; cgh-1(tn691)</i>	MTG570	976.3	2.18	2.52
	N2	MTG7-C	533.3	2.21	2.52
3	AB1	MTG23	176.1	2.11	2.36
	CB4856	MTG8	1077.2	2.15	2.43
	GXW1	MTG24	1044.5	2.17	2.42
	JU1400	MTG472	673.7	2.17	2.43
	KR314	MTG25	971.6	2.16	2.38
	N2	MTG7	1346.2	2.17	2.45

2.12 RNA-Seq Data Analysis

Processing of RNA-seq data was performed by Dr. Tatiana Maroilley, a current post-doctoral fellow in the Tarailo-Graovac lab, on Galaxy (<https://usegalaxy.eu/> - (Afgan *et al.* 2022)), using the guidelines previously described by Batut *et al.* (2021) to execute the quality control, genome alignment and read counts, as Quality control was done with FastQC (Galaxy Version 0.74+galaxy0) (S. 2010). MultiQC was also used to aggregate the FastQC reports (Ewels *et al.* 2016). Low quality sequences and adaptors were trimmed using Trimmomatic (Galaxy Version 0.39+galaxy0) (Bolger *et al.* 2014a). Alignment to reference *C. elegans* genome (WS286) and read

counts was performed with RNA STAR (Galaxy Version 2.7.10b+galaxy4) (Dobin *et al.* 2013). Final .bam files were visualized using IGV for pictured assessment of RNA expression. Significance (*p*-value) of potential different gene expression was assessed using limma (Galaxy Version 3.50.1+galaxy0), available on Galaxy platform (Liu *et al.* 2015). This was performed by comparing each natural isolate against N2 control. limma converts read counts into log₂-counts-per-million (logCPM) and models the mean-variance relationship between samples. Output contains adjusted *p*-values using Benjamini and Hochberg's false discovery rate control at a threshold value of 0.05 (BH 1995) (Benjamini and Hochberg 1995). All three replicates of each background (WT, *mat-1(ye121)* and *cgh-1(tn691)*) were grouped for increasing statistical power.

2.13 Grantham Scores

For assessing the impact of missense variants and the amino acid changes, specifically, I also made use of Grantham Scores. Grantham scores evaluate the chemical properties of each amino acid, including its polarity and molecular volume, and attributes a score for every specific change. The higher the score, the more “damaging” the substitution. Lowest score involves isoleucine/leucine substitutions (score = 5), while highest involves a cysteine/tryptophan substitution (score = 215). Table for score verification can be found in the original source (Grantham 1974).

Chapter 3: *mat-1* and *cgh-1* Display Phenotypic Variability In Different Natural Isolate Backgrounds

As previously reported in literature, essentiality of *mat-1* and *cgh-1* does not seem to be static. Instead, RNAi showed that depletion of gene products for both of these genes resulted in lower embryonic lethality in the CB4856 (Hawaiian) strain when compared to the lab cultivated N2 strain (Vu *et al.* 2015). These two “essential” genes, initially considered of utmost importance for survival, have shown to be partially dispensable in one background (CB4856), while crucial in the other (N2) (Vu *et al.* 2015). To understand this better, I decided to expand on this study by using temperature sensitive alleles of the *mat-1* and *cgh-1* and six strains of different genetic backgrounds.

3.1 RNA interference for *mat-1* and *cgh-1*

To confirm the previous findings, I first used the RNAi with clones for both target genes in the N2 and CB4856 backgrounds. RNAi by feeding was performed using *Y110A7A.d (mat-1)* and *C07H6.5 (cgh-1)* clones provided by Dr. David Hansen’s lab and using clones carrying the empty L4440 plasmid as controls. L4 worms were used for the hatch rate assay conducted at 20°C. In accordance with the earlier report presented by Vu et al (2015), N2 showed reduced embryonic survival for both *cgh-1* and *mat-1* clones, while CB4856 presented viable development. Hatch rates for CB4856 strain was close to 100% and recapitulated the WT results; meanwhile, N2 presented a hatch rate of less than 2% for *mat-1*, and close to 78% for *cgh-1* (Figure 3). The results also shed light on the phenotype of *mat-1* and *cgh-1* knockdown in the other wild isolates, for which literature lacked information. Interesting enough, knockdown of *cgh-1* with RNAi shows variable phenotype among the six natural isolates. Even though RNAi can be biased for not offering a complete gene impairment, these results could indicate a potential effect of genetic

background on both essential genes. On the other hand, divergent response to RNAi in different genetic backgrounds has already been observed in studies comparing N2 and CB4856 strains. It is known that penetrance of RNAi relies strongly on the target gene, however, the exact mechanisms behind such changes remain unclear (Bell *et al.* 2023). Further in this project, the phenotypic variability of *mat-1* and *cgh-1* was studied using temperature sensitive alleles, which allowed for thorough investigation of the lethal phenotypes associated with the deficiency of these genes.

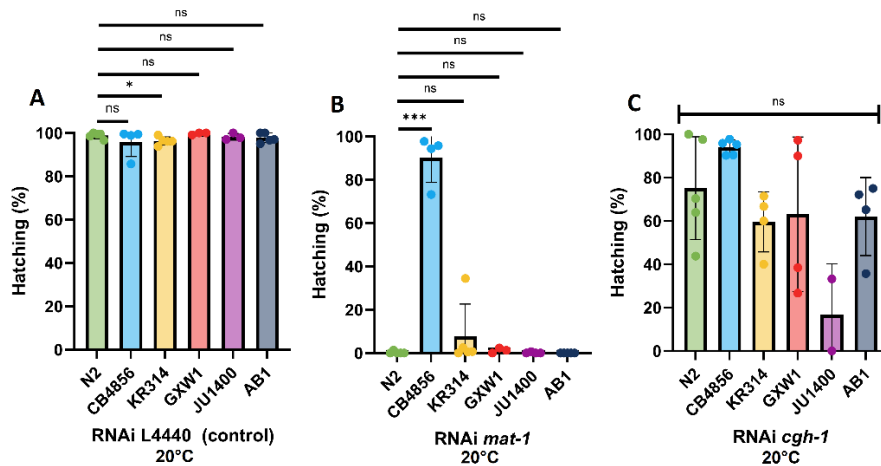


Figure 3. RNAi results for comparison of knockdown of *mat-1* (B) or *cgh-1* (C). For control, the L4440 empty vector was used, and all strains showed good viability, as expected (A). Statistical significance was calculated using unpaired t-tests with Welch’s correction. *p*-values are shown using “*” (0.01 to 0.05 - Significant), “***” (0.001 to 0.01 – Very significant), “****” (0.0001 to 0.001 – Extremely significant), or “*****” (<0.0001 – Extremely significant). Non-significant results are represented by “ns”.

3.2 The Use of Temperature-Sensitive Alleles for Phenotype Investigation

Considering that the target genes of this work have been deemed crucial for survival, I here use temperature-sensitive alleles, which provide a conditional phenotype. Traditional approaches used to study gene function often involve the creation of loss-of-function alleles. However, these methods may prohibit the identification of subtle differences in the phenotype not observable under such extreme perturbations. In contrast to a complete loss of function, temperature-sensitive alleles allow the observation of a wider phenotypic spectrum, which may point to mild modifiers that might not be identified otherwise. Temperature-sensitive alleles were first identified in bacteria,

and usually contain missense point variants (Edgar *et al.* 1964). At a lower, permissive temperature, protein would still be functional, however, once put into restrictive conditions, the higher temperatures are capable of impairing protein function, often by disrupting folding and 3D structure. In other words, the protein produced under restrictive conditions is mostly inactive (Edgar *et al.* 1964). To examine the differences in essentiality of *mat-1* and *cgh-1* across various genetic backgrounds, I employed *mat-1(ye121)* and *cgh-1(tn691)* temperature-sensitive alleles in six diverse *C. elegans* backgrounds and tested their viability at both permissive (15°C) and restrictive (>22°C) conditions.

3.2.1 Knock-in of Temperature-Sensitive Alleles in the Six Natural Isolates

An initial step of this project was performed by Dr. Xiao Li, a postdoctoral fellow in the Tarailo-Graovac Lab, who prepared the strains carrying the temperature-sensitive alleles for *mat-1* and *cgh-1*, *mat-1(ye121)* and *cgh-1(tn691)*, respectively. Using CRISPR/Cas9, Dr. Li performed the initial knock-in of the temperature-sensitive alleles in the reference strain N2 along with 5 other natural isolates for comparison of their phenotypes across these different genetic backgrounds. The N2 background is consistently used as a reference genome in *C. elegans* experiments and is applied as control due to its minimal genetic variation. Using CRISPR/Cas9, Dr. Li was able to recreate the temperature-sensitive hypomorphic variants in all six geographically diverse strains (Table 6).

Table 6. Strains carrying temperature-sensitive alleles for *mat-1* and *cgh-1* created by Dr. Xiao Li using CRISPR/Cas9. Strains received a unique MTG# ID for in-house reference upon its creation but will be referred to in this document by their wild isolate name and genotype, to facilitate reading.

	<i>mat-1 (ye121)</i>	<i>cgh-1 (tn691)</i>
N2	MTG561 - N2; <i>mat-1(ye121)</i>	MTG558 - N2; <i>cgh-1(tn691)</i>
AB1	MTG562 - AB1; <i>mat-1(ye121)</i>	MTG567 - AB1; <i>cgh-1(tn691)</i>
CB4856	MTG563 - CB4856; <i>mat-1(ye121)</i>	MTG568 - CB4856; <i>cgh-1(tn691)</i>
GXW1	MTG564 - GXW1; <i>mat-1(ye121)</i>	MTG569 - GXW1; <i>cgh-1(tn691)</i>
JU1400	MTG565 - JU1400; <i>mat-1(ye121)</i>	MTG559 - JU1400; <i>cgh-1(tn691)</i>
KR314	MTG566 - KR314; <i>mat-1(ye121)</i>	MTG570 - KR314; <i>cgh-1(tn691)</i>

3.3 Analyzing the Phenotype of *mat-1(ye121)* and *cgh-1(tn691)* across different genetic backgrounds

3.3.1 Hatch Rate Assays

After the knock-in of the temperature-sensitive alleles in the six natural isolates, the viability of these strains at restrictive and permissive temperatures was assessed by primarily analyzing their hatch rates and embryonic arrest, in the so-called Hatch Rate Assays: five worms per strain were individually placed onto NGM plates and checked daily for quantification of progeny; finally, hatch rate per worm was calculated (n=5) and average hatch rate per strain was assessed. Full description of the experiment can be found in the Materials and Methods Chapter 2. When comparing the different natural isolates, divergencies in these ratios are a strong indicative of plasticity of essentiality possibly due to variants in their genetic background. To confirm that any given phenotypic variability was due to each genetic background, and little influenced by external factors, it must be considered that strains were analysed simultaneously and under the exact same conditions (incubation temperature and time, agar plates' batch, food nutrition, etc). Additionally, to ratify that the phenotype observed under restrictive conditions was uniquely due to presence of temperature-sensitive alleles, experiments with the WT natural isolate strains were performed under the same restrictive conditions (23°C, 24°C, 25°C, and 25.3°C) and showed no significant decrease in hatch rates (Figure 4). Therefore, any reduction in fitness observed in the knock-in strains for *mat-1(ye121)* or *cgh-1(tn691)* is likely explained by the decrease in function of the genes induced by temperature rising. Additionally, any variation in the lethal phenotype across the strains is likely attributed to their unique genetic background.

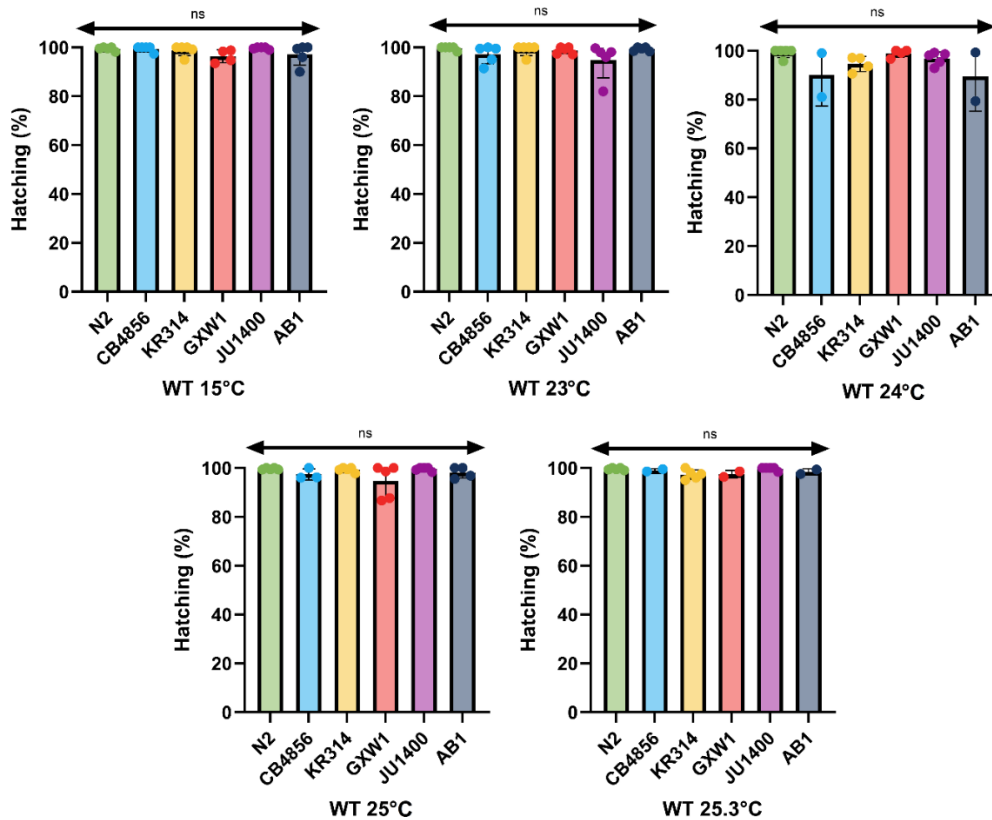


Figure 4. Hatch rates for WT isolates under restrictive conditions (23°C, 24°C, 25°C, and 25.3°C, respectively), showing no significant (ns) impairment of survival. Non-significant results are represented by “ns”.

3.3.1.1 Hatch rates at restrictive temperatures reveal phenotypic variability for *mat-1(ye121)*

At lower temperatures, it is expected that protein is functional and little to no impairment on survival is observed. This was noted for hatch rate assays performed for *mat-1(ye121)* strains at 15°C. When grown at this permissive condition, all six strains showed little or no embryonic lethality and hatch rates were close to 100% (Figure 5). However, once temperature is increased, I observed considerable phenotypic variability for *mat-1(ye121)* strains. Experiments were performed at restrictive conditions ranging from 22.4°C to 24°C.

Findings at 22.4°C confirmed phenotypic variability for *mat-1(ye121)* strains (Figure 5). At this condition, GXW1; *mat-1(ye121)* and JU1400; *mat-1(ye121)* have moderate reduction in

hatch rates, presenting 85% and 65% hatch rates, respectively. This effect was shown to be stronger for CB4856; *mat-1(ye121)* and AB1; *mat-1(ye121)* strains, both depicting 41% and 21% hatch rates, respectively. KR314; *mat-1(ye121)* strain was the one with best hatch rate, close to 100%, however, this was followed by N2; *mat-1(ye121)* (96%), showing no significant difference between these two backgrounds. The significant impairment of CB4856; *mat-1(ye121)* and AB1; *mat-1(ye121)* in comparison with control N2; *mat-1(ye121)* could suggest presence of enhancers in those genetic backgrounds, that may or may not be a common variant.

Slightly increasing temperature to 23°C recapitulated most of those results (Figure 5). As expected, all hatch rates were reduced, in accordance with the effects of temperature in the knocked-in allele. This time, GXW1; *mat-1(ye121)* and JU1400; *mat-1(ye121)* had even more severe phenotype, with average hatch rates of 8% and 4%, respectively. Again, CB4856; *mat-1(ye121)* (1.8%) and AB1; *mat-1(ye121)* (0.44%) presented the lower hatch rates, corroborating the idea of a common enhancer present in both backgrounds. N2; *mat-1(ye121)* and KR314; *mat-1(ye121)* also presented reduced hatch rate, however milder than the other strains (Figure 5). However, this specific temperature revealed significant difference between these two strains. When compared to N2 background, KR314; *mat-1(ye121)* had a much better hatch rate, which could indicate the presence of a unique modifier (potentially a suppressor) in this strain (Figure 5).

Finally, at 24°C, hatch rates also strongly suggest the presence of genetic modifiers in these backgrounds (Figure 5). Interestingly, GXW1; *mat-1(ye121)* and AB1; *mat-1(ye121)* showed no survival, while CB4856; *mat-1(ye121)* and JU1400; *mat-1(ye121)* presented less than 6% of hatching. The strain with the best hatch rate was KR314; *mat-1(ye121)*, followed by N2; *mat-1(ye121)*. Given that hatch rates for N2; *mat-1(ye121)* and KR314; *mat-1(ye121)* are significantly different, this could indicate the presence of a suppressor of *mat-1* in the KR314 background.

Additionally, KR314; *mat-1(ye121)* strain was the only one of the six strains which showed propagation abilities at this same restrictive condition, meaning it was able to reach adulthood and reproduce into further generations (as described in the section 3.3.2.1 of this chapter). On the other hand, the strong lethality observed in the other backgrounds might be an indicator of presence of enhancers in those strains, producing a more severe phenotype.

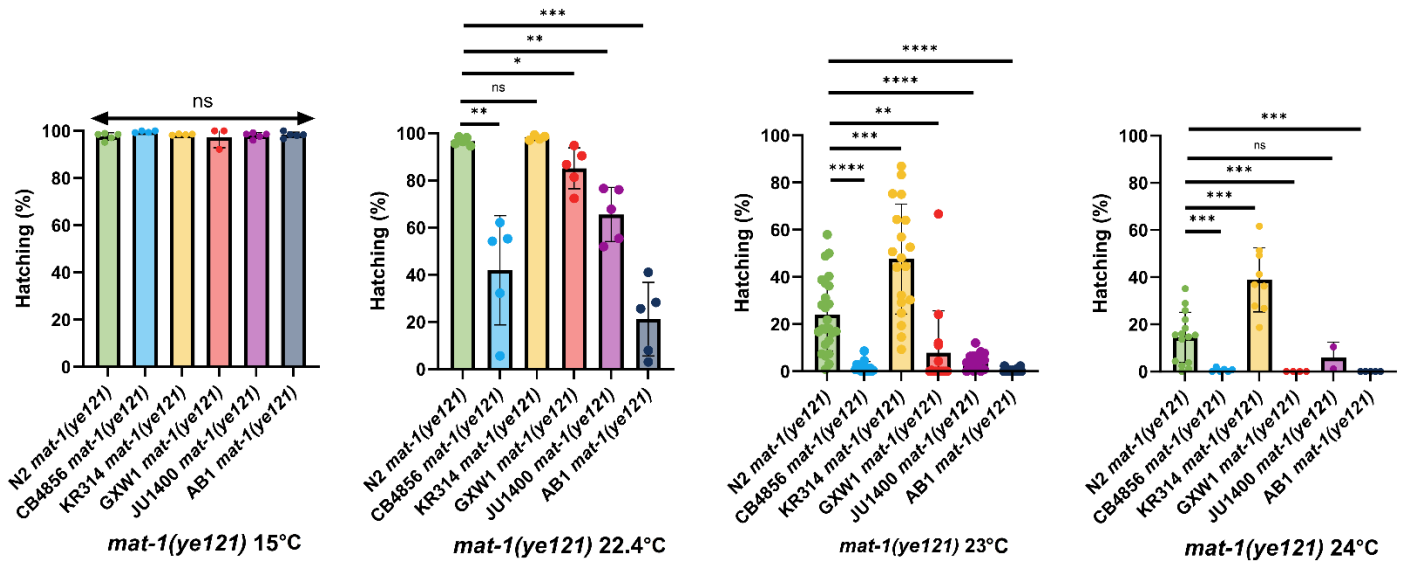


Figure 5. Hatch rates for *mat-1(ye121)* strains depict phenotypic variability for embryonic lethality. Phenotype is not observed under permissive conditions (15°C, left), in which egg hatching is close to 100%. As temperature is risen, lethal phenotype becomes more evident and is greatly variable according to each genetic background. Statistical significance was calculated using unpaired t-tests with Welch’s correction. *p*-values are shown using “*” (0.01 to 0.05 - Significant), “***” (0.001 to 0.01 – Very significant), “****” (0.0001 to 0.001 – Extremely significant), or “*****” (<0.0001 – Extremely significant). Non-significant results are represented by “ns”.

Impairment of survival of CB4856; *mat-1(ye121)* was a surprising result considering RNAi targeting the same gene and background showed opposite phenotype. The idea that an off-target effect of CRISPR/Cas9 could have caused the divergent phenotype in the CB4856; *mat-1(ye121)* strain was first considered, but then disregarded after thorough investigation (detailed description on Chapter 4). Simply, the nature of each genetic impairment (RNAi knockdown versus missense point variant) might be the potential cause for such discrepancy. Natural genetic variation is found

for nearly any measurable trait in *C. elegans*, including when using RNAi. Response to RNAi is known to vary according to the genetic background it acts on. In CB4856, specifically, RNAi seems strongly ineffective in targeting germline genes: loss of function of *ppw-1* was deemed responsible for blocking (or at least delaying) that response (Tijsterman *et al.* 2002). A recent study compared differential gene expression after knockdown of two germline-essential genes (*par-1* and *pos-1*) across strains with variable RNAi sensitivity, including CB4856, and observed strong strain-specific phenotypes (Bell *et al.* 2023). Additionally, wild isolates appear to vary in efficacy of germline RNAi depending on the target gene, but the detailed mechanisms remain to be elucidated.

3.3.1.2 Hatch rates at restrictive temperatures reveal phenotypic variability for *cgh-1(tm691)*

To start with, phenotype of *cgh-1(tm691)* strains was assessed at both permissive and restrictive conditions. Once placed at permissive condition of 15°C, all six strains showed little or no impairment of survival with hatch rates of over 90% (Figure 6). From there, lethal phenotype was assessed in a range of different restrictive conditions from 23°C to 25.3°C (Figure 6). Phenotypes observed at 23°C showed severe reduction of hatch rate for most strains, due to perturbation of *cgh-1* at this condition: phenotypic variability was noted, and all strains but KR314; *cgh-1(tm691)* presented less than 16% of average hatch rate. KR314; *cgh-1(tm691)*, having a hatch rate of 45%, showed significant difference in comparison to N2; *cgh-1(tm691)* (9.8%), a potential sign of presence of suppressors in the first background.

A similar trend was observed when increasing temperature to a more restrictive condition of 25°C. AB1; *cgh-1(tm691)*, CB4856; *cgh-1(tm691)*, and JU1400; *cgh-1(tm691)* had hatch rates of less than 6%, while GXW1; *cgh-1(tm691)* showed no hatch rate. N2; *cgh-1(tm691)* presented average 1.2% of hatch rate, while KR314; *cgh-1(tm691)* was the strain with the best performance,

even if mildly reaching 7%.

Finally, a slightly higher temperature of 25.3°C also recapitulated the phenotype observed at other restrictive conditions: all strains showed severe impairment of hatch rate, being best hatch rate still attributed to KR314; *cgh-1(tn691)* (9%). All other strains presented less than 6% of hatch rate, and CB4856; *cgh-1(tn691)* specifically, showed a very similar hatch rate in comparison with N2; *cgh-1(tn691)* (3.8% and 3.5%, respectively). Altogether, results obtained here confirmed the phenotypic variability of *cgh-1* across the six natural isolates and could also suggest the presence of suppressors in the KR314 background, or even intricate interaction networks of both suppressors and enhancers present in these strains. Further fitness of the strains was assessed using propagation assays and brood size analysis, as described later in this chapter.

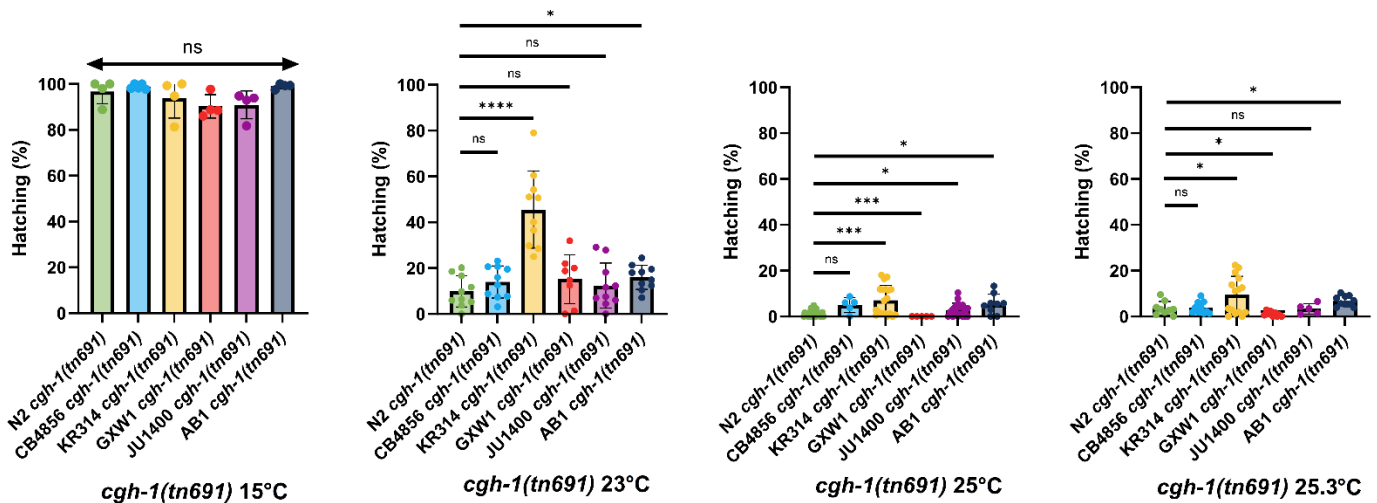


Figure 6. Hatch rates for *cgh-1(tn691)* strains depict phenotypic variability for embryonic lethality. Phenotype is not observed under permissive conditions (15°C, left), in which egg hatching is close to 100%. As temperature is risen, lethal phenotype becomes more evident and is greatly variable according to each genetic background. Statistical significance was calculated using unpaired t-tests with Welch's correction. *p*-values are shown using "*" (0.01 to 0.05 - Significant), "***" (0.001 to 0.01 - Very significant), "****" (0.0001 to 0.001 - Extremely significant), or "*****" (<0.0001 - Extremely significant). Non-significant results are represented by "ns".

