THE GENETIC ARCHITECTURE UNDERLYING THE CAENORHABDITIS ELEGANS RESPONSE TO GRASSLAND SOIL BACTERIA AND ITS EFFECTS ON FITNESS

by

VINOD KURUMATHURMADAM NAMBOOTHIRIPAD MONY

M.S., PURDUE UNIVERSITY, 2007

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2013

Abstract

Soil nematode communities are important components of the micro fauna in grassland ecosystems and their interaction with soil microbes affects important ecological processes such as decomposition and nutrient recycling. To study genetic mechanisms underlying ecologically important traits involved in the response of nematode communities to soil microbes, we employed genomic tools available for the model nematode, *Caenorhabditis elegans*. Previous work identified 204 *C. elegans* genes that were differentially expressed in response to growth on four different bacteria: *Bacillus megaterium*, *Pseudomonas sp*., *Micrococcus luteus* and *Escherichia coli.* For many of the genes the degree of differential gene expression between two bacterial environments predicted the magnitude of the effect of the loss of gene function on life-history traits in those environments. Mutations can have differential effects on fitness in variable environments, which can influence their maintenance in a population. Our fitness assays revealed that bacterial environments had varying magnitude of stress, defined as an environment in which the wild-type has a relatively low fitness. We performed fitness assays as part of a comprehensive analysis of life history traits on thirty five strains that contained mutations in genes involved in the *C. elegans* response to *E. coli, B. megaterium*, *Pseudomonas sp*. We found that many of the mutations had conditionally beneficial effects and led to increased fitness when nematodes bearing them were exposed to stressful bacteria. We compared the relative fitness of strains bearing these mutations across bacterial environments and found that the deleterious effects of many mutations were alleviated in the presence of stressful bacteria.

Although transcriptional profiling studies can identify genes that are differentially regulated in response to environmental stimuli, how the expressed genes provide functional specificity to a particular environment remains largely unknown. We focused on defense and metabolism genes involved in *C. elegans*-bacterial interactions and measured the survivorship of loss-of-function mutants in these genes exposed to different bacteria. We found that genes had both bacteria-specific and bacteria-shared responses. We then analyzed double mutant strains and found bacteria-specific genetic interaction effects. Plasticity in gene interactions and their environment-specific modulation have important implications for host phenotypic differentiation and adaptation to changing environments.

THE GENETIC ARCHITECTURE UNDERLYING THE CAENORHABDITIS ELEGANS RESPONSE TO GRASSLAND SOIL BACTERIA AND ITS EFFECTS ON FITNESS

by

VINOD KURUMATHURMADAM NAMBOOTHIRIPAD MONY

M.S., PURDUE UNIVERSITY, 2007

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2013

Approved by:

Major Professor
Michael Herman

Copyright

Abstract

Soil nematode communities are important components of the micro fauna in grassland ecosystems and their interaction with soil microbes affects important ecological processes such as decomposition and nutrient recycling. To study genetic mechanisms underlying ecologically important traits involved in the response of nematode communities to soil microbes, we employed genomic tools available for the model nematode, *Caenorhabditis elegans*. Previous work identified 204 *C. elegans* genes that were differentially expressed in response to growth on four different bacteria: *Bacillus megaterium*, *Pseudomonas sp*., *Micrococcus luteus* and *Escherichia coli.* For many of the genes the degree of differential gene expression between two bacterial environments predicted the magnitude of the effect of the loss of gene function on life-history traits in those environments. Mutations can have differential effects on fitness in variable environments, which can influence their maintenance in a population. Our fitness assays revealed that bacterial environments had varying magnitude of stress, defined as an environment in which the wild-type has a relatively low fitness. We performed fitness assays as part of a comprehensive analysis of life history traits on thirty five strains that contained mutations in genes involved in the *C. elegans* response to *E. coli, B. megaterium*, *Pseudomonas sp*. We found that many of the mutations had conditionally beneficial effects and led to increased fitness when nematodes bearing them were exposed to stressful bacteria. We compared the relative fitness of strains bearing these mutations across bacterial environments and found that the deleterious effects of many mutations were alleviated in the presence of stressful bacteria.

Although transcriptional profiling studies can identify genes that are differentially regulated in response to environmental stimuli, how the expressed genes provide functional specificity to a particular environment remains largely unknown. We focused on defense and metabolism genes involved in *C. elegans*-bacterial interactions and measured the survivorship of loss-of-function mutants in these genes exposed to different bacteria. We found that genes had both bacteria-specific and bacteria-shared responses. We then analyzed double mutant strains and found bacteria-specific genetic interaction effects. Plasticity in gene interactions and their environment-specific modulation have important implications for host phenotypic differentiation and adaptation to changing environments.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my Advisor Michael Herman for doing as much as he could to help me develop my professional and scientific skills. When I first met him, he told me to “keep away the fear of statistics”. My background was such that I was not confident enough that I would be able to do research, analyze it and contribute ideas. Thanks to his guidance, now I am confident that I would be able to work and conduct research on my own. Although, quite sometimes I might not have come up to his expectations, I appreciate his efforts to help me whether to better my presentation skills during any meetings, journal clubs or encouraging me to find research ideas or patiently correcting even grammatical errors in my thesis.

I want to thank for helpful comments from my committee members, Theodore Morgan, Anthony Joern and Jeremy Marshall especially in developing the second chapter in this thesis. I also wanted to thank Brian Darby for helping with statistics and explaining many concepts in ecology.

I would take this opportunity to thank past and current members of Herman lab, especially Joe Coolon for setting the stage for my research and Connor Brass for his friendship. I also thank Tony Grace and Byju Govindan for their immense support and guidance.

Dedication

I want to dedicate this dissertation to a few important people in my life.

My Grand Mother, although not with me physically, I still get inspiration remembering her love. My mother, who is the most committed and caring person I know and a great pillar of strength. My father, who inspires me with humility, simplicity and love. My brother, who was the inspiration and support behind my consideration to come to USA for higher studies. I also wanted to mention my mother-in-law from whom countless blessings have come on my way. And importantly, Sri Sri Ravishankar, my dearest friend, teacher and guide.

Sunil Prabhu and Priyamvada Prabhu for their care like my local guardians in Kansas.

My friends, Hanumantha Shanker and Predeesh Chandran, for their cheer and friendship.

My extended-family, Hari Kurumathur, Shalini Kurumathur and Sharika Kurumthur for their affection and inspiration.

My cousins, Neetha Parameswaran, Nisha Parameswaran and Maya Neelakandhan, for their love and care.

Lastly, but most importantly, Abhaya, my wife, whose gentle care and motivation allowed me to sail through all testing times. Many times her love was taken for granted and I wanted to take this opportunity to thank her.

Chapter 1 - Fitness consequences of mutations in variable environments

Introduction

Mutations are the ultimate source of genetic variation. Mutations can be broadly classified as deleterious, causing reduction in fitness, neutral, with little or no detectable fitness affect and beneficial, that increase fitness by helping organisms adapt to the environment (Eyre-Walker and Keightley 2007). Each generation populations will harbor new mutations (Agrawal and Whitlock 2012). The vast majority of spontaneously occurring mutations are suggested to have deleterious effects (Keightley and Lynch 2003). The probability of acquiring beneficial mutations and cost of the fidelity of DNA replication are suggested to be the selective forces that constrain the mutation rate to fall to zero (Sniegowski, Gerrish et al. 2000). Selection weeds out some of the deleterious mutations from a population at same rate as they are generated (Haldane 1937). The rate at which deleterious alleles are removed by selection and mutations generated reaches equilibrium at mutation-selection balance. "Mutation load" refers to the fitness reduction due to the presence of deleterious mutations segregating at mutation-selection balance (Agrawal and Whitlock 2012). A large fraction of genetic variation measured by quantitative genetic assays could be contributed by this pool of deleterious alleles which may not be important in adaptive evolution (Lynch, Latta et al. 1998).

Evolutionary consequences of deleterious mutations

Recurrent introduction of small-effect deleterious mutations in populations has been associated with many evolutionary consequences. It is implicated in many aspects of genome evolution, like deviations from the neutral expectation of molecular evolution (Ohta 1992), organelle genome evolution (Lynch and Blanchard 1998), synonymous codon usage bias in unicellular organisms (Bulmer 1991), preservation of duplicate genes (Force, Lynch et al. 1999), evolution of degenerate Y-chromosome and dosage compensation (Charlesworth and Charlesworth 1998). At the organism and population levels, it is suggested to be involved in the maintenance of genetic variation in populations by mutation-selection balance (Haldane 1937), extinction of small populations (Lynch, Conery et al. 1995) evolution of self-fertilization and inbreeding depression (Lande and Schemske 1985), evolution of diploidy (Otto and Goldstein 1992), sex and recombination (Kondrashov 1988) and aging (Partridge and Barton 1993).

Spontaneous deleterious mutations with environment-specific effects have also been found to contribute to genotype-environment interactions (GEI) for fitness traits (Fry, Heinsohn et al. 1996).

Environmental effects on fitness of specific mutations

The most direct influence of deleterious mutations in a population is reduction in fitness (Agrawal and Whitlock 2012). Though the small effect mutations have been invoked to explain many aspects of evolution, they are not considered to be the "stuff of adaptation" (Agrawal and Whitlock 2012). However, in order to understand the nature of genetic variations within populations it is important to know the rate of new mutations and their fitness effects (Halligan and Keightley 2009). Studies using mutation accumulation lines (MA lines) in various organisms have provided insight into the rates and fitness effects of spontaneous deleterious mutations (Drake, Charlesworth et al. 1998; Lynch, Blanchard et al. 1999; Halligan and Keightley 2009). Although studies with MA lines can reveal the overall effect of random mutations and their fitness effects, it does not provide clues about the functional significance of genes involved. Studying the effects of a large number of unknown mutations may not be helpful to identify how specific gene functions are influenced by mutations. Also, the absence of a mutation-free reference genotype makes it difficult to differentiate the specific effects of new mutations. Empirical investigation of fitness consequences of specific mutations can be performed only once their functional effects are known. Also, the accurate measurement of single mutation effects is only possible when it causes a large effect (higher or lower than 1% compared to wild-type) on fitness (Eyre-Walker and Keightley 2007).

Major aspects to be considered while studying how populations evolve are the genotype of the organisms, how the phenotype is expressed from the genotype and how phenotypic fitness varies according to environment (Elena and de Visser 2003). Phenotypic plasticity, that explains changes in expressed phenotype as a function of environment, could be due to differences in the environment-specific expression of alleles as well as interactions among loci (Scheiner 1993). The fitness consequences of a mutation will ultimately depend on the genetic background in which it appears and the environment in which its effect manifests or both (Remold and Lenski 2001). In order to understand how selective forces might act on mutations of specific genes and

how fitness effect of mutations changes in variable environments, mutations of single genes need to be analyzed.

No single genotype maximize fitness in all environments (Fry, Heinsohn et al. 1996). A large amount of genetic variation could be maintained due to genotype-environment interactions. In heterogeneous environments, different allelic variants could be selected at a locus which can potentially lead to sympatric speciation (Demeeus, Michalakis et al. 1993; Fry, Heinsohn et al. 1996). The equilibrium frequency of deleterious alleles in a population will depend on the strength of selection (Haldane 1937). If the deleterious fitness effect of a mutation is exacerbated in a specific environment, selection is stronger and more effective at removing them from the population (MacLellan, Kwan et al. 2012). Along with changes in the strength of selection, predictions relating to mutation load, inbreeding depression and genetic variance in fitness also change (Agrawal and Whitlock 2010). Environment specific effects of deleterious alleles can cause mutational collapse of fitness in marginal habitats that lead to the evolution of ecological specialization (Kawecki, Barton et al. 1997). Populations of organisms might respond to variations in the environment and evolve into niche specialists or generalists (Elena and Sanjuan 2003). A population that is adapted to be a specialist will be under constant stabilizing selection where selection purges out deleterious mutations (Elena and Sanjuan 2003).

Classification of mutations based on their environment specific fitness effects

Environmental variability can be generally classified as stressful and non-stressful depending on its effect on fitness (Martin and Lenormand 2006). An environment can be considered to be stressful when absolute fitness of wild-type is reduced when compared to its absolute fitness in some other reference context (Martin and Lenormand 2006; Agrawal and Whitlock 2010). Similarly, depending on its interaction with different environments, the fitness of a particular genotype can be classified into three categories (Kondrashov and Houle 1994; Elena and de Visser 2003): 1) If there is a mutation that disrupts an essential function of the organism, its fitness is expected to be reduced in both stressful and benign environments making it unconditionally deleterious. Even when unconditionally deleterious, the relative effect on fitness due to the mutation may be different in each environment (Elena and de Visser 2003). 2) If the mutation affects the organisms in such a way that its effect matches with requirements of a specific environment, it could be conditionally neutral, being deleterious in some environments,

but neutral in other environments. 3) The mutation could be conditionally beneficial, providing a beneficial effect in one environment, but deleterious effect in another (Kondrashov and Houle 1994; Elena and de Visser 2003). If the mutation is unconditionally deleterious, it will be removed from the population and may not have any long term impact in terms of its evolutionary consequence (Elena and de Visser 2003). But conditionally beneficial mutations can cause the maintenance of genetic polymorphism and in the extreme case could lead to ecological specialization in marginal habitats (Kondrashov and Houle 1994; Kawecki, Barton et al. 1997; Elena and de Visser 2003). Genotypes might also show trade-offs in fitness related traits across different environments (Fry 1996) since without trade-offs a single genotype would be expected to have high fitness among all environments (Elena and Lenski 2003). Antagonistic pleiotropic effects on fitness, as in the case of conditionally beneficial mutations, with harmful effects in one environment, but beneficial effects in another, is one of the mechanisms that has been suggested to account for the occurrence of such trade-offs (Elena and Lenski 2003). Conditionally neutral mutations cannot be identified easily, being indistinguishable from the wild-type in benign environments (Kondrashov and Houle 1994).

Although the above classification explains possible effects of the environment on mutations in qualitative terms, quantitative changes to fitness effects also have evolutionary consequences (Elena and de Visser 2003; Kishony and Leibler 2003). If in a stressful environment the deleterious effect of a mutation remains unchanged from that of a benign environment, it does not contribute to plasticity in fitness effects. But for some unconditionally deleterious mutations, magnitude of fitness reduction might be different across environments (Kondrashov and Houle 1994). Stressful environments can cause further aggravation of deleterious effect of mutations relative to a benign environment. Different scenarios have been suggested that might lead to this aggravated deleterious effect of mutations due to stress (Agrawal and Whitlock 2010). Under normal conditions, organisms have "margins of safety" in physiological functions with more capacity for such functions. Various physiological response/repair systems might buffer the effect of mutations. But if mutations affect such functions, they can cause an aggravated effect in stressful environments. Stressful environments may also cause increased differences in fitness by affecting the ecological relationship between individuals, for example, increased competition for resources cause disadvantages for mutants. Organisms could be more accustomed to benign environments where extended periods of

exposure and selection might have led to removal of deleterious effect mutations. However, if stressful environments were a hitherto unexposed environment, more mutations might exhibit fitness defects.

The fitness defect of deleterious mutations can also be alleviated by a stressful environment (Kishony and Leibler 2003). Recent studies have shown that stress can cause buffering of mutational effects (Casanueva, Burga et al. 2012). If in benign environment, the buffering systems are not active, the mutation might cause fitness defect. But since stressful environments induce buffering effects, it can reduce the deleterious effect of a mutation. Also, organisms may not reach their full genetic potential under stressful conditions. Then the fitness reduction caused by a mutation would be much less in a stressful environment as compared to a benign environment where the fitness is substantially high (Hoffmann and Merila 1999; Kishony and Leibler 2003). Martin and Lenormand, using a fitness landscape model, have provided several predictions for the effect of mutations in stressful environments (Martin and Lenormand 2006). They predict that mutations should exhibit more variable fitness effects in stressful environments than in a benign environment (Martin and Lenormand 2006; MacLellan, Kwan et al. 2012). Also, unconditionally deleterious mutations should have higher fitness defects in stressful environments (Martin and Lenormand 2006; Wang, Sharp et al. 2009). Lastly, some of the deleterious alleles might become beneficial in a stressful environment (Martin and Lenormand 2006; MacLellan, Kwan et al. 2012).

Studies involving the effect of environment on fitness of mutations

Various studies using mutations in specific genes that tested their fitness effects in stressful versus non-stressful environments have provided mixed results. Remold and Lenski (2001) performed a study in which 26 random insertion mutations in *E. coli* were exposed to limiting nutrient conditions (maltose instead of glucose) and low temperature (28 degrees instead of 37 degrees) as stressful environments. They found that the variance in fitness effects due to mutations were higher in a limited nutrient environment, but did not change with low temperature. Thus, depending on the environment, selection intensities on mutations can vary. This resource dependent variance in fitness effect might be due to canalization, where phenotype becomes insensitive to effect of mutations, in selected environments (Remold and Lenski 2001). They also found that 3 mutations (12%) showed improved fitness in the low resource

environment and became conditionally beneficial. If the genotype is at its fitness peak (as in benign environment), there is less chance of any mutation to be beneficial (Orr 1998; Remold and Lenski 2001). Absence of selection in a novel environment (as in stressful environment) might provide more chances for exposing the beneficial effects of mutations whereas in a benign environment most mutations tend to be detrimental. Another study with 216 *E. coli* genotypes carrying random insertion mutations with one, two or three mutations was performed in the presence and absence of a plasmid parasite as the stress element (Cooper, Lenski et al. 2005). They found that on average, the plasmid parasite aggravated the harmful effects of deleterious mutations. This also depended on specific genotypes with reduced the severity of effect for some mutations, but aggravated effect for others. This study shows that parasite loads can increase the severity of mutations. Muller's ratchet effect predicts stochastic loss of fittest genotype due to mutation load in finite populations (Lynch, Burger et al. 1993). Since recombination allows the genome of the progeny to be different from parents, this theory proposes a basis for the evolution of sex. If parasites can increase the deleterious effect of mutations, it can provide stronger reasons for recombination to be advantageous in populations.. Another study in which 65 random mutations in *E. coli* were tested on a variety of stressful environments that included low temperature, chemicals and poor nutrients showed an average alleviation of deleterious effect of mutations (Kishony and Leibler 2003). Similarly, an average alleviation of defect caused by mutations was also found in a study with 526 single deletion lines of yeast interacting with a stressful environment like high temperature, poor nutrition, chemicals and salinity (Jasnos, Tomala et al. 2008).

Studies involving environmental effects on fitness of mutations in multi-cellular organisms found mostly aggravating or unchanging affects. In the fruit fly, EMS mutagenesis was performed on males and many quantitative traits including viability were measured in their out-bred heterozygous F3 progeny under poor nutrient conditions (Yang, Tanikawa et al. 2001). They found that the decline of viability in poor nutrient conditions was lower than in better nutrition conditions. An experiment using 20 mutations with visible phenotypic effects in adult fruit flies, compared the effects of high quality and low quality food and found that the average offspring survival was 42% lower in the low quality environment (Wang, Sharp et al. 2009). They also found variance in the fitness effect of mutations in low quality environment was 1.77 times larger than in high quality environment. Interestingly, a few mutations also appeared to be

beneficial in low quality environments. Another study in fruit flies that used eight visible mutations and measured larval viability in the presence or absence of bacterial pathogen, *Pseudomonas aeruginosa*, found that the mutations tend to be more deleterious in presence of the pathogen (Young, Yourth et al. 2009). But they also found that this effect was mainly due to three genes and the remainder did not have any change in fitness effect in presence or absence of the pathogen. Lastly, a study using 9 recessive mutations with visible phenotypes was tested for two fitness components, male mating success and productivity in standard diet and novel (stressful) diet (MacLellan, Kwan et al. 2012). They did not observe any difference in the average deleterious effects of mutations for both fitness parameters in stressful or benign environments. Specifically, they found two mutations with increased deleterious effect for productivity in the stressful environment.

Summary of environmental effects on fitness of mutations

The effect of stress on the deleterious effects of mutations was mixed; with some aggravating, some not changing and some alleviating effects. This trend remained irrespective of whether the mutations studies where performed on single-cellular or multi-cellular organisms or the kind of stress applied. Since the genes affected by the mutations are selected at random, one cannot be sure whether these alleles are experiencing any kind of selection pressures. Although random and unbiased selection of mutations will help to understand the general effect of stress on them, it is likely that many of their functions are completely irrelevant to the specific stress applied. The fitness trait examined was growth rate for many of the experiments with single-cellular organisms (Kishony and Leibler 2003; Jasnos, Tomala et al. 2008) and larval survival or mean productivity for most of the experiments with multi-cellular organisms (mainly fruit fly) (Yang, Tanikawa et al. 2001; Wang, Sharp et al. 2009; Young, Yourth et al. 2009; MacLellan, Kwan et al. 2012). So in many of these experiments, there are also chances that the effect of environment on mutations may not be completely reflected by the phenotypes studied.

To understand how natural selection acts on various alleles in a population, we need to know not only their phenotypic effect, but also the environmental effect on fitness components (MacLellan, Kwan et al. 2012). It is not practical to test all the alleles in a genome for their functional effects in different environments and it would be hard to pin point the targets of selection in any particular environment, so it warrants testing selected genotypes in a relevant

environment. Since fitness effect of genes depend on the environments in which they are expressed (Jackson, Linder et al. 2002), studying the effect of mutations on such genes in those environments will help to get a more realistic picture of how those functions evolve. Genes transcriptionally activated in a particular environment could be an important fraction that will undergo greatest changes in fitness in such environments and thereby selection pressures when mutated. By applying increased magnitude of stress that is relevant to the functions of such genes, any specific trends in relative fitness effect of mutations can be identified. Since there is a greater probability of selective influences from the environment on those mutations, predictions and testable hypothesis can be made regarding their fitness effects. For example if the stress causes an aggravating or alleviating effect on fitness in one environment, it can be tested to find similar trends in more stressful or less stressful environments. In order to understand how stresses affect mean selection of gene functions and also to get at the mechanisms by which these effects occur, mutations in individual genes have to be tested in relevant environments.

Epistasis in biological systems

It has been observed that biological systems in which the functions of genes are perturbed undergo large-scale coordinated changes in organismal phenotype (Garfield and Wray 2010). The high level of modularity of genetic systems is exemplified by the fact that in many model systems, gene interaction networks have been mapped (Costanzo, Baryshnikova et al. 2010)-(Yeast); (Lee, Lehner et al. 2008)- (*C. elegans*); (Wang, Marshall et al. 2012)-(*Arabidopsis thaliana*); (Costello, Dalkilic et al. 2009)-(*Drosophila melanogaster*)). Comparisons of networks of conserved genes and pathways derived in different species have shown that although there is extensive rewiring of individual interactions, genetic interactomes are governed by general design principles with conserved features of gene interactions and conserved functional crosstalk between biological processes (Koch, Costanzo et al. 2012; Ryan, Roguev et al. 2012). Ryan et al suggested a hierarchical modularity in the evolution of genetic interactions with conservation highest within protein complexes, lower within biological processes for example; transcription, mRNA processing, translation; and lowest between them (Ryan, Roguev et al. 2012). Thus, in order to gain better insight into the functional effect of environmentally relevant genes and the effect of mutations on them, they will have to be addressed in the context of integrated systems.

When the phenotypic effect of genotype depends on the underlying genetic background it is referred as epistasis (Carlborg and Haley 2004). The term epistasis has been used in two different scenarios (Whitlock, Phillips et al. 1995; Phillips 2008). A gene is referred to be epistatic to another when its allele masks the effect of alleles at the other locus, suggesting the possibility that the genes involved might belong to a pathway that hierarchically modulates a biological function (Phillips 2008). In quantitative genetics, epistasis means any statistical deviation from additive effects of interacting loci in their contributions to the phenotype (Phillips 2008). There could be substantial epistatic gene action in which particular alleles can interact and produce phenotypes that vary drastically relative to their combination with alternate alleles. Epistasis of quantitative genetics is a function of frequency of alleles in a population. Thus, if the alleles that show epistatic gene action happen to be rare, they may not contribute much to the variance component (Whitlock, Phillips et al. 1995). So the absence of epistatic variance cannot be taken as the evidence of absence of epistatic gene action (Phillips 2008). Moreover, since the fitness functions of allelic combinations define the fitness of a phenotype, the evolutionary potential will depend on epistatic gene action rather than epistatic variance (Whitlock, Phillips et al. 1995). Since we are considering the effects of interaction of mutations on specific genes and the environmental influence on them, epistatic gene action is discussed further.

Evolutionary consequences of epistasis among deleterious mutations

Interactions between mutations have been suggested to play important roles in evolutionary processes (You and Yin 2002). Epistatic gene interaction among deleterious mutations changes the fitness consequences of the mutation which, in turn, has been shown to affect the mutation load (Kimura and Maruyama 1966). The nature of interactions among the mutations also affects genetic drift and the fixation of deleterious alleles (Phillips, Otto et al. 2000). When gene interactions are pervasive and variable, the fitness consequences of each mutation will depend on prior mutations (Phillips, Otto et al. 2000). With variation in epistatic effects of mutations, topography of adaptive landscape will constitute multiple peaks and valleys (Whitlock, Phillips et al. 1995; Phillips 2008). Epistatic effects among mutation also play an important role in models of the evolution of sex and recombination (Kondrashov 1988; de Visser and Elena 2007). In the mutational deterministic hypothesis, when the combined effect of two or more deleterious mutations is more severe than their independent effects, selection favors recombination, that breaks them apart (Peters and Lively 2000). Similarly, negative linkage

disequilibrium with combinations of advantageous and deleterious alleles, decreases the additive variance for fitness and slows down selection. Recombination can speed up the response to selection by breaking this combination and reducing the frequency of intermediate genotypes (Peters and Lively 2000). In the Dobzhansky-Muller hybrid incompatibility model of reproductive isolation, mutant alleles evolved separately in two genotypes can have an epistatic interaction when they are combined, generating a sterile or lethal hybrid (Coyne 1992). Effects of deleterious mutations and their interactions are suggested to be important in the evolution of diploidy (Kondrashov and Crow 1991). By examining protein evolution, a recent study has shown that 90 percent of all amino-acid substitutions have neutral or beneficial effect only in the genetic backgrounds in which they occur. In a different background or species, these might be deleterious pointing to the fact that epistatic effects are important in protein evolution (Breen, Kemena et al. 2012).

Classification of epistasis based on fitness effects

The way in which the mutations interact is important to understand their fitness effect and how they are influenced by selective forces. A commonly used classification of epistasis is whether they are synergistic or antagonistic. Synergistic epistasis occurs when combined effects of two alleles are more severe than expected based on their independent effects and antagonistic epistasis is when it is less severe (Chiu, Marx et al. 2012). The overall effect of epistasis based upon this definition depends on the type of mutations and the system involved. For example, deleterious effect mutations with increased deleterious effect on fitness that causes a negative epistatic effect will be termed synergistic. But if the interaction renders a positive effect by reducing the detrimental effect then it is termed antagonistic (Chiu, Marx et al. 2012). Since these terms are context-dependent, Phillips suggested positive epistasis when phenotypic effect is higher than expected by individual effects and negative epistasis when it is lower than expected to be used in all cases irrespective of effect of individual mutations (Phillips 2008). Phillips also suggested the term ‘sign epistasis’ to explain the change in relative direction of the effect of individual loci. If two mutations with deleterious effects increase fitness, it results in an adaptive valley in the fitness landscape, the evolutionary implications of which would be different compared to a scenario where their effects are in the same direction (Phillips 2008).

Studies involving the effect of epistatsis among mutations

Conducting *in silico* simulations of growth cycles in T7 phage particles that carried multiple mutations suggested that mildly deleterious mutations interacted synergistically in poor-growth environments, but antagonistically in rich environments. But irrespective of the environments, severely deleterious mutations interacted antagonistically (You and Yin 2002). A survey of epistasis in a wide variety of organisms found that antagonistic epistasis is common for organisms with simpler and compact genomes (like RNA viruses) and synergistic interactions are common for complex genomes (like multi-cellular eukaryotes) (Sanjuan and Elena 2006). Two separate studies; one using metabolic control theory of enzymatic flux (Szathmary 1993) and another using network modeling of pathways and functions (Sanjuan and Nebot 2008) also suggested antagonistic epistasis to be prevalent in prokaryotes and synergistic epistasis in eukaryotes. A fitness landscape model that could predict the distribution of epistatic effects from the distribution of single mutation effects has also been proposed in microbes (Martin, Elena et al. 2007). But theoretical studies point to the fact that in the presence of recombination, evolution might favor antagonistic epistasis rather than synergistic epistasis so that there is increased buffering against the effect of deleterious mutations (Desai, Weissman et al. 2007).

In one study, 47 genotypes of Vesicular Stomatitis Virus were generated carrying pairs of nucleotide substitution mutations. Their separate and combined deleterious effects on fitness revealed both synergistic and antagonistic interactions (Sanjuan, Moya et al. 2004). In yeast, 639 random crosses were generated among a large group of single gene deletions and growth curves of resulting progeny were assayed (Jasnos and Korona 2007). They found average antagonistic effect with reduced growth defects for the double deletions compared to the combined effects of single deletions. Another study in fruit flies used five visible mutations to put together various combinations of mutations eventually generating 32 homozygous lines with multiple mutations and measured productivity and competitive male mating success as fitness traits (Whitlock and Bourguet 2000). A strong average synergistic effect was found for productivity but not for male mating success. In *C. elegans*, EMS mutagenesis was performed on a collection of lines that were previously mutagenized and also on un-mutagenized background. There was not a significant difference between fitness effects of the new mutations on mutagenized and un-mutagenized backgrounds, suggesting that specific environmental conditions might be needed for the epistatic effects to arise (Peters and Keightley 2000).

Studies involving the environmental effect on epistatic interactions

Studies have shown that the environment influences the way mutations interact. In one study involving Tobacco etch virus (RNA virus), it was found that the fitness effect of mutations depends on the genetic background. The interaction between mutations was modulated based on the degree of genetic divergence between primary (Tobacco) and alternative hosts (other species in tobacco family) of this virus (Lalic and Elena 2013). Another study that involved a catalytic RNA species, *Azoarcus* group 1 ribozyme, measured the fitness effects of the interaction between point mutations in three different environments. They found that relative fitness of the catalytic RNA changes; being neutral in its *native* environment (with typical ribozyme reaction conditions), negative in one *new* environment (with a chemical change in its substrate) and positive in another environment that mimics thermal stress (Hayden and Wagner 2012). In *E. coli*, 27 genotypes that contained pairs of mutations were generated by recombination of single mutants. Determination of their combined and separate fitness effects revealed both synergistic as well as antagonistic interactions (Elena and Lenski 1997). Single and double-gene deletions in yeast were then exposed to benign and stressful environments and the growth rate was measured (Jasnos, Tomala et al. 2008). In both benign and stressful environments, the growth rate of the double deletions was higher than expected from single deletion effects indicating antagonistic epistasis.

Synergistic epistasis between deleterious mutations along with increased deleterious effect due to parasites have been proposed to be create a scenario(where recombination is advantageous) that tries to explain the evolution of sex (Cooper, Lenski et al. 2005). But in a study with *E. coli* containing known numbers of transposon-insertion mutations using plasmids as the parasite did not detect any synergy (Cooper, Lenski et al. 2005). In another study, doing reciprocal crosses among 20 deleterious mutations with visible phenotypic effect, 10 pairs of double deletions were generated in fruit flies. The viability of double deletions strains was determined and compared with combined effects of single mutants in environments having high or low nutrition quality (Wang, Sharp et al. 2009). There was a positive epistatic (antagonistic) effect when the mutations were combined with the double deletions showing lower than expected reduction in fitness irrespective of the environment.

Summary of fitness effects of epistatic interactions on mutations

Epistasis causes changes in the selection pressure on a focal allele depending on the effect of another locus in the genome (Wang, Sharp et al. 2009). If there is antagonistic epistasis between deleterious alleles, the strength of selection against the focal allele is reduced. On the contrary, if it is synergistic, there will be more intense selection against it. Environmental perturbations also can influence this interaction and the fitness effect can vary accordingly. Studies using multiple mutations have shown both synergistic and antagonistic interactions in prokaryotes as well as eukaryotes. Though both cases of epistasis is being found in various studies, average effect of epistasis affecting fitness could be close to zero, although with a large variance (de Visser and Elena 2007). Mean epistasis has been suggested to be correlated to the effect of individual mutations or genetic robustness (Gros, Le Nagard et al. 2009). Evolution of epistasis might be based on a balance between strength of selection and intensity of drift (Gros, Le Nagard et al. 2009).

Since most of the above studies used random mutations, there is less chance to find any kind of directional epistasis (Agrawal and Whitlock 2012). Also there is no specific trend of environmental effect on interaction between mutations. Epistatic interactions might be dependent on the genomic context as well as the specific biological functions of the interacting genes. In a model of yeast metabolism, epistasis was found to be more common among genes that are involved in same metabolic function (Segre, DeLuna et al. 2005). But the direction of their interaction effects was varied with synergistic, antagonistic or independent effects. Another study in yeast using *in silico* simulations based on a metabolic network model found condition-specific effects of interactions showing plasticity in interaction effects (Harrison, Papp et al. 2007). Since there is enormous possibility of combinations among genes in a genome a more judicious approach would be studying the effects of mutations on probable or known gene interactions and their dependence on the environment. A group of functionally interacting loci such as those involved in enzymatic mechanisms, signal transduction or developmental pathways should be analyzed to get better insights into how epistasis influences mutation-loads (Rice 1998). We explored the interaction of *C. elegans* with grassland soil bacteria to investigate the effects of stress on individual mutations in bacteria-responsive host genes. Furthermore, in order to understand how environment influences the gene interactions we have performed epistasis analysis among genes specifically involved in defense and metabolism functions.

Chapter 2 - *C. elegans* genomic fitness in response to bacterial stress

Introduction

Mutations are the ultimate source of genetic variation. Recurrent mutations with small deleterious effects more than which could be removed by natural selection (mutation-selection balance) contribute to, "mutation load", which cause a reduction in fitness of a population (Agrawal and Whitlock 2012). The study of the fitness effect of such mutations is important in order to understand many aspects of genome evolution such as the origins of genome complexity (Lynch and Conery 2003) (Lynch and Blanchard 1998), population level phenomena such as the maintenance of standing genetic variation that affects fitness (Charlesworth and Hughes 1996), the evolution of sex and recombination (Kondrashov 1988) and the extinction of small populations (Lynch, Conery et al. 1995). Mutation accumulation studies have contributed to our understanding of the rate that small-effect deleterious mutations occur and their general effect on organismal fitness (Vassilieva and Lynch 1999; Vassilieva, Hook et al. 2000; Shaw, Geyer et al. 2002). Studies using mutation accumulation lines (MA lines) mainly helped to determine the genomic rate of new mutations affecting quantitative traits and distribution of their fitness effects. By analyzing MA lines it was also possible to study interaction between different mutations, degree of their dominance and environmental dependence of their fitness effects. But, mutation accumulation lines might have many random mutations within a single line and one cannot be certain about the fitness effect of specific mutations. In addition, such studies neither identify the genes that are targets of selection in a particular environment nor the selective influence of the environment on specific mutants. Inferences about the selective consequences of mutations cannot be made without an empirical investigation of fitness effects of such mutants in specific environments.

In order to understand how populations evolve, it is important to know their genetic make-up, how the underlying genotype realizes its phenotypic characters that affect fitness (molecular and biochemical functions) and how the phenotype is influenced by various environments (Elena and de Visser 2003). An environment can be classified as stressful when it reduces the fitness of the wild-type compared to another benign environment (Martin and Lenormand 2006; Agrawal and Whitlock 2010). Mutations of particular genes of known

functions can be tested in populations to study their fitness effect in various environments (Young, Yourth et al. 2009; Laffafian, King et al. 2010; MacLellan, Kwan et al. 2012). This allows the investigation of the effects that environmental changes impose on selection of specific gene functions and the identification of genes that experience greatest selection pressure in a given environment.

Since most organisms are exposed to different forms of biotic or abiotic stress in their natural environments, it is important to study the consequences of stressful environments on the fitness of organisms with mutations in specific genes to know the selection pressure imposed by stress. If stress causes a reduction in the fitness of a mutant, the mutation will be more quickly removed from the population. Such elimination of mutations by selection will reduce mutation load in the population influencing the mutation-selection balance. On the contrary, if the stress alleviates the fitness consequence of a mutation, it could be better adapted in the stressful environment as compared to a benign environment (MacLellan, Kwan et al. 2012). In general, gene deletions that cause the disruption of gene functions are considered to cause further reduction in fitness under stress (Elena and de Visser 2003; Martin and Lenormand 2006). Organisms might have "margins of safety" in physiological functions and may not be affected by mutations with minor effects. This excess capacity in functions may not be attained in stressful conditions (Agrawal and Whitlock 2010). However, studies using single gene deletions have previously showed that stressful environments can have variable effects on mutants ranging from reduced fitness (Young, Yourth et al. 2009), to no change (MacLellan, Kwan et al. 2012) or increased fitness (Kishony and Leibler 2003). In most studies of the fitness effects of single gene deletions, the individual mutations chosen for study had visible phenotypes (Young, Yourth et al. 2009; Laffafian, King et al. 2010; MacLellan, Kwan et al. 2012) that were not related to the environment tested and might not be directly subjected to the forces of selection. Thus, the inferences from these experiments can only be applied to the effects a stress on random mutations that may or may not be relevant to the genes involved in the organismal response to the environment being studied. In order to gain better insight into the mechanisms by which stressful environments affect selection on mutants, identifying the stressful environment that affects selection and finding estimates of selection for individual mutants in multiple environments is needed (Agrawal and Whitlock 2010).

To understand the adaptive significance of traits and gene functions, organisms have to be studied in environments that are most likely to have shaped their evolution (Jackson, Linder et al. 2002). Studying the fitness effect of mutations in such environments will provide better insights on population evolution. Nematodes are the most abundant invertebrates found in soil ecosystems and their interactions with soil microbes, especially bacteria, affects various ecological processes such as decomposition and nutrient recycling. Our prior research at Konza Prairie Biological Station has shown that nematode and microbial communities respond to environmental perturbations (Jones, Todd et al. 2006; Coolon, Jones et al. 2009; Coolon, Jones et al. 2013). These changes in bacterial flora could have large consequences for nematode genomic fitness in such environments. In order to understand evolutionarily conserved mechanisms of gene functions in native soil nematodes responding to bacterial environments, we have used the genetic model nematode *Caenorhabditis elegans* to model soil nematode-bacterial interactions. We previously used transcriptomic analysis to identify 204 differentially regulated genes in response to soil bacteria isolated from prairie soils (Coolon, Jones et al. 2009). Since fitness effects of genes are dependent on the environment in which they are expressed (Jackson, Linder et al. 2002), we reasoned that these genes would be most relevant, perhaps experiencing direct or indirect selection pressures while responding to the respective bacterial environments. In order to understand the functional significance of these genes we examined fitness consequences of loss-of-function mutations in 21 of the 204 genes in four different bacterial environments (Coolon, Jones et al. 2009). Almost 25% of the genes differentially regulated when exposed to grassland soil bacteria were found to be involved in defense and metabolic functions. A considerable portion of the *C. elegans* genetic repertoire encoding above functions might be involved in the response to bacteria as they can be either a food item or a pathogen. Thus the fitness effect of mutations in defense and/or metabolism genes will be influenced by these bacterial environments making it likely that they play important roles in selection shaping the evolutionary history of these genes. So, in the current study we have also included 14 additional genes that were annotated to be involved in defense or metabolism functions. We speculated that investigating the fitness consequences of mutants of bacterial response genes would allow us to have clear insights of how selection pressures act on specific gene functions in response to similar environments.