3.3.2 Long-Term Propagation Assays

Considering that gene essentiality not only refers to genes indispensable for survival, but also to ones that when disturbed may cause impairment in reproduction, a second approach was used to further study the phenotype of *mat-1(ye121)* and *cgh-1(tn691)*. The plasticity of essentiality, in this case, can be further investigated by performing propagation assays: long-term observance of population growth and survival. In this experiment, worms are incubated for a two-weeks period, without disturbance, which is enough time to assess presence versus absence of propagation. After that period, plates are scored for its capacity to propagate into further generations and classified into three categories: “No” (only presence of parental worm and unhatched eggs observed). “Mild” (<10 F1 generation worms) “Yes” (presence of several worms, multiple generations, and little or no unhatched eggs observed) (Figure 7). Considering the use of temperature-sensitive alleles in this project, propagation assays were performed at different temperatures in a range from 15°C to 25.3°C, to expand the phenotype inquiry. Full description of experiment can be found in the Materials and Methods Section 2.4.

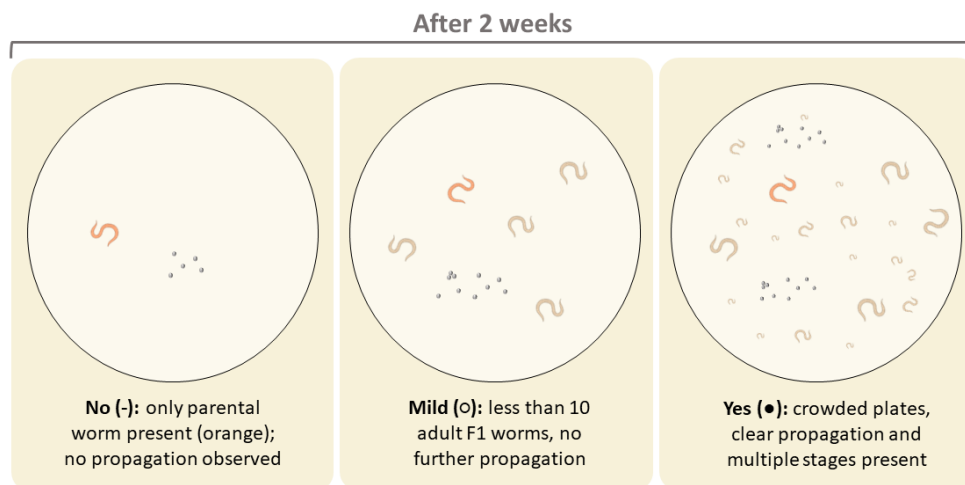


Figure 7. Propagation assay. After final observation, plates are classified into three categories according to the level of propagation present.

To serve as a control and certify that phenotype for propagation assays was a product of the hypomorphic variants in *mat-1* and *cgh-1* here being studied, the same conditions were applied to the wild-type natural isolates. Overall, as depicted in Table 7, all six wild isolates showed propagation abilities in every temperature that they were tested under.

Table 7. Propagation of wild-type strains at restrictive temperature. Propagation observed after two weeks of uninterrupted incubation at correspondent temperature. Results were classified accordingly as “No propagation” (-), “Mild propagation” (○), and “True propagation” (●). Notably, the higher the temperature, the lower the fitness and propagation rate of each strain.

		Temperature						
		15°C	20°C	21.7°C	23°C	24°C	25°C	25.3°C
Genetic Background	AB1	●	●	●	●	●	●	●
	CB4856	●	●	●	●	●	●	●
	GXW1	●	●	●	●	●	●	●
	JU1400	●	●	●	●	●	●	●
	KR314	●	●	●	●	●	●	●
	N2	●	●	●	●	●	●	●

3.3.2.1 Propagation of *mat-1(ye121)*

The long-term propagation of *mat-1* was tested in a range of temperatures varying from 20°C to 25.3°C. At the lowest end, at 20°C, all *mat-1(ye121)* strains presented crowded plates and were able to propagate. Once temperature was slightly risen to 21.7°C, however, impairment of hatch rate and propagation could be observed for AB1; *mat-1(ye121)*, while others remained viable. In accordance with hatch rate assays, KR314; *mat-1(ye121)* was the strain that predominantly had the best propagation observed throughout different restrictive temperatures, also overcoming propagation of N2; *mat-1(ye121)*. Both KR314; *mat-1(ye121)* and N2; *mat-1(ye121)* were able to propagate at 23.2°C, with the presence of crowded plates after a two-week period, while other backgrounds failed to expand. Once the temperature was slightly increased to 23.7°C, N2; *mat-1(ye121)* seemed to lose its propagation capacity, and only KR314; *mat-1(ye121)*

presented propagation, even if mild. At this temperature, KR314; *mat-1(ye121)* plates presented few F1 adult worms, but no further growth, while all remaining strains did not show any signs of propagation. A similar result was observed at the slightly higher condition of 24°C: KR314; *mat-1(ye121)* showed mild propagation, while most other backgrounds were completely impaired. However, KR314; *mat-1(ye121)* was still considered as the best performance, since all the plates for this background showed consistent results. Finally, a higher temperature of either 25°C or 25.3°C seemed to severely impact all *mat-1(ye121)* strains, and no growth was observed in any of the six backgrounds. This was taken as the absolute restrictive temperature, in which *mat-1* function was fully hindered, or at the minimum, not functional enough to allow survival of these organisms (Table 8).

Table 8. Propagation of *mat-1(ye121)* strains at restrictive temperature. Propagation observed after two weeks of uninterrupted incubation at correspondent temperature. Results were classified accordingly as “No propagation” (-), “Mild propagation” (○), and “True propagation” (●). Notably, the higher the temperature, the lower the fitness and propagation rate of each strain. N/A represents strains not tested at the specific condition noted.

		Temperature							
		20°C	21.7°C	23°C	23.2°C	23.7°C	24°C	25°C	25.3°C
Genetic Background	AB1 <i>mat-1(ye121)</i>	●	-	-	-	-	-	-	N/A
	CB4856; <i>mat-1(ye121)</i>	●	●	●	-	-	-	-	N/A
	GXW1; <i>mat-1(ye121)</i>	●	●	●	-	-	-	-	N/A
	JU1400; <i>mat-1(ye121)</i>	●	●	●	-	-	-	-	N/A
	KR314; <i>mat-1(ye121)</i>	●	●	●	●	○	○	-	-
	N2; <i>mat-1(ye121)</i>	●	●	●	●	-	-	-	-

3.3.2.2 Propagation of *cgh-1(tm691)*

For *cgh-1*, propagation was tested in a variety of restrictive temperatures ranging from 20°C to 25.3°C. Primarily, at temperatures equal or under 23.7°C, all *cgh-1(tm691)* strains were deemed capable of propagating after observation of crowded plates after the two-week period. This includes 20°C and 21.7°C experiments, as depicted in Table 9. Once the temperature was raised

to 24°C, KR314; *cgh-1(tn691)* was able to propagate, while others only retained mild propagation capacity. At 25°C, AB1; *cgh-1(tn691)*, N2; *cgh-1(tn691)*, JU1400; *cgh-1(tn691)*, GXW1; *cgh-1(tn691)*, and CB4856; *cgh-1(tn691)* did not present any propagation, while KR314; *cgh-1(tn691)* still showed mild propagation. Similarly, at the final temperature of 25.3°C, only KR314; *cgh-1(tn691)* strain was able to propagate into a mild stage, while all others remained fully compromised (Table 9).

In sum, the analysis of both hatch rates and propagation of *mat-1(ye121)* and *cgh-1(tn691)* phenotypes here shown confirm the phenotypic variability of these conditional alleles in different genetic backgrounds. The presence of natural and unique variations in each of these backgrounds may be the key to understanding such phenotypic divergence and is the topic of discussion in the next chapters of this thesis. With the use of bioinformatics tools and multi-omics datasets, I have investigated these six wild isolates and attempted to identify potential modifiers responsible for enhancing – or suppressing – the phenotypes here discussed.

Table 9. Propagation of *cgh-1(tn691)* strains at restrictive temperature. Propagation observed after two weeks of uninterrupted incubation at correspondent temperature. Results were classified accordingly as “No propagation” (-), “Mild propagation” (○), and “True propagation” (●). Notably, the higher the temperature, the lower the fitness and propagation rate of each strain.

		Temperature						
		20°C	21.7°C	23°C	23.7°C	24°C	25°C	25.3°C
Genetic Background	AB1; <i>cgh-1(tn691)</i>	●	●	●	●	○	-	-
	CB4856; <i>cgh-1(tn691)</i>	●	●	●	●	○	-	-
	GXW1; <i>cgh-1(tn691)</i>	●	●	●	●	○	-	-
	JU1400; <i>cgh-1(tn691)</i>	●	●	●	●	○	-	-
	KR314; <i>cgh-1(tn691)</i>	●	●	●	●	●	○	○
	N2; <i>cgh-1(tn691)</i>	●	●	●	●	○	-	-

Chapter 4: Genomics Analyses Identify Candidate Modifier

Variants

4.1 Whole Genome Sequencing Analysis for CRISPR/Cas9 Off-Target Effects

CRISPR/Cas9 tends to be extremely precise due to the use of specific guide RNAs; little or no off-target effect is usually observed for gene editing in *C. elegans* (Iyer *et al.* 2015; Au *et al.* 2019). However, our hatching rates obtained for CB4856; *mat-1(ye121)* under restrictive conditions did not match the RNAi data described in the literature (Vu *et al.* 2015). Even considering that our RNAi experiment followed the previously reported results, the possibility that CRISPR/Cas9 caused off-target editing during the knock-in of the temperature-sensitive allele was considered as a possible explanation for the observed divergence. To investigate this issue, we submitted the *mat-1* and *cgh-1* CRISPRed strains for WGS Illumina srWGS. Sample preparation was performed with the Qiagen Blood & Cell Culture DNA Mini Kit for DNA extraction. An in-house user-friendly Galaxy-based pipeline was used for the analyses of the WGS data (Maroille *et al.* 2023). Detailed methods description is found in Chapter 2 of this thesis. Off-target CRISPR effect could be identified by comparing the WGS data from CRISPRed strains to the ones obtained for wild-type natural isolate strains.

The variants were filtered focusing on the ones present in protein-coding regions. Finally, comparing the filtered variants to the CRISPR construct design allowed the discerning between the ones that could have been caused by off-target editing and the others that are product of spontaneous events. The variants initially cured from this analysis were not further investigated as potential off-target effects, either for being in a gene with no known relation to *mat-1* or *cgh-1*, or for being present in hypervariable regions (HVRs) of the genome (Chida 2023). A search for these variants in the WT background of each strain helped with filtering the ones that are simply natural

variations or potential effect of spontaneous variants. A search for variants affecting genes known to interact with *mat-1* and *cgh-1* also did not lead to any major candidates. Additionally, any variant present in other backgrounds or strains treated with a different CRISPR construct were also disregarded as potential sequencing errors. Finally, variants present in HVRs was disregarded as these regions are difficult to map and lead to low-quality reads. No variant that could potentially be attributed to an off-target CRISPR/Cas9 effect was identified. Therefore, the differential phenotype of CB4856; *mat-1(ye121)* in comparison to RNAi experiments for *mat-1* is more likely explained by the divergent molecular effects caused by these two methods. Considering that the use of temperature-sensitive alleles for *mat-1* was not explored by Vu et al (2015), the phenotype here observed for *mat-1(ye121)* allele in the CB4856 background could not be compared to previous literature.

Even though no clear off-target effect for *mat-1(ye121)* could be detected, a different outcome was seen for *cgh-1(tm691)* strains. Initially, three missense SNVs in N2; *cgh-1(tm691)* were identified and further investigated. These are the variants present in genes *zip-6*, *ZK546.5*, and *soap-1*. The first one, *zip-6* (IV:11,215,167G > T) caused an amino acid change from glutamine to lysine, having a Grantham score of 53 (conservative). This gene has a regulatory interaction with *hsp-3*, *hbl-1* and five other genes that physically interact with *cgh-1*. The second variant, in *ZK546.5* (II:4,937,417 C > A), also caused an amino acid change from glutamine to lysine, having the same Grantham score of 53 points. This is an essential gene involved in transcription. However, this variant was present in both N2; *cgh-1(tm691)* and *mat-1(ye121)* backgrounds, therefore likely due to spontaneous event rather than CRISPR off-target effect, considering that both strains were edited using different constructs/targets. No direct interaction of *ZK546.5* with *mat-1* or *cgh-1* was therefore found. The third SNV, however, was concerning,

considering that *soap-1* is an essential gene with common interactions with *cgh-1* (Figure 8B). This missense variant was present on chromosome V (V:7,207,103 A > G), causing a leucine to serine amino acid change with Grantham score of 145 (radical), and it was present exclusively in the N2; *cgh-1(tn691)* background (Figure 8A). The nature of this variant does not correspond to potential CRISPR/Cas9 off-target effects, considering that small insertions and deletions are usually the hallmark of endonuclease-induced (double-strand break) damage (Iyer *et al.* 2015). A similar approach using WGS for identifying off-target effects was applied to mice and human stem cell clones and found a very low incidence of true off-target editing, insufficient to create any concerns (Veres *et al.* 2014; Iyer *et al.* 2015).

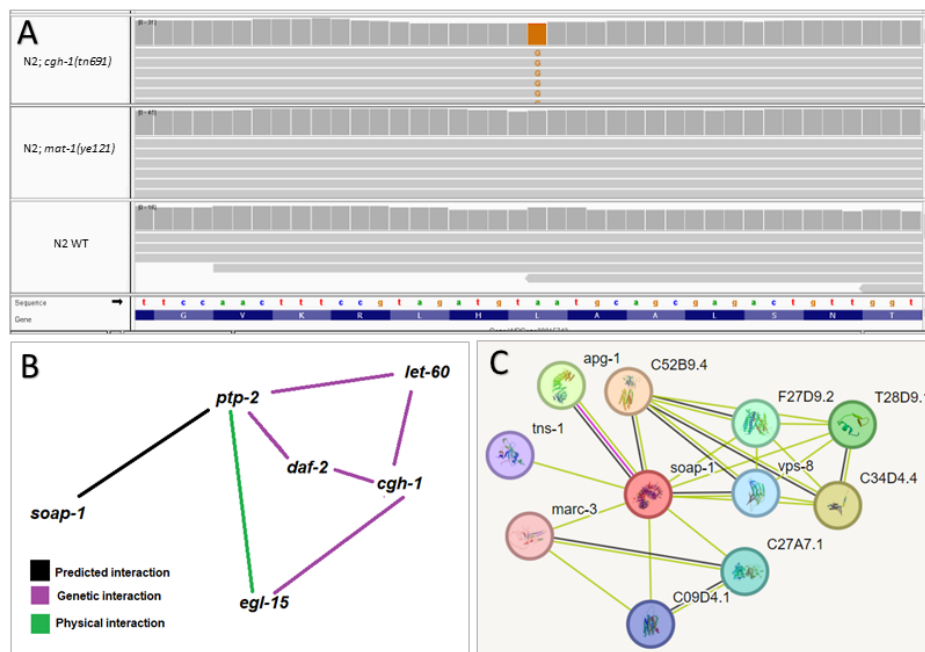


Figure 8. *soap-1* (V:7207103 A > G) information. (A) depicts IGV screenshot confirming presence of this variant exclusively in the N2; *cgh-1(tn691)* background. (B) depicts summarized interactions of *soap-1* and its indirect linkage to *cgh-1*. (C) STRING *soap-1* protein interactions search (<https://string-db.org/> Accessed on September 5th, 2023).

However, this missense variant in *soap-1* became of interest due to its gene essentiality and potential impact on the phenotype that was studied in this project. To rule out the possibility of its influence in the *cgh-1(tn691)* phenotype, a new N2 strain, not containing the *soap-1* variant, was

CRISPRed for the *cgh-1* allele. After careful analysis using hatch rate assays at both permissive and restrictive temperatures (15°C, 23°C, 25°C, and 25.3°C) (Figure 9), no significant difference was detected between the N2; *cgh-1(tn691)* either carrying the variant in *soap-1* (MTG558) or not (MTG709). Therefore, no phenotypic association between *soap-1* (V:7,207,103 A > G) and *cgh-1(tn691)* was confirmed. More likely, the SNVs present in *zip-6*, *ZK546.5*, and *soap-1* are due to spontaneous events.

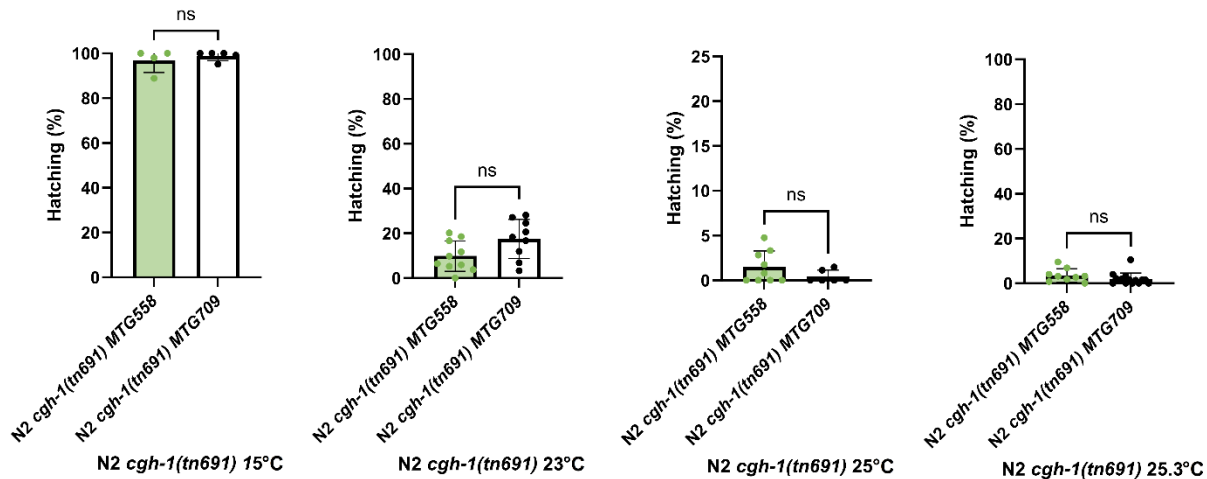


Figure 9. Investigation of *soap-1* as a potential interaction with *cgh-1* and influence on the *cgh-1(tn691)* phenotype. Little or no significance difference was observed between the N2; *cgh-1(tn691)* strain carrying the *soap-1* variant (MTG558) or the newly created N2; *cgh-1(tn691)* strain without the presence of *soap-1* variant (MTG709). This confirmed the lack of association between this specific SNV in *soap-1* and the phenotype observed in N2; *cgh-1(tn691)*. Statistical significance was calculated using unpaired t-tests with Welch's correction. Non-significant results are represented by "ns".

4.2 WGS Investigation of Genetic Modifiers for *mat-1* and *cgh-1*

The data here presented supports our hypothesis that detrimental variants in *mat-1* and *cgh-1* result in variable outcomes in different genetic backgrounds (Chapter 3). By comparing the hatching rates from the assays outlined in Chapter 3, I noted that there is a clear difference in survival between these natural isolates, supporting the presence of genetic modifiers. Based on the phenotypes observed, candidate modifiers were searched for using WGS data from the six natural

isolates, considering primarily variants that could act as a suppressor based on similar function or direct interaction with our target genes. Then, these candidate modifiers were investigated using CRISPR/Cas9 to test their capacity to alter the phenotype of *mat-1(ye121)* and *cgh-1(tn691)* (Figure 10). Importantly, part of my project also involved identifying and testing structural variants (SVs) as possible modifiers for the two target genes. The identification of these variants and analysis of WGS data was performed by Afiya Chida, an MSc graduate from our lab, who had a key role in this project (Chida 2023).

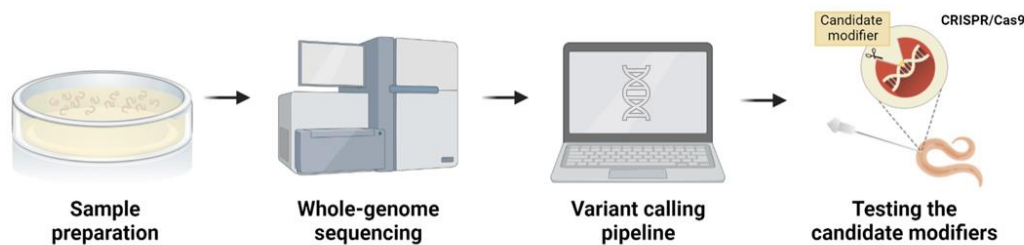


Figure 10. Genetic modifiers identification workflow.

4.2.1 Prioritization of SNVs as Candidate Modifiers for *mat-1* and *cgh-1*

4.2.1.1 Initial Bioinformatics Analysis

Afiya Chida applied computational methods to better understand the natural isolate genomes and identify variants in genes that may be functionally associated with *mat-1* and *cgh-1*. srWGS was performed for all six natural isolate strains, and initial analysis was focused on SNVs (Chida 2023). Back in 2009, Dixon et al underlined how more complex variants (such as deletions or overexpression of genes) are perhaps an uncommon form of genetic modifiers, and that the study of simpler genetic interactions between single nucleotide polymorphisms (SNPs) could provide a more true-like view of genetic networks (Dixon *et al.* 2009). Here, the variants were prioritized based on their association with our target genes, similar function, and their location,

focusing on non-synonymous variants within the protein coding region predicted to result in “moderate” to “high” effects - a classification based on the “physical” impact of each variant in the final protein.

Afiya Chida’s analysis led to an extensive list of variants (Chida 2023). For instance, an initial analysis of the CB4856 strain resulted in over 22,000 variants. Targeted analysis, based on known/previously reported genetic interactions (WormBase) for *mat-1* or *cgh-1* led to the identification of four variants in genes related to *mat-1* or *cgh-1* (Table 10) that we decided to investigate further. All the candidate variants are missense. To test whether these variants have any effect on the *mat-1*(*ye121*) and *cgh-1*(*tn691*) phenotypes, I made use of CRISPR/Cas9 in the N2 background containing the corresponding temperature-sensitive allele (for *mat-1* or *cgh-1*) and assessed any changes to the phenotypes at both the permissive and restrictive temperatures. To do so, a new hatching rate assay was performed, and its results compared to the original background and the other natural isolates that originally carried the variant at hand. If the candidate modifier acts as a suppressor, the new phenotype should present an increase in hatch rate and propagation. Conversely, an enhancing activity should show the opposite phenotype – an increase in the embryonic lethality. If no differences in the phenotype are observed, then the candidate gene can be considered as a non-modifier of the primary variant.

Table 10. Variants of interest, related to *mat-1* and *cgh-1*, identified in initial WGS and bioinformatics analysis.

GOI	Candidate Genetic Modifier	Allele ID	Type of Variant	Amino acid change
<i>mat-1</i>	<i>such-1</i> (III:11517531 C>G)	<i>mod476</i>	Missense Variant Moderate Impact	Cys > Ser
	<i>such-1</i> (III:11517714 G>C)	<i>mod483</i>	Missense Variant Moderate Impact	Asn > Lys
<i>cgh-1</i>	<i>hbl-1</i> (X:5824753 C>T)	<i>mod482</i>	Missense Variant Moderate Impact	Thr > Met
	<i>nhl-2</i> (III:4897421 T>G)	<i>mod473</i>	Missense Variant Moderate Impact	Glu > Asp

4.2.1.2 Machine Learning Application for Prioritization of Candidate Modifiers

Subsequently, in addition to the initial bioinformatics analysis performed by Afiya Chida, the search for modifiers was expanded with the use of machine learning (ML). For facilitating the search for potential candidate modifiers, an in-house ML model developed by a PhD student at Tarailo-Graovac lab, Tahsin Rahit, was used. The model, named as ModSpy (Rahit et al., manuscript in preparation), analyses gene ontology and genetic interactions according to a target gene and results in a list of prioritized variants. The final list of variants contained a probability score of each variant being a modifier for the target gene. Candidate variants were selected based on both the ML predictions and position in the ranking, and further information gathered in WormBase. Detailed analysis can be found in Chapter 2.10. I initially searched for variants that could be acting as suppressors in the KR314, and enhancers in the CB4856 and AB1, based on phenotype presented for *mat-1(ye121)* or *cgh-1(tn691)* alleles (Chapter 3). Selection criteria included association with target gene (gene function and gene ontology listed on WormBase), impact of variant (Grantham score), and y-probability score given by the ML model. This led to prioritization of one candidate enhancer and one candidate suppressor for each target gene, as depicted in Table 11. Backgrounds containing each of the candidate modifiers studied here can be found in Table 12.

Table 11. Candidate modifiers for *mat-1* and *cgh-1* identified using machine learning approach.