Bacterial environments imposed differential fitness effects on wild-type *C. elegans* with *B. megaterium* being most stressful, *E. coli* intermediate and *Pseudomonas sp*., most benign environment. In order to study how stressful environment affects the fitness of mutants, we have explored their fitness effects in the various bacterial environments. We asked what would be the fitness consequence of mutants when they are exposed to varying magnitudes of stress and whether increased stress could further reduce the fitness of mutants or alleviate the reduced fitness. We found conditionally beneficial effects of mutations in bacteria-responsive genes in stressful bacterial environments. We determined the relative fitness of mutants in order to compare their fitness effects across stressful vs. less stressful bacterial comparisons. In each bacterial comparison some mutants displayed increased relative fitness when grown on stressful bacteria compared to less stressful bacteria. We have also undertaken a comprehensive analysis of life-history traits of mutants in bacterial environments to understand their contribution to fitness consequences of mutants. Additionally we have found that the increase in fitness in response to stressful bacteria displays a trade-off with post-reproductive lifespan.

Materials and Methods

C. elegans and bacteria strains and maintanence

The following mutant strains were used. Bristol (*N2*), *cpr-5* (*ok2344*), *cyp-34A9* (*ok2401*), *lbp-5* (*tm1618*), *clec-50* (*ok2455*), *lec-6* (*tm2552*), *lec-8* (*tm1477*), *lec-9* (*tm1206*), *lys-1* (*ok2445*), *lys-2* (*tm2398*), *lys-4* (*tm2938*), *lys-5* (*tm2439*), *lys-10* (*tm2558*), *mtl-1* (*tm1770*), *hex-1* (*tm1992*). Remaining 21 strains involved in the analysis was obtained from Coolon, Jones et al. (2009). Growth and maintenance conditions were as described (Brenner 1974; Sulston and Hodgkin 1988). Use of bacteria was as for *E. coli* (OP50), *Pseudomonas sp*. (NCBI's GenBank database accession number-EU704696) and *Bacillus megaterium* (EU704698).

Fitness and post reproductive lifespan (TD_{50}) assays

Demographic measures were collected for individual worms in the three bacterial environments. Using life table analysis, fitness/lambda (λ) and other population parameters were calculated. Mutant functional tests were performed by plating eggs on to the test bacteria and then placing progeny from this generation onto the test bacteria, one L4 hermaphrodite (P_0

worm) per plate was incubated at 20^0C with at least 13 replicates per treatment per strain. Re-plating the original worm was done daily till they stopped laying eggs (8-10 days). Progeny per day was counted (age specific reproduction or m_x). Survival of the P_0 worm was monitored as well as the survival of all the progeny of the first day reproductive period (averaged for 3 replications) to determine the age specific survival (l_x); where x=age class. The age to reproductive maturity was found by monitoring the P_0 worm every 12 hr and recording the earliest time point at which eggs were laid (3 replications). Net Reproductive Rate (R_0) was calculated as sum of l_x times m_x ($\Sigma l_x m_x$). Generation time (GT) was calculated by $(\Sigma x l_x m_x)/(\Sigma l_x m_x)$. Lambda ($\lambda$) was determined from R_0 and GT by calculating $\lambda = e^{(\ln R_0/GT)}$, and λ was used as a measure of absolute fitness as in Coolon, Jones et al. (2009). Replicate populations and subsequent life table calculations were used as replicates for statistical tests and each treatment by strain combination was repeated at least 13 times. For the 21 strains assayed by Coolon, Jones et al. (2009), each of the treatments included at least 5 replications.

For the post reproductive lifespan, longevity assays are performed as previously described (Tan, Mahajan-Miklos et al. 1999; Tan and Ausubel 2000) and TD_{50} was calculated from survivorship curves as time to death for 50% of individuals in a population. Briefly, worms were synchronized by bleaching to collect eggs and hatched in M9 overnight. Worms were then grown to L4 on *E. coli* (OP50) to standardize test populations, and then transferred to the test bacteria (10 worms per plate) and were maintained at 25^0C. Surviving worms were then re-plated daily and the fraction surviving was determined every 24 hours. Worms were determined to be dead when they no longer responded to touch with platinum wire. All longevity assays were conducted in at least ten independent replicate experiments in each bacterium.

Statistical analysis

Hypothesis testing of *a priori* contrasts was done using MIXED procedure in SAS (SAS Institute Inc., Cary, North Carolina, USA). The statistical model used for these tests is shown.

Model: $Y=\mu$ + genotype + error

Tests were done separately for each bacterium with each hypothesis tested using test statements, and Y equal to any of the measured life history values (i.e. λ, R_0, GT, TD_{50}, age of reproductive maturity etc.) For the comparisons among relative measures, the difference between

each relative measure was calculated in respective bacterial comparisons. Confidence intervals of the difference were assessed to be positive, negative or including zero.

Linear regressions and correlation analysis

Liner regressions were performed in Excel using Log_2(fold change in expression) as the independent variable and Log_2(fold change Lambda) as dependent variable. The linear regressions were performed for 22 genes used in the functional tests in 3 environmental comparisons and included 31 instances (data points) of differential gene expression.

For each of the relative life history measures, correlation plots were depicted with life history measure of stressful bacteria on Y-axis and benign bacteria on X-axis.

Results

Fitness of wild-type and mutant nematodes in different bacterial environments

We isolated three bacterial species from Konza prairie grassland soils to use in transcriptomic analysis, two of which, *Pseudomonas sp.* and *Bacillus megaterium,* were found to be associated with native soil nematodes (Coolon, Jones et al. 2009). The third, *Micrococcus luteus*, was abundant and culturable from grassland soil samples. In this study we focused on these bacteria as well as *Escherichia* (OP50), which is the normal *C. elegans* lab diet. From previous analysis we found that fitness (λ) of wild-type *C. elegans* (N2) is significantly higher when fed *Pseudomonas sp* and lower when fed *B. megaterium* compared to the standard lab diet of *E. coli* OP50 (Coolon, Jones et al. 2009), which had an intermediate fitness. We confirmed this environmental effect by also conducting life table analysis and selected these bacterial species as benign, stressful and intermediately stressful environment respectively (Figure 2.1).

Our prior transcriptomic analysis showed that 25 percent of all differentially expressed genes responding to the different bacterial environments involved metabolism and/or defense functions (Coolon, Jones et al., 2009). For a bacterivorous nematode species, bacteria serve as prey as well as a source of potential pathogens. For *C. elegans*, this dual role of bacteria has been found to be true in various lab settings (Darby 2005) and also in its natural environment (Felix and Braendle 2010; Felix and Duveau 2012). We reasoned that these genes involved in

metabolism and/or defense functions could be of immediate relevance in responses to their environments. Thus, in addition to 21 genes that were functionally analyzed by Coolon et al (2009), the current study included 14 additional genes from our transcriptomic analysis that were specifically involved in metabolism and defense functions. Life table analysis was performed on mutants of these genes to study their fitness effects in all three bacterial environments (Table 2.1). None of the mutants had a higher fitness than did the wild-type on *Pseudomonas sp*. (benign environment), with 74 percent of mutants having a significantly ($p<0.05$) lower fitness. In *E. coli* (intermediately stressful) 37 percent of the mutants had significantly lower fitness and 17 percent had higher fitness. In *B. megaterium* (most stressful) 45 percent of mutants had significantly lower fitness and 40 percent had higher fitness. Among functional classes of genes, we do not see over-representation of any gene classes with trends of increased or decreased fitness common to all bacteria. Also, there was no trend where all the mutants belonging to any functional group classification have similar fitness effects in any two bacteria. But there are genes belonging to different functional groups (*dpy-14*, *lbp-5*, *mtl-1*, *clec-50*, *lec-9* and *ctl-1*) that have significantly lower fitness in all three environments. Since *Pseudomonas sp* and *E. coli* are gram negative species *B. megaterium,* gram positive, we asked whether there were mutants with similar effects on *Pseudomonas sp* and *E. coli*, but different from *B. megaterium*. Other than a few genes that belong to different functional groups (*dpy-17*, *rol-6*, *sqt-2*, *acdh-1*, *c23h5.8*) that have lower fitness (which is even lesser than those found responding to all the bacteria), there is no such trend. Also, except for one mutant (*nkb-3*), we do not see any mutants with an increased fitness in the intermediately stressful environment but with a lower fitness in the most stressful environment. Although we observed mutants with deleterious effects in each environment, a general trend appears to be that many mutants display increased fitness in stressful environments (Table 2.1).

Mutants display higher relative fitness in stressful environments

Mutations are referred to as "unconditionally deleterious" when they have deleterious fitness effects in multiple environments. Similarly, they are referred as "conditionally beneficial" when they have deleterious fitness effect in one environment but beneficial effect in another environment (Kondrashov and Houle 1994; Elena and de Visser 2003). Along with many unconditionally deleterious mutations we have also seen conditionally beneficial mutations that,

interestingly, exhibited beneficial effect in stressful environments. In some studies involving single-celled organisms, mutations have been found to display alleviating the deleterious effect when exposed to stressful environments (Kishony and Leibler 2003; Jasnos, Tomala et al. 2008).

In order to study the effect of mutations on fitness in response to an increased magnitude of bacterial stress, we compared all pair-wise stressful vs. less stressful bacterial environments. Since the fitness of different mutants cannot be directly compared between bacterial environments, we used fitness of the mutant relative to fitness of wild-type animals (relative fitness). Relative fitness (mutant/wild-type fitness) in each environment was calculated and we plotted each pair-wise comparison with the more stressful bacterial environment on the Y-axis and the less-stressful bacterial environment on the X-axis (Figure 2.2a, 2.2b, 2.2c). If there is no difference of the relative fitness of each mutant in each bacterial environment, the points should lie on the 1:1 reference line indicated on each plot. However, in each comparison the correlation laid above the 1:1 reference line, indicating that they are more fit in the more stressful environment as compared to less-stressful environment. This effect is most pronounced in *Pseudomonas sp.-B. megaterium* comparison that has the greatest difference in absolute fitness for the wild-type, where the majority of the points were above the 1:1 line (Figure 2.2a). The correlation of mutant fitness in this bacterial pair is high (R^2=0.503) indicating that many mutants have similar trend of fitness in both environments, but relatively higher fitness in *B. megaterium*. In the *Pseudomonas sp.-E. coli* comparison, many genes fell along 1:1 line indicating that in a comparison involving more benign environments , many of the mutants display lower fitness (relative fitness below 1) (Figure 2.2b). In the *E. coli-B. megaterium* comparison, the points were only slightly above the 1:1 reference line which might be due to the fact that they are the environments with least difference in absolute fitness for wild-type (Figure 2.2c). We analyzed the difference in relative fitness of mutants in response to stressful bacteria as compared to less stressful bacteria for each cross-environmental comparison (Table 2.2, Table 2.4). The overall mean of relative fitness was significantly higher in *B. megaterium* compared to *Pseudomonas sp* (B-P) and *E. coli* compared to *Pseudomonas sp.* (E-P) but not in *B. megaterium* compared to *E. coli* (B-E) (Table 2.3) where relative fitness was higher in *B. megaterium* but not significant. We observed that 65 percent of mutants showed significantly ($p<0.05$) increased fitness in *B. megaterium* compared to *Pseudomonas sp.* (B-P), 37 percent in *E. coli* compared to *Pseudomonas sp.* (E-P) and 31 percent in *B. megaterium* compared to *E. coli* (B-E). In the B-E

comparison, slightly higher number of genes showed lower fitness (34 percent) (Table 2.2). Only two mutants, *cey-2* and *elo-5* showed increased relative fitness in all bacterial comparisons. Also, all the mutants that showed increased relative fitness in E-P comparisons retained it in the B-P comparison, which had the greatest difference in absolute fitness. We categorized the number of mutants that have relative fitness values significantly increased, decreased or similar in each stressful vs. less stressful bacterial comparisons and performed a chi-square test to determine significant differences from null hypothesis in which the probability of each category was equivalent (Table 2.3). We found a significant deviation for the number of mutants from the expected probability for B-P and E-P category but not for B-E. This shows an over-representation of mutants with higher, lower or similar effects in relative fitness in B-P and E-P and not in B-E. Thus relative fitness comparisons showed that the mutants exhibit higher relative fitness in more stressful environments and the same mutants retain this increased fitness in even more stressful environment comparisons (B-P comparison has more drastic fitness difference for wild-type compared to E-P and B-E).

Differential fitness of mutants does not correlate with conditional expression

The effect of genes on organismal fitness might vary with environments. Variation in fitness effects of mutations in different environments might reflect the effect of environment-dependent gene expression, which is termed "conditional expression" (Martin and Lenormand 2006). We previously observed that differential fitness of mutants in specific genes was significantly correlated with differences in levels of gene expression in each environmental comparison (Coolon, Jones et al. 2009). We performed a similar analysis using 22 genes with 31 instances of significant ($q<0.05$) differential gene expression (Figure 2.3). Although the slope of the best-fit line was negative (-0.056), it was not significantly different from zero ($p=0.357$). Also, there was no correlation ($r=0.17$) between the gene expression and fitness of the mutants. This difference from earlier observations could be due to the introduction of a new set of genes that had variable effects on the correlation of gene expression- and function, in this case fitness. In our earlier study we have seen that instead of differential fitness (λ), differential TD_{50} of mutants have more correlation to differential expression (Coolon, Jones et al. 2009). So, fitness could be a more robust and comprehensive trait that does not fluctuate drastically with the quality of environment (bacterial species) since the genes we selected are already responsive to

bacteria in general. Another difference is that in the present study we examined only three bacteria, whereas the previous study included four bacteria, thus providing fewer instances of differential expression.

Increased brood size and decreased generation time contribute to the increased fitness of mutants in stressful environments

Total brood size and generation time are major components of fitness (λ) calculated by cohort life table analysis. Observations of total brood size of the mutants in each bacteria and generation time calculated from the life tables were used to assess their contribution to fitness (Table 2.5). Although the total brood size of wild-type in *Pseudomonas sp* and *E. coli* were similar, increased fitness in *Pseudomonas sp* could be attributed to the reduced generation time in *Pseudomonas sp.* compared to *E. coli*. In *B. megaterium* total brood size is lower and generation time is longer, making it the least fit environment. Although the absolute values of brood size of mutants compared to respective wild-type values were significantly lower in all bacteria, the percentage of mutants that had lower brood sizes were reduced in both stressful environments (54 % in *Pseudomonas sp.*, 42% in *E. coli* and 48% in *B. megaterium*) and some of the mutants show increased brood sizes in *B. megaterium* (20%). On the contrary, 74% of mutants had increased generation times in *Pseudomonas sp.* compared to wild-type, but generation time was reduced in stressful environments (23% and 37% in *E. coli* and *B. megaterium* respectively). Also, more mutants showed reduced generation times in *E. coli* (17%) and *B. megaterium* (40%). Except for *rol-6,* all the mutants that had an increased total brood size also had increased fitness in *B. megaterium*, indicating that increased fitness of mutants in *B. megaterium* is almost always associated with increased brood size. However, not all the mutants with reduced generation time have increased fitness in *E. coli* and *B. megaterium* Only 2 out of 6 and 8 out of 14 mutants that reduced generation are found to increase fitness in *E. coli* and *B. megaterium*.

In order to understand whether any of these components indicated specific trend corresponding to increased fitness, we plotted relative brood size (brood size in mutants/N2 brood size) (Figure 2.4a, 2.4b, 2.4c) and relative generation time (generation time in mutant/N2 generation time) (Figure 2.5a, 2.5b, 2.5c) each cross-environmental comparison. We observed a high correlation in each cross-environment comparison for relative brood size (R^2=0.68 in B-P,

0.47 in E-P and 0.45 in B-E), demonstrating that that the mutants tend to have similar adjustments in brood size irrespective of the changes in environment. However, the mutants show considerable variability in generation time in the different bacterial environments demonstrated by the low correlation for generation time in all cross environmental comparisons (R^2=0.31 in B-P, 0.16 in E-P and 0.002 in B-E. We also examined the difference in relative brood size (Table 2.7, Table 2.4) and generation time was analyzed (Table 2.8, Table 2.4) to determine whether changes in brood size or generation time of the mutants contribute to relative fitness difference in those environments. The overall mean of relative brood size was significantly higher in B-P and E-P but lower (but not significant) in B-E (Table 2.4). For relative generation time, the overall mean was significantly lower in B-P and E-P but higher (but marginally significant) in B-E. We found that the direction of changes in relative brood size corresponded to that for relative fitness for 18 mutants in 32 instances (both increased or both decreased) in respective cross-environment comparisons. There were only five instances (involving five mutants) that did not show this pattern. For three of the mutants (*cyp-34A9*, *nkb-3*, *mtl-1* and *lec-9*) the change in brood size corresponded to the change in relative fitness in all the environment comparisons, and for seven others (*cpr-5*, *lys-2*, *mtl-1* in B-P and E-P, *dpy-14*, *hex-1*, *rol-6* in B-P and B-E, *lbp-5* in B-E and E-P) the change in brood size corresponded to the change in relative fitness in two of the three environment comparisons. For generation time, we found that the direction of changes in relative generation time corresponded to that for relative fitness for 20 mutants in 42 instances. There were 10 instances (with 10 mutants) that did not match this pattern. The relative generation time for mutants, *cyp-34A9*, *elo-5*, *lys-1*, *mtl-1* corresponded with relative fitness changes in all three environments and 12 instances where the changes corresponded in two environment comparisons (*cpr-5*, *lys-2*, *mtl-1* in B-P and E-P, *c23h5.8*, *hex-1*, *mtl-2* and *sqt-2* in B-P and B-E, *clec-50*, *nkb3*, *lbp-5*, *lec-9* in B-E and E-P). We did not see any specific gene functional classes having similar patterns of brood size or generation time that corresponded with the fitness changes. There were two genes; *cyp-34A9* and *mtl-1* that showed patterns of both changes in brood size and generation time in all three environments that corresponded to expected relative fitness changes. Thus we have seen that both changes in brood size and generation time can bring about increase in fitness in stressful environments and this is not specific to any gene functional classes. Also the increased fitness displayed by mutants is not due to any specific bacterial effect since we saw all the bacterial

comparisons showing trends of increased brood size and decreased generation time leading to increased fitness.

Similar to the chi-square test performed for relative fitness, the number of mutants that have effects on relative brood size and generation time were categorized as significantly increased, decreased or similar in each stressful vs. less stressful bacterial comparisons and tested for differences from expected probability of on third in each category (Table 2.3). We found a significant deviation from expected probability for all bacterial comparisons in brood size. For generation time there was significant deviation in B-P and E-P category but not for B-E.

In order to confirm that our relative fitness measures correspond to changes in brood size and generation time as expected from their calculation of absolute fitness values, we performed a regression analysis using relative brood size and generation time as predictor for relative fitness in each bacteria. We found significant positive coefficient slope for brood size and significant negative coefficient slope for generation time as expected (Table 2.6).