GOI	Candidate Genetic Modifier	Allele ID	Type of Variant	Amino acid change
<i>mat-1</i>	<i>bub-3</i> (II:13728035 G>C)	<i>mod495</i>	Missense Variant Moderate Impact	Val > Leu
	<i>Y73B6A.1</i> (IV: 6663170 C>T)	<i>mod516</i>	Missense Variant Moderate Impact	Ala > Thr
<i>cgh-1</i>	<i>ddx-15</i> (III:5591654 C>G)	<i>mod496</i>	Missense Variant Moderate Impact	Arg > Thr
	<i>glp-1</i> (III:9097961 A>C)	<i>mod497</i>	Missense Variant Moderate Impact	Glu > Ala

After careful selection, the variants were tested with the use of CRISPR/Cas9 (further discussed in Chapter 5), followed by phenotype assessment using the same methodology as given for *mat-1* and *cgh-1* phenotype investigation (Chapter 3). If the candidate variant is influencing the phenotype of target gene, an increase (suppression) or decrease (enhancement) of the lethal phenotype would be observed. Therefore, if double mutant carrying target allele and candidate modifier simply replays the phenotype of the single mutant, the secondary variant would be disregarded as a modifier for the primary one.

Table 12. Candidate modifier prioritized in this study and their presence or absence in each of the six natural isolates' backgrounds AB1, CB4856, GXW1, JU1400, KR314 and N2. "■" represents the backgrounds originally carrying the correspondent variant.

			Natural Isolate Background						
			AB1	CB4856	GXW1	JU1400	KR314	N2	
Gene of interest	<i>mat-1</i>	Candidate modifier	<i>such-1(mod476)</i>	■	■	■	■	■	
			<i>such-1(mod483)</i>		■				
			<i>bub-3(mod495)</i>					■	
			<i>Y73B6A.1(mod516)</i>	■	■				
	<i>cgh-1</i>		<i>hbl-1(mod482)</i>		■				
			<i>nhl-2(mod473)</i>				■	■	
			<i>ddx-15(mod496)</i>					■	
			<i>glp-1(mod497)</i>		■				

4.3 SVs as Potential Candidate Modifiers for *mat-1* and *cgh-1*

Given the origin of the natural isolate strains, I considered the possibility that modification of *mat-1* and *cgh-1* phenotypes might not only be caused by SNVs, but also by SVs, variants in the genome of at least 50 nucleotides (Feuk *et al.* 2006). It is known that these larger variants, such as insertions, deletions, duplications, inversions and translocations, may also play a role as genetic modifiers, even though these are less represented in literature (Wright *et al.* 2018; Rylaarsdam and Guemez-Gamboa 2019). That being said, our bioinformatics team aimed on analysing the genomic

data obtained through srWGS, and long-reads WGS (lrWGS) (PacBio –SMRT developed Pacific BioSciences) to define the SVs in these natural isolate strains. When searching for larger variants using srWGS, although task was mostly successful, some regions that displayed substantial hypervariability were encountered, intervening with the confident identification of SVs as described by Lee *et al.*, (2021). Although CeNDR provides WGS data (aligned reads in .bam files) for this species, those can lack information on more complex variants (SVs), due to manipulation of reads in the .bam files. Therefore, we applied lrWGS for a deeper investigation on these genomes as an experimental validation strategy for whatever was detected with srWGS, and to increase chances to explore HVRs.

Samples were prepared for long read sequencing using the Promega Wizard® HWM DNA extraction kit and following the protocol suggested by the manufacturer. lrWGS was performed for all the WT natural isolates used in this study (detailed information listed on Chapter 2.6). Afiya Chida performed variant analysis to prioritize the ones that affect the genes that could act as modifiers, considering if the SV is affecting protein-coding regions of genes that have similar function to *mat-1* or *cgh-1* or that are known to interact with them. Variants that affect protein production potentially have a stronger modifier effect. PCR reactions were used to confirm the presence of the candidate SVs, and Sanger sequencing was applied for further confirmation. Altogether, this work helped identifying SVs unique to each of these wild isolates and further enriched the understanding of the species and its natural variation (Chida 2023). Overall, analyses of the SVs have not identified any variants that affect genes that could be deemed as genetic modifiers of *mat-1*; however, a complex genomic rearrangement (CGR) in *alg-2* present in CB4856 strain, may be of interest as a potential modifier for *cgh-1*. This gene, *alg-2*, is involved in miRISC complex and interacts with *cgh-1*, *nhl-2* and *hbl-1* (Hammell *et al.* 2009). Exploration

of this SV present in *alg-2* was initially considered to be investigated using RNAi. However, RNA-seq data analysis still showed regular gene expression of *alg-2* in CB4856 and no significant difference in comparison with expression in N2, making its exploration impracticable. *p*-values for *alg-2* and IGV visualization of gene expression can be found in Table 13 and Figure 11, respectively.

Table 13. *p*-values for *alg-2* expression from AB1, CB4856, KR314, GXW1 and JU1400 in comparison with N2 control. Values are not significant and suggest no relevant differential gene expression for these genes in the aforementioned genetic backgrounds.

	<i>p</i> -value against N2 Control				
	AB1	CB4856	KR314	GXW1	JU1400
<i>alg-2</i>	0.06	0.66	0.27	0.61	0.34

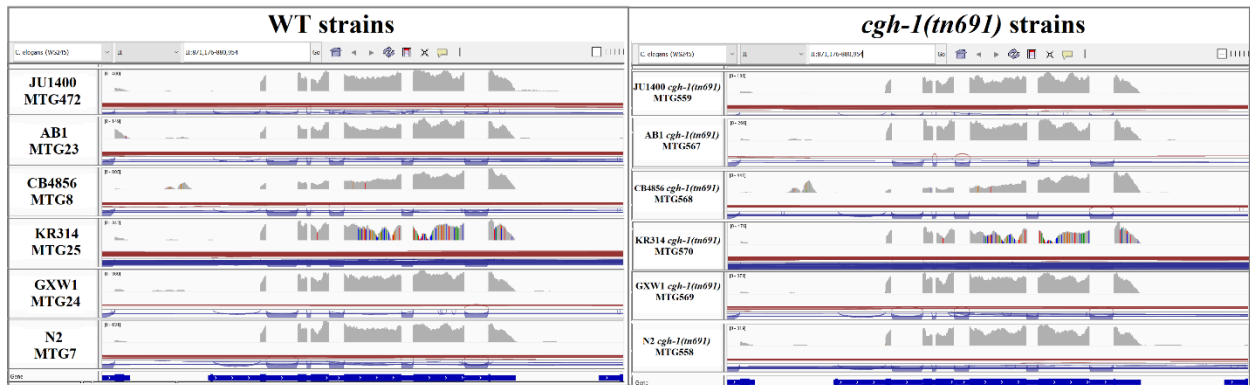


Figure 11. IGV visualization of gene expression of *alg-2* in JU1400, AB1, CB4856, KR314, GXW1 and N2 WT strains (left) and *cgh-1(tn691)* strains (right). Peaks (gray) represent reads coverage within exon and reflect the mRNA expression in each region. Absence of peaks represent lack of reads coverage. Variants are colorful and color-coded according to the type of variant or nucleotide change. No significant difference in gene expression was observed in IGV, either comparing the natural isolates against each other, or comparing each natural isolate before and after CRISPR knock-in of *cgh-1(tn691)*.

4.4 Total mRNA-sequencing and Transcriptomics Analysis

It is known that the genetic background influences gene expression at both mRNA and protein levels in *C. elegans* (Kamkina *et al.* 2016). Whether *cgh-1* and *mat-1* expression levels differ between all six strains involved in this study could further explain the phenotypic variability

observed in the aforementioned experiments. It is not new that expression levels might differ across different genetic backgrounds. In fact, varying gene expression across different *C. elegans* natural isolates has already been reported for many of its 20,000 genes (Kamkina *et al.* 2016; Zhang *et al.* 2022). Natural genetic variation between N2 and CB4856 for example have shown 1,532 genes to have divergent expression at the mRNA level, 712 of these having higher expression in CB4856, and other 820 having higher expression in N2. Additionally, 129 proteins were also found to be differently expressed between these two strains, most being related to insulin-signaling and stress-response pathways (Kamkina *et al.* 2016). Other differences for these two wild isolates at the translation level have been described using recombinant inbred lines and quantitative traits locus (QTL) mapping. Out of the 44 genes tested by Singh *et al.* (2016) through this method, half of them showed a significant difference in protein abundance. More specifically, a QTL on chromosome II was shown to influence protein levels of the phosphatidylserine receptor protein PSR-1.

In another study, both N2 and CB4856 and inbred lines were used for insights into cryptic genetic variation and regulatory variation in gene expression: according to their response to heat stress. Snoek *et al.*, (2017) identified 3305 genes in which transcription was affected. Furthermore, more modern technologies such as ultralong-read direct RNA sequencing have been applied to analyse the complexity of *C. elegans* transcriptome, allowing the identification of novel splicing isoforms (Li *et al.* 2020). Here, transcriptomics was applied for scrutinizing *mat-1* and *cgh-1* phenotypes and comparing their expression across the natural isolates. Worms were synchronized and collected aiming at 70% of an adult population, a stage in which their development is complete, and mRNA abundance is higher, taking into consideration *cgh-1* and *mat-1* highest expression in the germline. RNA was extracted using a Trizol-chloroform protocol followed by in-column

DNase treatment (ZYMO Research RNA Clean & Concentrator kit). Whole transcriptomics sequencing was performed by Sick Kids in Toronto (Illumina NovaSeq 6000). Detailed protocol can be found in the Materials & Methods section in this document (Chapter 2.11).

4.4.1 RNA-Seq Analysis Reveals No Differential *mat-1* and *cgh-1* Expression Across All Six Natural Isolates

Processing of RNA-seq data was performed by Dr. Tatiana Maroilley, a current post-doctoral fellow in the Tarailo-Graovac lab on Galaxy (usegalaxy.eu/ - (Afgan *et al.* 2022)), using the guidelines previously described by Batut *et al.*, (2021) to execute the quality control, genome alignment, and read counts. Detailed data processing can be found on Chapter 2. This final step generated .bam files that I visually assessed for the gene expression of *mat-1* and *cgh-1*. As depicted in Figure 12 (left) and Figure 13 (left), no initial difference can be observed in gene expression for either *mat-1* or *cgh-1* between all six natural isolates in the WT background. The same can be seen in the strains with CRISPR knock-in of *cgh-1(tn691)* and *mat-1(ye121)* alleles, as depicted in Figure 12 (right) and Figure 13 (right), respectively.

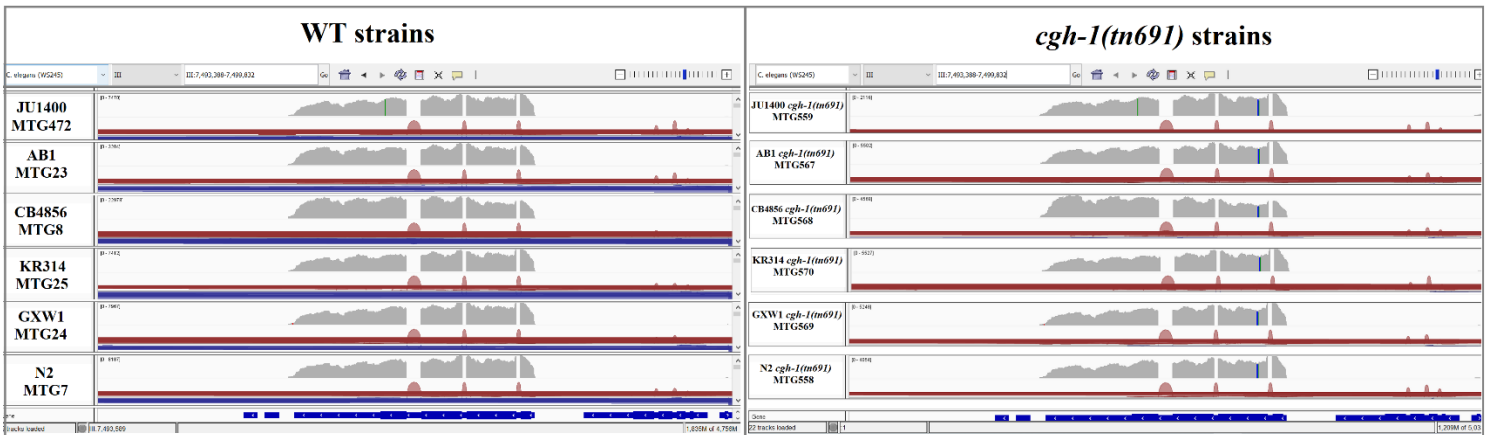


Figure 12. IGV visualization of gene expression of *cgh-1* in JU1400, AB1, CB4856, KR314, GXW1 and N2 WT strains (left) and *cgh-1(tn691)* strains (right). Peaks (gray) represent reads coverage within exon and reflect the mRNA expression in each region. Absence of peaks represent lack of reads coverage. Variants are colorful and color-coded according to the type of variant or nucleotide change. No visual difference in gene expression was observed in IGV, either comparing the natural isolates against each other, or comparing each natural isolate before and after CRISPR knock-in of *cgh-1(tn691)*.

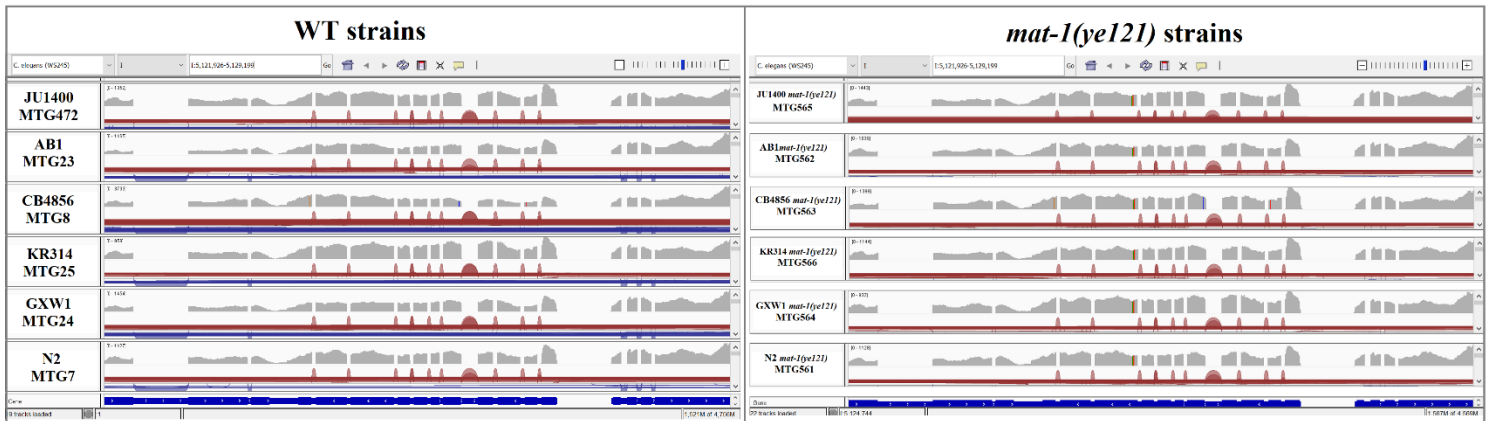


Figure 13. IGV visualization of gene expression of *mat-1* in JU1400, AB1, CB4856, KR314, GXW1 and N2 WT strains (left) and *mat-1(ye121)* strains (right). Peaks (gray) represent reads coverage within exon and reflect the mRNA expression in each region. Absence of peaks represent lack of reads coverage. Variants are colorful and color-coded according to the type of variant or nucleotide change. No visual difference in gene expression was observed in IGV, either comparing the natural isolates against each other, or comparing each natural isolate before and after CRISPR knock-in of *mat-1(ye121)*.

The *p*-values for gene expression of *mat-1* and *cgh-1* were also calculated using limma, as explained in the next section (Liu *et al.* 2015). Those did not offer any significant difference in comparison with the control N2 (Table 14). This corroborates the IGV analysis that no differential gene expression for these genes is present in these natural isolate backgrounds, which suggests that modifier effect cannot be explained by variability in *mat-1* and *cgh-1* expression.

4.4.2 Natural Isolates Differ in Universal Gene Expression

Significance (*p*-value) of potential different gene expression between the natural isolates was assessed using limma (Galaxy Version 3.50.1+galaxy0), also available on Galaxy platform, by comparing each natural isolate against N2 control (Liu *et al.* 2015). limma converts read counts into log₂-counts-per-million (logCPM) and models the mean-variance relationship between samples. Output contains adjusted *p*-values using Benjamini and Hochberg's false discovery rate control at a threshold value of 0.05 (BH 1995) (Benjamini and Hochberg 1995). For increasing

statistical power in this step, all three replicates of each background (WT, *mat-1(ye121)* and *cgh-1(tn691)*) were grouped, considering no differences in mRNA expression should be expected within the group (no significant differential gene expression for *mat-1* and *cgh-1* was observed, and all strains were cultured under the same permissive conditions).

Table 14. *p*-values for *mat-1* and *cgh-1* expression from AB1, CB4856, KR314, GXW1 and JU1400 in comparison with N2 control. Values are not significant (*p*-value < 0.05) and suggest no relevant differential gene expression for these genes in the aforementioned genetic backgrounds.

		<i>p</i> -value against N2 Control				
		AB1	CB4856	KR314	GXW1	JU1400
Gene	<i>mat-1</i>	0.65	0.19	0.57	0.41	0.96
	<i>cgh-1</i>	0.30	0.26	0.89	0.70	0.16

This analysis also revealed unique differential gene expression patterns for each of the natural isolates (Figure 14). Even though most of the genes were equally expressed in N2 and other isolates, a significant number had either reduced or increased expression in the other strains (Figure 14). Specifically, most of the genes that were differentially expressed were downregulated in comparison with N2. The strain with closest expression pattern to N2 was GXW1, having the lowest number of both upregulated and downregulated genes. GXW1 presented strong lethal phenotype for both *mat-1* and *cgh-1* tests (Chapter 3), which could be a reflection of its less variable genome. Next, KR314 was the second lowest differential expression numbers, and AB1, the third. Strains with highest number of differentially expressed genes were JU1400 and CB4856.

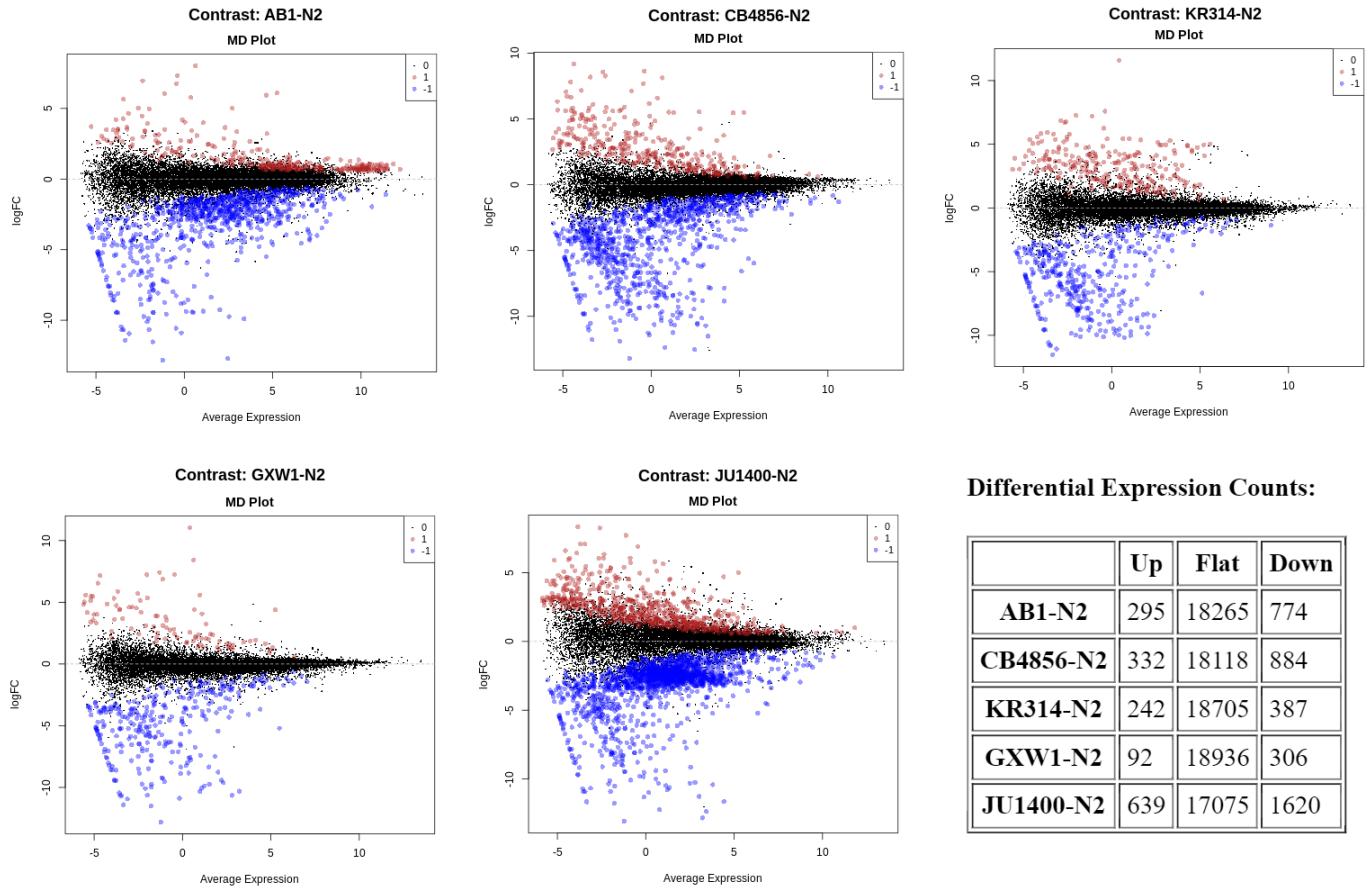


Figure 14. MD plots of each natural isolate against N2 control strain depicting differential expression for both over- (red) or under-expression (blue). Genes with little or no significant differential expression can be seen in black (flat). Table on bottom right translates data from MD plots into differential expression.

Additionally, a brief analysis of the candidate modifiers investigated in this project revealed no significant differential expression for most genes and strains (Table 15), with the exception of *hbl-1* in the AB1 background, in which WT strain seem to have reduced expression of this gene (Figure 15). The lack of differential expression in the candidate modifier genes in any particular background neither confirms nor disregards their activity as modifiers for *mat-1* and *cgh-1*, making the phenotypic experimental validation (Chapter 5) more reliable for such conclusions.

Table 15. *p*-values for candidate modifier genes investigated in this study as potential modifiers of *mat-1* and *cgh-1*. Only differences with a *p*-value < 0.05 are considered statistically significant.

		<i>p</i> -value against N2 Control				
		AB1	CB4856	KR314	GXW1	JU1400
Gene	<i>such-1</i>	0.82	0.08	0.65	0.33	0.53
	<i>bub-3</i>	0.82	0.16	0.91	0.99	0.58
	<i>Y73B6A.1</i>	0.19	0.08	0.30	0.57	0.12
	<i>hbl-1</i>	0.01	0.06	0.75	0.93	0.24
	<i>nhl-2</i>	0.31	0.21	0.43	0.47	0.35
	<i>glp-1</i>	0.29	0.63	0.77	0.57	0.16
	<i>ddx-15</i>	0.35	0.44	0.88	0.43	0.63

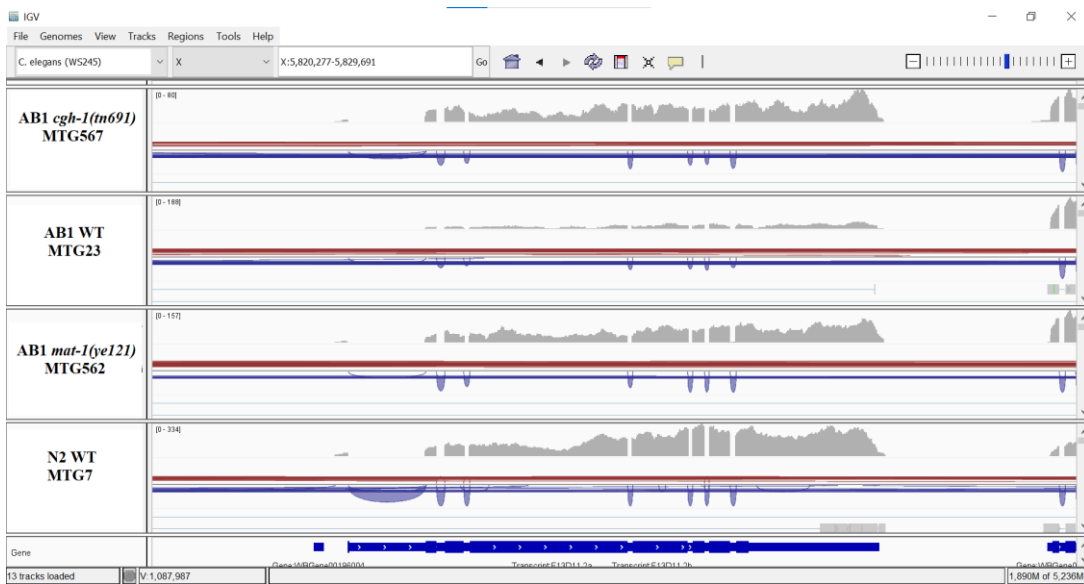


Figure 15. AB1 strains and IGV visualization of *hbl-1* gene expression, showing reduced expression in the WT background. Peaks (gray) represent reads coverage within exon and reflect the mRNA expression in each region. Absence of peaks represent lack of reads coverage. Variants are colorful and color-coded according to the type of variant or nucleotide change.