Increased brood size in early age-classes contributes to the increased fitness of mutants in response to stressful bacteria

Earlier studies have shown that increased brood size in *C. elegans* may not be favorable in natural environments since it is accompanied by longer generation time (Hodgkin and Barnes 1991). Hodgkin and Barnes (1991) also speculate that *C. elegans* in natural environments, being an r-selected species, would favor higher earlier brood sizes to increase fitness rather than higher late brood sizes. Jenkins, McColl et al (2004) found that in *C. elegans* early fitness traits rather than life time fertility explained the genetic trade-offs of increased lifespan to fitness introduced by long lived mutants of *daf-2* and *age-1*. In order to know whether there was any specific contribution of early or late brood size to the increased fitness of the mutants found in stressful environments, we decided to dissect the patterns of total brood size in stressful vs. less-stressful environments into separate age-classes. We found that the period of most intense reproduction fell in a window of first 4 days with first day being the day of onset of reproduction in each bacterial environment (Table 2.9). We also found that second day reproduction was the highest in all the bacterial environments for the wild-type. Among the three bacteria, the best fit environment (*Pseudomonas* sp.) has the highest brood size on the second day followed by less fit

environment (*E. coli*) and then by the least fit environment (*B. megaterium*). The brood size on the first day also showed the same trend. However, on the third and fourth day, brood size declined in *Pseudomonas sp.* and increased for *E. coli* and *B. megaterium* showing that in the most benign environment, worms tend to lay eggs earlier. For the mutant brood sizes, we observed that most had a significantly lower brood size compared to wild-type for the first (60%) and second day (63%) in response to *Pseudomonas sp*., which mostly corresponds to their reduced fitness. The proportion of mutants with reduced brood size declined in both *E. coli* and *B. megaterium* for first day (23% and 6%) and second day (31% and 46%). Also all the mutants except *lec-8* (for *B. megaterium*) that showed increased brood size in first and second day for *E. coli* and *B. megaterium* also showed increased fitness in corresponding environments. For *Pseudomonas sp*, although there are mutants that increased brood size in both third and fourth day (13 mutants), none had increased fitness. For *E. coli*, out of the eight mutants that increased brood size on the third and fourth days, only two mutants had increased brood size in the third day (*mtl-2* and *gei-7*) had increased fitness. Similarly for *B. megaterium*, only six out of the 12 mutants that have increased brood size on the third and fourth days showed increased fitness. This shows that early brood size increase translates to higher fitness more than late brood size increases.

The cross-environment correlation of relative brood size of each day (mutant brood size of day x/N2 brood size of day x) was plotted for each of these days (Figures 2.6-2.9). There was a great deal of variation for mutant relative brood size across the cross-environment correlations for first, third and forth days. However, the relative brood sizes for the second day was highly correlated in each cross-environment comparison (R^2=0.65 in B-P, 0.58 in E-P and 0.57 in B-E). By analyzing the pattern of changes in relative brood sizes of mutants in each day, we found that first and second day brood sizes were subject to more variation compared to the third and fourth days (Table 2.10). All the mutants, except for *pab-2*, *sqt-2* and *gld-1* (in E-P comparison) with increased brood size in the first and second days (10 mutants in total) showed corresponding fitness increases in respective bacterial comparisons. Interestingly, *nkb-3*, *cpr-5*, *lec-9*, *lys-2* and *mtl-1* mutants have increased relative brood sizes on the second day for B-P and E-P which appears to account for their increase in total relative brood size that in turn corresponded to increased relative fitness in these environments. Finally, all the above mutants have increased

relative brood size also for the first day in E-P comparison suggesting that they allocate more reproductive output to the earlier age classes to increase fitness.

Mutants with reduced age of reproductive maturity have increased fitness in stressful bacteria

In an experimental evolution study using heterogeneous populations of *C. elegans*, directly selecting for early reproduction Anderson, Reynolds et al 2011, found that increased early reproduction resulted in less late reproduction, suggesting a trade-off, but with no change in total fitness or lifespan (Anderson, Reynolds et al. 2011). Since we found the increased fitness of mutants was accompanied by an increase in early fecundity, we asked whether the change in fitness we observed might be brought about by fluctuations in time to reach reproductive maturity. To determine the age to reproductive maturity we monitored wild-type and mutant worms every 12 hr and recorded the earliest time point at which eggs were laid as the age of reproductive maturity (Table 2.11). Although the age of reproductive maturity for wild-type is comparable in *Pseudomonas sp*. and *E. coli*, we found that in the most stressful environment, the worms take longer time to start laying eggs. So in the more benign environment, worms not only put more eggs into earlier age-classes, but also laid these eggs earlier than in stressful environments. Although almost 50% of the mutants have significantly delayed reproductive maturity compared to wild-type in all the bacteria, 31% of mutants also showed reduced age of reproductive maturity in the most stressful environment. Nine of these 11 mutants (except *dpy-17* and *ctl-1*) also showed increased fitness in *B. megaterium*, demonstrating that early reproductive maturity can also contribute to increased fitness in stressful environment.

To better visualize the differences in age of reproductive maturity across bacterial environments, we plotted the relative age of reproductive maturity (age of reproductive maturity of mutant/age of reproductive maturity of wild-type) for each cross-environment (Figure 2.10a, 2.10b, 2.10c). We observed that the relative age of reproductive maturity shows between bacterial were not well correlated. In each of the comparisons, we found the relative age of reproductive maturity of mutants was greater in the less stressful environments compared to high-stress environments (points below 1:1 line). In the cross-bacterial comparisons of difference in relative age of reproductive maturity (Table 2.12), we found 83 percent instances in B-P, 61 percent in B-E and 70 percent in E-P with corresponding patterns of relative fitness (lower age of

reproductive maturity corresponding to higher fitness and vice versa.). The majority of these instances (19 in B-P, 11 in B-E and 5 in E-P) the relative age of reproductive maturity was lower in stressful bacteria compared with the less stressful counterpart showing that mutants with reduced age of reproductive maturity has increased fitness.

Increased fitness of mutants in stressful environments display trade-offs with post-reproductive lifespan

Two major components involved in trade-offs for reproduction are cost for survival and cost for future reproduction (Stearns 1989). The antagonistic pleiotropic theory of aging suggests that fitness effect of a gene could be traded-off with its lifespan functions (Williams 1957). One of the suggested reasons of senescence is that natural selection would maximize selection on genes that contribute to youthful vigor and reproductivity, which would cause decreased vigor later on in life. Thus it seems likely that selection on genes the function to promote increased early reproduction would also cause rapid aging and reduced longevity. Previous work in *C. elegans* demonstrated that mutants with extended lifespan such as *age-1* and *daf-2* have reduced fitness, especially in stressful environments (Walker, McColl et al. 2000; Jenkins, McColl et al. 2004; Chen, Senturk et al. 2007). Thus we wondered whether the converse might be true and hypothesized that the mutants that showed increased fitness in stressful environments would have a trade-off with other life-history traits specifically, lifespan. Post-reproductive lifespan was measured for each of the 14 mutants in each environment as time to death for 50% of the population (TD_{50}) using survivorship curves (Table 2.13). In wild-type animals, lifespan was longer in the least fit environment and shortest in the most benign environment showing that there might be trade-off between fitness and lifespan. In response to all three bacteria, more than 40% of mutants displayed significantly shorter lifespans compared to wild-type animals. Mutations in *cpr-5* and *lec-9* in *Pseudomomas sp*., *lec-9* in *E. coli* and *lec-9, cpr-5* and *cyp-34A9* in *B. megaterium* showed significantly increased lifespans, but have decreased fitness in corresponding environments. Conversely, five of six mutants with higher fitness in *E. coli* and 10 of 14 mutants with higher fitness in *B. megaterium* had shorter lifespans in corresponding environments. Relative TD_{50} was calculated (mutant TD_{50}/wild-type TD_{50}) and changes in relative TD_{50} of mutants in each cross-bacterial comparisons (Table 2.14) were compared with relative fitness. If there is a trade-off between lifespan and fitness, mutants that show an increase

in relative TD_{50} in a cross-bacterial comparison should show lower fitness in same bacterial comparison and vice versa. There are 21 instances among 15 mutants that show this trend, of which 17 instances are those with reduced relative TD_{50} in the stressful bacterium versus the non-stressful bacterium comparison but with increased fitness in the same bacterial comparison. Thus we find that the increased fitness of some of the mutants in stressful environments could have a trade-off with their lifespan functions.

Discussion

Our previous transcriptomic analysis provided us with *C. elegans* genes that are differentially regulated in response to grassland soil bacteria. This included genes that are differentially regulated in pairwise comparison of four different bacteria, *E. coli*, *Pseudomonas sp.*, *B. megaterium* and *Micrococcus luteus*. We speculated that these would be the genes that have most potential of being acted upon by forces of selection when exposed to those bacterial environments. In addition, *Pseudomonas sp.* and *B. megaterium* has been isolated in association with native soil nematodes and thus increasing the chance of being involved in nematode-bacterial interactions. For the current analysis we selected one more stressful environment (*B. megaterium*) and one less stressful environment (*Pseudomonas sp.*) compared to the standard *C. elegans* laboratory diet *E. coli* OP50. We estimated the fitness of loss-of- function mutants in genes that are found to have differential expression in corresponding environments (conditional expression) as well as in other bacterial environments. Fitness changes of mutants in different environments were analyzed by calculating relative fitness in each environment, showing that mutants tended to have increased fitness in the presence of more stressful bacterial environments in all cross-bacterial comparisons.

Although our criteria of choosing the genes in this study was their environment-responsiveness, our primary objective of using mutations was not to investigate their functions specific to bacterial environments. We deliberately picked bacteria-responsive genes so that we could identify any fitness fluctuations in their mutants in terms of the magnitude of stress applied. We assume that these would constitute a fraction of genes that experience selection pressures in bacterial environments as opposed to any random list of genes. Here we have seen a

general trend where animals bearing mutations in these genes display increased fitness in more stressful bacterial environments. In many cases we see the same mutants displaying a higher relative fitness as they are exposed to a more stressful environment. Since this is not specific to any functional categories tested or for the broad categorization of defense and metabolism functions, it appears this is a general trend irrespective of the putative gene functions for these bacteria-responsive genes. Our cross-bacterial correlation for mutant relative fitness is high, especially for B-P, and moderate, for E-P and B-E. Similarly, all the cross-bacterial correlations for total brood size and second day brood size were also high. This might point to the fact that, although they are affected by the magnitude of stress; functions of such genes might be conserved in different bacterial environments with specific trends in fitness.

Although we did not see a strong gene expression-function correlation, there were instances of mutants that affect specific genes (*lys-4*, *mtl-2* and *gei-7*) that showed a reduction of fitness in environments where they are expressed at higher levels. Thus we cannot completely rule out the possibility that the fitness of these mutants explained by "conditional expression". Specifically, among the genes whose mutants showed increased fitness in *E. coli* and *B. megaterium*, 40 percentage of instances corresponded to low expression in respective environments. For all the remaining genes whose mutation showed increased fitness, we have seen lifespan to be lower than the wild-type showing extensive trade-offs. This shows that conditionally beneficial mutations with pleiotropic functions could display increased fitness due to trade-offs. Interestingly, we also see 20 percent of instances with mutants showing increased fitness corresponding to higher gene expression in respective environments. Fitness increases in the mutants in response to stress were accompanied by an increase in total brood size along with reduced age of reproductive maturity. This also conforms to the predictions of a major life-history trade-off theory that explain the cost of reproduction (Stearns 1989); the genetic costs represented by antagonistic pleiotropy (Williams 1957). Presence of antagonistically pleiotropic mutations, having opposing fitness effects in two environments, is an important factor that might lead to ecological specialization (Martin and Lenormand 2006).

Mutations in two genes (*lec-6* and *lec-8*) did not have any fitness defect in any of the bacterial environments. Also, mutations in only 8 percent of genes showed unconditionally deleterious effect with reduced fitness in all the bacterial environments (Table 2.1). Fitness of all other mutants was variable at least in one environment. Mutants with unconditionally deleterious

effect are suggested to be selectively eliminated from the population and do not have any long term consequence in evolution (Elena and de Visser 2003). Since for other mutants, the fitness defect was variable in different environments, the strength of selection also will be changed in those environments. This shows that not only the quality (species of bacteria) of stress, but the quantity of stress (magnitude of stress) should be considered while assessing the effects of mutations. Among these mutants, doing cross-bacterial comparisons with relative fitness, we have found that the deleterious effect could be alleviated in stressful environments (Table 2.2). As the magnitude of stress increased, many of the mutants showed beneficial effects with their fitness higher than the wild-type in *E. coli* and *B. megaterium* (Table 2.1). In another study that used random mutations in *E. coli*, conditionally beneficial mutations have been reported in stressful environment due to low nutrition conditions (Remold and Lenski 2001). They found 12 percent of mutants showing higher fitness in stressful environments. But we saw higher number of mutants with conditionally beneficial effect in both stressful bacteria *E. coli* (17%) and *B. megaterium* (40%). Conditionally beneficial mutations with antagonistically pleiotropic effect in fitness could be due to trade-offs (Elena and Lenski 2003). In our study, we have found trade-offs for conditionally beneficial mutations with post-reproductive lifespan.

Not all the genes may be equally important in all bacterial environments. An organism might be most fit in a particular environment because its genetic machinery is fine-tuned in such a way that most of its genes might be optimally used. In such case, most mutations should be detrimental to its fitness in that environment (for example, *Pseudomonas sp.*). A recent study has shown that when mutants are exposed to a stressful environment, the stress responses especially triggered by heat shock factor 1 (HSF-1), a master regulator of environmental stress response, can induce buffering of mutational effects (Casanueva, Burga et al. 2012). Since stress responses are not induced in a benign environment, the fitness reduction is not buffered causing relatively higher fitness defects in it compared to a stressful environment. Also when exposed to stressful environments mutants are found undergoing early reproductive maturity along with trade-offs with post-reproductive lifespan. Function of certain gene products could be in favor of longer lifespan or more somatic growth and maintenance. Mutants in these genes could have a higher fitness under stressful environment since those gene functions could be costly under stressful conditions. For example, genes like *elo-5* (Kniazeva, Crawford et al. 2004) and *pab-2* (Ciosk, DePalma et al. 2004) are required for rapid growth and somatic development and mutants in

these genes were found to have higher fitness in stressful environments, *E. coli* and *B. megaterium* but not in a benign environment, *Pseudomonas sp.* (Table 2.1).

Effects of alleviation of fitness defects in response to stressful environments were found previously in at least two prior studies (Kishony and Leibler 2003; Jasnos, Tomala et al. 2008). In the first study (Kishony and Leibler 2003) 65 random mutations were induced by chemical mutagens in *E. coli* and their growth rates of mutants were measured in various environments that included stressful environments. In the second study (Jasnos, Tomala et al. 2008), various single and double deletion Yeast strains are measured for their growth rate in several stress-inducing environments. Thus both of these studies involved unicellular organisms and fitness was measured as increased growth rates in culture and there was no scope of further dissection of their effects on other life history traits for trade-offs. One explanation for the occurrence of increased fitness in stressful conditions is that the mutation and stressful conditions may not always affect the same pathways. If they are acting in separate pathways, a positive epistatic (antagonistic epistasis) effect induced by the mutation in a pathway that interacts with stress-regulated pathway can cause increased fitness in those mutants. This is especially possible in our study since we have selected genes that are involved specifically in bacterial interactions. So a mutant that affects one pathway is more likely to interact with a stress-elicited pathway. Exposed to a stress, a specific or a promiscuous interaction partner (like chaperones) can also buffer the deleterious effect of mutation which may not happen in benign environment where effect of mutations manifest completely.

Similarly, a positive epistatic effect, in which the fitness of double gene deletions are higher than fitness predicted from effects of single gene deletions, was recorded in many studies (Jasnos and Korona 2007; Jasnos, Tomala et al. 2008; Wang, Sharp et al. 2009). Also in yeast, genes that belonged to a variety of functional classes shared this tendency (Jasnos and Korona 2007). This shows that two deleterious mutations in an organism can interact antagonistically. They suggest that the initial cause for reduced fitness, whether genetic stress (mutation) or environmental stress, will have the highest effect in terms of its magnitude (Jasnos, Tomala et al. 2008). With additional stress, whether environmental stress or genetic stress (a second mutation), the fitness reduction is relatively lower. This will be not happen if the mutants are exposed to benign environment. Understanding the nature of epistatic interactions in double mutant strains

among these genes and looking for synergistic effects in stressful vs. non-stressful conditions can provide more insights in this regard.

In regressive evolution, traits decrease or degenerate over the course of evolution (Jeffery 2009). A classic example of this is loss of pigmentation in the cave fish, *Astyanax mexicanus* which has been attributed to mutations in single genes in at least two cases affecting genes, *aca2* and *Mc1r* (Jeffery 2009). There are two competing hypothesis that seek to explain this phenomenon; neutral mutation and selection (Gross, Borowsky et al. 2009). A trait no longer under selection can accumulate mutations and may be lost by non-adaptive drift. The second hypothesis suggests that mutations can have other pleiotropic advantages which can be selected for; although, there is no specific evidence for such advantages found in the case of the cave fish (Jeffery 2009). Our finding that some of the mutations display fitness advantage in stressful environments could be one way that can result in a selection driven gene loss from the population. By trading-off with alternate functions (as post-reproductive lifespan in our study), mutations in some genes might exhibit fitness advantage in a stressful habitat.

Epistasis analysis performed for the bacteria-responsive genes using loss-of-function mutants indicate that some of the *lys*, *lec* and *protease* genes interact to affect a different life history trait, survivorship (Chapter 3). We cannot rule out the possibility that many of the above genes interact in response to the bacterial species tested, also affecting fitness. Mutations in genes that interact in response to a particular bacterial environment could have similar fitness effects when exposed to that bacterium. Thus it can be argued that the genes selected for the study may not represent an 'independent' random sample of bacteria-responsive genes and that mutants of many genes display similar fitness effects because they belong to same pathway. In other words, depending upon the bacterial environment, the fitness effect exhibited by the mutant of any given gene might be dependent on the response of its interacting partner. Based on the gene interactions determined in our epistasis experiments affecting survivorship, we can have some expectations regarding fitness effects for mutants. For example, *cpr-5*, *clec-50*, *lec-9* and *lys-4* belonged to a pathway affecting survivorship in response to *E. coli* and *B. megaterium,* whereas *lec-9* and *cpi-1* belonged to a pathway in response to *Pseudomonas sp*. In *B. megaterium*, mutants in all the above genes showed reduced fitness. Similarly in *Pseudomonas sp*. mutants of *lec-9* and *cpi-1* also showed reduced fitness. But in *E. coli* only mutants of *clec-50* and *lec-9* showed reduced fitness whereas *cpr-5* and *lys-4* mutants did not show significantly

reduced fitness. The genes interacting to affect survivorship response might not be the same as those affecting fitness and this could be a reason that all the mutants did not display the predicted fitness effects in *E. coli*. Thus, in the case of environment-responsive genes, the knowledge of gene interactions, albeit to a small extent, can have predictive capacity in terms of fitness effect of mutants just as we have seen in the case of "conditional expression". Although studying the effect of individual gene mutations will help us to understand their specific fitness effect in a particular environment, with additional knowledge about the gene interaction network, the magnitude and direction of fitness effect might be made more predictive for a similar environment.

Figures and tables

Table 2.1 Absolute fitness of mutants in bacterial environments

Wild-type *C. elegans* (*N2*) and mutant strains were grown on the three bacteria and absolute fitness (λ) of the individuals in a population were measured. The genes were ordered based on molecular functions based on GO terms. Red shades indicates significant (P<0.05) increase relative to wild-type and yellow indicates significant (P<0.05) decrease relative to wild-type.