Chapter 5: Investigation of Candidate Genetic Modifiers for *mat-1* and *cgh-1*

5.1 Knock-in of Candidate Modifiers With CRISPR/Cas9

The selected candidate modifiers were primarily tested in the N2 background carrying the gene of interest (*mat-1* or *cgh-1*). N2 is the reference genome and offers a genetic background with little variation, due to its lab-conditioned domestication (Brenner 1974). The CRISPR/Cas9 incorporation was performed by microinjection targeting a specific structure in the hermaphrodite worm's gonads, the rachis (Figure 16), which contains premature germline cells (Pazdernik and Schedl 2013). This way, we optimize the transfection and the incorporation of the new variant in the next generation. To help with the identification of the worms that were successfully transfected, we make use of the co-CRISPR technique. Our injection mix also contains CRISPR constructs targeting *dpy-10*, which affects the organismal morphology and creates a visible phenotype to screen for potential heterozygotes. This additional variant is dominant and presents as “dumpy” phenotype in homozygous worms and a “roller” phenotype in heterozygous worms. Usually, successful CRISPR worms will be identified in the F1 generation having the roller phenotype. Rollers are then plated individually and genotyped for the variant of interest with the use of PCR and Sanger sequencing. When a heterozygous worm for the target gene is identified, the F2 generation is also genotyped to isolate the ones homozygous for the variant of interest and wild type for the *dpy-10* phenotype. CRISPR/Cas9 was shown to be prompt and effective for validating genetic modifier SNVs in *C. elegans* (Jean *et al.* 2021; Stuart 2022). Detailed CRISPR/Cas9 approach, gRNA sequences and genotyping strategy can be found in the Materials & Methods section of this thesis.

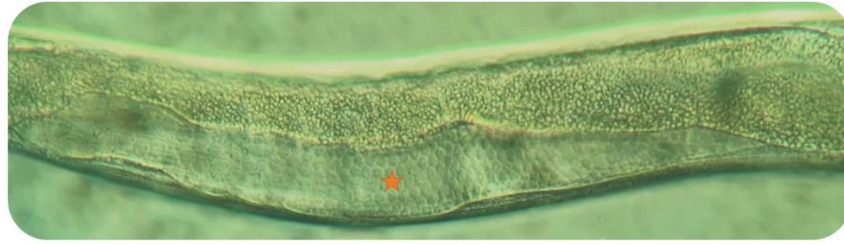


Figure 16. *C. elegans* target region for CRISPR/Cas9 microinjection highlighted in orange (rachis). The rachis is rich in germline nuclei that will later become oocytes.

5.2 Candidate modifiers for *mat-1*

5.2.1 *The SUPpressor of spindle CHECKpoint defect gene, such-1*

such-1, an APC5-like gene, is a component of the APC complex and acts as a positive regulator of mitotic metaphase/anaphase transition. SUCH-1 is a paralog of GFI-3, both being similar to yeast Apc5p and to human APC5 (Tarailo *et al.* 2007a). Both *such-1* and *gfi-3* are co-expressed in the germline, but do not have overlapping expression in other tissues. Impairment of either *such-1* or *gfi-3* alone does not result in a notable phenotype, but lack of both components at the same time leads to mitotic arrest (Stein *et al.* 2010).

A specific *such-1* variant, *h1960*, was shown to delay anaphase onset in germline and early embryo and to suppress the impairment of *mdf-1*, a component of the SAC. This variant induces alternative splicing and removal of 32 amino acids from the final protein, and delays mitosis in two-fold, allowing for proper progression of cell division even with disrupted SAC function (Tarailo *et al.* 2007a). Additionally, a gain-of-function variant in *such-1*, (*av9gf*), was shown to suppress *mat-3(or180ts)* embryonic lethality, another APC component and ortholog of APC8, in a study of *mat-3* temperature-sensitive alleles (Stein *et al.* 2010). In fact, this same gain-of-function variant, *such-1(av9gf)* was also found to weakly suppress other three APC/C alleles in addition to *mat-3(or180ts)*: *emb-30(or420ts)*, *mat-1(ax212ts)*, and *emb-1(hc62ts)*. However, even though the one-cell mitotic arrest was successfully bypassed, only a small percentage of embryos were

capable of hatching (Stein *et al.* 2010). The study of another *such-1* allele, (*t1668*), also helped revealing a potential antagonism between APC/C and SAC, in which APC negatively regulates SAC to promote mitotic exit (Bezler and Gönczy 2010).

The evidence of *such-1* suppression of other genes involved in the mitotic cell cycle, and its confirmed interaction with *mat-1* (due to presence in the same complex), were crucial for selecting this gene for modifier investigation. In the genetic backgrounds present in this study, two variants were identified in the *such-1* gene: III:11,517,531 C>G (*such-1(mod476)*), and III:11,517,714 G>C (*such-1(mod483)*). The first one, was present in five of the six natural isolates (AB1, CB4856, GXW1, JU1400, and KR314), while the second allele was unique to Hawaiian strain CB4856. Both variants were tested in the N2; *mat-1(ye121)* background, but no significant increase in hatch rate was observed for either of them. Additionally, no difference in propagation was noted for either *mod476* or *mod483* alleles, disregarding these variants as suppressors of *mat-1(ye121)* (Figure 17).

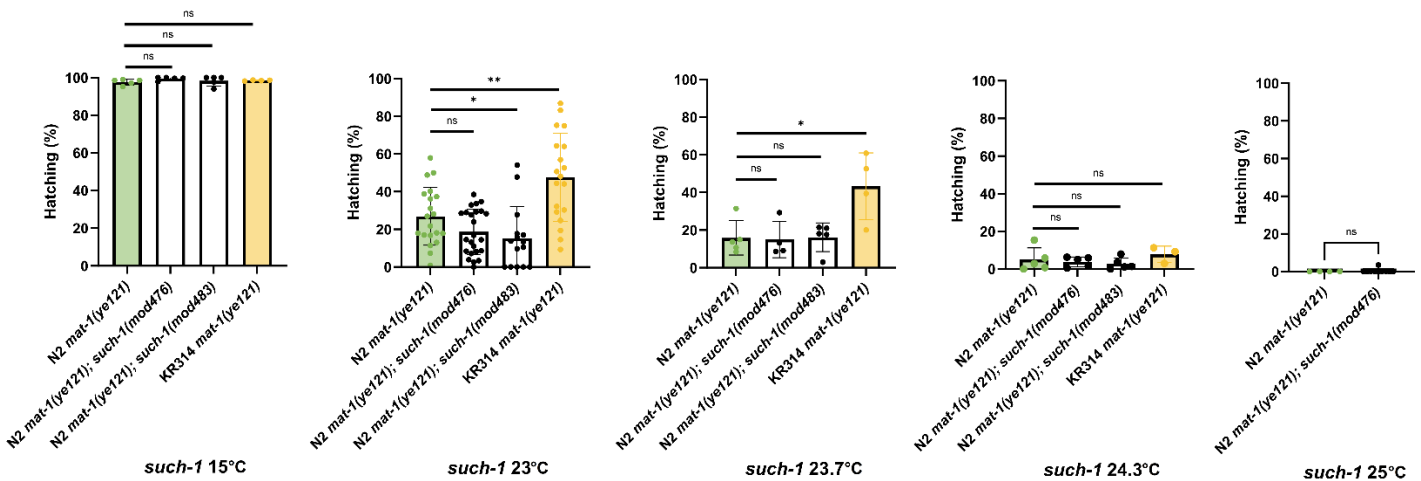


Figure 17. Hatch rate assays for investigation of *such-1* as a modifier for *mat-1*. No significant difference in hatch rate was seen when comparing N2; *mat-1(ye121)* with double mutants for *such-1* variants at either permissive (15°C) or restrictive conditions. Statistical significance was calculated using unpaired t-tests with Welch's correction. *p*-values are shown using “*” (0.01 to 0.05 - Significant), “**” (0.001 to 0.01 - Very significant), “***” (0.0001 to 0.001 - Extremely significant), or “****” (<0.0001 - Extremely significant). Non-significant results are represented by “ns”.

5.2.2 The yeast *BUB* homolog gene, *bub-3*

The gene *BUB3* was initially identified in yeast, in a screen for genes involved in mitotic arrest after exposure to microtubule depolymerizing drugs, and it is an important component of the SAC, together with *MAD1*, *MAD2*, *MAD3*, and *BUB1* (Hoyt *et al.* 1991; Li and Murray 1991; Weiss and Winey 1996). The SAC is essential for monitoring proper chromosome segregation and allowing for cell division continuity (Cleveland *et al.* 2003). It temporarily delays metaphase-to-anaphase transition until proper alignment of chromosomes and their attachment to mitotic spindles is reached, avoiding chromosome instability. This is achieved by the temporary inhibition of APC/C for prevention of anaphase onset, which is done by inhibition of the APC/C activator, *Cdc20* (Musacchio and Salmon 2007). The SAC genes are highly conserved in the eukaryotic kingdoms and have orthologs in *C. elegans* named *mdf-1*, *mdf-2*, *san-1*, *bub-1*, and *bub-3*, respectively (Kitagawa and Rose 1999; Oegema *et al.* 2001; Nystul *et al.* 2003; Stein *et al.* 2007; Tarailo *et al.* 2007a). Other higher eukaryotes contain additional elements in the SAC complex, including Rod (ROugh-Deal), Zw10 (Zeste-White 10) and CENP-F (Chan *et al.* 2000; Yang *et al.* 2005; Kops *et al.* 2005). In *C. elegans*, SAC also includes *hcp-1*, *hcp-2*, and *rod-1* (Moore *et al.* 1999; Desai *et al.* 2003; Tarailo *et al.* 2007b).

Studies for knockdown of SAC elements using RNAi identified that, although depletion of *bub-3* on its own does not show a phenotype, the combination of *bub-3*(RNAi) with *mdf-1*(RNAi), *mdf-2*(RNAi) or *san-1*(RNAi) seem to show a modification of the primary phenotype, leading to decreased viability (Tarailo *et al.* 2007b). The same is not observed in yeast, and no synthetic lethality of *BUB3* with any of the other SAC components had been observed (Tarailo *et al.* 2007b). During SAC signaling, Bub3 is involved in another important activity: the mitotic checkpoint complex (MCC), together with Bub1, BubR1, Mad2 and Cdc20. Bub3 contributes to

the binding of MCC to APC/C. This is responsible for the direct inhibition of APC/C, preventing premature exit from mitosis (Overlack *et al.* 2017).

The variant identified in *bub-3* was unique to KR314 strain and located at II:13,728,035. This missense variant causes a G>C nucleotide change, which leads to substitution from a valine to a leucine (Grantham score = 32). By being present in the background with the best hatch and propagation results, this variant was primarily considered as a potential suppressor for *mat-1*. However, there was no significant difference observed for either hatch rate or propagation for this variant, when knocked-into the N2; *mat-1(ye121)* background. Hatch rate experiments were performed at three conditions (22.4°C, 23°C and 24°C, Figure 18), and propagation was tested in a range from 20°C to 25.3°C, only showing propagation up to 23°C (Table 16). Altogether, the results suggest that this specific variant in *bub-3* is not suppressing *mat-1(ye121)* phenotype, or its effect may be masked by the presence of another, unknown, variant.

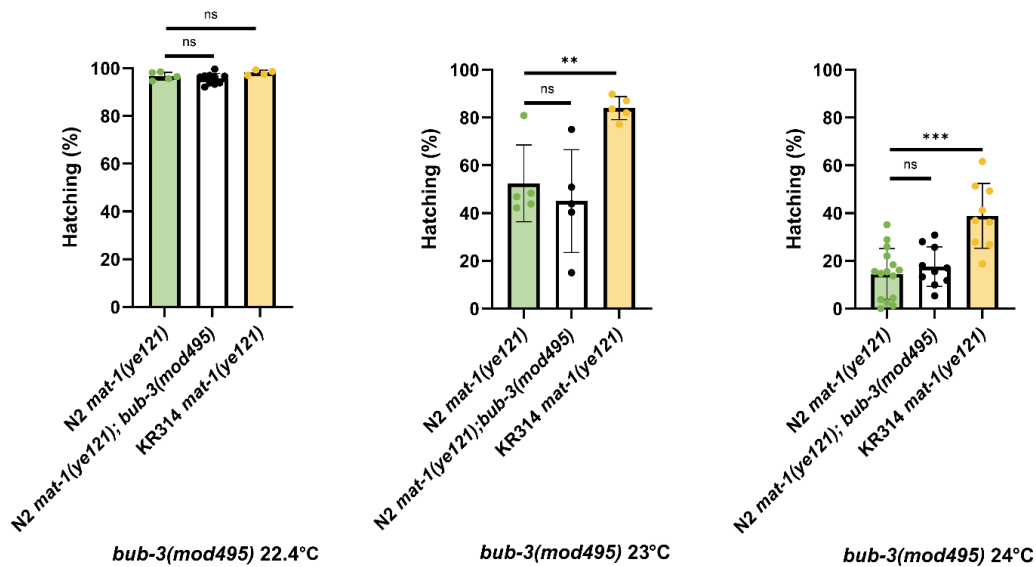


Figure 18. Hatch rate assays for investigation of *bub-3* as a modifier for *mat-1*. No significant difference in hatch rate was seen when comparing N2; *mat-1(ye121)* with N2; *mat-1(ye121); bub-3(mod495)* at either temperature tested (22.4°C, 23°C, and 24°C). No difference was observed in propagation assays when comparing those two genotypes, disregarding this variant as a suppressor for *mat-1*. Statistical significance was calculated using unpaired t-tests with Welch’s correction. *p*-values are shown using “*” (0.01 to 0.05 - Significant), “***” (0.001 to 0.01 – Very significant), “****” (0.0001 to 0.001 – Extremely significant), or “*****” (<0.0001 – Extremely significant). Non-significant results are represented by “ns”.

5.2.3 *Y73B6A.1*

Y73B6A.1 is a mitotic kinase predicted to be involved in mitotic cell cycle and to enable histone H3T3 activity by phosphorylation of H3. Its homology is compared to haspin, a group of proteins required for metaphase chromosome cohesion and alignment during mitosis (Manning 2005; Moura *et al.* 2018). During cell division, accumulation of the chromosomal passenger complex (CPC) at the centromeric region involves two pathways that rely on either H3 histone (regulated by haspin), and H2A histone (regulated by BUB1) (Barbosa *et al.* 2022). Impairment of haspin disrupts cohesion binding and sister chromatid association, blocking normal chromosomal alignment and activating SAC, leading to prometaphase mitotic arrest (Dai and Higgins 2005; Dai *et al.* 2006).

In this project, the variant present in *Y73B6A.1* and prioritized as a potential enhancer of *mat-1* is common between CB4856 and AB1 backgrounds, located at IV:6,663,170, and causing a C>T nucleotide change, leading to an amino acid substitution from alanine to threonine (Grantham score = 58). Its prioritization as a potential enhancer was according to results obtained by *mat-1(ye121)* investigation listed in Chapter 3, in which CB4856 and AB1 backgrounds presented lower hatch rates than the control N2 at semi-restrictive conditions (such as 23°C), suggesting the presence of an enhancer in these backgrounds. Therefore, *Y73B6A.1* was selected due to presence of a common variant in these two natural isolates.

The variant was studied in the N2; *mat-1(ye121)* background. No significant difference was observed for hatch rates in comparison with the original background at either the permissive condition of 15°C or the restrictive conditions of 23°C and 24°C (Figure 19). Additionally, no significant difference was observed for long-term propagation at either 20°C, 23°C, 24°C or 25°C (Table 16). Altogether, there is no evidence of interaction between the studied *Y73B6A.1* (IV:6,663,170 C>T) variant and *mat-1*.

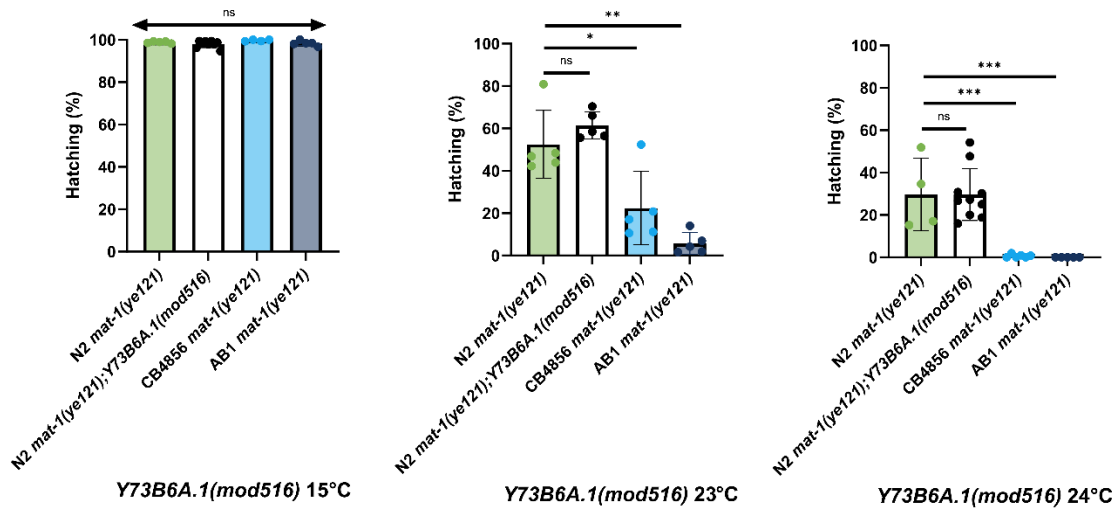


Figure 19. Hatch rate assays for investigation of *Y73B6A.1* as a modifier for *mat-1*. No significant difference in hatch rate was seen when comparing N2; *mat-1(ye121)* with N2; *mat-1(ye121); Y73B6A.1(mod516)* at 15°C, 23°C, and 24°C. Statistical significance was calculated using unpaired t-tests with Welch’s correction. *p*-values are shown using “*” (0.01 to 0.05 - Significant), “***” (0.001 to 0.01 – Very significant), “****” (0.0001 to 0.001 – Extremely significant), or “*****” (<0.0001 – Extremely significant). Non-significant results are represented by “ns”.

5.3 Candidate modifiers for *cgh-1*

5.3.1 The HunchBack Like (fly gap gene related) gene, *hbl-1*

The gene *hbl-1*, previously referred as *lin-57*, is involved in negative regulation of transcription by RNA polymerase II, larval development, and oviposition. Its protein is located in the nucleus and expressed in multiple tissues including muscle, germline, and neurons. Its 982-amino-acid protein is a homologue of the *Drosophila* hunchback gene (Fay *et al.* 1999). It encodes a zinc finger transcription factor member of the worm’s heterochronic gene pathway, which is responsible for properly timing postembryonic development (Abrahante *et al.* 2003). Therefore, impairment of *hbl-1* leads to locomotion, egg laying and morphogenesis issues. More specifically, loss of function of *hbl-1* results in premature differentiation of hypodermis, seam cell fusion, and adult alae synthesis (Abrahante *et al.* 2003). RNAi knockdown of *hbl-1* also revealed severe progeny arrest, and the few worms that reached adulthood showed defects in vulva morphology,

as long as egg-laying defective phenotypes. These development issues were primarily associated with the abnormal organization of hypodermal cells (Fay *et al.* 1999).

HBL-1 expression is negatively regulated during larval development by *mir-48*, *mir-241*, and *mir-84*, genes part of the *let-7* family of micro RNAs (miRNAs). This specifically occurs during L2-L3 transition of larval stage, by binding complementary elements in the 3'UTR of *hbl-1* transcripts (Abbott *et al.* 2005). *hbl-1* seems to act downstream of *let-7* miRNAs. However, *let-7* family miRNAs are also regulated by HBL-1, representing a negative feedback loop. Additionally, *hbl-1* presents interactions with other genes mentioned in this work, such as *cgh-1* and *nhl-2*. The impairment of *cgh-1* in *mir-48*; *mir-84* loss-of-function mutants increases developmental delay, but is suppressed by *hbl-1* RNAi, suggesting that CGH-1 acts along with *mir-48* and *mir-84* to downregulate HBL-1 (Hammell *et al.* 2009). Overall expression of *hbl-1* seems to be highest at embryogenesis, and decreased after (Fay *et al.* 1999).

Here, I investigated the missense variant in *hbl-1* (X:5,824,753 C>T) present uniquely in the CB4856 background. This variant was knocked-into the N2; *cgh-1(tn691)* background for phenotype investigation and assessment of its potential interaction with *cgh-1*. As depicted in Figure 20, no significant difference was seen in hatching rates when comparing single versus double mutants (N2; *cgh-1(tn691)*; *hbl-1(mod482)*), at either the permissive temperature of 15°C, or at restrictive conditions of 23°C, 25°C, and 25.3°C. Further phenotype investigation with propagation assays also did not result in any improvement in strain hatch rate and propagation (Table 17). Altogether, this suggests no interaction of *hbl-1* (X:5,824,753 C>T) with *cgh-1*.

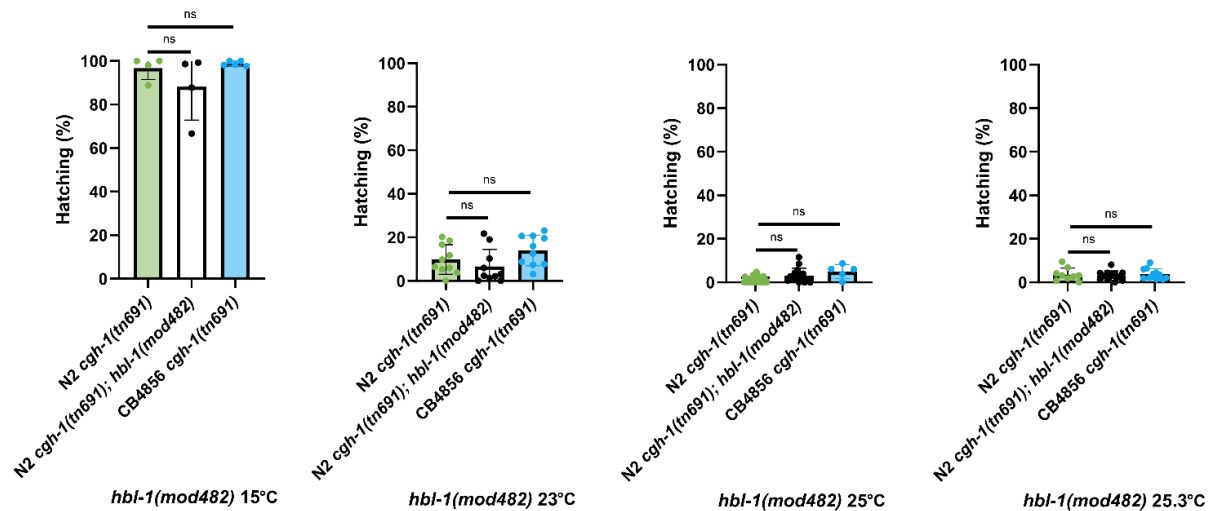


Figure 20. Hatch rate assays for investigation of *hbl-1* as a modifier for *cgh-1*. No significant difference in hatch rate was seen when comparing N2; *cgh-1(tn691)* with N2; *cgh-1(tn691); hbl-1(mod482)* at either permissive (15°C) or restrictive (23°C, 25°C, and 25.3°C) conditions. Statistical significance was calculated using unpaired t-tests with Welch’s correction. *p*-values are shown using “*” (0.01 to 0.05 - Significant), “**” (0.001 to 0.01 – Very significant), “***” (0.0001 to 0.001 – Extremely significant), or “****” (<0.0001 – Extremely significant). Non-significant results are represented by “ns”.

5.3.2 The NHL (ring finger b-box coiled coil) domain containing gene, *nhl-2*

The gene *nhl-2* enables DEAD/H-box RNA helicase binding activity, establishment of cell polarity, and it is also a positive regulator of development. It is expressed in several structures, including germline, and it is an ortholog of human *TRIM45*. It is known to physically interact with *cgh-1*, and also presents genetic interactions with *hbl-1* (Zhong and Sternberg 2006; Hammell *et al.* 2009). Knock-out of *nhl-2* in combination with *cgh-1* was also shown to be more detrimental than single mutants and showed embryonic and larval death, suggesting their combined requirement for organism viability (Hammell *et al.* 2009). NHL-2 is localized in P-bodies in cytoplasm, acts as a cofactor for miRISC and functions in association with CGH-1 in somatic cells during larval development. Together, they promote the post-transcriptional regulation of some miRNA targets, including *hbl-1* (Hammell *et al.* 2009). NHL-2 protein contains multiple domains, including B-box domains, known for RNA-binding and protein:protein interactions (Matthews and

Sunde 2002; Hall 2005). The loss of NHL-2 does not impair CGH-1-miRISC binding, but NHL-2 could potentially facilitate this association (Hammell *et al.* 2009).

The suppression effect of *nhl-2* in impaired development has been reported through a loss-of-function variant in *nhl-2* deemed capable of suppressing a hypomorphic variant in *lin-41(ma104)*. On the other hand, null *nhl-2* alleles were shown to enhance the phenotype of single mutants for *let-7* family miRNAs (*mir-48*, *mir-241*, or *mir-84*). This effect, however, was completely bypassed by RNAi depletion of *hbl-1*, suggesting that *nhl-2* functions in association with *let-7* family miRNAs to inhibit HBL-1 (Hammell *et al.* 2009).

Among the natural isolates studied in this project, two of them, KR314 and JU1400, shared a common missense SNV in *nhl-2* (III:4,897,421 T>G). This variant was tested as a potential modifier for *cgh-1* by being knocked-in in the N2; *cgh-1(tn691)* background. No significant difference was observed for hatch rates at 15°C, 23°C, or 25°C as depicted in Figure 21. Some mild significance was observed for 25.3°C, but this became irrelevant when no improvement in propagation was detected at restrictive-condition propagation assays. Therefore, *nhl-2* (III:4,897,421 T>G) was disregarded as a modifier for *cgh-1*.