	Absolute Fitness			Functional GO term specification
Gene	*Pseudomonas sp.*	*E. coli*	*B. megaterium*	
N2	4.19	3.28	2.57	
cey-2	2.83	3.08	2.79	Y-box domain
cey-4	3.57	3.51	2.95	Y-box domain
cpi-1	3.69	3.25	2.99	Cysteine Protease inhibitor
cpr-5	2.78	3.23	2.32	Cysteine Protease
cyp-34A9	2.96	3.25	2.19	Cytochrome P450
cyp-37A1	3.64	3.59	2.85	Cytochrome P450
dpy-14	1.85	1.89	0.96	Collagen
dpy-17	3.20	2.84	2.69	Collagen
rol-6	3.11	2.82	2.56	Collagen
sqt-2	3.72	2.97	3.39	Collagen
daf-22	3.41	3.37	2.62	Fatty acid beta oxidation
acdh-1	3.78	2.99	3.01	Fatty acid beta oxidation
elo-5	4.07	4.11	4.18	Polyunsaturated FA metabolism
dhs-28	2.43	2.23	1.86	Dehydroxystroid dehydrogenase
fat-2	4.23	3.27	3.18	Fatty acyl desaturase
lbp-5	2.88	3.07	2.15	Fatty acid binding protein
clec-50	2.47	2.90	2.21	C-type lectin
lec-6	4.30	3.27	2.43	Lectin
lec-8	4.02	3.22	2.74	Lectin
lec-9	2.75	3.12	2.04	Lectin
lys-1	2.67	3.26	2.34	Lsozyme
lys-2	2.80	3.23	2.40	Lysozyme
lys-4	3.97	3.19	2.30	Lysozyme
lys-5	4.01	3.19	2.31	Lysozyme
lys-10	4.37	3.32	2.30	Lysozyme
mtl-1	2.32	3.14	1.83	Metallothionein
mtl-2	4.09	3.77	3.75	Metallothionein
pab-2	4.29	4.14	3.20	Poly adenylate binding
hsp-12.6	3.72	3.10	3.00	Heat shock protein
C23H5.8	3.07	2.72	3.30	Not known
ctl-1	2.77	2.91	2.29	Catalase, antioxidant
nkb-3	2.24	3.53	2.06	Na+ K+ ATPase
gei-7	3.77	3.52	3.27	Glyoxalate cycle
gld-1	3.53	3.15	2.78	Germline development
hex-1	4.04	3.15	2.19	Hexosaminidase

Figure 2.1 Absolute fitness of wild-type in bacterial environments

Wild-type *C. elegans* (*N2*) was grown on the three bacteria and absolute fitness (λ) of the individuals in a population was measured. Standard error is indicated as error bars.

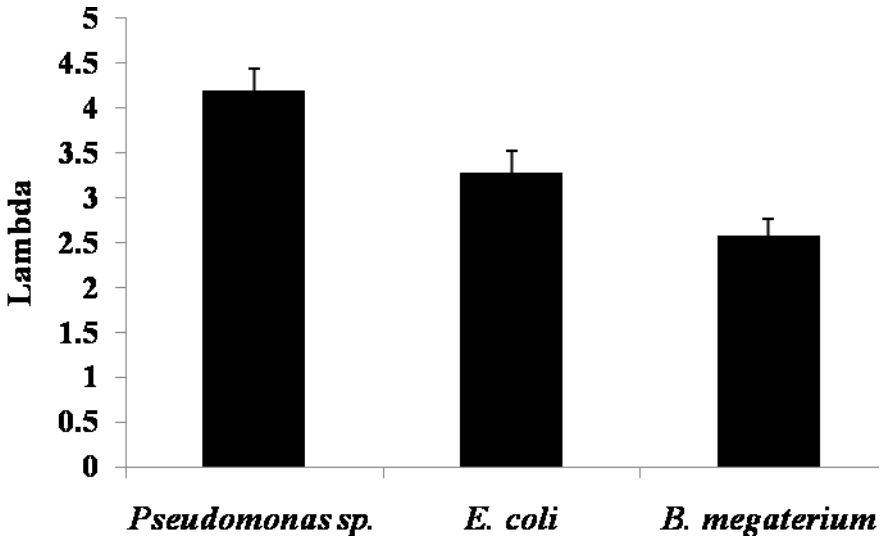
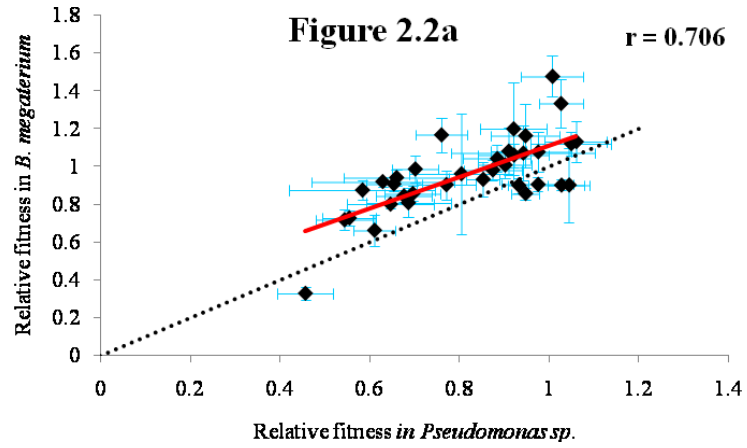
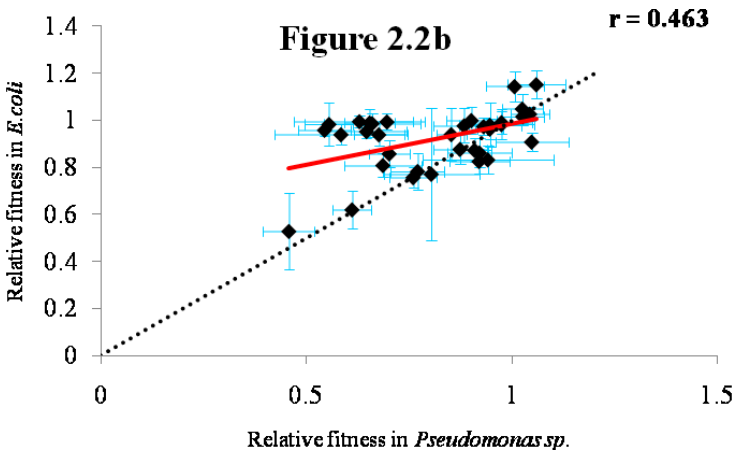


Figure 2.2 Relative fitness of mutants in stressful vs. less stressful bacteria

Relative fitness of mutants was calculated (mutant absolute fitness/wild-type fitness) in each bacterium. Confidence intervals (α=0.05) is indicated as error bars. Pearson's Correlation coefficient (r) is shown. Dashed line represents 1:1 reference line with equal relative fitness.





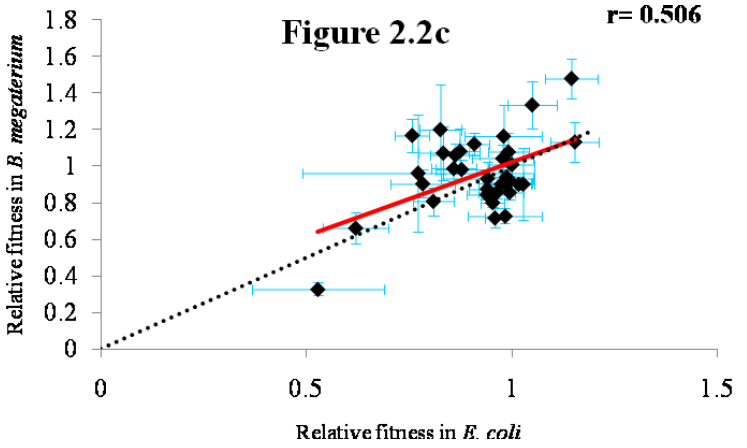


Table 2.2 Comparisons of relative fitness of mutants for stressful vs. less stressful bacteria

Difference between relative fitness of each bacterial comparison is calculated. Red shade indicates that the confidence interval (α=0.05) of the difference is positive; yellow indicates negative and no shading indicates confidence interval contain zero. B, *B. megaterium*.; E, *E. coli* OP50; P, *Pseudomonas* sp.

Gene	B-P	B-E	E-P
acdh-1			
c23h5.8			
cey-2			
cey-4			
clec-50			
cpi-1			
cpr-5			
ctl-1			
cyp-37A1			
cyp-34A9			
daf-22			
dhs-28			
dpy-14			
dpy-17			
elo-5			
F55F3.3			
fat-2			
gei-7			
gld-1			
hex-1			
hsp-12.6			
lbp-5			
lec-6			
lec-8			
lec-9			
lys-1			
lys-10			
lys-2			
lys-4			
lys-5			
mtl-1			
mtl-2			
pab-2			
rol-6			
sqt-2			

Table 2.3 χ^2 test for random probabilities of relative life history measures

The difference in relative life history categories is found for each bacterial comparison. Number of relative life history measures categorized as significantly higher, lower or similar in each comparison is used as levels in the χ^2 test of independence.

Life history measure	χ^2 test of 1/3 probabilities (higher/similar/lower)					
	B-P		B-E		E-P	
	χ2	pvalue	χ2	pvalue	χ2	pvalue
Relative fitness	17.2	0.0002	0.057	0.9718	14.1	0.0008
Relative Brood	9.66	0.008	9.66	0.008	36.4	<0.0001
Relative GT	9.66	0.008	5.2	0.074	18.57	<0.0001

Table 2.4 Summary of relative life history measures of mutants in stressful vs. less stressful bacteria.

The mean difference in relative life history measure of mutants is calculated for each bacterial comparison. Standard error and *P*-value for a t-test significantly different from zero is shown.

Life history category	B-P			B-E			E-P		
	mean	se	pvalue	mean	se	pvalue	mean	se	pvalue
Rel fitness	0.141	0.0111	<0.0001	0.00132	0.012	0.4561	0.128	0.0127	<0.0001
Rel Brood	0.1465	0.0248	<0.0001	-0.00952	0.0279	0.3668	0.1413	0.0287	<0.0001
Rel GT	-0.084	0.00888	<0.0001	0.0187	0.0108	0.0421	-0.0909	0.0103	<0.0001

Figure 2.3 Regression of differential fitness on differential expression

Differential expression was used as a predictor for fitness of mutants. Linear regressions were performed using Log_2 transformed fold change in gene expression from microarray experiments as the independent variable and Log_2 transformed fold change in mutant fitness as dependent variable in bacterial comparisons. There were 22 genes with 31 instances of differential gene expression. Equation shows the slope and intercept of the regression and r is the Pearson's correlation coefficient.

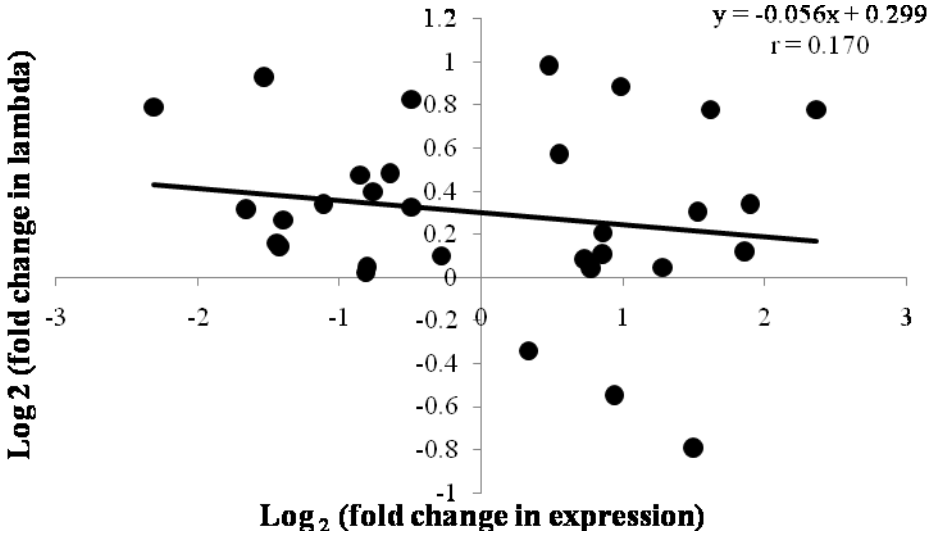


Table 2.5 Absolute total brood size and generation time of mutants exposed to different bacterial environments

Wild-type *C. elegans* (*N2*) and mutant strains were grown on the three bacteria and total brood size and generation time were determined. Generation time was determined as GT = $(\Sigma x l_x m_x)/(\Sigma l_x m_x)$(in days) using life tables. Red shade indicates significant (P<0.05) increase relative to wild-type and yellow indicates significant (P<0.05) decrease relative to wild-type. B, *B. megaterium*.; E, *E. coli* OP50; P, *Pseudomonas* sp.

	Total Brood size			Generation Time		
Gene	P	E	B	P	E	B
N2	291.82	296.43	212.87	3.89	4.78	5.69
acdh-1	290.83	283.20	184.83	4.32	5.14	4.80
C23H5.8	313.40	212.60	239.20	5.13	5.35	4.58
cey-2	328.00	348.20	289.00	5.56	5.20	5.52
cey-4	356.60	313.80	363.40	4.62	4.58	5.45
clec-50	68.18	195.15	145.77	4.45	4.96	6.64
cpi-1	219.44	242.00	209.11	4.15	4.71	4.99
cpr-5	127.82	276.15	163.08	4.54	4.80	6.09
ctl-1	120.20	177.00	73.00	4.17	4.84	4.77
cyp-34A9	174.62	300.54	158.46	4.71	4.82	6.49
cyp-37A1	311.40	301.00	298.60	4.44	4.46	5.42
daf-22	256.67	267.80	184.33	4.41	4.60	5.30
dhs-28	138.33	129.60	30.00	5.51	5.99	5.80
dpy-14	38.40	53.60	0.60	4.93	5.44	6.33
dpy-17	173.00	125.83	142.83	4.19	4.90	4.19
elo-5	342.60	276.00	247.00	4.16	3.97	3.85
nkb-3	31.20	198.00	72.40	4.26	4.19	5.91
fat-2	328.40	238.20	283.40	4.02	4.61	4.89
gei-7	302.33	299.20	200.67	4.31	4.54	4.36
gld-1	215.00	222.40	165.80	4.24	4.71	5.00
hex-1	258.69	302.85	150.00	4.04	4.97	6.37
hsp-12.6	135.40	136.40	109.80	3.72	4.34	4.27
lbp-5	137.36	268.54	161.39	4.53	4.97	6.57
lec-6	264.40	290.25	181.00	3.79	4.84	6.03
lec-8	244.60	302.60	268.00	3.96	4.92	5.54
lec-9	132.09	288.15	154.77	4.76	4.98	7.04
lys-1	156.33	280.15	206.39	4.96	4.79	6.23
lys-10	240.64	313.23	171.77	3.67	4.79	6.21
lys-2	137.89	284.69	188.62	4.72	4.84	6.11
lys-4	285.27	280.08	203.69	4.10	4.87	6.42
lys-5	278.60	297.80	178.40	4.04	4.82	6.19
mtl-1	95.00	242.69	110.46	5.21	4.76	7.30
mtl-2	346.00	361.20	269.50	4.16	4.44	4.26
pab-2	304.00	300.20	279.40	3.94	4.01	4.84
rol-6	211.60	162.00	257.60	4.72	4.74	5.91
sqt-2	339.00	233.80	218.00	4.44	5.01	4.20

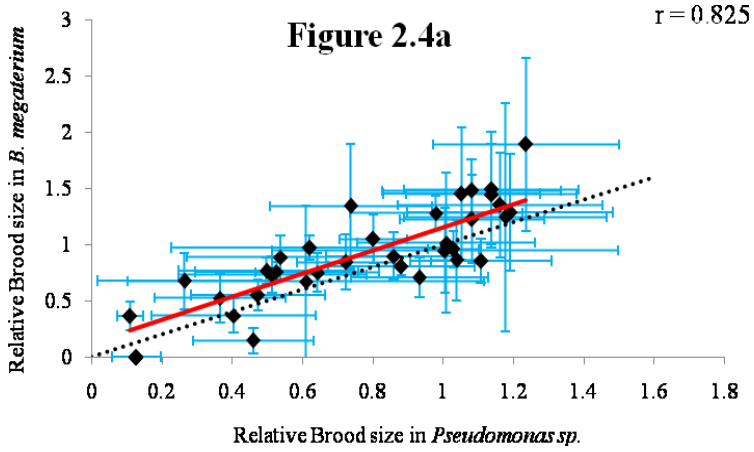
Table 2.6 Regression of relative fitness to total brood size and generation time (GT).

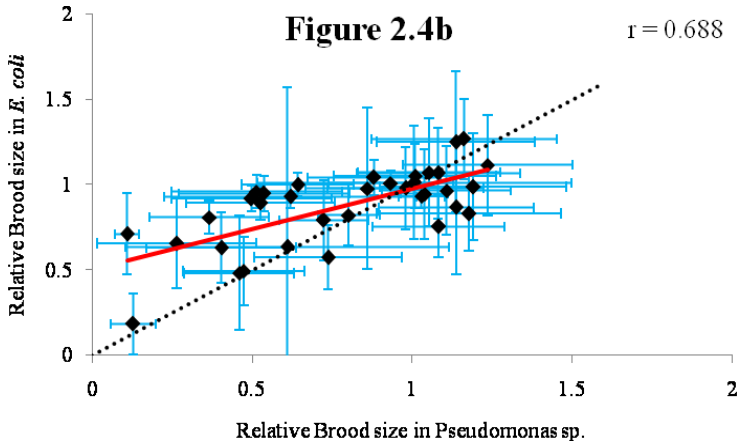
Relative brood size and generation time were used as a predictor for relative fitness of mutants. Linear regressions were performed using brood size or generation time as independent variable and mutant fitness as dependent variable in each bacterium. The standard error and *P*-value of coefficient of slope is shown.

Regression	*B megaterium*			*E. coli*			*Pseudomonas sp.*		
	slope	se	pvalue	slope	se	pvalue	slope	se	pvalue
Rel fit to Brood	0.23	0.01037	<0.0001	0.24	0.007896	<0.0001	0.31	0.01007	<0.0001
Rel fit to GT	-0.87	0.03407	<0.0001	-0.97	0.021489	<0.0001	-0.75	0.02753	<0.0001

Figure 2.4 Relative Brood size of mutants in stressful vs. less stressful bacteria

Relative Brood size of mutants was calculated (mutant brood size/wild-type brood size) in each bacterium. Confidence intervals (α=0.05) is indicated as error bars. Pearson's Correlation coefficient (r) is shown. Dashed line represents 1:1 reference line with equal relative brood size.