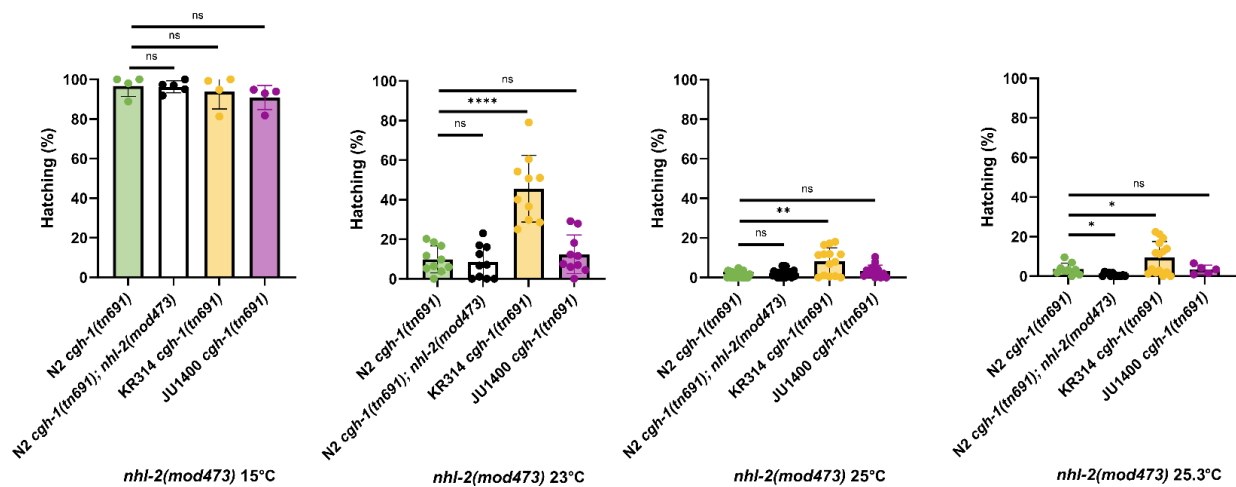


Figure 21. Hatch rate assays for investigation of *nhl-2* as a modifier for *cgh-1*. No significant difference in hatch rate was seen when comparing N2; *cgh-1(tn691)* with N2; *nhl-2(mod473)* *cgh-1(tn691)* at either permissive (15°C) or restrictive (23°C, 25°C, and 25.3°C) conditions. Statistical significance was calculated using unpaired t-tests with

Welch's correction. *p*-values are shown using “*” (0.01 to 0.05 - Significant), “**” (0.001 to 0.01 – Very significant), “***” (0.0001 to 0.001 – Extremely significant), or “****” (<0.0001 – Extremely significant). Non-significant results are represented by “ns”.

5.3.3 The DEAD box helicase homolog gene, *ddx-15*

The *ddx-15* gene is predicted to enable RNA activity and RNA helicase activity. It is present in the nuclear foci and involved in alternative splicing (Tariq *et al.* 2013; Pollo *et al.* 2023). *ddx-15* represses RNA editing at the 3' UTR end, therefore, its impairment leads to severe increase in editing. At the protein level, worm DDX-15 has almost 75% similarity in comparison with human protein (Tariq *et al.* 2013). In mice, it has a role in antimicrobial response and intestinal inflammation due to regulation of α -defensins, and in humans, reduced protein levels of Ddx15 is associated with ulcerative colitis (UC) (Wang *et al.* 2021). It functions as an RNA virus sensor and influences the production of immune system cytokines in dendritic cells (Lu *et al.* 2014; Xing *et al.* 2021).

The SNV identified in *ddx-15* is a missense variant causing a G>C substitution, which results in arginine to threonine amino acid change. This variant is unique to the KR314 background, and it is located at III:5,591,654 (Grantham score = 71). It was the top candidate for modification of *cgh-1* in the list of variants unique to KR314 background given by the machine learning model (details for candidate variant selection are explained in Chapter 4). However, hatch rate and propagation assays were not suggestive of interaction between *ddx-15* and *cgh-1*, given that there was no significant difference for hatch rates between N2; *cgh-1(tn691)* and N2; *cgh-1(tn691); ddx-15(mod496)*, as shown in Figure 22. Additionally, no difference in propagation was observed (Table 17), making an interaction between *ddx-15* (III:5,591,654 G>C) and *cgh-1* unlikely.

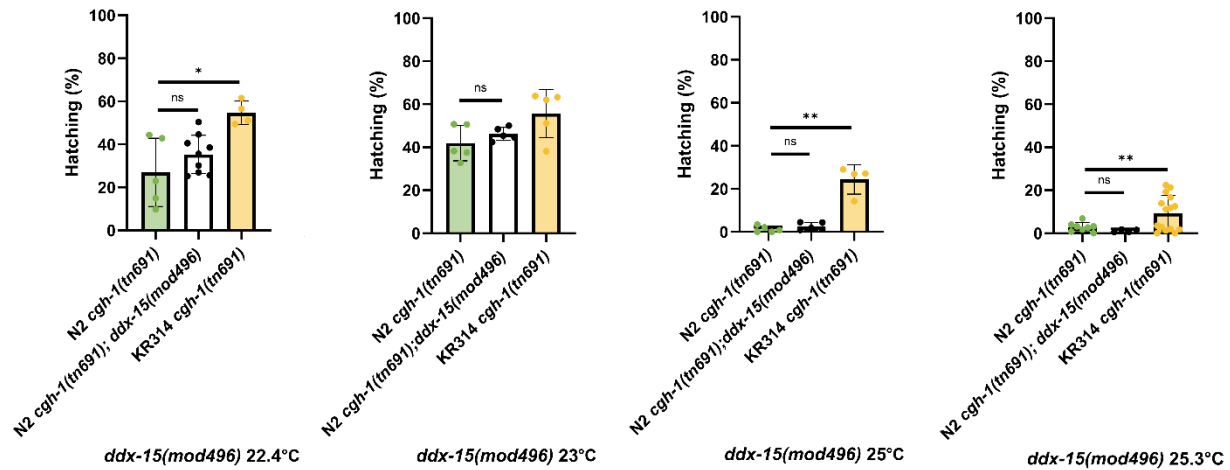


Figure 22. Hatch rate assays for investigation of *ddx-15* as a modifier for *cgh-1*. No significant difference in hatch rate was seen when comparing N2; *cgh-1(tm691)* with N2; *ddx-15(mod496) cgh-1(tm691)* at either of the tested conditions (22.4°C, 23°C, 25°C, and 25.3°C). *p*-values are shown using “*” (0.01 to 0.05 - Significant), “**” (0.001 to 0.01 – Very significant), “***” (0.0001 to 0.001 – Extremely significant), or “****” (<0.0001 – Extremely significant), according to the significance degree (from least to most significant, respectively). Non-significant results are represented by “ns”.

5.3.4 The abnormal Germ Line Proliferation gene, *glp-1*

The *glp-1* gene is involved in RNA polymerase II DNA binding activity and is a regulator of gene expression. Its transcription factor activity is important in the germline for control of cell division and decision between mitosis or meiosis fate. Variants in *glp-1* cause germline cells that would naturally divide through mitosis to undergo meiosis instead (Austin and Kimble 1987). *glp-1* is a transmembrane protein that allow the cell-cell interactions and cell fate control in the germline. It is essential for distal tip cell (DTC) control in the germline, and a lack of *glp-1* has the same effect of killing DTC: germline polarity is lost, and all cells enter meiosis instead (Crittenden *et al.* 1994). GLP-1 controls stem cell population through spatial regulation of *gld-1* and *gld-2*. Consecutive expression of *glp-1* or impairment of both *gld-1* and *gld-2* result in tumorous germline and general, instead of selected, cell proliferation (Hansen *et al.* 2004a; b).

glp-1 is also important for early embryogenesis and specification of certain cell fates and

cell proliferation through a maternal effect. Therefore, disruption of *glp-1* can cause embryonic lethality. Additionally, it is suggested to act in cell-cell interactions between descendants of AB and P₁ that normally produce pharyngeal cells, meaning its impairment leads to pharynx location and size defects (Priess *et al.* 1987). *glp-1* seems to be temporally and spatially restricted to anterior descendants of 2- to 28-cell embryos (Evans *et al.* 1994). Only the AB descendant cells require GLP-1 function, which seems to act as a receptor for P₂ and MS cells (Shelton and Bowerman 1996).

Two variants of interest were identified in *glp-1*, one unique to KR314 background, and other unique to CB4856. Both were missense SNVs. Based on previous literature and potential synthetic lethality of *glp-1* and *cgh-1* double mutants (identified in RNAi study), I focused on testing the variant in CB4856 as a potential enhancer of *cgh-1(tn691)* (Byrne *et al.* 2007). This variant, located at III:9,097,961, causes a substitution of A>C, resulting in amino acid change from glutamine to alanine (Grantham score = 107 – “moderately radical”, high impact amino acid change). The variant was knocked-in in the N2; *cgh-1(tn691)* background for assessing its modifying potential. However, no significant difference in hatch rates for this strain was observed in comparison with the original N2; *cgh-1(tn691)* background, at neither of the conditions tested (21.7 °C, 23°C, and 25 °C) (Figure 23). Additionally, no divergence in propagation was noted for the N2; *cgh-1(tn691); glp-1(mod497)* in comparison with N2; *cgh-1(tn691)*. Together, the results suggest that the presence of this missense variant in *hbl-1* is not capable of altering the phenotype of *cgh-1*.

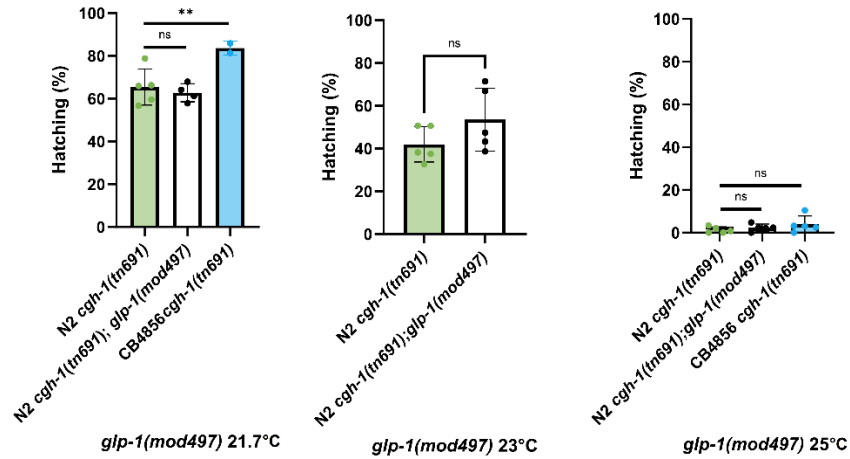


Figure 23. Hatch rate assays for investigation of *glp-1* as a modifier for *cgh-1*. No significant difference in hatch rate was seen when comparing N2; *cgh-1(tn691)* with N2; *cgh-1(tn691); glp-1(mod497)* at either of the tested conditions (21.7°C, 23°C, and 25°C). *p*-values are shown using “*” (0.01 to 0.05 - Significant), “**” (0.001 to 0.01 – Very significant), “***” (0.0001 to 0.001 – Extremely significant), or “****” (<0.0001 – Extremely significant), according to the significance degree (from least to most significant, respectively). Non-significant results are represented by “ns”.

5.4 Propagation assays for candidate modifier strains

Table 16. Propagation of *mat-1(ye121)* candidate modifier strains at restrictive temperature. Propagation observed after two weeks of uninterrupted incubation at correspondent temperature. Results were classified between “No propagation” (-), “Mild propagation” (○), and “True propagation” (●). Notably, the higher the temperature, the lower the fitness and propagation rate of each strain. N/A represents strains not tested at the specific condition noted.

		Temperature						
		20°C	23°C	23.2°C	23.7°C	24°C	25°C	25.3°C
Genetic Background	AB1; <i>mat-1(ye121)</i>	●	○	-	-	-	-	N/A
	CB4856; <i>mat-1(ye121)</i>	●	●	-	-	-	-	N/A
	GXW1; <i>mat-1(ye121)</i>	●	●	-	-	-	-	N/A
	JU1400; <i>mat-1(ye121)</i>	●	●	-	-	-	-	N/A
	KR314; <i>mat-1(ye121)</i>	●	●	●	○	○	-	-
	N2; <i>mat-1(ye121)</i>	●	●	●	-	-	-	-
	N2; <i>mat-1(ye121); such-1(mod476)</i>	●	●	○	-	-	-	N/A
	N2; <i>mat-1(ye121); such-1(mod483)</i>	●	●	N/A	N/A	-	-	N/A
	N2; <i>mat-1(ye121); bub-3(mod495)</i>	●	●	N/A	N/A	-	-	-
N2; <i>mat-1(ye121); Y73B6A.1(mod516)</i>	●	●	N/A	N/A	-	-	N/A	

All candidate modifiers were tested not only for alterations in hatch rate, but also for improvement or worsening of long-term propagation. Strains were tested in a range of restrictive conditions varying from 20°C to 25.3°C. None of the variants tested for suppression or enhancement of *mat-1* or *cgh-1* showed differences in propagation in comparison with control, at any of the given conditions. Together with hatch rate assays described above, these results were fundamental for ruling out the studied variants as modifiers for the target genes in this study, *mat-1* and *cgh-1*.

Table 17. Propagation of *cgh-1(tn691)* candidate modifier strains at restrictive temperature. Propagation observed after two weeks of uninterrupted incubation at correspondent temperature. Results were classified between “No propagation” (-), “Mild propagation” (○), and “True propagation” (●). Notably, the higher the temperature, the lower the fitness and propagation rate of each strain. N/A represents strains not tested at the specific condition noted.

		Temperature					
		20°C	23°C	23.7°C	24°C	25°C	25.3°C
Genetic Background	AB1; <i>cgh-1(tn691)</i>	●	●	●	○	-	-
	CB4856; <i>cgh-1(tn691)</i>	●	●	●	○	-	-
	GXW1; <i>cgh-1(tn691)</i>	●	●	●	○	-	-
	JU1400; <i>cgh-1(tn691)</i>	●	●	●	○	-	-
	KR314; <i>cgh-1(tn691)</i>	●	●	●	●	○	○
	N2; <i>cgh-1(tn691)</i>	●	●	●	○	-	-
	N2; <i>cgh-1(tn691); hbl-1(mod482)</i>	●	●	●	○	-	-
	N2; <i>nhl-2(mod473) cgh-1(tn691)</i>	●	●	●	○	-	-
	N2; <i>ddx-15(mod496) cgh-1(tn691)</i>	●	●	N/A	○	-	-
	N2; <i>cgh-1(tn691) glp-1(mod497)</i>	●	●	N/A	○	-	-

Chapter 6: Discussion

6.1 Phenotypic Variability of *mat-1* and *cgh-1* is Affected by Different Genetic Backgrounds

Whether gene essentiality is immutable, or variable, has been a recurrent topic in latest decades of scientific research. In *C. elegans*, most of essential genes were only probed in N2 reference genetic background. More recently, a growing evidence of phenotypic variability in different natural isolates have been discovered (Vu *et al.* 2015; Zhang *et al.* 2022). In my thesis project, I studied the phenotype of two essential genes, *mat-1* and *cgh-1*. These two genes were specifically picked for investigating gene essentiality in *C. elegans* as: 1) they were already shown to be essential for survival in the N2 background; and 2) both have available temperature-sensitive alleles, a characteristic that provides a conditional phenotype and therefore allows the thorough investigation of requirement of these gene products in different genetic backgrounds. Evidence of phenotypic variability for *mat-1* and *cgh-1* knockdown had been reported by Vu et al (2015), who compared the RNAi knockdown phenotype of over 1,400 genes between the reference genome N2 and the widely diverse strain CB4856. For both genes, the phenotype observed was more severe in N2 background, and suppressed in the CB4856 background. This is suggestive of putative genetic factors in each background that influence the phenotypic outcomes. These genetic modifiers could be enhancing the phenotype in the first background (N2) or suppressing it in the second (CB4856).

With that in mind, I here expanded the investigation of plasticity of essentiality by studying the phenotype of *mat-1* and *cgh-1* in N2 and CB4856 backgrounds, in addition to another four diverse wild isolates (AB1, Australia; GXW1, China; KR314, Canada; JU1400, Spain). The prioritization of each background was based on their geographical location, aiming on having the most diverse genetic repertoire. For this investigation, temperature-sensitive alleles for *mat-1* and

cgh-1 were knocked-in using CRISPR/Cas9 in all six natural isolate backgrounds. Using a temperature-sensitive allele allows the study of lethal variants once conditions are shifted to restrictive temperature, but still provides an easy maintenance of stocks and long-term worm culture when placed at lower, permissive conditions. The temperature-sensitive alleles investigated in this project are *mat-1(ye121)* and *cgh-1(tn691)*, two missense SNVs that lead to amino acid change and protein instability only at higher temperatures (Rappleve *et al.* 2002).

The use of a stable, inherited variant in place of RNAi was decided upon, based on achieving a more reliable/heritable perturbation of these target genes. RNAi is more unstable and may lead to false negative results if knock-down is not properly achieved (Tischler *et al.* 2008). It is only capable of impairing mRNA products, but not any protein that is produced before the organism's exposing to RNAi treatment (Bell *et al.* 2023). Secondly, it can also result in false-positive conclusions due to siRNAs ability to bind to non-complementary sequences (it allows a few base mismatches) (Echeverri *et al.* 2006). Finally, dsRNA introduction in specific mammalian cell types was shown to induce immune response, which probably interferes with the observation of phenotypes (Bridge *et al.* 2003; Robbins *et al.* 2006). Additionally, some *C. elegans* wild isolates are naturally less sensitive to RNAi in the germline due to unique polymorphisms (Tijsterman *et al.* 2002).

In fact, a divergent phenotype was observed for CB4856 background when comparing RNAi and temperature-sensitive allele experiments in this project (Chapter 3). Impairment of survival of CB4856; *mat-1(ye121)* was a surprising result considering RNAi targeting the same gene and background showed opposite feature, (shown in section 3.1 and previous literature) (Vu *et al.* 2015). While for RNAi, the Hawaiian background is still capable of surviving, the same is not observed with the knock-in of *mat-1(ye121)* or *cgh-1(tn691)*. Instead, the impairment of *mat-*

I or *cgh-1* using the temperature-sensitive alleles led to a strong lethal phenotype that could not be noted for RNAi. The idea that an off-target effect of CRISPR/Cas9 could have caused the divergent phenotype in the CB4856; *mat-1(ye121)* strain was first considered, but then disregarded after thorough investigation (detailed description on Chapter 4).

Natural genetic variation is found for nearly any measurable trait in *C. elegans*, including when using RNAi (Paaby *et al.* 2015; Andersen and Rockman 2022). Response to RNAi is known to vary according to the genetic background it acts on (Bell *et al.* 2023). In CB4856, specifically, RNAi seems strongly ineffective in targeting germline genes: loss of function of *ppw-1* was deemed responsible for blocking (or at least delaying) that response (Tijsterman *et al.* 2002; Chou *et al.* 2023). A recent study compared differential gene expression after knockdown of two germline-essential genes (*par-1* and *pos-1*) across strains with variable RNAi sensitivity, including CB4856, and observed strong strain-specific phenotypes (Bell *et al.* 2023). Additionally, wild isolates appear to vary in efficacy of germline RNAi depending on the target gene, but the detailed mechanisms remain to be elucidated (Paaby *et al.* 2015).

For the comparison of phenotype of *mat-1* and *cgh-1* using the temperature-sensitive alleles, I performed two key experiments: hatch rate assays and propagation assays. The first consists in observing and quantifying the total progeny of each worm and the number of eggs that hatched. The final hatch rate is obtained by dividing the number of hatched eggs over the total number of progenies. The second consists in observing long-term propagation of worms and assessing if strains are able to achieve adulthood and reproduce into multiple generations. The combination of both assays is necessary for full understanding of the phenotype of *mat-1* and *cgh-1*, given that essential genes are not only the ones necessary for hatching, but also ones required for development and reproduction success (Jordan *et al.* 2002). Each assay was performed at a

range of different temperatures from 15°C to 25.3°C. Lower temperatures offer a permissive condition, in which MAT-1 and CGH-1 are still functioning. With temperature increase, proteins become unstable and non-functional, allowing for the investigation of the effect of each genetic background in the presence of each perturbation.

Both hatch rate and propagation assays uncovered great phenotypic variability for *mat-1* and *cgh-1*. To begin with, hatch rate assays were performed at the permissive temperature of 15°C. At that condition, no impairment in protein is expected, and indeed, there was no reduction in hatch rate for any of the natural isolates, containing either *mat-1* or *cgh-1* alleles. However, when temperature was raised to more restrictive conditions, what was observed was a variable expressivity of the lethal phenotype. Some backgrounds, such as AB1 and GXW1, were severely impacted and had lower hatch rates. On the other hand, KR314 background presented the highest hatch rate for both *mat-1* and *cgh-1* experiments in any given restrictive condition (>20°C).

A similar trend was observed for propagation assays, also performed at both permissive and restrictive conditions. For *mat-1(ye121)* strains, temperatures as low as 21.7°C were already restrictive enough to impair propagation of AB1. At a higher condition of 23.2°C, almost all the strains had inhibited propagation, with the exception of KR314 and N2 backgrounds. However, once temperature was increased, only KR314 was able to survive, even if mildly. For *cgh-1(tn691)* strains, impairment of propagation initiated at 24°C, in which KR314 background was still able to propagate, while others only presented mild population growth. At any temperature higher than that, KR314 remained with the best performance, showing mild propagation, while other wild isolates failed to grow in population size.

Taken together, both hatch rate assays and propagation assays have a common result: the presence of phenotypic variability, and the finding that the KR314 background was the most

“resistant” to the lethal phenotype of *mat-1* and *cgh-1*. This is a suggestive of the presence of unique variants acting as suppressors in this specific background. In addition, some other isolates, such as AB1, CB4856 and GXW1, showed worst performance in comparison with reference N2, also raising the possibility of presence of enhancers in these backgrounds. To confirm that any phenotypic variability here observed was solely caused by differences in each genetic background, and not the effect of natural sensitivity to temperature or other environmental factors, the same experiments and conditions were replicated using all six WT strains as controls. No lethal phenotype was observed in any of the temperatures tested for hatch rates or propagation assays. Additionally, during each experiment and stock maintenance, worms were rigorously cultured using similar conditions, environment, and food, to avoid any external influence in the phenotype of the target genes. Therefore, any variable expressivity observed here, most likely is due to differences in the genetic background of each wild isolate.

6.2 Natural Genetic Variation and its Influence on Gene Expression

Natural genetic variation in gene expression influences the phenotypes and it is an important factor for evolutionary processes (Oleksiak *et al.* 2002). Variation can be extensive among organisms with diverse genotypes, and in *C. elegans*, it is estimated that variation in gene expression can be up to 70% (Viñuela *et al.* 2012; Volkens *et al.* 2013). The question of whether gene expression variation is translated into proteome variation is less understood, but was assessed by Kamkina *et al.* (2016) by comparing the N2 and CB4856 backgrounds. They identified a list of over 1,500 genes differentially expressed between N2 and CB4856, which were enriched for genetic variations and Expression Quantitative Trait Loci (eQTLs), directly linking the gene expression divergencies to the presence of genetic variants. Additionally, they observed that the strongest mRNA expression divergence was during L4 stage (Kamkina *et al.* 2016). Using stable

isotope labeling by amino acids in cell culture (SILAC) and mass spectrometry, they investigated if gene expression variation would result in protein abundance variation. Out of the ~3,000 proteins measured, they identified 129 proteins with significantly different expression at both protein and gene levels, being most related to insulin-signaling and stress-response pathways. Most of them were instead found to be differentially expressed only at the protein level. They also observed an association between presence of eQTLs and differential mRNA and protein levels expression: the majority of genes changed in both levels were enriched for eQTLs, while genes with no differential expression had little presence of these elements. eQTLs are genomic loci, or SNPs, that explain variation in expression levels of mRNAs (Nica and Dermitzakis 2013). Finally, they concluded that the functional implication of the differentially expressed proteins may be related to aging, and therefore influence N2 and CB4856 lifespan (Kamkina *et al.* 2016).

Here, comparison of RNA-Seq data from each wild isolate with reference N2 also showed differential gene expression and unique expression patterns for each background (Chapter 4). Another explanation for the phenotypic variability observed for *mat-1* and *cgh-1* (Chapter 3) could be the natural differential gene expression of the given natural isolates. Additionally, potential genetic modifiers could be acting by altering expression of the target genes. To investigate that, all six WT wild isolates were sent for whole RNA-Seq. Data analysis revealed a substantial natural differential gene expression for the wild isolates in comparison with the reference N2. However, no significant difference was observed for either the expression of *mat-1* or *cgh-1*, among all of the natural isolates. That, in accordance with the aforementioned experiments, reinforces the hypothesis of presence of genetic modifiers in these isolates that do not necessarily modulate the expression of these genes, yet are responsible for the plasticity in gene essentiality for *mat-1* and *cgh-1*.

6.3 Identifying Genetic Modifiers – An Ongoing Challenge

Based on the phenotypic investigation of *mat-1* and *cgh-1*, there could be the presence of enhancers or suppressors in each genetic background, or even the combination of both suppressors and enhancers in the same background, creating a unique and intricate network. In this project, an attempt to identify the potential modifiers for *mat-1* and *cgh-1* was made, combining modern bioinformatics analysis (in collaboration with colleagues from bioinformatics team in the Tarailo-Graovac lab), and reverse genetic techniques (CRISPR/Cas9). Investigation of candidate variants was conducted by combining multiple strategies (Chapter 4) and led to prioritization of eight candidates: four for *mat-1*, and four for *cgh-1*. The genes were selected based on known interactions with *mat-1* or *cgh-1*, presence in the same biological process/pathway, or functional and biological relevancy. The selected candidate modifiers for *mat-1* were in *such-1* (two variants), *bub-3*, and *Y73B6A.1*, while candidate modifiers for *cgh-1* were in *hbl-1*, *nhl-2*, *ddx-15*, and *glp-1*.