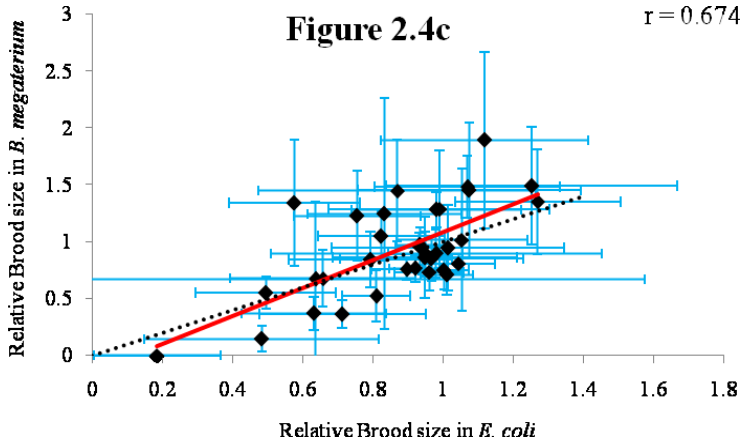
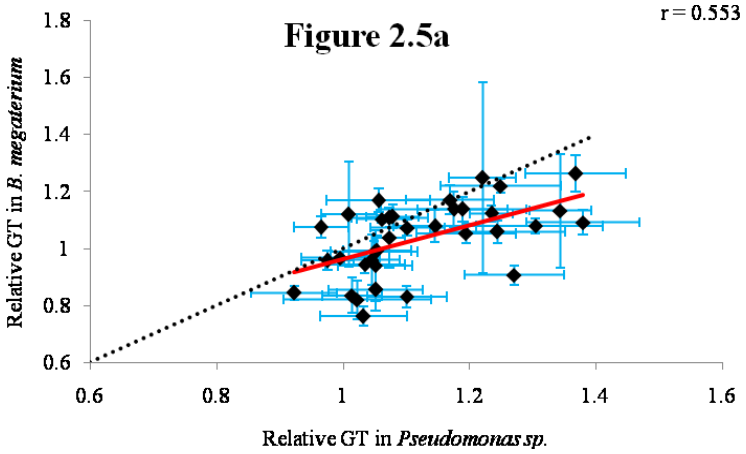
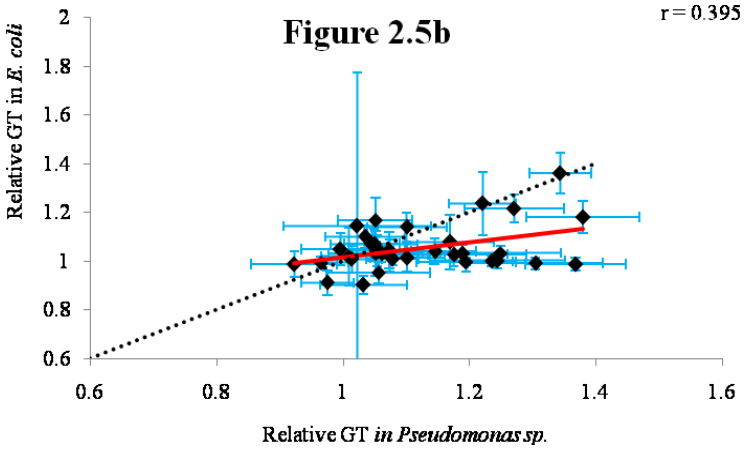


Figure 2.5 Relative Generation Time of mutants in stressful vs. less stressful bacteria

Relative Generation Time (GT) of mutants was calculated (mutant GT/wild-type GT) in each bacterium. Confidence intervals (α=0.05) is indicated as error bars. Pearson's Correlation coefficient (R^2) is shown. Dashed line represents 1:1 reference line with equal relative GT.





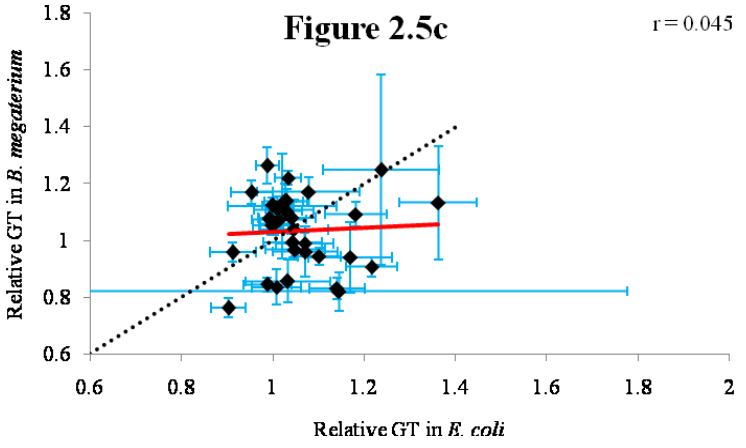


Table 2.7 Comparisons of relative brood size of mutants for stressful vs. less stressful bacteria

Difference between relative brood sizes of each bacterial comparison is calculated. Red shade indicates that the confidence interval (α=0.05) of the difference is positive; yellow indicates negative and no shading indicates confidence interval contain zero. B, *B. megaterium*.; E, *E. coli* OP50; P, *Pseudomonas* sp.

Gene	B-P	B-E	E-P
acdh-1			
c23h5.8			
cey-2			
cey-4			
clec-50			
cpi-1			
cpr-5			
ctl-1			
cyp-37A1			
cyp-34A9			
daf-22			
dhs-28			
dpy-14			
dpy-17			
elo-5			
F55F3.3			
fat-2			
gei-7			
gld-1			
hex-1			
hsp-12.6			
lbp-5			
lec-6			
lec-8			
lec-9			
lys-1			
lys-10			
lys-2			
lys-4			
lys-5			
mtl-1			
mtl-2			
pab-2			
rol-6			
sqt-2			

Table 2.8 Comparisons of relative Generation Time of mutants for stressful vs. less stressful bacteria

Difference between relative Generation Times of each bacterial comparison is calculated. Red shade indicates that the confidence interval (α=0.05) of the difference is positive; yellow indicates negative and no shading indicates confidence interval contain zero. B, *B. megaterium*.; E, *E. coli* OP50; P, *Pseudomonas* sp.

Gene	B-P	B-E	E-P
acdh-1			
c23h5.8			
cey-2			
cey-4			
clec-50			
cpi-1			
cpr-5			
ctl-1			
cyp-37A1			
cyp-34A9			
daf-22			
dhs-28			
dpy-14			
dpy-17			
elo-5			
F55F3.3			
fat-2			
gei-7			
gld-1			
hex-1			
hsp-12.6			
lbp-5			
lec-6			
lec-8			
lec-9			
lys-1			
lys-10			
lys-2			
lys-4			
lys-5			
mtl-1			
mtl-2			
pab-2			
rol-6			
sqt-2			

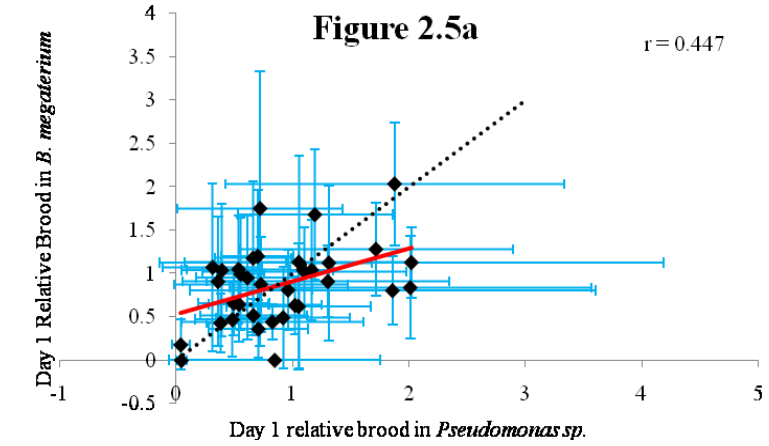
Table 2.9 Age specific absolute brood size of mutants exposed to different bacterial environments

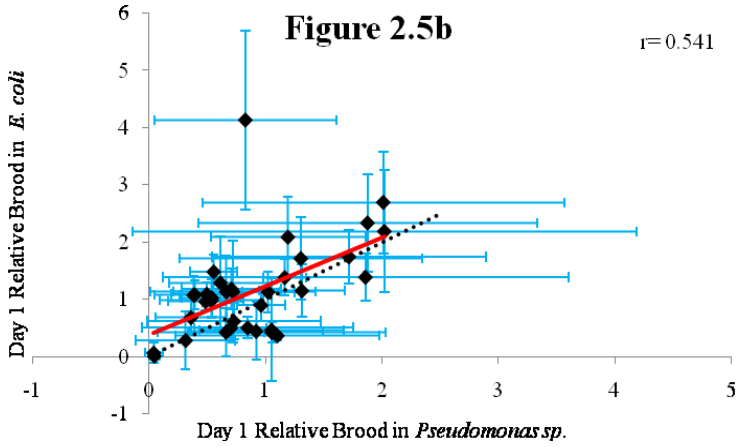
Wild-type *C. elegans* (*N2*) and mutant strains were grown on the three bacteria and brood size per day was assayed. Red shade indicates significant (P<0.05) increase relative to wild-type and yellow indicates significant (P<0.05) decrease relative to wild-type. B, *B. megaterium*.; E, *E. coli* OP50; P, *Pseudomonas* sp.

	Brood size - Day 1			Brood size - Day 2			Brood size - Day 3			Brood size - Day		
Gene	P	E	B	P	E	B	P	E	B	P	E	B
N2	60.4	40.57	24.2	194.5	139.6	108.3	36.96	95.53	78.17	1.19	16.2	4.83
acdh-1	42.3	11.40	34.0	152.5	107.8	72.00	87.83	134.0	71.60	7.00	25.0	18.2
C23H5.8	34.4	10.20	33.6	205.0	123.8	153.2	73.80	72.60	51.20	0.20	5.80	1.20
cey-2	59.6	62.40	68.8	191.6	155.8	148.0	73.40	128.8	70.60	3.20	1.00	1.60
cey-4	57.4	58.60	39.2	198.8	170.8	137.4	100.0	84.40	171.0	0.40	0.00	15.2
clec-50	22.7	28.00	16.8	34.46	79.39	60.91	10.55	64.23	64.36	0.36	22.0	23.1
cpi-1	33.8	24.13	42.2	124.7	86.25	101.3	54.11	95.25	59.33	4.67	26.0	4.11
cpr-5	25.4	44.54	13.5	61.91	134.6	72.92	36.36	80.39	57.69	4.09	13.4	13.2
ctl-1	24.4	13.60	0.20	77.60	90.20	46.60	30.00	71.40	30.75	1.00	1.80	2.67
cyp-34A9	36.4	65.31	29.0	78.23	148.2	55.77	53.08	74.77	62.92	5.85	9.54	7.38
cyp-37A1	62.6	72.60	27.6	203.8	168.8	132.0	44.80	58.80	115.8	0.20	0.80	22.2
daf-22	19.8	33.20	30.0	105.5	100.4	71.17	109.6	90.80	71.00	26.0	32.0	48.5
dhs-28	1.67	1.40	5.83	71.33	34.00	12.17	55.50	60.80	3.50	9.50	29.8	9.83
dpy-14	2.20	0.20	0.00	31.00	29.00	0.40	8.33	22.00	0.25	1.00	3.00	0.00
dpy-17	11.1	9.50	33.1	88.50	65.50	69.00	82.40	58.25	49.25	6.75	13.6	10.2
elo-5	39.6	56.00	55.8	215.8	173.8	173.0	79.40	43.40	17.60	7.80	2.80	0.60
nkb-3	24.4	110.0	15.4	5.40	46.60	21.00	1.00	33.80	25.80	0.40	7.60	10.2
fat-2	55.4	37.40	27.6	212.8	136.2	131.4	59.60	63.80	111.6	0.60	0.60	12.2
gei-7	23.1	29.80	55.1	161.1	112.2	103.0	115.6	127.6	49.40	2.00	26.6	1.00
gld-1	21.8	29.60	17.8	113.2	123.2	131.2	76.40	65.20	16.20	3.40	4.00	0.60
hex-1	68.1	48.23	22.6	157.0	144.3	64.85	31.00	94.92	58.54	1.92	12.8	3.38
hsp-12.6	40.4	45.40	30.2	90.80	71.40	74.60	4.20	15.40	4.80	0.00	3.80	0.20
lbp-5	34.5	44.85	15.4	59.27	124.5	63.08	38.18	85.23	60.39	5.27	11.6	20.6
lec-6	66.0	37.25	35.4	183.6	150.5	107.0	14.60	72.75	37.80	0.00	24.5	0.80
lec-8	41.6	33.00	42.2	172.0	143.8	172.2	31.00	91.60	53.60	0.00	32.6	0.00
lec-9	31.6	50.23	18.0	61.82	130.4	63.31	30.55	81.39	59.15	6.09	21.3	11.3
lys-1	33.2	47.00	28.9	73.33	127.0	108.2	48.11	93.31	62.46	1.22	11.8	6.62
lys-10	77.0	60.69	26.0	154.8	145.9	90.15	6.46	79.31	50.85	1.82	20.5	4.23
lys-2	24.4	50.77	28.9	83.44	120.3	94.69	29.56	92.85	58.77	0.44	17.0	6.00
lys-4	63.6	39.77	24.1	176.8	135.5	98.62	44.18	83.54	63.69	0.64	15.0	17.0
lys-5	58.6	38.00	27.8	195.4	146.6	94.20	24.60	88.20	56.20	0.00	23.2	0.20
mtl-1	30.6	43.85	14.8	38.62	114.0	48.54	20.92	68.39	33.46	2.54	12.7	10.6
mtl-2	46.8	31.60	36.8	196.1	166.2	136.3	102.0	141.8	91.17	1.00	17.2	5.17
pab-2	56.2	47.40	42.8	212.2	202.2	117.2	34.60	50.20	100.8	1.00	0.40	18.6
rol-6	21.2	16.20	28.6	125.2	97.40	117.0	64.00	50.50	89.60	1.20	9.75	22.0
sqt-2	23.2	31.80	12.4	163.2	166.8	124.2	133.8	35.00	72.40	18.4	0.20	14.6

Figure 2.6 Relative Day 1 Brood size of mutants in stressful vs. less stressful bacteria

Relative day 1 Brood size of mutants was calculated (mutant brood size/wild-type brood size) in each bacterium. Confidence intervals (α=0.05) is indicated as error bars. Pearson's Correlation coefficient (r) is shown. Dashed line represents 1:1 reference line with equal relative brood size.





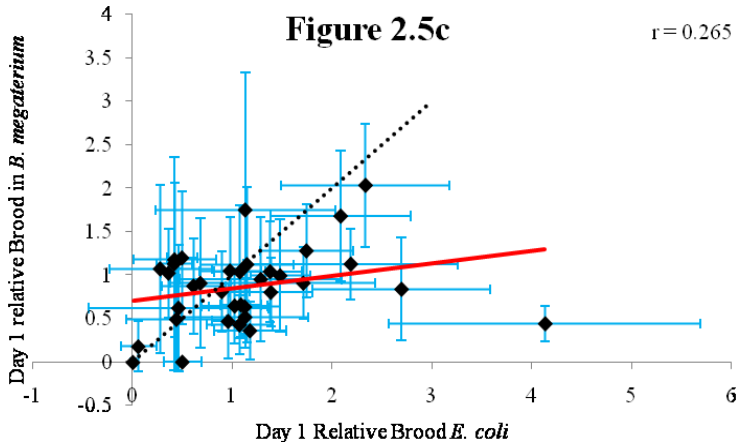
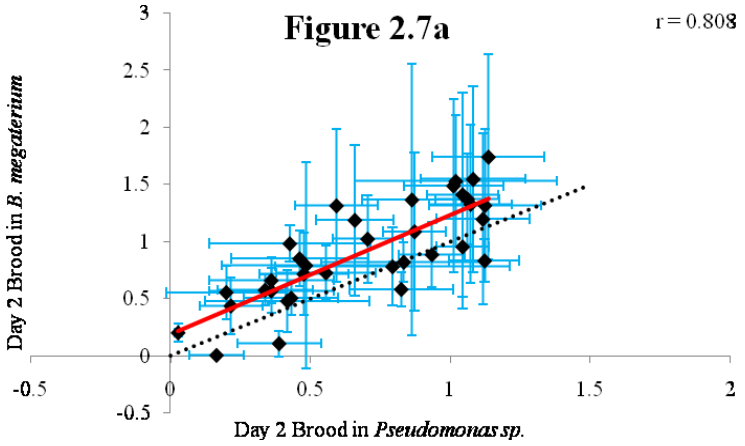
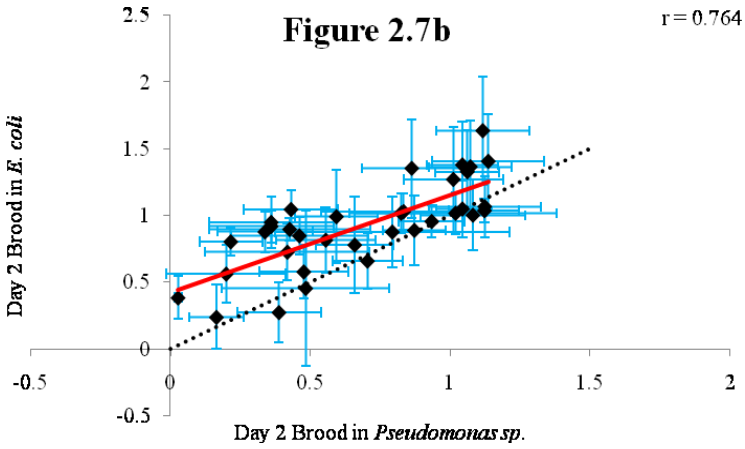


Figure 2.7 Relative Day 2 Brood size of mutants in stressful vs. less stressful bacteria

Relative day 2 Brood size of mutants was calculated (mutant brood size/wild-type brood size) in each bacterium. Confidence intervals (α=0.05) is indicated as error bars. Pearson's Correlation coefficient (r) is shown. Dashed line represents 1:1 reference line with equal relative brood size.





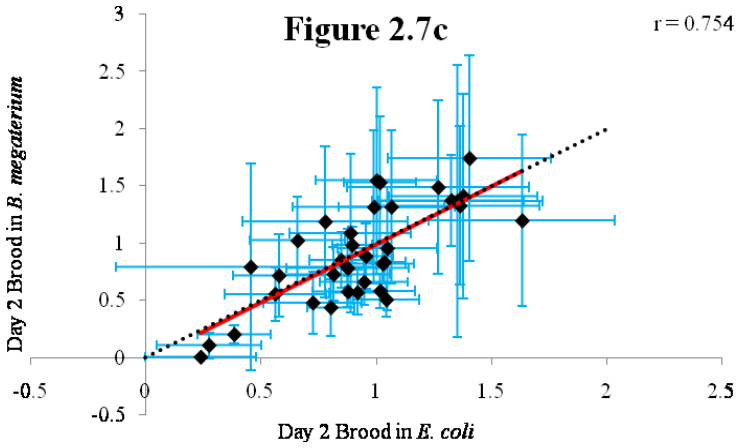
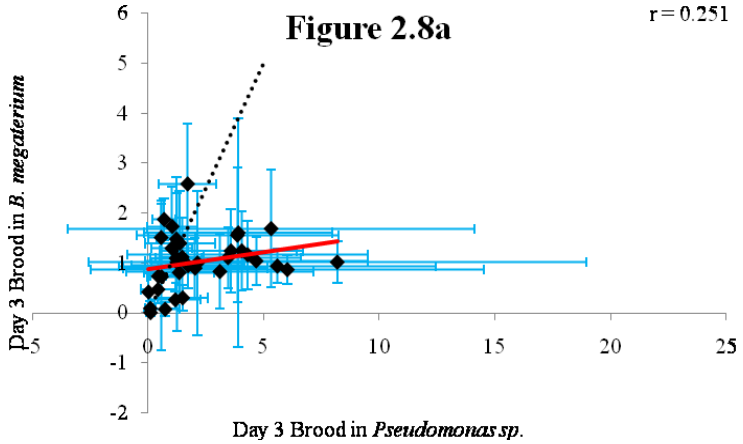
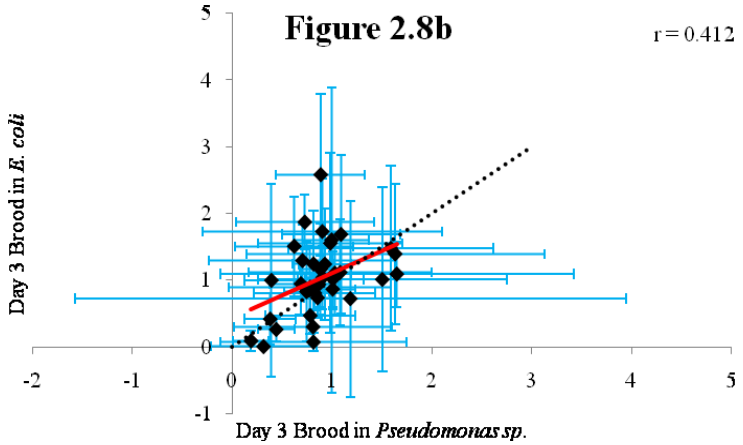


Figure 2.8 Relative Day 3 Brood size of mutants in stressful vs. less stressful bacteria

Relative day 3 Brood size of mutants was calculated (mutant brood size/wild-type brood size) in each bacterium. Confidence intervals (α=0.05) is indicated as error bars. Pearson's Correlation coefficient (r) is shown. Dashed line represents 1:1 reference line with equal relative brood size.





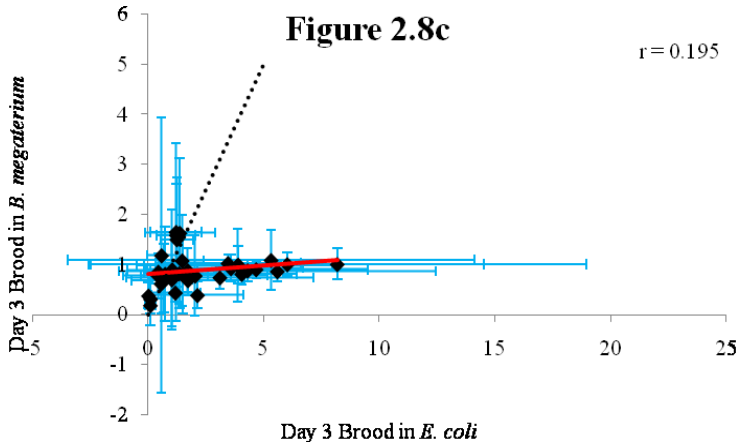
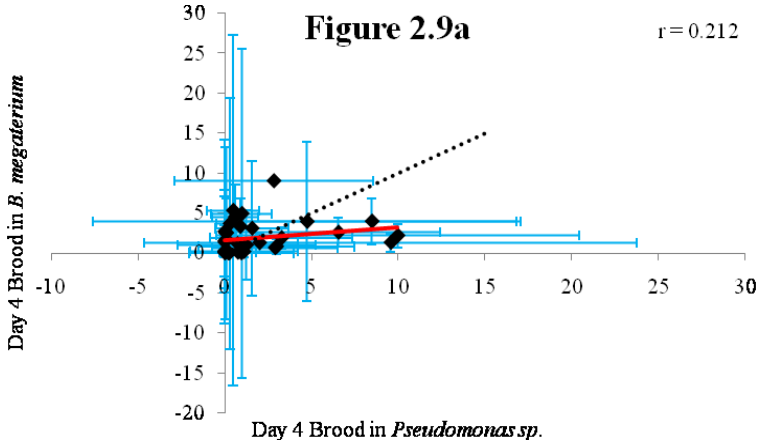
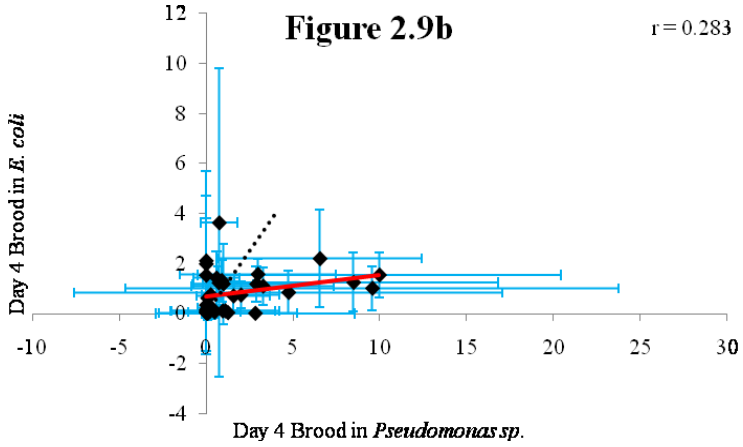


Figure 2.9 Relative Day 4 Brood size of mutants in stressful vs. less stressful bacteria

Relative day 4 Brood size of mutants was calculated (mutant brood size/wild-type brood size) in each bacterium. Confidence intervals (α=0.05) is indicated as error bars. Pearson's Correlation coefficient (r) is shown. Dashed line represents 1:1 reference line with equal relative brood size.





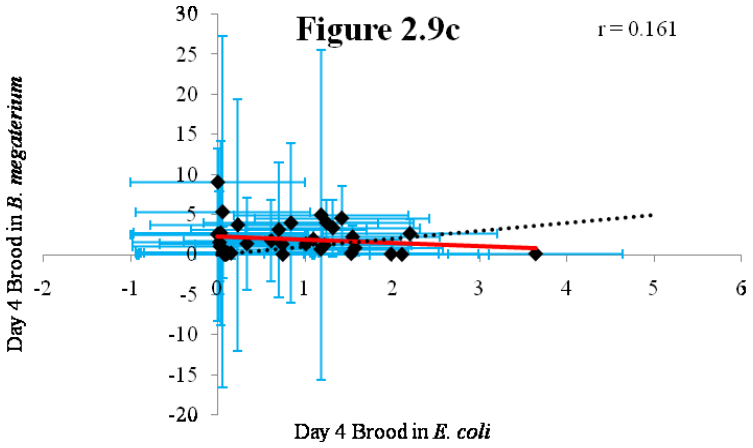


Table 2.10 Comparisons of relative Brood size of mutants for stressful vs. less stressful bacteria on first, second, third and fourth day.

Difference between relative Brood sizes for each day is calculated for each bacterial comparison. Red shade indicates that the confidence interval (α=0.05) of the difference is positive; yellow indicates negative and no shading indicates confidence interval contain zero. B, *B. megaterium*.; E, *E. coli* OP50; P, *Pseudomonas* sp.

Gene	Rel Brood-Day 1			Rel Brood-Day 2			Rel Brood-Day 3			Rel Brood-Day 4		
	B-P	B-E	E-P	B-P	B-E	E-P	B-P	B-E	E-P	B-P	B-E	E-P
acdh-1												
c23h5.8												
cey-2												
cey-4												
clec-50												
cpi-1												
cpr-5												
ctl-1												
cyp-37A1												
cyp-34A9												
daf-22												
dhs-28												
dpy-14												
dpy-17												
elo-5												
F55F3.3												
fat-2												
gei-7												
gld-1												
hex-1												
hsp-12.6												
lbp-5												
lec-6												
lec-8												
lec-9												
lys-1												
lys-10												
lys-2												
lys-4												
lys-5												
mtl-1												
mtl-2												
pab-2												
rol-6												
sqt-2												

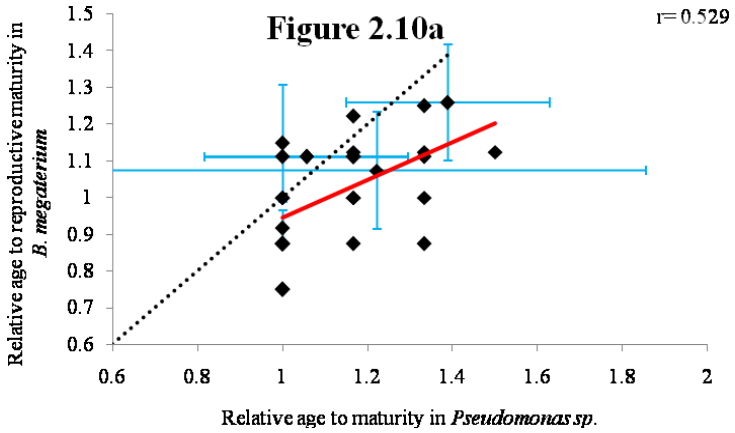
Table 2.11 Absolute age of reproductive maturity of mutants exposed to different bacterial environments

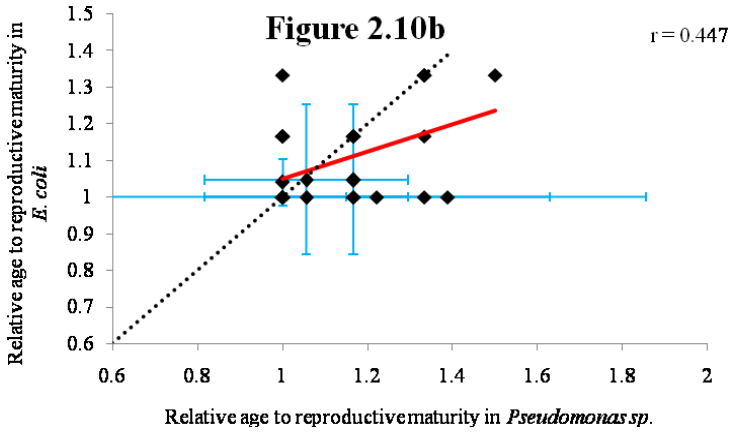
Wild-type *C. elegans* (*N2*) and mutant strains were grown on the three bacteria and the time taken to lay first eggs was assayed. Red shade indicates significant (P<0.05) increase relative to wild-type and yellow indicates significant (P<0.05) decrease relative to wild-type. B, *B. megaterium*.; E, *E. coli* OP50; P, *Pseudomonas* sp.

Age of reproductive maturity (days)			
Gene	P	E	B
N2	3.00	3.19	4.17
acdh-1	3.00	3.50	3.50
C23H5.8	4.00	4.00	3.50
cey-2	4.50	4.00	4.50
cey-4	3.50	3.50	4.00
clec-50	4.00	3.50	5.00
cpi-1	3.00	3.29	3.83
cpr-5	3.50	3.50	4.50
ctl-1	3.50	3.50	3.50
cyp-34A9	3.50	3.67	5.00
cyp-37A1	3.50	3.50	4.00
daf-22	3.00	3.00	4.00
dhs-28	4.00	4.00	4.00
dpy-14	4.00	4.00	5.00
dpy-17	3.00	3.00	3.50
elo-5	3.00	3.00	3.00
nkb-3	4.00	3.50	4.50
fat-2	3.00	3.50	3.50
gei-7	3.00	3.00	3.50
gld-1	3.00	3.50	4.00
hex-1	3.17	3.67	5.00
hsp-12.6	3.00	3.50	3.50
lbp-5	3.50	3.67	5.00
lec-6	3.00	3.50	5.17
lec-8	3.00	3.50	4.50
lec-9	3.50	3.67	5.50
lys-1	3.50	3.50	5.00
lys-10	3.00	3.50	5.00
lys-2	3.67	3.50	4.83
lys-4	3.17	3.50	5.00
lys-5	3.17	3.50	5.00
mtl-1	4.17	3.50	5.67
mtl-2	3.00	3.00	3.00
pab-2	3.00	3.00	3.50
rol-6	3.50	3.50	4.50
sqt-2	3.00	4.00	3.00

Figure 2.10 Relative Age of reproductive maturity of mutants in stressful vs. less stressful bacteria.

Relative age of reproductive maturity of mutants was calculated (mutant age/wild-type age in each bacterium. Confidence intervals (α=0.05) is indicated as error bars. Pearson's Correlation coefficient (r) is shown. Dashed line represents 1:1 reference line with equal relative age.





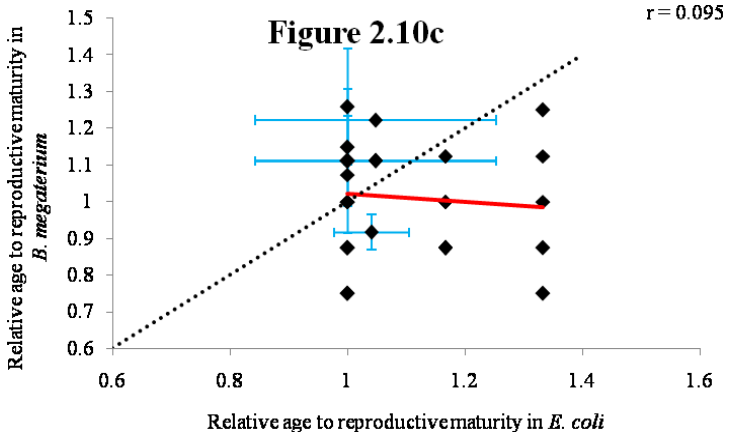


Table 2.12 Comparisons of relative age specific reproductive maturity of mutants for stressful vs. less stressful bacteria

Difference between relative age specific reproductive maturities of each bacterial comparison is calculated. Red shade indicates that the confidence interval (α=0.05) of the difference is positive; yellow indicates negative and no shading indicates confidence interval contain zero. B, *B. megaterium*.; E, *E. coli* OP50; P, *Pseudomonas* sp.

Gene	B-P	B-E	E-P
acdh-1			
c23h5.8			
cey-2			
cey-4			
clec-50			
cpi-1			
cpr-5			
ctl-1			
cyp-37A1			
cyp-34A9			
daf-22			
dhs-28			
dpy-14			
dpy-17			
elo-5			
F55F3.3			
fat-2			
gei-7			
gld-1			
hex-1			
hsp-12.6			
lbp-5			
lec-6			
lec-8			
lec-9			
lys-1			
lys-10			
lys-2			
lys-4			
lys-5			
mtl-1			
mtl-2			
pab-2			
rol-6			
sqt-2			

Table 2.13 Absolute post reproductive lifespan of mutants exposed to different bacterial environments

Wild-type *C. elegans* (*N2*) and mutant strains were grown on the three bacteria and the time of death of 50% of the population (TD_{50}) was found out. Red shade indicates significant ($P<0.05$) increase relative to wild-type and yellow indicates significant ($P<0.05$) decrease relative to wild-type. B, *B. megaterium*.; E, *E. coli* OP50; P, *Pseudomonas* sp.

TD50 (days)			
Gene	P	E	B
N2	8.94	9.59	12.93
acdh-1	5.50	5.00	10.40
C23H5.8	6.00	7.80	8.90
cey-2	7.50	6.10	7.00
cey-4	5.90	5.60	3.70
clec-50	7.50	9.05	12.13
cpi-1	7.74	11.57	13.05
cpr-5	11.36	12.18	17.04
ctl-1	3.90	6.20	8.50
cyp-34A9	9.30	8.70	15.70
cyp-37A1	8.50	8.00	9.50
daf-22	8.20	6.30	15.00
dhs-28	7.30	6.70	10.20
dpy-14	3.10	2.40	4.10
dpy-17	3.00	4.00	12.30
elo-5	5.00	5.50	9.50
nkb-3	5.00	3.10	5.50
fat-2	11.40	9.90	13.70
gei-7	7.60	5.70	14.30
gld-1	4.30	5.60	5.50
hex-1	11.90	10.00	12.50
hsp-12.6	6.60	5.70	9.50
lbp-5	10.00	9.30	12.90
lec-6	8.00	7.67	8.75
lec-8	9.14	9.29	12.90
lec-9	11.19	12.68	17.67
lys-1	9.20	8.70	11.30
lys-10	9.11	8.41	13.55
lys-2	7.90	9.22	12.00
lys-4	9.15	8.25	12.63
lys-5	7.00	10.25	10.50
mtl-1	8.10	8.20	12.10
mtl-2	8.00	6.10	13.80
pab-2	7.70	6.60	8.90
rol-6	7.70	3.10	10.20
sqt-2	4.20	6.90	7.20

Table 2.14 Comparisons of relative TD_{50} maturity of mutants for stressful vs. less stressful bacteria

Difference between relative TD_{50} of each bacterial comparison is calculated. Red shade indicates that the confidence interval (α=0.05) of the difference is positive; yellow indicates negative and no shading indicates confidence interval contain zero. B, *B. megaterium*; E, *E. coli* OP50; P, *Pseudomonas* sp.

Gene	B-P	B-E	E-P
acdh-1			
c23h5.8			
cey-2			
cey-4			
clec-50			
cpi-1			
cpr-5			
ctl-1			
cyp-37A1			
cyp-34A9			
daf-22			
dhs-28			
dpy-14			
dpy-17			
elo-5			
F55F3.3			
fat-2			
gei-7			
gld-1			
hex-1			
hsp-12.6			
lbp-5			
lec-6			
lec-8			
lec-9			
lys-1			
lys-10			
lys-2			
lys-4			
lys-5			
mtl-1			
mtl-2			
pab-2			
rol-6			
sqt-2			

Chapter 3 - *C. elegans* response to grassland soil bacteria involves environment-specific rewiring of gene interactions

Introduction

Although a single gene could be an important component that affects a trait, there are perhaps no cases where a trait is entirely controlled by a single gene. Biological processes are the outcome of interaction of gene products that can be mapped into biochemical, morphogenetic or signal-transduction pathways. Recently it has become clear that most biological process are controlled by pathways that are highly regulated, often branched with feedback and feed forward mechanisms (Greenspan 2001; Wang and Sherwood 2011). The interaction of alleles at different loci that effect a phenotype is termed epistasis (Cheverud and Routman 1995). In the broadest sense, epistasis refers to the genotype influence on phenotype that is dependent on genetic background (Carlborg and Haley 2004). Since most biological functions involve genes that interact with each other in a hierarchy, epistasis is ubiquitous in biological systems (Phillips 2008).

The term epistasis is being used in two ways (Moore 2005). Bateson's definition of epistasis explains how a phenotype is affected due to interaction of two genes, which indicates that they exist in genetic pathways (Phillips 2008) while in quantitative genetic models, epistasis is referred as statistical deviations from additive effects of alleles in their effects on phenotype (Fenster, Galloway et al. 1997; Phillips 2008). In the first case epistasis is used to explain the expression of a phenotype and in second case it refers to the contribution of alleles at different loci to phenotypic variation found in natural populations (Fenster, Galloway et al. 1997). Epistatic gene action in a genetic pathway depends only on the combination of alleles whereas epistatic variance in a quantitative trait is a function of allele frequencies (Whitlock, Phillips et al. 1995). If an allelic pair have a fitness effect different from the effect when they are combined with other alleles, it may not always result in sufficient frequencies in a population to mount substantial epistatic variance. But, since genotypic fitness is defined based on fitness of the combination of alleles, evolutionary potential depends on epistatic gene action rather than epistatic variance (Whitlock, Phillips et al. 1995).

Epistasis, in general, has been suggested to play important role in evolution. Various models of evolution propose epistasis to cause the formation of multiple adaptive peaks across the fitness landscape (Whitlock, Phillips et al. 1995; Phillips 2008). Epistasis has been proposed to be an important factor for the evolution and maintenance of sex and recombination (Kondrashov 1988).When negative epistasis (negative linkage disequilibrium) generates associations of favorable and deleterious alleles together on a chromosome, sex and recombination can bring together favorable alleles and deleterious alleles on to the distinct chromosomes (Otto and Gerstein 2006). Thus the response to selection is improved, causing sexual populations to quickly purge deleterious alleles and allow the frequency of favorable alleles to increase more than in asexual populations (Otto and Gerstein 2006). Reproductive isolation due to the Dobzhansky-Muller model of hybrid incompatibility results when epistatic interactions between alleles from different populations with negative fitness consequences come together (Wu and Ting 2004). Canalization, which describes the stability of complex developmental processes by reducing phenotypic variance and genetic buffering against mutations requires genetic interaction among a network of genes that are redundant and robust (Moore 2005). Recently, work in *C. elegans* has shown that incomplete penetrance of mutations in isogenic individuals could be due to stochastic variation in the activity of interaction partners (Burga, Casanueva et al. 2011)

Genetic architecture refers to genes, including their epistatic relations and interactions between genes and environments, that affect the expression of a trait (Wade, Winther et al. 2001). How a genome responds to an environment depends on its genetic architecture. Studies have shown that the effect of epistasis has to be considered to understand how organisms adapt to changing environments (Hayden and Wagner 2012; Flynn, Cooper et al. 2013; Lalic and Elena 2013). In an example of recessive epistasis in the well established pathway underlying coat color variation in mammals, *Melanocortin 1 receptor* (*Mc1r*) locus operates downstream of *Agouti* (Phillips 2008). It has been shown that the adaptive light color pattern in beach mice that camouflages in its environment results from this non-additive interaction between structural changes in the *Mc1r* locus and regulatory changes in *Agouti* (Steiner, Weber et al. 2007). In addition, when two genomes interact intimately, especially in a host-pathogen co-evolution system, epistasis has been found to be an important component of complex traits such as host

disease resistance (Kover and Caicedo 2001) and contribute to phenotypic differentiation in resistance (Wegner, Berenos et al. 2008).