All variants were initially tested in the N2 background carrying the primary variant of interest (*mat-1(ye121)* or *cgh-1(tn691)*) and compared to the single mutant N2 background. The minimal number of variants present in N2 background facilitates the investigation of the candidate modifier without influence of other variants. All candidate modifiers were analysed by replicating the hatch rate and propagation assays, at same restrictive and permissive conditions utilized for studying *mat-1* and *cgh-1* phenotypes. If a variant is found to be a suppressor, it would lead to increase in hatch rate and amelioration of propagation. On the other hand, variants acting as enhancers will lead to reduction of hatch rate and propagation abilities. If no significant difference is observed between single and double mutants, the variant would then be disregarded as a genetic modifier for *mat-1* and *cgh-1*.

Of all variants tested for potential modification of *mat-1* and *cgh-1* phenotypes, none have shown a significant difference in comparison with respective controls. Some of these variants were present in genes that are already known to interact with *mat-1* or *cgh-1*, such as *hbl-2* and *nhl-2*. Still, the specific SNVs here studied were not sufficient to induce a phenotype modification on their own. One hypothesis is that the effect of these variants may be masked by presence of other modifiers in each background, or need to act in conjunction with other modifiers to produce a significant phenotype. Therefore, the conclusion is that none of the variants investigated in this project were modifying *mat-1* or *cgh-1*, or alternatively, could not produce significant results on its own, but may be able to modify the phenotypes when in a modifier network. In fact, the presence of intricate genetic interactions networks is an important factor to consider. Previously literature has shown that some genes may be modified by more than one variant, leading to extensive and complex interactions (Riordan and Nadeau 2017; Stuart 2022).

Mapping genetic interactions in outbred populations is a challenging task, which is facilitated by the use of inbred organisms such as worms and yeast (Dixon *et al.* 2009). Genetic interactions may be classified into “negative” and “positive” interactions. The first, often refers to enhancement of the primary phenotype, and even synthetic lethality, in which no phenotype is observed in single mutants that operate in parallel biological processes, whereas impairment of both elements blocks viability. The second group are positive interactions that manifest as suppressors that have ability to alleviate certain traits (Dixon *et al.* 2009).

Genetic modifiers, in general, may represent interaction networks that evolved to coordinate gene function during development and adulthood. This effective coordination is not a fixed feature, but rather a mutable activity according to genetic background and environmental exposures, which turns each background into a beautifully intricate interaction network (Riordan

and Nadeau 2017). Interestingly, no specific mechanism is associated with the appearance of modifiers. In fact, any sequence change capable of altering a gene's function may be considered as a modifier, which is dependant on genetic background and environmental contexts (Riordan and Nadeau 2017). These usually represent variants affecting protein coding regions, but can also extend to UTRs, promoter regions, introns and other regulatory regions (Rodríguez *et al.* 2013).

The size of genetic interaction networks is astonishing. In yeast, it is assumed that complete network may contain over ~100,000 interactions (Tong *et al.* 2004). Essential genes can exhibit five times more genetic interactions than nonessential genes, and may act as hubs for intricate networks (Davierwala *et al.* 2005). Additionally, synthetic lethal genetic interactions are more common in genes from different pathways, while physical interactions are more common between genes from a shared pathway (Dixon *et al.* 2009). However, the presence of at least one essential gene in a pathway can increase the event of synthetic lethal genetic interactions (Boone *et al.* 2007; Bandyopadhyay *et al.* 2008).

To date, most large-scale studies investigated genetic interactions between two alleles. In reality, many relevant interactions are likely to involve more complex interactions of three or more genes. Additionally, modifiers can also influence the phenotype of multigenic traits, in which they are predicted to modify a least one, and perhaps more genes involved in that characteristic (Riordan and Nadeau 2017). The improvement in reverse genetic techniques and computational tools may facilitate the examination of such issue (Dixon *et al.* 2009). Taken together, all these factors explain the challenges of understanding genetic interactions and identifying genetic modifiers, which would be fundamental for assertive prediction of genetic modifiers or the specific conditions (genetic and environmental) that could induce their modifying effect (Riordan and Nadeau 2017).

6.4 The Use of WGS and Reverse Genetics

The adoption of WGS in this project was an important strategy for the candidate genetic modifiers identification, as its efficiency had been proved in mutagenesis screens, but not yet established for natural isolates investigation. More traditional methods such as mapping loci and targeted genomic sequencing are usually more time consuming and have been slowly replaced by WGS in modifier screens, due to its increasing affordability. However, despite its efficiency, most researchers are still resistant to using WGS and rely on more conventional strategies, most likely due to its high cost and necessary bioinformatics training.

More recently, the use of WGS for modifier identification was exemplified by Jean et al (2021), who identified intragenic modifiers for the temperature-sensitive allele *zyg-1(it25)*. The use of WGS, in that case, was crucial for identifying a reversion of the *zyg-1(it25)* allele in one of the EMS-created strains, and also allowed for pinpointing the contribution of each variant placed in different genetic backgrounds. In their work, they have also demonstrated that WGS is more cost and time effective than Sanger sequencing, and allows for a more thorough investigation of extragenic variants, as it provides a complete picture of the genomic context (Jean *et al.* 2021). Combination of WGS with CRISPR/Cas9 for modifier identification has been proven useful for both intragenic and extragenic variant search in mutagenesis strains in a faster and more efficient manner. However, given the complexity of natural isolates' genomes, this strategy requires further analytical and experimental development for use in this specific scenario: while typical EMS screens will induce the appearance of ~400 variants per genome, natural isolates backgrounds carry instead thousands of variants, which severely increases the challenge in identifying a genetic modifier.

Here, the use of WGS allied to reverse genetics was applied for investigation of natural isolates and the effect of each genetic background in the plasticity of essentiality of *mat-1* and *cgh-*

I and allowed for a broader interpretation of each wild isolates' background. Once phenotypic experiments for *mat-1* and *cgh-1* were performed, the resulting data served as a direction for the WGS analysis and candidate modifier prioritization, taking into consideration variants unique to the strains with the most outstanding phenotypes. CRISPR/Cas9, then, allowed for a precise and rapid testing of each candidate variant. Initial testing was performed with variants present in genes that were already known to modify *mat-1* or *cgh-1*. These novel variants could also be acting as modifiers, or have no effect (as shown), which could have only been discovered through experimental analysis.

Even though the identification of a true modifier was not achieved in this project, the application of such modern strategy has proved useful for obtaining a collection of variants that could be further investigated in the future for their potential modifying effect. This also highlights the importance of developing more efficient strategies for modifier identification in natural isolates, since they offer a much larger number of variants to start from, and one only knows the effect of each one after experimental validation. Even if efficient for identification of genetic modifiers in mutagenesis-derived lab strains, the use of WGS alone has been proved to be more challenging in wild isolates. Most models for bioinformatics studies were developed or trained using laboratory strains dataset, therefore indicating that new tools are necessary to assertively predict modifiers in natural isolates, a current work-in-progress in this scientific field. Riordan and Nadeau (2017) highlight how there is still a vast knowledge about genetic network biology that remains to be elucidated, which hampers the accurate prediction of genes with modifying potential and the conditions under their which activities rise.

Finally, WGS analysis for all six natural isolates in this project proved useful for identification (performed by Afiya Chida) and validation of previously unappreciated SNVs and

SVs (Chida, 2023). Most of the identified SVs were not prioritized as candidate modifiers, as no linkage to *mat-1* and *cgh-1* was noted. One variant in particular, in *alg-2*, was considered as a potential modifier for *cgh-1*, but not further investigated as no significant differential gene expression was observed in RNA-Seq data analysis (Chapter 4.3).

6.5 Future Directions

In continuity to this project, the exploration of RNA-seq data might be useful for potential modifier indicator, as differential gene expression may be a clue on the phenotype explanation for each wild isolate. Integration of WGS and RNA-Seq data may be useful since genetic variation is closely linked to differential gene expression, and these six natural isolates have shown unique expression patterns that could be further explored (Chapter 5) (Kamkina *et al.* 2016). Additionally, exploration of modifiers may be continued with remaining variants curated from WGS and ML analysis, focusing on variants that were not explored here yet, but show potential for modification of *mat-1* or *cgh-1*. Besides, the investigation of variants present in non-coding regions was not explored in this project, but could offer a new range of possibilities for investigation of phenotypic variability (e.g. intronic variants capable of altering protein splicing). Modifier variants, other more common in coding regions, may also affect UTRs and promoter regions (Rodríguez *et al.* 2013; Jean *et al.* 2021).

Another strategy for modifier identification could be outcrossing the strains with best hatch rate with the reference genome N2. This could be achieved, in this case, by crossing KR314 and N2, using a restrictive condition to select progeny carrying the potential modifier. Each new generation would be backcrossed to parental N2, until the remaining offspring carries a minimal number of variants that includes potential suppressors. Outcrossing, even though time-consuming, provides a final population with reduced number of variants that facilitates the curation of genetic

modifiers. These populations could be sent for WGS and variants be prioritized using the same strategy applied here to create a final – and shorter – candidate variant list.

6.6 Final Considerations – The Relevance of Investigating Genetic Modifiers of Gene Essentiality

Dixon et al (Dixon *et al.* 2009) reviewed the systematic mapping of genetic interactions at the time and concluded that the effects of different environmental and genetic background conditions in genetic interactions still remained to be better understood. Indeed, studying the same gene and phenotype under different conditions and genetic backgrounds may be fundamental for elucidating stable and plastic genetic interactions. For that purpose, *C. elegans* represents a great model organism, due to its easy manipulation, diverse natural genetic backgrounds, and inbred population.

C. elegans possesses a substantial diversity among its wild isolates and each of those genetic backgrounds might be used to probe conditional essentiality of the genes. The work here presented was of relevance for a better understanding of gene essentiality and genetic modification of that process. The advance of gene editing tools, (e.g. CRISPR/Cas9), allowed us to easily study this issue and prove the concept of plasticity of essentiality caused solely due to the influence of the genetic background. Digging into this helps with elucidating fundamental processes, genetic interactions, and demystify the concept of a static essentiality. Also, the work here presented can open new directions for genotype-phenotype studies and enlighten the possible consequences of disruption of the essential regulatory machinery of genes.

Above all, this project may cooperate to answer the question raised by Hodgkin (2001) a few years after *C. elegans* genome was sequenced for the very first time: “What are all these genes doing?”. Even after their functions were discovered, there is still much left to be elucidated about

their interactions. Besides assisting to the knowledge on *mat-1* and *cgh-1* phenotypes, I hope my work here will also provide grounds for new and intriguing discoveries on *C. elegans* natural phenotypic variability and genetic modification.

Additionally, the nematode *C. elegans* contains many orthologous genes to those considered important or involved in human disorders. In fact, *cgh-1* is an ortholog of human *DDX6* which was recently associated with a rare condition causing intellectual disability and dysmorphic features (Balak *et al.* 2019). *C. elegans* easy manipulation makes them the perfect tool for developing and tailoring a method for identification of genetic modifiers that might be translated into an efficient tool for genetic diseases prognosis in the long run. Genetic modifiers were already reported for human conditions such as in cystic fibrosis and spinal muscular dystrophy, and may represent a great ally in fighting these conditions and promoting health (Oprea *et al.* 2008; Emond *et al.* 2012).

References

- Abbott A. L., E. Alvarez-Saavedra, E. A. Miska, N. C. Lau, D. P. Bartel, *et al.*, 2005 The let-7 MicroRNA family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in *Caenorhabditis elegans*. *Dev. Cell* 9: 403–414.
<https://doi.org/10.1016/J.DEVCEL.2005.07.009>
- Abrahante J. E., A. L. Daul, M. Li, M. L. Volk, J. M. Tennessen, *et al.*, 2003 The *Caenorhabditis elegans* hunchback-like gene *lin-57/hbl-1* controls developmental time and is regulated by microRNAs. *Dev. Cell* 4: 625–637. [https://doi.org/10.1016/S1534-5807\(03\)00127-8](https://doi.org/10.1016/S1534-5807(03)00127-8)
- Afgan E., A. Nekrutenko, B. A. Grünig, D. Blankenberg, J. Goecks, *et al.*, 2022 The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Res.* 50: W345–W351. <https://doi.org/10.1093/NAR/GKAC247>
- Albert Hubbard E. J., and T. Schedl, 2019 Biology of the *Caenorhabditis elegans* Germline Stem Cell System. *Genetics* 213: 1145–1188. <https://doi.org/10.1534/GENETICS.119.300238>
- Amsterdam A., R. M. Nissen, Z. Sun, E. C. Swindell, S. Farrington, *et al.*, 2004 Identification of 315 genes essential for early zebrafish development. *Proc. Natl. Acad. Sci. U. S. A.* 101: 12792–12797. <https://doi.org/10.1073/PNAS.0403929101>
- Andersen E. C., and M. V. Rockman, 2022 Natural genetic variation as a tool for discovery in *Caenorhabditis* nematodes. *Genetics* 220. <https://doi.org/10.1093/GENETICS/IYAB156>
- Au V., E. Li-Leger, G. Raymant, S. Flibotte, G. Chen, *et al.*, 2019 CRISPR/Cas9 Methodology for the Generation of Knockout Deletions in *Caenorhabditis elegans*. *G3 Genes|Genomes|Genetics* 9: 135–144. <https://doi.org/10.1534/G3.118.200778>
- Austin J., and J. Kimble, 1987 *glp-1* Is required in the germ line for regulation of the decision between mitosis and meiosis in *C. elegans*. *Cell* 51: 589–599. [https://doi.org/10.1016/0092-8674\(87\)90128-0](https://doi.org/10.1016/0092-8674(87)90128-0)
- Baba T., T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, *et al.*, 2006 Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2: 2006.0008. <https://doi.org/10.1038/MSB4100050>
- Baker R. E., R. Harris, and K. Zhang, 1998 Mutations synthetically lethal with *cep1* target *S. cerevisiae* kinetochore components. *Genetics* 149: 73.
<https://doi.org/10.1093/GENETICS/149.1.73>
- Balak C., M. Benard, E. Schaefer, S. Iqbal, K. Ramsey, *et al.*, 2019 Rare De Novo Missense

- Variants in RNA Helicase DDX6 Cause Intellectual Disability and Dysmorphic Features and Lead to P-Body Defects and RNA Dysregulation. *Am. J. Hum. Genet.* 105: 509–525. <https://doi.org/10.1016/J.AJHG.2019.07.010>
- Bandyopadhyay S., R. Kelley, N. J. Krogan, and T. Ideker, 2008 Functional Maps of Protein Complexes from Quantitative Genetic Interaction Data. *PLoS Comput Biol* 4: 1000065. <https://doi.org/10.1371/journal.pcbi.1000065>
- Barbosa J., C. E. Sunkel, and C. Conde, 2022 The Role of Mitotic Kinases and the RZZ Complex in Kinetochore-Microtubule Attachments: Doing the Right Link. *Front. Cell Dev. Biol.* 10. <https://doi.org/10.3389/FCELL.2022.787294>
- Barrière A., and M. A. Félix, 2005a Natural variation and population genetics of *Caenorhabditis elegans*. *WormBook* 1–19. <https://doi.org/10.1895/wormbook.1.43.1>
- Barrière A., and M. A. Félix, 2005b High Local Genetic Diversity and Low Outcrossing Rate in *Caenorhabditis elegans* Natural Populations. *Curr. Biol.* 15: 1176–1184. <https://doi.org/10.1016/J.CUB.2005.06.022>
- Bartha I., J. Di Iulio, J. C. Venter, and A. Telenti, 2018 Human gene essentiality. *Nat. Rev. Genet.* 19: 51–62. <https://doi.org/10.1038/nrg.2017.75>
- Bartholomew D. P., R. A. Hawkins, and J. A. Lopez, 2012 Hawaii Pineapple: The Rise and Fall of an Industry. *HortScience* 47: 1390–1398. <https://doi.org/10.21273/HORTSCI.47.10.1390>
- Batut B., M. van den Beek, M. A. Doyle, and N. Soranzo, 2021 RNA-Seq Data Analysis in Galaxy. *Methods Mol. Biol.* 2284: 367–392. https://doi.org/10.1007/978-1-0716-1307-8_20/FIGURES/6
- Bell A. D., H. T. Chou, F. Valencia, and A. B. Paaby, 2023 Beyond the reference: gene expression variation and transcriptional response to RNA interference in *Caenorhabditis elegans*. *G3 Genes, Genomes, Genet.* 13: 1–12. <https://doi.org/10.1093/g3journal/jkad112>
- Benjamini Y., and Y. Hochberg, 1995 Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Source J. R. Stat. Soc. Ser. B* 57: 289–300.
- Berthelet S., J. Usher, K. Shulist, A. Hamza, N. Maltez, *et al.*, 2010 Functional Genomics Analysis of the *Saccharomyces cerevisiae* Iron Responsive Transcription Factor Aft1 Reveals Iron-Independent Functions. *Genetics* 185: 1111–1128. <https://doi.org/10.1534/GENETICS.110.117531>
- Bezler A., and P. Gönczy, 2010 Mutual Antagonism Between the Anaphase Promoting Complex

- and the Spindle Assembly Checkpoint Contributes to Mitotic Timing in *Caenorhabditis elegans*. *Genetics* 186: 1271–1283. <https://doi.org/10.1534/GENETICS.110.123133>
- Blomen V. A., P. Májek, L. T. Jae, J. W. Bigenzahn, J. Nieuwenhuis, *et al.*, 2015 Gene essentiality and synthetic lethality in haploid human cells. *Science* 350: 1092–1096. <https://doi.org/10.1126/SCIENCE.AAC7557>
- Boag P. R., A. Nakamura, and T. K. Blackwell, 2005 A conserved RNA-protein complex component involved in physiological germline apoptosis regulation in *C. elegans*. *Development* 132: 4975–4986. <https://doi.org/10.1242/DEV.02060>
- Bolger A. M., M. Lohse, and B. Usadel, 2014a Trimmomatic: a flexible trimmer for {Illumina} sequence data. *Bioinformatics* 30: 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Bolger A. M., M. Lohse, and B. Usadel, 2014b Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120. <https://doi.org/10.1093/BIOINFORMATICS/BTU170>
- Boone C., H. Bussey, and B. J. Andrews, 2007 Exploring genetic interactions and networks with yeast. *Nat. Rev. Genet.* 8: 437–449. <https://doi.org/10.1038/nrg2085>
- Bosch-Guiteras N., and J. van Leeuwen, 2022 Exploring conditional gene essentiality through systems genetics approaches in yeast. *Curr. Opin. Genet. Dev.* 76. <https://doi.org/10.1016/J.GDE.2022.101963>
- Boutros M., A. A. Kiger, S. Armknecht, K. Kerr, M. Hild, *et al.*, 2004 Genome-Wide RNAi Analysis of Growth and Viability in *Drosophila* Cells. *Science* (80-.). 303: 832–835. https://doi.org/10.1126/SCIENCE.1091266/SUPPL_FILE/BOUTROS.SOM.TABLES8.PDF
- Brenner S., 1974 THE GENETICS OF CAENORHABDITIS ELEGANS. *Genetics* 77: 71–94. <https://doi.org/10.1093/GENETICS/77.1.71>
- Bridge A. J., S. Pebernard, A. Ducraux, A. L. Nicoulaz, and R. Iggo, 2003 Induction of an interferon response by RNAi vectors in mammalian cells. *Nat. Genet.* 34: 263–264. <https://doi.org/10.1038/NG1173>
- Byrne A. B., M. T. Weirauch, V. Wong, M. Koeva, S. J. Dixon, *et al.*, 2007 A global analysis of genetic interactions in *Caenorhabditis elegans*. *J. Biol.* 6: 1–27. <https://doi.org/10.1186/JBIOL58/FIGURES/9>

- Cacheiro P., V. Muñoz-Fuentes, S. A. Murray, M. E. Dickinson, M. Bucan, *et al.*, 2020 Human and mouse essentiality screens as a resource for disease gene discovery. *Nat. Commun.* 2020 11: 1–16. <https://doi.org/10.1038/s41467-020-14284-2>
- Chan G. K. T., S. A. Jablonski, D. A. Starr, M. L. Goldberg, and T. J. Yen, 2000 Human Zw10 and ROD are mitotic checkpoint proteins that bind to kinetochores. *Nat. Cell Biol.* 2: 944–947. <https://doi.org/10.1038/35046598>
- Chen S., Y. E. Zhang, and M. Long, 2010 New Genes in *Drosophila* Quickly Become Essential. *Science* 330: 1682–1685. <https://doi.org/10.1126/science.1196380>
- Chen R., L. Shi, J. Hakenberg, B. Naughton, P. Sklar, *et al.*, 2016 Analysis of 589,306 genomes identifies individuals resilient to severe Mendelian childhood diseases. *Nat. Biotechnol.* 2016 34: 531–538. <https://doi.org/10.1038/nbt.3514>
- Chida A. R., 2023 Identifying the Genetic Factors in Natural Genome Backgrounds That May Modulate Phenotypic Outcomes in *C. elegans*. <https://doi.org/10.11575/PRISM/41620>
- Cho S. W., S. Kim, J. M. Kim, and J. S. Kim, 2013 Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* 31: 230–232. <https://doi.org/10.1038/NBT.2507>
- Chou H. T., F. Valencia, J. C. Alexander, A. D. Bell, D. Deb, *et al.*, 2023 Diversification of small RNA pathways underlies germline RNAi incompetence in wild *C. elegans* strains, (P. Wittkopp, Ed.). *Genetics*. <https://doi.org/10.1093/GENETICS/IYAD191>
- Chowdhury R., A. Chowdhury, and C. D. Maranas, 2015 Using Gene Essentiality and Synthetic Lethality Information to Correct Yeast and CHO Cell Genome-Scale Models. *Metab.* 2015, Vol. 5, Pages 536-570 5: 536–570. <https://doi.org/10.3390/METABO5040536>
- Cingolani P., A. Platts, L. L. Wang, M. Coon, T. Nguyen, *et al.*, 2012 A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*. 6: 80–92. https://doi.org/10.4161/FLY.19695/SUPPL_FILE/KFLY_A_10919695_SM0001.ZIP
- Cleveland D. W., Y. Mao, and K. F. Sullivan, 2003 Centromeres and kinetochores: From epigenetics to mitotic checkpoint signaling. *Cell* 112: 407–421. [https://doi.org/10.1016/S0092-8674\(03\)00115-6](https://doi.org/10.1016/S0092-8674(03)00115-6)
- Cohen-Fix O., J. M. Peters, M. W. Kirschner, and D. Koshland, 1996 Anaphase initiation in *Saccharomyces cerevisiae* is controlled by the APC-dependent degradation of the anaphase

inhibitor Pds1p. *Genes Dev.* 10: 3081–3093. <https://doi.org/10.1101/gad.10.24.3081>

Cong L., F. A. Ran, D. Cox, S. Lin, R. Barretto, *et al.*, 2013 Multiplex genome engineering using CRISPR/Cas systems. *Science* 339: 819–823. <https://doi.org/10.1126/SCIENCE.1231143>

Consortium* T. C. elegans S., 1998 Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* (80-.). 282: 2012–2018. https://doi.org/10.1126/SCIENCE.282.5396.2012/SUPPL_FILE/C-ELEGANS.XHTML

Cook D. E., S. Zdraljevic, J. P. Roberts, and E. C. Andersen, 2017 CeNDR, the *Caenorhabditis elegans* natural diversity resource. *Nucleic Acids Res.* 45: D650–D657. <https://doi.org/10.1093/NAR/GKW893>

Corsi A. K., B. Wightman, and M. Chalfie, 2015 A Transparent Window into Biology: A Primer on *Caenorhabditis elegans*. *Genetics* 200: 387–407. <https://doi.org/10.1534/GENETICS.115.176099>

Costanzo M., B. VanderSluis, E. N. Koch, A. Baryshnikova, C. Pons, *et al.*, 2016 A global genetic interaction network maps a wiring diagram of cellular function. *Science* (80-.). 353. https://doi.org/10.1126/SCIENCE.AAF1420/SUPPL_FILE/COSTANZO-SM.PDF

Coulson A., Y. Kozono, B. Lutterbach, R. Shownkeen, J. Sulston, *et al.*, 1991 YACs and the *C. elegans* genome. *Bioessays* 13: 413–417. <https://doi.org/10.1002/BIES.950130809>

Crittenden S. L., E. R. Troemel, T. C. Evans, and J. Kimble, 1994 GLP-1 is localized to the mitotic region of the *C. elegans* germ line. *Development* 120: 2901–2911. <https://doi.org/10.1242/DEV.120.10.2901>

Crombie T. A., S. Zdraljevic, D. E. Cook, R. E. Tanny, S. C. Brady, *et al.*, 2019 Deep sampling of Hawaiian *Caenorhabditis elegans* reveals high genetic diversity and admixture with global populations. *Elife* 8. <https://doi.org/10.7554/ELIFE.50465>

Dai J., and J. M. G. Higgins, 2005 Haspin: a mitotic histone kinase required for metaphase chromosome alignment. *Cell Cycle* 4: 665–668. <https://doi.org/10.4161/CC.4.5.1683>

Dai J., B. A. Sullivan, and J. M. G. Higgins, 2006 Regulation of mitotic chromosome cohesion by Haspin and Aurora B. *Dev. Cell* 11: 741–750. <https://doi.org/10.1016/J.DEVCEL.2006.09.018>

Danecek P., J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, *et al.*, 2021 Twelve years of SAMtools and BCFtools. *Gigascience* 10: 1–4. <https://doi.org/10.1093/gigascience/giab008>

Davierwala A. P., J. Haynes, Z. Li, R. L. Brost, M. D. Robinson, *et al.*, 2005 The synthetic

- genetic interaction spectrum of essential genes. *Nat. Genet.* 37: 1147–1152.
<https://doi.org/10.1038/NG1640>
- Davis E. S., L. Wille, B. A. Chestnut, P. L. Sadler, D. C. Shakes, *et al.*, 2002 Multiple subunits of the *Caenorhabditis elegans* anaphase-promoting complex are required for chromosome segregation during meiosis I. *Genetics* 160: 805–813.
<https://doi.org/10.1093/genetics/160.2.805>
- Desai A., S. Rybina, T. Müller-Reichert, A. Shevchenko, A. Shevchenko, *et al.*, 2003 KNL-1 directs assembly of the microtubule-binding interface of the kinetochore in *C. elegans*. *Genes Dev.* 17: 2421–2435. <https://doi.org/10.1101/GAD.1126303>
- Dicarlo J. E., J. E. Norville, P. Mali, X. Rios, J. Aach, *et al.*, 2013 Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* 41: 4336–4343.
<https://doi.org/10.1093/NAR/GKT135>
- Dickinson M. E., A. M. Flenniken, X. Ji, L. Teboul, M. D. Wong, *et al.*, 2016 High-throughput discovery of novel developmental phenotypes. *Nat.* 2016 5377621 537: 508–514.
<https://doi.org/10.1038/nature19356>
- Dixon S. J., M. Costanzo, A. Baryshnikova, B. Andrews, and C. Boone, 2009 Systematic mapping of genetic interaction networks. *Annu. Rev. Genet.* 43: 601–625.
<https://doi.org/10.1146/annurev.genet.39.073003.114751>
- Dobin A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, *et al.*, 2013 STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15–21.
<https://doi.org/10.1093/bioinformatics/bts635>
- Doitsidou M., R. J. Poole, S. Sarin, H. Bigelow, and O. Hobert, 2010 *C. elegans* Mutant Identification with a One-Step Whole-Genome-Sequencing and SNP Mapping Strategy. *PLoS One* 5: e15435. <https://doi.org/10.1371/JOURNAL.PONE.0015435>
- Dowell R. D., O. Ryan, A. Jansen, D. Cheung, S. Agarwala, *et al.*, 2010 Genotype to phenotype: A Complex problem. *Science* (80-.). 328: 469.
https://doi.org/10.1126/SCIENCE.1189015/SUPPL_FILE/DOWELL_SOM.PDF
- Doyle M. A., R. B. Gasser, B. J. Woodcroft, R. S. Hall, and S. A. Ralph, 2010 Drug target prediction and prioritization: using orthology to predict essentiality in parasite genomes. *BMC Genomics* 11. <https://doi.org/10.1186/1471-2164-11-222>
- Echeverri C. J., P. A. Beachy, B. Baum, M. Boutros, F. Buchholz, *et al.*, 2006 Minimizing the

- risk of reporting false positives in large-scale RNAi screens. *Nat. Methods* 3: 777–779.
<https://doi.org/10.1038/NMETH1006-777>
- Edgar R. S., G. H. Denhardt, and R. H. Epstein, 1964 A Comparative Genetic Study of Conditional Lethal Mutations of Bacteriophage T4d. *Genetics* 49: 635.
<https://doi.org/10.1093/GENETICS/49.4.635>
- Emond M. J., T. Louie, J. Emerson, W. Zhao, R. A. Mathias, *et al.*, 2012 Exome sequencing of extreme phenotypes identifies DCTN4 as a modifier of chronic *Pseudomonas aeruginosa* infection in cystic fibrosis. *Nat. Genet.* 2012 44: 886–889.
<https://doi.org/10.1038/ng.2344>
- Evans T. C., S. L. Crittenden, V. Kodoyianni, and J. Kimble, 1994 Translational control of maternal *glp-1* mRNA establishes an asymmetry in the *C. elegans* embryo. *Cell* 77: 183–194. [https://doi.org/10.1016/0092-8674\(94\)90311-5](https://doi.org/10.1016/0092-8674(94)90311-5)
- Ewels P., M. Magnusson, S. Lundin, and M. Källér, 2016 MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32: 3047–3048.
<https://doi.org/10.1093/bioinformatics/btw354>
- Fay D. S., H. M. Stanley, M. Han, and W. B. Wood, 1999 A *Caenorhabditis elegans* homologue of hunchback is required for late stages of development but not early embryonic patterning. *Dev. Biol.* 205: 240–253. <https://doi.org/10.1006/DBIO.1998.9096>
- Feuk L., A. R. Carson, and S. W. Scherer, 2006 Structural variation in the human genome. *Nat. Rev. Genet.* 2006 7: 85–97. <https://doi.org/10.1038/nrg1767>
- Fire A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, *et al.*, 1998 Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nat.* 1998 391: 806–811. <https://doi.org/10.1038/35888>
- Flibotte S., M. L. Edgley, I. Chaudhry, J. Taylor, S. E. Neil, *et al.*, 2010 Whole-genome profiling of mutagenesis in *Caenorhabditis elegans*. *Genetics* 185: 431–441.
<https://doi.org/10.1534/GENETICS.110.116616>
- Fraser A., 2015 Essential Human Genes. *Cell Syst.* 1: 381–382.
<https://doi.org/10.1016/J.CELS.2015.12.007>
- Frézal L., and M. A. Félix, 2015 *C. elegans* outside the Petri dish. *Elife* 2015.
<https://doi.org/10.7554/ELIFE.05849.001>
- Funabiki H., H. Yamano, K. Kumada, K. Nagao, T. Hunt, *et al.*, 1996 Cut2 proteolysis required

- for sister-chromatid separation in fission yeast. *Nat.* 1996 3816581 381: 438–441.
<https://doi.org/10.1038/381438a0>
- Garrison E., and G. Marth, 2012 Haplotype-based variant detection from short-read sequencing
- Gasch A. P., B. A. Payseur, and J. E. Pool, 2016 The Power of Natural Variation for Model Organism Biology. *Trends Genet.* 32: 147–154. <https://doi.org/10.1016/J.TIG.2015.12.003>
- Gawronski J. D., S. M. S. Wong, G. Giannoukos, D. V. Ward, and B. J. Akerley, 2009 Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for *Haemophilus* genes required in the lung. *Proc. Natl. Acad. Sci. U. S. A.* 106: 16422.
<https://doi.org/10.1073/PNAS.0906627106>
- Gluecksohn-Waelsch S., 1963 Lethal Genes and Analysis of Differentiation. *Sci.* 142 1269–1276.
- Golden A., P. L. Sadler, M. R. Wallenfang, J. M. Schumacher, D. R. Hamill, *et al.*, 2000 Metaphase to Anaphase (*mat*) Transition–Defective Mutants in *Caenorhabditis elegans*. *J. Cell Biol.* 151: 1469–1482. <https://doi.org/10.1083/JCB.151.7.1469>
- Grantham R., 1974 Amino acid difference formula to help explain protein evolution. *Science* (80-.). 185: 862–864. <https://doi.org/10.1126/SCIENCE.185.4154.862>
- Greenwald I. S., and H. R. Horvitz, 1982 DOMINANT SUPPRESSORS OF A MUSCLE MUTANT DEFINE AN ESSENTIAL GENE OF CAENORHABDITIS ELEGANS. *Genetics* 101: 211–225. <https://doi.org/10.1093/GENETICS/101.2.211>
- Gumienny T. L., E. Lambie, E. Hartweg, H. R. Horvitz, and M. O. Hengartner, 1999 Genetic control of programmed cell death in the *Caenorhabditis elegans* hermaphrodite germline. *Development* 126: 1011–1022. <https://doi.org/10.1242/DEV.126.5.1011>
- Haldane J. B. S., 1941 The relative importance of principal and modifying genes in determining some human diseases. *J. Genet.* 41: 149–157.
<https://doi.org/10.1007/BF02983018/METRICS>
- Hall T. M. T., 2005 Multiple modes of RNA recognition by zinc finger proteins. *Curr. Opin. Struct. Biol.* 15: 367–373. <https://doi.org/10.1016/J.SBI.2005.04.004>
- Ham T. J. Van, K. L. Thijssen, R. Breitling, R. M. W. Hofstra, R. H. A. Plasterk, *et al.*, 2008 *C. elegans* Model Identifies Genetic Modifiers of α -Synuclein Inclusion Formation During Aging. *PLOS Genet.* 4: e1000027. <https://doi.org/10.1371/JOURNAL.PGEN.1000027>
- Hammell C. M., I. Lubin, P. R. Boag, T. K. Blackwell, and V. Ambros, 2009 *nhl-2* Modulates

- MicroRNA Activity in *Caenorhabditis elegans*. *Cell* 136: 926–938.
<https://doi.org/10.1016/J.CELL.2009.01.053>
- Hansen D., L. Wilson-Berry, T. Dang, and T. Schedl, 2004a Control of the proliferation versus meiotic development decision in the *C. elegans* germline through regulation of GLD-1 protein accumulation. *Development* 131: 93–104. <https://doi.org/10.1242/DEV.00916>
- Hansen D., E. J. A. Hubbard, and T. Schedl, 2004b Multi-pathway control of the proliferation versus meiotic development decision in the *Caenorhabditis elegans* germline. *Dev. Biol.* 268: 342–357. <https://doi.org/10.1016/J.YDBIO.2003.12.023>
- Harris T. W., N. Chen, F. Cunningham, M. Tello-Ruiz, I. Antoshechkin, *et al.*, 2004 WormBase: a multi-species resource for nematode biology and genomics. *Nucleic Acids Res.* 32. <https://doi.org/10.1093/NAR/GKH066>
- Hart T., M. Chandrashekhar, M. Aregger, Z. Steinhart, K. R. Brown, *et al.*, 2015 High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell* 163: 1515–1526. <https://doi.org/10.1016/J.CELL.2015.11.015>
- Hodgkin J., and T. Doniach, 1997 Natural Variation and Copulatory Plug Formation in *Caenorhabditis elegans*. *Genetics* 146: 149–164.
<https://doi.org/10.1093/GENETICS/146.1.149>
- Hodgkin J., 2001 What does a worm want with 20,000 genes? *Genome Biol.* 2: comment2008.1. <https://doi.org/10.1186/GB-2001-2-11-COMMENT2008>
- Hou J., J. van Leeuwen, B. J. Andrews, and C. Boone, 2018 Genetic Network Complexity Shapes Background-Dependent Phenotypic Expression. *Trends Genet.* 34: 578. <https://doi.org/10.1016/J.TIG.2018.05.006>
- Hou J., G. Tan, G. R. Fink, B. J. Andrews, and C. Boone, 2019 Complex modifier landscape underlying genetic background effects. *Proc. Natl. Acad. Sci. U. S. A.* 116: 5045–5054. <https://doi.org/10.1073/PNAS.1820915116/-/DCSUPPLEMENTAL>
- Hoyt M. A., L. Totis, and B. T. Roberts, 1991 *S. cerevisiae* genes required for cell cycle arrest in response to loss of microtubule function. *Cell* 66: 507–517. [https://doi.org/10.1016/0092-8674\(81\)90014-3](https://doi.org/10.1016/0092-8674(81)90014-3)
- Hubstenberger A., S. L. Noble, C. Cameron, and T. C. Evans, 2013 Translation Repressors, an RNA Helicase, and Developmental Cues Control RNP Phase Transitions during Early Development. *Dev. Cell* 27: 161–173. <https://doi.org/10.1016/J.DEVCEL.2013.09.024>

- Hwang W. Y., Y. Fu, D. Reyon, M. L. Maeder, S. Q. Tsai, *et al.*, 2013 Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* 31: 227–229.
<https://doi.org/10.1038/NBT.2501>
- Iyer V., B. Shen, W. Zhang, A. Hodgkins, T. Keane, *et al.*, 2015 Off-target mutations are rare in Cas9-modified mice. *Nat. Methods* 12: 479. <https://doi.org/10.1038/nmeth.3408>
- Jean F., S. Stasiuk, T. Maroilley, C. Diao, A. Galbraith, *et al.*, 2021 Whole genome sequencing facilitates intragenic variant interpretation following modifier screening in *C. elegans*. *BMC Genomics* 22: 1–15. <https://doi.org/10.1186/S12864-021-08142-8/TABLES/6>
- Jiang Q., Y. Y. Ho, L. Hao, C. N. Berrios, and A. Chakravarti, 2011 Copy Number Variants in Candidate Genes Are Genetic Modifiers of Hirschsprung Disease. *PLoS One* 6: e21219.
<https://doi.org/10.1371/JOURNAL.PONE.0021219>
- Jiang W., D. Bikard, D. Cox, F. Zhang, and L. A. Marraffini, 2013 RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* 31: 233–239.
<https://doi.org/10.1038/NBT.2508>
- Johnsen R. C., and D. L. Baillie, 1991 Genetic analysis of a major segment [LGV(left)] of the genome of *Caenorhabditis elegans*. *Genetics* 129: 735–752.
<https://doi.org/10.1093/GENETICS/129.3.735>
- Johnson K. R., Q. Y. Zheng, and K. Noben-Trauth, 2006 Strain background effects and genetic modifiers of hearing in mice. *Brain Res.* 1091: 79–88.
<https://doi.org/10.1016/J.BRAINRES.2006.02.021>
- Jordan I. K., I. B. Rogozin, Y. I. Wolf, and E. V. Koonin, 2002 Essential genes are more evolutionarily conserved than are nonessential genes in bacteria. *Genome Res.* 12: 962–968.
<https://doi.org/10.1101/GR.87702>
- Jorgensen E. M., and S. E. Mango, 2002 The art and design of genetic screens: *Caenorhabditis elegans*. *Nat. Rev. Genet.* 3: 356–369. <https://doi.org/10.1038/nrg794>
- Juo P., and J. M. Kaplan, 2004 The anaphase-promoting complex regulates the abundance of GLR-1 glutamate receptors in the ventral nerve cord of *C. elegans*. *Curr. Biol.* 14: 2057–2062. <https://doi.org/10.1016/J.CUB.2004.11.010>
- Jurado P., E. Kodama, Y. Tanizawa, and I. Mori, 2010 Distinct thermal migration behaviors in response to different thermal gradients in *Caenorhabditis elegans*. *Genes, Brain Behav.* 9: 120–127. <https://doi.org/10.1111/J.1601-183X.2009.00549.X>

- Kabir M., A. Barradas, G. T. Tzotzos, K. E. Hentges, and A. J. Doig, 2017 Properties of genes essential for mouse development. *PLoS One* 12: e0178273.
<https://doi.org/10.1371/JOURNAL.PONE.0178273>
- Kaletta T., and M. O. Hengartner, 2006 Finding function in novel targets: *C. elegans* as a model organism. *Nat. Rev. Drug Discov.* 5: 387–399. <https://doi.org/10.1038/NRD2031>
- Kamath R. S., A. G. Fraser, Y. Dong, G. Poulin, R. Durbin, *et al.*, 2003 Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* 421: 231–237.
<https://doi.org/10.1038/NATURE01278>
- Kamath R. S., and J. Ahringer, 2003 Genome-wide RNAi screening in *Caenorhabditis elegans*. *Methods* 30: 313–321. [https://doi.org/10.1016/S1046-2023\(03\)00050-1](https://doi.org/10.1016/S1046-2023(03)00050-1)
- Kamkina P., L. B. Snoek, J. Grossmann, R. J. M. Volkers, M. G. Sterken, *et al.*, 2016 Natural genetic variation differentially affects the proteome and transcriptome in *Caenorhabditis elegans*. *Mol. Cell. Proteomics* 15: 1670–1680. <https://doi.org/10.1074/mcp.M115.052548>
- Kim D. U., J. Hayles, D. Kim, V. Wood, H. O. Park, *et al.*, 2010 Analysis of a genome-wide set of gene deletions in the fission yeast *Schizosaccharomyces pombe*. *Nat. Biotechnol.* 28: 617–623. <https://doi.org/10.1038/nbt.1628>
- Kim W., R. S. Underwood, I. Greenwald, and D. D. Shaye, 2018 OrthoList 2: A New Comparative Genomic Analysis of Human and *Caenorhabditis elegans* Genes. *Genetics* 210: 445–461. <https://doi.org/10.1534/GENETICS.118.301307>
- Kim H. M., Y. Hong, and J. Chen, 2022 A Decade of CRISPR-Cas Genome Editing in *C. elegans*. *Int. J. Mol. Sci.* 23. <https://doi.org/10.3390/IJMS232415863/S1>
- Kingdom R., and C. F. Wright, 2022 Incomplete Penetrance and Variable Expressivity: From Clinical Studies to Population Cohorts. *Front. Genet.* 13.
<https://doi.org/10.3389/FGENE.2022.920390>
- Kitagawa R., and A. M. Rose, 1999 Components of the spindle-assembly checkpoint are essential in *Caenorhabditis elegans*. *Nat. Cell Biol.* 1: 514–521.
<https://doi.org/10.1038/70309>
- Koch R., H. G. A. M. Van Luenen, M. Van Der Horst, K. L. Thijssen, and R. H. A. Plasterk, 2000 Single nucleotide polymorphisms in wild isolates of *Caenorhabditis elegans*. *Genome Res.* 10: 1690–1696. <https://doi.org/10.1101/GR.GR-1471R>
- Koo B. M., G. Kritikos, J. D. Farelli, H. Todor, K. Tong, *et al.*, 2017 Construction and Analysis

- of Two Genome-Scale Deletion Libraries for *Bacillus subtilis*. *Cell Syst.* 4: 291-305.e7.
<https://doi.org/10.1016/J.CELS.2016.12.013>
- Kops G. J. P. L., B. A. A. Weaver, and D. W. Cleveland, 2005 On the road to cancer: aneuploidy and the mitotic checkpoint. *Nat. Rev. Cancer* 5: 773–785. <https://doi.org/10.1038/NRC1714>
- Kraemer B. C., J. K. Burgess, J. H. Chen, J. H. Thomas, and G. D. Schellenberg, 2006 Molecular pathways that influence human tau-induced pathology in *Caenorhabditis elegans*. *Hum. Mol. Genet.* 15: 1483–1496. <https://doi.org/10.1093/HMG/DDL067>
- Kukhtar D., 2021 CRISPR-Cas9 to model retinitis pigmentosa caused by mutations in splicing factors in *C. Elegans*
- Kutscher L. M., and S. Shaham, 2014 Forward and reverse mutagenesis in *C. elegans*. *WormBook* 1–26. <https://doi.org/10.1895/WORMBOOK.1.167.1>
- Kuwabara P. E., and N. O’Neil, 2001 The use of functional genomics in *C. elegans* for studying human development and disease. *J. Inherit. Metab. Dis.* 24: 127–138.
<https://doi.org/10.1023/A:1010306731764>
- Kuzmin E., B. VanderSluis, W. Wang, G. Tan, R. Deshpande, *et al.*, 2018 Systematic analysis of complex genetic interactions. *Science* (80-.). 360: aao1729.
<https://doi.org/10.1126/SCIENCE.AAO1729>
- Langerak S., A. Trombley, J. R. Patterson, D. Leroux, A. Couch, *et al.*, 2019 Remodeling of the endoplasmic reticulum in *Caenorhabditis elegans* oocytes is regulated by CGH-1. *Genesis* 57. <https://doi.org/10.1002/DVG.23267>
- Leaver M., P. Domínguez-Cuevas, J. M. Coxhead, R. A. Daniel, and J. Errington, 2009 Life without a wall or division machine in *Bacillus subtilis*. 457.
<https://doi.org/10.1038/nature07742>
- Lee D., S. Zdraljevic, L. Stevens, Y. Wang, R. E. Tanny, *et al.*, 2021 Balancing selection maintains hyper-divergent haplotypes in *Caenorhabditis elegans*. *Nat. Ecol. Evol.* 2021 56 5: 794–807. <https://doi.org/10.1038/s41559-021-01435-x>
- Leeuwen J. Van, C. Pons, G. Tan, J. Z. Wang, J. Hou, *et al.*, 2020 Systematic analysis of bypass suppression of essential genes. *Mol. Syst. Biol.* 16: e9828.
<https://doi.org/10.15252/MSB.20209828>
- Lehner B., C. Crombie, J. Tischler, A. Fortunato, and A. G. Fraser, 2006 Systematic mapping of genetic interactions in *Caenorhabditis elegans* identifies common modifiers of diverse

- signaling pathways. *Nat. Genet.* 2006 38: 896–903. <https://doi.org/10.1038/ng1844>
- Lehrbach N. J., F. Ji, and R. Sadreyev, 2017 Next-Generation Sequencing for Identification of EMS-Induced Mutations in *Caenorhabditis elegans*. *Curr. Protoc. Mol. Biol.* 117. <https://doi.org/10.1002/CPMB.27>
- Li R., and A. W. Murray, 1991 Feedback control of mitosis in budding yeast. *Cell* 66: 519–531. [https://doi.org/10.1016/0092-8674\(81\)90015-5](https://doi.org/10.1016/0092-8674(81)90015-5)
- Li H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25: 1754–1760. <https://doi.org/10.1093/BIOINFORMATICS/BTP324>
- Li R., X. Ren, Q. Ding, Y. Bi, D. Xie, *et al.*, 2020 Direct full-length RNA sequencing reveals unexpected transcriptome complexity during *Caenorhabditis elegans* development. *Genome Res.* 30: 287–298. <https://doi.org/10.1101/GR.251512.119>
- Liu R., A. Z. Holik, S. Su, N. Jansz, K. Chen, *et al.*, 2015 Why weight? Modelling sample and observational level variability improves power in RNA-seq analyses. *Nucleic Acids Res.* 43. <https://doi.org/10.1093/nar/gkv412>
- Liu J. qing, and T. Li, 2019 CRISPR-Cas9-mediated loss-of-function screens. *Front. Life Sci.* 12: 1–13. <https://doi.org/10.1080/21553769.2019.1670739>
- Lu H., N. Lu, L. Weng, B. Yuan, Y. Liu, *et al.*, 2014 DHX15 Senses Double-Stranded RNA in Myeloid Dendritic Cells. *J. Immunol.* 193: 1364–1372. <https://doi.org/10.4049/jimmunol.1303322>
- Luo H., Y. Lin, F. Gao, C. T. Zhang, and R. Zhang, 2014 DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements. *Nucleic Acids Res.* 42. <https://doi.org/10.1093/NAR/GKT1131>
- Luo H., Y. Lin, T. Liu, F. L. Lai, C. T. Zhang, *et al.*, 2021 DEG 15, an update of the Database of Essential Genes that includes built-in analysis tools. *Nucleic Acids Res.* 49: D677–D686. <https://doi.org/10.1093/NAR/GKAA917>
- Maine E. M., 2008 Studying gene function in *Caenorhabditis elegans* using RNA-mediated interference. *Brief. Funct. Genomic. Proteomic.* 7: 184–194. <https://doi.org/10.1093/BFGP/ELN019>
- Manning G., 2005 Genomic overview of protein kinases *. 13. <https://doi.org/10.1895/wormbook.1.60.1>

- Maroille T., X. Li, M. Oldach, F. Jean, S. J. Stasiuk, *et al.*, 2021 Deciphering complex genome rearrangements in *C. elegans* using short-read whole genome sequencing. *Sci. Rep.* 11: 1–11. <https://doi.org/10.1038/s41598-021-97764-9>
- Maroille T., K. M. T. Hassan Rahit, A. R. Chida, F. Cotra, V. R. A. Barbosa, *et al.*, 2023 MOM: A user-friendly Galaxy workflow to detect modifiers from genome sequencing data using *C. elegans*, (J. Claycomb, Ed.). *G3* (Bethesda). <https://doi.org/10.1093/G3JOURNAL/JKAD184>
- Martínez-Carranza E., H. Barajas, L. D. Alcaraz, L. Servín-González, G. Y. Ponce-Soto, *et al.*, 2018 Variability of bacterial essential genes among closely related bacteria: The case of *Escherichia coli*. *Front. Microbiol.* 9: 1059. <https://doi.org/10.3389/FMICB.2018.01059/BIBTEX>
- Matthews J. M., and M. Sunde, 2002 Zinc fingers--folds for many occasions. *IUBMB Life* 54: 351–355. <https://doi.org/10.1080/15216540216035>
- McGrath P. T., M. V. Rockman, M. Zimmer, H. Jang, E. Z. Macosko, *et al.*, 2009 Quantitative Mapping of a Digenic Behavioral Trait Implicates Globin Variation in *C. elegans* Sensory Behaviors. *Neuron* 61: 692–699. <https://doi.org/10.1016/J.NEURON.2009.02.012>
- Meyer K. J., and M. G. Anderson, 2017 Genetic modifiers as relevant biological variables of eye disorders. *Hum. Mol. Genet.* 26: R58–R67. <https://doi.org/10.1093/HMG/DDX180>
- Moore L. L., M. Morrison, and M. B. Roth, 1999 Hcp-1, a Protein Involved in Chromosome Segregation, Is Localized to the Centromere of Mitotic Chromosomes in *Caenorhabditis elegans*. *J. Cell Biol.* 147: 471. <https://doi.org/10.1083/JCB.147.3.471>
- Mori H., T. Baba, K. Yokoyama, R. Takeuchi, W. Nomura, *et al.*, 2015 Identification of essential genes and synthetic lethal gene combinations in *Escherichia coli* K-12. *Methods Mol. Biol.* 1279: 45–65. https://doi.org/10.1007/978-1-4939-2398-4_4/FIGURES/8
- Morley J. F., H. R. Brignull, J. J. Weyers, and R. I. Morimoto, 2002 The threshold for polyglutamine-expansion protein aggregation and cellular toxicity is dynamic and influenced by aging in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U. S. A.* 99: 10417–10422. <https://doi.org/10.1073/PNAS.152161099>
- Moura D. S., I. Campillo-Marcos, M. Vázquez-Cedeira, and P. A. Lazo, 2018 VRK1 and AURKB form a complex that cross inhibit their kinase activity and the phosphorylation of histone H3 in the progression of mitosis. *Cell. Mol. Life Sci.* 75: 2591.

<https://doi.org/10.1007/S00018-018-2746-7>

- Musacchio A., and E. D. Salmon, 2007 The spindle-assembly checkpoint in space and time. *Nat. Rev. Mol. Cell Biol.* 8: 379–393. <https://doi.org/10.1038/NRM2163>
- Navarro R. E., E. Y. Shim, Y. Kohara, A. Singson, and T. K. Blackwell, 2001 *cgh-1*, a conserved predicted RNA helicase required for gametogenesis and protection from physiological germline apoptosis in *C. elegans*. *Development* 128: 3221–3232. <https://doi.org/10.1242/DEV.128.17.3221>
- Ng L., A. Rüsç, L. L. Amma, K. Nordström, L. C. Erway, *et al.*, 2001 Suppression of the deafness and thyroid dysfunction in *Thrb*-null mice by an independent mutation in the *Thra* thyroid hormone receptor alpha gene. *Hum. Mol. Genet.* 10: 2701–2708. <https://doi.org/10.1093/HMG/10.23.2701>
- Nica A. C., and E. T. Dermitzakis, 2013 Expression quantitative trait loci: Present and future. *Philos. Trans. R. Soc. B Biol. Sci.* 368. <https://doi.org/10.1098/rstb.2012.0362>
- Nijman S. M. B., 2011 Synthetic lethality: General principles, utility and detection using genetic screens in human cells. *FEBS Lett.* 585: 1–6. <https://doi.org/10.1016/J.FEBSLET.2010.11.024>
- Nystul T. G., J. P. Goldmark, P. A. Padilla, and M. B. Roth, 2003 Suspended animation in *C. elegans* requires the spindle checkpoint. *Science* 302: 1038–1041. <https://doi.org/10.1126/SCIENCE.1089705>
- Oegema K., A. Desai, S. Rybina, M. Kirkham, and A. A. Hyman, 2001 Functional analysis of kinetochore assembly in *Caenorhabditis elegans*. *J. Cell Biol.* 153: 1209–1225. <https://doi.org/10.1083/JCB.153.6.1209>
- Okahata M., A. Ohta, H. Mizutani, Y. Minakuchi, A. Toyoda, *et al.*, 2016 Natural variations of cold tolerance and temperature acclimation in *Caenorhabditis elegans*. *J. Comp. Physiol. B Biochem. Syst. Environ. Physiol.* 186: 985–998. <https://doi.org/10.1007/S00360-016-1011-3/FIGURES/5>
- Oleksiak M. F., G. A. Churchill, and D. L. Crawford, 2002 Variation in gene expression within and among natural populations. *Nat. Genet.* 32: 261–266. <https://doi.org/10.1038/ng983>
- Oprea G. E., S. Kröber, M. L. McWhorter, W. Rossoll, S. Müller, *et al.*, 2008 *Plastin 3* is a protective modifier of autosomal recessive spinal muscular atrophy. *Science* 320: 524–527. <https://doi.org/10.1126/SCIENCE.1155085>

- Overlack K., T. Bange, F. Weissmann, A. C. Faesen, S. Maffini, *et al.*, 2017 BubR1 Promotes Bub3-Dependent APC/C Inhibition during Spindle Assembly Checkpoint Signaling. *Curr. Biol.* 27: 2915-2927.e7. <https://doi.org/10.1016/j.cub.2017.08.033>
- Paaby A. B., A. G. White, D. D. Riccardi, K. C. Gunsalus, F. Piano, *et al.*, 2015 Wild worm embryogenesis harbors ubiquitous polygenic modifier variation. *Elife* 4. <https://doi.org/10.7554/ELIFE.09178>
- Papp B., C. Pál, and L. D. Hurst, 2004 Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. *Nat.* 2004 4296992 429: 661–664. <https://doi.org/10.1038/nature02636>
- Pazdernik N., and T. Schedl, 2013 Introduction to germ cell development in *Caenorhabditis elegans*. *Adv. Exp. Med. Biol.* 757: 1–16. https://doi.org/10.1007/978-1-4614-4015-4_1
- Peter M., A. Castro, T. Lorca, C. Le Peuch, L. Magnaghi-Jaulin, *et al.*, 2001 The APC is dispensable for first meiotic anaphase in *Xenopus* oocytes. *Nat. Cell Biol.* 3: 83–87. <https://doi.org/10.1038/35050607>
- Peters J. M., 2002 The anaphase-promoting complex: Proteolysis in mitosis and beyond. *Mol. Cell* 9: 931–943. [https://doi.org/10.1016/S1097-2765\(02\)00540-3](https://doi.org/10.1016/S1097-2765(02)00540-3)
- Pierre S. E. St., L. Ponting, R. Stefancsik, and P. McQuilton, 2014 FlyBase 102--advanced approaches to interrogating FlyBase. *Nucleic Acids Res.* 42. <https://doi.org/10.1093/NAR/GKT1092>
- Pollo S. M. J., A. Leon-Coria, H. Liu, D. Cruces-Gonzalez, C. A. M. Finney, *et al.*, 2023 Transcriptional patterns of sexual dimorphism and in host developmental programs in the model parasitic nematode *Heligmosomoides bakeri*. *Parasites and Vectors* 16: 1–24. <https://doi.org/10.1186/S13071-023-05785-2/TABLES/4>
- Priess J. R., H. Schnabel, and R. Schnabel, 1987 The *glp-1* locus and cellular interactions in early *C. elegans* embryos. *Cell* 51: 601–611. [https://doi.org/10.1016/0092-8674\(87\)90129-2](https://doi.org/10.1016/0092-8674(87)90129-2)
- Prior T. W., A. R. Krainer, Y. Hua, K. J. Swoboda, P. C. Snyder, *et al.*, 2009 A Positive Modifier of Spinal Muscular Atrophy in the SMN2 Gene. *Am. J. Hum. Genet.* 85: 408–413. <https://doi.org/10.1016/J.AJHG.2009.08.002>
- Qian W., D. Ma, C. Xiao, Z. Wang, and J. Zhang, 2012 The Genomic Landscape and Evolutionary Resolution of Antagonistic Pleiotropy in Yeast. *Cell Rep.* 2: 1399–1410. <https://doi.org/10.1016/J.CELREP.2012.09.017>

- Rahit K. M. T. H., and M. Tarailo-Graovac, 2020 Genetic {Modifiers} and {Rare} {Mendelian} {Disease}. *Genes (Basel)*. 11: 239. <https://doi.org/10.3390/genes11030239>
- Rancati G., J. Moffat, A. Typas, and N. Pavelka, 2017 Emerging and evolving concepts in gene essentiality. *Nat. Rev. Genet.* 2017 19: 34–49. <https://doi.org/10.1038/nrg.2017.74>
- Rappleye C. A., A. Tagawa, R. Lyczak, B. Bowerman, and R. V. Aroian, 2002 The anaphase-promoting complex and separin are required for embryonic anterior-posterior axis formation. *Dev. Cell* 2: 195–206. [https://doi.org/10.1016/S1534-5807\(02\)00114-4](https://doi.org/10.1016/S1534-5807(02)00114-4)
- Richaud A., G. Zhang, D. Lee, J. Lee, and M. A. Félix, 2018 The local coexistence pattern of selfing genotypes in *Caenorhabditis elegans* natural metapopulations. *Genetics* 208: 807–821. <https://doi.org/10.1534/GENETICS.117.300564/-/DC1>
- Riordan J. D., and J. H. Nadeau, 2017 From Peas to Disease: Modifier Genes, Network Resilience, and the Genetics of Health. *Am. J. Hum. Genet.* 101: 177–191. <https://doi.org/10.1016/j.ajhg.2017.06.004>
- Robbins M. A., M. Li, I. Leung, H. Li, D. V. Boyer, *et al.*, 2006 Stable expression of shRNAs in human CD34+ progenitor cells can avoid induction of interferon responses to siRNAs in vitro. *Nat. Biotechnol.* 24: 566–571. <https://doi.org/10.1038/NBT1206>
- Rodríguez J. M., S. Wolfrum, M. Robblee, K. Y. Chen, Z. N. Gilbert, *et al.*, 2013 Altered expression of Raet1e, a major histocompatibility complex class 1-like molecule, underlies the atherosclerosis modifier locus Ath11 10b. *Circ. Res.* 113: 1054–1064. <https://doi.org/10.1161/CIRCRESAHA.113.302052>
- Rousset F., J. Cabezas-Caballero, F. Piastra-Facon, J. Fernández-Rodríguez, O. Clermont, *et al.*, 2021 The impact of genetic diversity on gene essentiality within the *Escherichia coli* species. *Nat. Microbiol.* 2021 6: 301–312. <https://doi.org/10.1038/s41564-020-00839-y>
- Rylaarsdam L., and A. Guemez-Gamboa, 2019 Genetic Causes and Modifiers of Autism Spectrum Disorder. *Front. Cell. Neurosci.* 13. <https://doi.org/10.3389/FNCEL.2019.00385>
- S. A., 2010 {FastQC}: a quality control tool for high throughput sequence data
- Sega G. A., 1984 A review of the genetic effects of ethyl methanesulfonate. *Mutat. Res.* 134: 113–142. [https://doi.org/10.1016/0165-1110\(84\)90007-1](https://doi.org/10.1016/0165-1110(84)90007-1)
- Shakes D. C., P. L. Sadler, J. M. Schumacher, M. Abdolrasulnia, and A. Golden, 2003 Developmental defects observed in hypomorphic anaphase-promoting complex mutants are linked to cell cycle abnormalities. *Development* 130: 1605–1620.

- <https://doi.org/10.1242/DEV.00385>
- Shelton C. A., and B. Bowerman, 1996 Time-dependent responses to glp-1-mediated inductions in early *C. elegans* embryos. *Development* 122: 2043–2050.
<https://doi.org/10.1242/dev.122.7.2043>
- Shi J., E. Wang, J. P. Milazzo, Z. Wang, J. B. Kinney, *et al.*, 2015 Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol.* 33: 661–667.
<https://doi.org/10.1038/NBT.3235>
- Sin O., H. Michels, and E. A. A. Nollen, 2014 Genetic screens in *Caenorhabditis elegans* models for neurodegenerative diseases ☆. <https://doi.org/10.1016/j.bbadis.2014.01.015>
- Singh K. D., B. Roschitzki, L. B. Snoek, J. Grossmann, X. Zheng, *et al.*, 2016 Natural Genetic Variation Influences Protein Abundances in *C. elegans* Developmental Signalling Pathways. *PLoS One* 11: e0149418. <https://doi.org/10.1371/JOURNAL.PONE.0149418>
- Snoek B. L., M. G. Sterken, R. P. J. Bevers, R. J. M. Volkers, A. van't Hof, *et al.*, 2017 Contribution of trans regulatory eQTL to cryptic genetic variation in *C. elegans*. *BMC Genomics* 18: 1–15. <https://doi.org/10.1186/S12864-017-3899-8/FIGURES/6>
- Stein K. K., E. S. Davis, T. Hays, and A. Golden, 2007 Components of the Spindle Assembly Checkpoint Regulate the Anaphase-Promoting Complex During Meiosis in *Caenorhabditis elegans*. *Genetics* 175: 107. <https://doi.org/10.1534/GENETICS.106.059105>
- Stein K. K., J. E. Nesmith, B. D. Ross, and A. Golden, 2010 Functional Redundancy of Paralogs of an Anaphase Promoting Complex/Cyclosome Subunit in *Caenorhabditis elegans* Meiosis. *Genetics* 186: 1285–1293. <https://doi.org/10.1534/GENETICS.110.123463>
- Stewart M. K., N. L. Clark, G. Merrihew, E. M. Galloway, and J. H. Thomas, 2005 High genetic diversity in the chemoreceptor superfamily of *Caenorhabditis elegans*. *Genetics* 169: 1985–1996. <https://doi.org/10.1534/genetics.104.035329>
- Stuart A. R., 2022 UNIVERSITY OF CALGARY Investigating Multiple Modifier Network Candidates as Suppressors of *zyg-1/PLK4* in *C. elegans*
- Sulston J. E., and H. R. Horvitz, 1977 Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev. Biol.* 56: 110–156. [https://doi.org/10.1016/0012-1606\(77\)90158-0](https://doi.org/10.1016/0012-1606(77)90158-0)
- Sulston J. E., and H. R. Horvitz, 1981 Abnormal cell lineages in mutants of the nematode *Caenorhabditis elegans*. *Dev. Biol.* 82: 41–55. [https://doi.org/10.1016/0012-1606\(81\)90427-](https://doi.org/10.1016/0012-1606(81)90427-)

- Swan K. A., D. E. Curtis, K. B. McKusick, A. V. Voinov, F. A. Mapa, *et al.*, 2002 High-throughput gene mapping in *Caenorhabditis elegans*. *Genome Res.* 12: 1100–1105. <https://doi.org/10.1101/gr.208902>
- Tabara H., A. Grishok, and C. C. Mello, 1998 RNAi in *C. elegans*: soaking in the genome sequence. *Science* 282: 430–431. <https://doi.org/10.1126/SCIENCE.282.5388.430>
- Taieb F. E., S. D. Gross, A. L. Lewellyn, and J. L. Maller, 2001 Activation of the anaphase-promoting complex and degradation of cyclin B is not required for progression from Meiosis I to II in *Xenopus* oocytes. *Curr. Biol.* 11: 508–513. [https://doi.org/10.1016/S0960-9822\(01\)00145-2](https://doi.org/10.1016/S0960-9822(01)00145-2)
- Talbert P. B., C. D. S. LeCiel, and S. Henikoff, 1994 Modification of the *Drosophila* Heterochromatic Mutation Brown(dominant) by Linkage Alterations. *Genetics* 136: 559. <https://doi.org/10.1093/GENETICS/136.2.559>
- Tamura M., T. Isojima, T. Kasama, R. Mafune, K. Shimoda, *et al.*, 2017 Novel DHCR7 mutation in a case of Smith-Lemli-Opitz syndrome showing 46,XY disorder of sex development. *Hum. genome Var.* 4. <https://doi.org/10.1038/HGV.2017.15>
- Tarailo-Graovac M., J. Wang, D. Tu, D. L. Baillie, A. M. Rose, *et al.*, 2010 Duplication of *cyb-3* (cyclin B3) suppresses sterility in the absence of *mdf-1/MAD1* spindle assembly checkpoint component in *Caenorhabditis elegans*. *Cell Cycle* 9: 4858–4865. <https://doi.org/10.4161/CC.9.24.14137>
- Tarailo-Graovac M., T. Wong, Z. Qin, S. Flibotte, J. Taylor, *et al.*, 2014 Cyclin B3 and dynein heavy chain cooperate to increase fitness in the absence of *mdf-1/MAD1* in *Caenorhabditis elegans*. *Cell Cycle* 13: 3089–3199. <https://doi.org/10.4161/15384101.2014.949491>
- Tarailo-Graovac M., J. Y. A. Zhu, A. Matthews, C. D. M. van Karnebeek, and W. W. Wasserman, 2017 Assessment of the ExAC data set for the presence of individuals with pathogenic genotypes implicated in severe Mendelian pediatric disorders. *Genet. Med.* 19: 1300. <https://doi.org/10.1038/GIM.2017.50>
- Tarailo M., R. Kitagawa, and A. M. Rose, 2007a Suppressors of Spindle Checkpoint Defect (such) Mutants Identify New *mdf-1/MAD1* Interactors in *Caenorhabditis elegans*. *Genetics* 175: 1665–1679. <https://doi.org/10.1534/GENETICS.106.067918>
- Tarailo M., S. Tarailo, and A. M. Rose, 2007b Synthetic Lethal Interactions Identify Phenotypic

- “Interologs” of the Spindle Assembly Checkpoint Components. *Genetics* 177: 2525.
<https://doi.org/10.1534/GENETICS.107.080408>
- Tariq A., W. Garncarz, C. Handl, A. Balik, O. Pusch, *et al.*, 2013 RNA-interacting proteins act as site-specific repressors of ADAR2-mediated RNA editing and fluctuate upon neuronal stimulation. *Nucleic Acids Res.* 41: 2581–2593. <https://doi.org/10.1093/NAR/GKS1353>
- Thompson O., M. Edgley, P. Strasbourger, S. Flibotte, B. Ewing, *et al.*, 2013 The million mutation project: A new approach to genetics in *Caenorhabditis elegans*. *Genome Res.* 23: 1749. <https://doi.org/10.1101/GR.157651.113>
- Thompson O. A., L. B. Snoek, H. Nijveen, M. G. Sterken, R. J. M. Volkers, *et al.*, 2015 Remarkably divergent regions punctuate the genome assembly of the *Caenorhabditis elegans* hawaiian strain CB4856. *Genetics* 200: 975–989.
<https://doi.org/10.1534/GENETICS.115.175950/-/DC1>
- Tijsterman M., K. L. Okihara, K. Thijssen, and R. H. A. Plasterk, 2002 PPW-1, a PAZ/PIWI protein required for efficient germline RNAi, is defective in a natural isolate of *C. elegans*. *Curr. Biol.* 12: 1535–1540. [https://doi.org/10.1016/S0960-9822\(02\)01110-7](https://doi.org/10.1016/S0960-9822(02)01110-7)
- Timmons L., D. L. Court, and A. Fire, 2001 Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene* 263: 103–112. [https://doi.org/10.1016/S0378-1119\(00\)00579-5](https://doi.org/10.1016/S0378-1119(00)00579-5)
- Timmons L., H. Tabara, C. C. Mello, and A. Z. Fire, 2003 Inducible systemic RNA silencing in *Caenorhabditis elegans*. *Mol. Biol. Cell* 14: 2972–2983. <https://doi.org/10.1091/MBE.E03-01-0858/ASSET/IMAGES/LARGE/MK0732279005.JPEG>
- Tischler J., B. Lehner, and A. G. Fraser, 2008 Evolutionary plasticity of genetic interaction networks. *Nat. Genet.* 40: 390–391. <https://doi.org/10.1038/NG.114>
- Tong A. H. Y., G. Lesage, G. D. Bader, H. Ding, H. Xu, *et al.*, 2004 Global Mapping of the Yeast Genetic Interaction Network. *Science* (80-.). 303: 808–813.
https://doi.org/10.1126/SCIENCE.1091317/SUPPL_FILE/TONG.SOM.PDF
- Tzur Y. B., A. E. Friedland, S. Nadarajan, G. M. Church, J. A. Calarco, *et al.*, 2013 Heritable custom genomic modifications in *Caenorhabditis elegans* via a CRISPR-Cas9 system. *Genetics* 195: 1181–1185. <https://doi.org/10.1534/GENETICS.113.156075/-/DC1>
- Uhlmann F., F. Lottspeltch, and K. Nasmyth, 1999 Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1. *Nat.* 1999 4006739 400: 37–42.

<https://doi.org/10.1038/21831>

- Uhlmann F., D. Wernic, M. A. Poupert, E. V. Koonin, and K. Nasmyth, 2000 Cleavage of cohesin by the CD clan protease separin triggers anaphase in yeast. *Cell* 103: 375–386. [https://doi.org/10.1016/S0092-8674\(00\)00130-6](https://doi.org/10.1016/S0092-8674(00)00130-6)
- Veres A., B. S. Gosis, Q. Ding, R. Collins, A. Ragavendran, *et al.*, 2014 Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. *Cell Stem Cell* 15: 27–30. <https://doi.org/10.1016/J.STEM.2014.04.020>
- Vergara I. A., M. Tarailo-Graovac, C. Frech, J. Wang, Z. Qin, *et al.*, 2014 Genome-wide variations in a natural isolate of the nematode *Caenorhabditis elegans*. *BMC Genomics* 15: 255. <https://doi.org/10.1186/1471-2164-15-255>
- Viñuela A., L. Basten Snoek, J. A. G. Riksen, and J. E. Kammenga, 2012 Aging uncouples heritability and expression-qtL in *caenorhabditis elegans*. *G3 Genes, Genomes, Genet.* 2: 597–605. <https://doi.org/10.1534/g3.112.002212>
- Volkers R. J. M., L. B. Snoek, C. J. va. H. Hubar, R. Coopman, W. Chen, *et al.*, 2013 Gene-environment and protein-degradation signatures characterize genomic and phenotypic diversity in wild *Caenorhabditis elegans* populations. *BMC Biol.* 11: 1–13. <https://doi.org/10.1186/1741-7007-11-93>
- Vu V., A. J. Verster, M. Schertzberg, T. Chuluunbaatar, M. Spensley, *et al.*, 2015 Natural Variation in Gene Expression Modulates the Severity of Mutant Phenotypes. *Cell* 162: 391–402. <https://doi.org/10.1016/J.CELL.2015.06.037>
- Wade Harper J., J. L. Burton, and M. J. Solomon, 2002 The anaphase-promoting complex: it's not just for mitosis any more. *Genes Dev.* 16: 2179–2206. <https://doi.org/10.1101/GAD.1013102>
- Wallenfang M. R., and G. Seydoux, 2000 Polarization of the anterior–posterior axis of *C. elegans* is a microtubule-directed process. *Nat.* 2000 4086808 408: 89–92.
- Wang H., H. Yang, C. S. Shivalila, M. M. Dawlaty, A. W. Cheng, *et al.*, 2013 One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153: 910–918. <https://doi.org/10.1016/J.CELL.2013.04.025>
- Wang T., K. Birsoy, N. W. Hughes, K. M. Krupczak, Y. Post, *et al.*, 2015 Identification and characterization of essential genes in the human genome. *Science* 350: 1096–1101.

- <https://doi.org/10.1126/SCIENCE.AAC7041>
- Wang Y., K. He, B. Sheng, X. Lei, W. Tao, *et al.*, 2021 The RNA helicase Dhx15 mediates Wnt-induced antimicrobial protein expression in Paneth cells. *Proc. Natl. Acad. Sci. U. S. A.* 118. <https://doi.org/10.1073/PNAS.2017432118/-/DCSUPPLEMENTAL>
- Weiss E., and M. Winey, 1996 The *Saccharomyces cerevisiae* spindle pole body duplication gene MPS1 is part of a mitotic checkpoint. *J. Cell Biol.* 132: 111–123. <https://doi.org/10.1083/JCB.132.1.111>
- Wicks S. R., R. T. Yeh, W. R. Gish, R. H. Waterston, and R. H. A. Plasterk, 2001 Rapid gene mapping in *Caenorhabditis elegans* using a high density polymorphism map. *Nat. Genet.* 28: 160–164. <https://doi.org/10.1038/88878>
- Winzeler E. A., D. D. Shoemaker, A. Astromoff, H. Liang, K. Anderson, *et al.*, 1999 Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* (80-.). 285: 901–906. https://doi.org/10.1126/SCIENCE.285.5429.901/SUPPL_FILE/1040380S1_THUMB.GIF
- Wright C. F., D. R. FitzPatrick, and H. V. Firth, 2018 Paediatric genomics: diagnosing rare disease in children. *Nat. Rev. Genet.* 2018 19: 253–268. <https://doi.org/10.1038/nrg.2017.116>
- Xing J., X. Zhou, M. Fang, E. Zhang, L. J. Minze, *et al.*, 2021 DHX15 is required to control RNA virus-induced intestinal inflammation. *Cell Rep.* 35: 109205. <https://doi.org/10.1016/J.CELREP.2021.109205>
- Yang Z., J. Guo, Q. Chen, C. Ding, J. Du, *et al.*, 2005 Silencing mitosis induces misaligned chromosomes, premature chromosome decondensation before anaphase onset, and mitotic cell death. *Mol. Cell. Biol.* 25: 4062–4074. <https://doi.org/10.1128/MCB.25.10.4062-4074.2005>
- Yoshiki A., and K. Moriwaki, 2006 Mouse phenome research: Implications of genetic background. *ILAR J.* 47: 94–102. <https://doi.org/10.1093/ILAR.47.2.94/ILAR-47-2-94FIG3.GIF>
- Zarocostas J., 2006 Serious birth defects kill at least three million children a year. *BMJ* 332: 256. <https://doi.org/10.1136/BMJ.332.7536.256-B>
- Zhan T., and M. Boutros, 2015 Towards a compendium of essential genes – From model organisms to synthetic lethality in cancer cells.

- <https://doi.org/10.3109/10409238.2015.1117053> 51: 74–85.
- <https://doi.org/10.3109/10409238.2015.1117053>
- Zhang R., H. Y. Ou, and C. T. Zhang, 2004 DEG: a database of essential genes. *Nucleic Acids Res.* 32: D271–D272. <https://doi.org/10.1093/NAR/GKH024>
- Zhang Z., and Q. Ren, 2015 Why are essential genes essential? - The essentiality of *Saccharomyces* genes. *Microb. Cell* 2: 280. <https://doi.org/10.15698/MIC2015.08.218>
- Zhang M., C. Wang, T. D. Otto, J. Oberstaller, X. Liao, *et al.*, 2018 Uncovering the essential genes of the human malaria parasite *Plasmodium falciparum* by saturation mutagenesis. *Science* 360. <https://doi.org/10.1126/SCIENCE.AAP7847>
- Zhang G., N. M. Roberto, D. Lee, S. R. Hahnel, and E. C. Andersen, 2022 The impact of species-wide gene expression variation on *Caenorhabditis elegans* complex traits. *Nat. Commun.* 2022 13: 1–13. <https://doi.org/10.1038/s41467-022-31208-4>
- Zhong W., and P. W. Sternberg, 2006 Genome-wide prediction of *C. elegans* genetic interactions. *Science* (80-.). 311: 1481–1484.
https://doi.org/10.1126/SCIENCE.1123287/SUPPL_FILE/ZHONGW.SOM.PDF
- Zhou Q., K. Rammohan, S. Lin, N. Robinson, O. Li, *et al.*, 2003 CD24 is a genetic modifier for risk and progression of multiple sclerosis. *Proc. Natl. Acad. Sci. U. S. A.* 100: 15041–15046. <https://doi.org/10.1073/PNAS.2533866100/ASSET/DD3824C7-AE48-44B5-B9B0-B8BD64F2E4AB/ASSETS/GRAPHIC/ZPQ0250332240005.JPEG>
- Zuryn S., S. Le Gras, K. Jamet, and S. Jarriault, 2010 A strategy for direct mapping and identification of mutations by whole-genome sequencing. *Genetics* 186: 427–430.
<https://doi.org/10.1534/GENETICS.110.119230>