Just as there are genes of adaptive significance, can we expect to find allelic interactions of adaptive significance? Understanding whether changes in the environment changes specificity in gene interaction partners is an important step towards finding the most adaptive allelic combinations involved in this interaction. In our approach of ecological genomics exploring mechanisms underlying genomic responses of organisms to natural environment (Ungerer, Johnson et al. 2008), genetic interactions are likely to play an important role. For the model nematode, *C. elegans*, the natural environment includes a heterogeneous mixture of bacteria which serve as both food source as well as pathogenic effects (Felix and Braendle 2010; Felix and Duveau 2012). *C. elegans* lifespan is a complex phenotype quantified in response to bacteria which is affected by many different genetic pathways (Darby 2005; Irazoqui, Urbach et al. 2010).

For over a decade, many laboratories have documented the genetic basis of *C. elegans* host-defense responses to bacteria. These efforts led to the discovery of evolutionarily conserved host signaling pathways that respond to bacterial infection and also regulate its metabolism. One of the most important signal cascades involved in the response to bacteria is the p38 Mitogen Activated Protein Kinase (p38 MAPK) pathway. This pathway involves NSY-1, a MAP kinase kinase kinase that phosphorylates SEK-1, a MAP kinase kinase that in turn phosphorylates the MAP kinase, PMK-1 (Kim, Feinbaum et al. 2002). *sek-1* and *pmk-1* mutants exhibit reduced survivorship when fed bacterial pathogens, *Pseudomonas aeruginosa* and *E. coli* OP50 (when grown at 25°C) (Troemel, Chu et al. 2006). Another important evolutionarily conserved pathway involved in regulating lifespan stress responses and pathogen resistance is the DAF-2 insulin-like growth factor receptor pathway (Irazoqui, Urbach et al. 2010; Partridge, Gravato-Nobre et al. 2010). Loss of function of *daf-2* triggers the activation of downstream transcription factor DAF-16 which leads to expression of antimicrobial peptides and other stress response factors (Irazoqui and Ausubel 2010). In response to *Pseudomonas aeruginosa* and *E. coli* OP50, the lifespan extension brought about by mutations in *daf-2* is completely suppressed by loss of *daf-16* function (Troemel, Chu et al. 2006). TOL-1, which belongs to Toll-like conserved family of pathogen signal receptors, was found to be required for defense responses to bacterial pathogen, *Salmonella enterica* (Tenor and Aballay 2008). *tol-1* mutants have been shown to have shortened

survivorship on *E. coli* OP50, but not on *Pseudomonas aeruginosa* (Pujol, Link et al. 2001). DBL-1 is non-canonical transforming growth factor-β (TGFβ) ligand that triggers the induction of *C. elegans* antimicrobial genes encoding caenacins in response to fungal pathogen *Dreshmeria coniospora*. *dbl-1* mutants have been shown to shorten lifespan in response to *E. coli* OP50 (Mallo, Kurz et al. 2002). The TGFβ pathway has also been shown to regulate *C. elegans* resistance to *Pseudomoas aeruginosa* and *S. marcescens* (Mallo, Kurz et al. 2002; Kurz and Ewbank 2003). The difference in survivorship of *C. elegans* in response to *E. coli* and *Pseudomonas aeruginosa* is drastic (Troemel, Chu et al. 2006). Since components of all of the above pathways are involved in modulating responses to these two bacteria (except *tol-1* for *Pseudomonas aeruginosa*), genetic mechanisms contributing to differential survivorship in bacteria is not the mere involvement of these components. Specific responses to these two bacteria could be either due to their modulated responses in terms of activity or interaction with other gene products.

Our prior transcriptomic analyses that explored the interaction of *C. elegans* with soil bacteria allowed us to find genes with differential expression patterns which could be correlated to their mutational phenotypes in respective bacterial environments (Coolon, Jones et al. 2009). Mutations in these genes were found to affect life history traits such as survivorship and intrinsic rate of population growth and thus have fitness consequences specific to bacterial environments.. We have seen genes that are specifically involved in responses to a bacterium as well as genes that are general modulators involved in responding to multiple bacteria. These results prompted us to investigate the functional basis of gene interactions in relation to the bacterial environment. To gain insights into the adaptive significance of gene interactions, we asked whether there are gene interactions that are specific to a particular bacterium and moreover whether the gene components involved are changed or modulated in response to specific bacteria. Using the genetic tools available with *C. elegans*, we generated double mutants and tested their survivorship in different bacterial environments to determine the epistatic relations of genes specific to the environment. We find that along with a change in bacterial environment, genetic components are changed and gene interactions are modulated according to the environment.

Materials and Methods

C. elegans and bacteria-Strains and maintanence

The following mutant strains were used. Bristol (*N2*), *asp-6* (*tm2213*), *cpi-1* (*ok1213*) *cpr-5* (*ok2344*), *cyp-34A9* (*ok2401*), *lbp-5* (*tm1618*), *clec-50* (*ok2455*), *fat-7* (*BX153*), *F57F5.1* (*ok3419*), *lec-6* (*tm2552*), *lec-8* (*tm1477*), *lec-9* (*tm1206*), *lys-1* (*ok2445*), *lys-2* (*tm2398*), *lys-4* (*tm2938*), *lys-10* (*tm2558*), *mtl-1* (*tm1770*), *hex-1* (*tm1992*) *pmk-1* (*km25*), *sek-1* (*km4*), *daf-2* (*e1368*) *daf-16* (*mu86*), *tol-1* (*nr2033*), *dbl-1* (*nk3*). Growth and maintenance conditions were as described (Brenner 1974; Sulston and Hodgkin 1988). Use of bacteria was as for *E. coli* (OP50), *Pseudomonas sp*. (NCBI's GenBank database accession number-EU704696) and *Bacillus megaterium* (EU704698).

Generation of double mutants

By crossing respective single mutants, double mutants were generated. F_2 populations were surveyed to find the double deletions by using PCR based markers. Primers used for genes are listed in Table 3.6.

Survivorship assays

For examining survivorship, longevity assays are performed as previously described (Tan, Mahajan-Miklos et al. 1999; Tan and Ausubel 2000), and survivorship was calculated from survivorship curves as Average Lifespan of individuals in a population. Briefly, worms were synchronized by bleaching to collect eggs and hatched in M9 overnight. Worms were then grown to L4 on *E. coli* (OP50) to standardize test populations, then transferred to the test bacteria (10 worms per plate) and were maintained at 25^0C. Surviving worms were then re-plated daily and the fraction surviving was determined every 24 hours. Worms were considered dead when they no longer responded to touch with platinum wire. All longevity assays were conducted in at least ten independent replicate experiments in each bacterium except for *fat-7 F57F5.1 , lec-6* for which there was 5 replications.

Statistical analysis

Hypothesis testing of *a proiri* contrasts was done using MIXED procedure in SAS (SAS Institute Inc., Cary, North Carolina, USA) using the following statistical model.

Model: $Y=\mu + genotype + error$

Tests were performed separately for each bacterium with each hypothesis tested using test statements, and Y equal to the measured survivorship.

Results

Wild-type and mutant survivorship when feeding on grassland soil bacteria

We previously identified 204 *C. elegans* genes that were differentially expressed in pairwise comparisons of nematodes grown on three bacterial species isolated from prairie soils (*Micrococcus luteus, Pseudomonas sp.* and *Bacillus megaterium*) and the standard lab food, *E. coli* OP50 (Coolon, Jones et al. 2009). Functional analyses of available mutants affecting 21 of the 204 genes reveal many significant gene by environment interactions (Coolon, Jones et al. 2009). This demonstrated that transcriptional profiling identified genes of functional importance in each bacterial environment and indicated that many general and specific effect genes are involved in *C. elegans* response to different bacterial environments. Gene ontology analysis indicated that a quarter of the genes were involved in defense or metabolism functions. Since bacteria could act as either pathogenic or metabolic sources for *C. elegans*, we focused further functional analyses on these genes. For this study we considered additional gene mutants from a list of an additional 215 genes that were also differentially expressed, but at a lower FDR cut-off ($q=0.05$ rather than $q=0.01$) in our previous study and were specifically involved in defense and metabolism functions. Among these, lectins, lysozymes and protease functions comprised 24 percent. We included 17 additional gene mutants that primarily corresponded to these functions (Table 3.1). Finally, we focused on nematode-bacterial interactions and thus did not include *Micrococcus luteus* in the present functional analysis since this bacteria was not found in association with native nematodes whereas *Pseudomonas sp*. and *Bacillus megaterium* was isolated in association with Rhabditid nematodes (*Pellioditis* sp. and *Oschieus sp*. respectively) (Coolon, Jones et al. 2009).

Adult lifespan (starting from the L4 stage) of wild-type and mutants was obtained by analyzing survivorship of nematodes grown on each bacterium; *E. coli*, *Pseudomonas sp*. and *B.*

megaterium. Wild-type lifespan was found to be the longest on *B. megaterium* (12.5 days) followed by *E. coli* (9.6 days) and the shortest on *Pseudomonas sp*. (8.1 days) (Figure 3.1). Survivorship of mutants that affected several genes (*clec-50*, *lec-9*, *cpi-1* and *cpr-5),* were significantly different than wild-type on all three bacteria, while other mutants differed from wild-type on only two bacteria (*lys-4* and *cyp-34A9* on *B. megaterium* and *Pseudomonas sp.)* and some only differed on one bacterium (Figure 3.1). In particular, *lys-10* and *mtl-1* mutants were specific to *E. coli*; *lbp-5*, *lec-8* and *hex-1* were specific to *Pseudomonas sp*.; and *asp-6*, *lys-1* and *lys-2* were specific to *B. megaterium*. It is interesting to note that though *C. elegans* employs multiple lysozymes in response to bacteria (*lys-4*, *lys-10* in *E. coli* and *lys-4*, *lys-1* and *lys-2* in *B megaterium*), it also uses same lysozyme in response to multiple bacteria (*lys-4* against *E. coli* and *B. megaterium*). Similarly, it employs multiple lectins (*lec-8*, *lec-9*, *clec-50*) in response to *Pseudomonas sp*.; but the same lectins (*lec-9*, *clec-50*) are employed in all the three bacterial environments. Thus there are specific host factors that are involved in a specific response to each bacterium and general host factors involved in defending and/or metabolizing multiple bacteria. It is possible that host responses to multiple bacteria involve the same pathways of general effect genes with similar lifespan effects (whether increasing or decreasing lifespan). For protease-related genes, *cpr-5* is involved in all bacterial environments; although *cpi-1* is involved in all bacterial environments, the effect on survivorship is environment specific; being increased on *E. coli*, but decreased on *Pseudomonas sp*. and *B. megaterium*. Although some of these genes affect survivorship in response to multiple bacteria, they could be regulated so that the effect is environment-specific. Thus, additional specificity in the response to different bacteria might be provided by genes with specific effects in each bacterial environment and also by environment specific-regulation of general effect genes (Figure 3.1).

Components of evolutionary conserved innate immunity pathways are involved in the C. elegans response to grassland soil bacteria

Since the major *C. elegans* innate immunity pathways, p38 MAPK, DAF-2 ILS, TOL-1 and DBL-1/TGFß were shown to be involved in modulating survivorship in response to multiple bacteria, we reasoned that they might also have similar effects in response to the soil bacteria. Thus we set about to test the effects of representative pathway mutants on survivorship to *Pseudomonas sp*. and *B. megaterium* as well as *E. coli* as a standard. We found that *daf-2*

mutants showed increased survivorship on each bacterium and that *daf-16*, *pmk-1* and *sek-1* mutants showed decreased survivorship, consistent with previous observations for *E. coli* and various pathogens (Figure 3.2). *tol-1* and *dbl-1* mutants showed bacteria specificity with *tol-1* mutants having decreased survivorship on *E. coli* but did not show any difference from wild-type on *Pseudomonas sp*. and *B. megateruim*; whereas *dbl-1* mutants had decreased survivorship on *E. coli* and *B. megaterium*, but no difference from wild-type on *Pseudomonas sp*. Thus we found that the DAF-2 insulin-like signaling and p38 MAPK pathways are general effect signaling cascades that are involved in responses to grassland soil bacteria as well as human pathogens.

Specific and general functional effects of Lectins, Lysozymes and Proteases in C. elegans responses to grassland bacteria

Lysozymes are enzymes that can digest the peptidoglycan component of bacterial cell walls and have been implicated in bacterial defense and digestion functions (Schulenburg and Boehnisch 2008). Many of these lysozymes have been shown to be involved in defending *C. elegans* against bacterial pathogens and were also found to be under the control of major immune pathways such as the DAF-2 insulin-like signaling, p38 MAPK and TGFβ pathways (Mallo, Kurz et al. 2002; Murphy, McCarroll et al. 2003; Troemel, Chu et al. 2006; Schulenburg, Hoeppner et al. 2008; Irazoqui, Troemel et al. 2010; Hahm, Kim et al. 2011). We tested four *lys* genes and found that *lys-4* and *lys-10* were often coexpressed (Table 3.1). Furthermore, mutations in both of these *lys* genes caused decreased survivorship on *E. coli* and *lys-4* mutations also caused decreased survivorship on *B. megaterium*.

Lectins are soluble or membrane-bound receptors involved in innate immune responses of metazoans to bacteria, viruses, fungi or other parasites (Dam and Brewer 2010). *C elegans* galectin, *lec-8*, has been found to be expressed in response to *Serratia marcescens* (Mallo, Kurz et al. 2002) and has also been found to provide resistance to *B. thuringiensis* by competitive exclusion of Cry5b toxin of *B. thuringiensis* (Ideo, Fukushima et al. 2009). Many *C. elegans* C-type lectin (*clec*) genes have been shown to be differentially regulated in response to various bacterial environments and their expression is regulated by several innate immunity pathways (Mallo, Kurz et al. 2002; O'Rourke, Baban et al. 2006; Wong, Bazopoulou et al. 2007) (Murphy, McCarroll et al. 2003; Alper, McBride et al. 2007; Irazoqui, Troemel et al. 2010). Each of the four *lec* genes that we tested were found to be coexpressed in at least two instances (Table 2.1).

Functional analyses revealed that *clec-50* mutants decreased survivorship and *lec-9* mutants increased survivorship in all the bacteria, and *lec-8* mutants increased survivorship on *Pseudomonas sp* (Figure 3.1).

Proteases and their inhibitors have been found in both hosts and pathogens and are involved in host-pathogen defense responses (Kopitar-Jerala 2012). In *C. elegans*, protease genes are differentially regulated in response to many bacterial environments (O'Rourke, Baban et al. 2006; Wong, Bazopoulou et al. 2007; Irazoqui, Troemel et al. 2010; JebaMercy, Pandian et al. 2011) and some proteases are regulated by the p38 MAPK (Troemel, Chu et al. 2006) and DAF-2 pathways (Kwon, Narasimhan et al. 2010). All four protease-related genes that we studied were coexpressed in at least two instances (Table 2.1). However, although the protease inhibitor, *cpi-1* was differentially expressed in similar environments as the protease genes, the pattern of regulation was reversed (Table 3.1). Mutations in *cpr-5* and *cpi-1* also showed differential functional effects in all bacterial environments (Figure 3.1).

Gene-interaction analysis reveals bacteria-specific rewiring of interaction partners

We have found several genes involved in defense and metabolism that have similar expression patterns in response to multiple bacteria (Table 3.1). We have also found that many of these genes are functionally important in regulating survivorship of *C. elegans* in response to *E. coli*, *Pseudomonas sp* and *B. megaterium* (Figure 3.1). Thus, many of these genes are co-expressed in multiple environments and also have similar functional responses to multiple bacteria. In order to determine whether these genes function within the same or different pathways and whether they are regulated in similar manner in multiple bacterial environments, we analyzed genetic interactions by constructing double loss-of-function mutants. We analyzed selected lysozymes, lectins and proteases as well as representative components of the conserved innate immune pathways described earlier. We examined survivorship of these double mutants on each of the bacteria to discover their interaction specificities in response to each bacterial environment.

Gene interaction analysis-rules and assumptions

Genetic pathway analysis is performed by comparing the phenotype of a double mutant with that of the single mutants. To properly interpret the results, one must determine the type of pathway being analyzed: biosynthetic (biochemical or morphogenetic) or regulatory. The genes

that we included in our analysis; lectins, proteases and lysozymes, do not seem to be involved in a biosynthetic or a substrate dependent metabolic pathway. In addition the major immunity pathway components (*pmk-1*, a kinase and *daf-16*, a transcriptional regulator) belong to known signal transduction pathways that regulate *C. elegans*-bacterial interactions. Thus, we assume that the pathway we are studying is a regulatory pathway. In a double mutant analysis, simple rules can be used to position the epistatic mutation to be in upstream or downstream gene of a regulatory pathway without detailed knowledge of the nature of mutations, the pathway or the molecular mechanism of regulation (Avery and Wasserman 1992). It is not always necessary that the protein products encoded by the two genes in the same pathway directly interact.

The phenotype that we have used to order gene functions within pathways is the average lifespan of nematodes in response to bacteria. Single mutants that are used in gene interaction analysis of a regulatory pathway will have either same phenotype or directly opposite phenotype (Ewbank 2006). For our analysis, we have used mutants that have average lifespan that are either longer or shorter than that of wild-type as opposite phenotypes (Figure 3.3). In a regulatory pathway, there are three possible outcomes of how two genes can be ordered when their null alleles are combined, depending upon the phenotype of double mutant (Wang and Sherwood 2011). If the double mutant phenotype is similar to one of the two mutants, the genes are interpreted to be in same pathway with the gene whose mutant phenotype is upheld (epistatic) being downstream of the gene whose mutant phenotype is masked (hypostatic). If the phenotype of double mutant is a sum of that of the single mutants, then the genes may not have a genetic interaction (additive). If the phenotype of double mutant is more than the sum of single mutants, the genes might be acting in parallel pathways that converge in a common function (synergistic) (Ewbank 2006; Wang and Sherwood 2011). For example, if in response to a bacterial environment a mutant of gene X (*x*) extends lifespan and mutant of gene Y (*y*) shortens lifespan significantly different from wild-type, gene X is considered to be a negative regulator of host lifespan, and Y, a positive regulator of host lifespan (Figure 3.3). If the double mutant *xy* has the same phenotype as y, the genes X and Y are in same pathway where gene Y is downstream of gene X. Thus, in response to this bacterial environment, negative regulation of host lifespan by gene X is mediated through gene Y or in other words, host lifespan is modulated by the gene X by negatively regulating gene Y which is a positive regulator (Figure 3.3).

From our gene expression studies and subsequent functional analysis using loss-of-function mutants, we have found that the *lys*, *lec* and *protease* genes are differentially regulated in response to grassland soil bacteria and have mutants with significant lifespan effects when exposed to these bacteria (Table 3.1, Figure 3.1). In addition, because our work as well as studies from other labs has shown that *lys*, *lec* and *protease* genes could be involved in *C. elegans*-bacterial interactions and many are regulated by one or more major immune signaling pathways, we decided to focus on them for our genetic interaction studies. Thus, double loss-of-function mutants were constructed between all combinations of the following: *lys* genes *lys4* and *lys-10*, *lec*tin genes *clec-50*, *lec-8* and *lec-9*, and protease-related genes *cpi-1* and *cpr-5*. Furthermore, we have shown that p38 MAPK pathway and *daf-2* insulin signaling pathway are also involved in regulating lifespan in all our bacterial environments (Figure 3.2). Thus in order to understand whether any of the above genes interact with the p38 MAPK or DAF-2 pathways; of *pmk-1/* p38 MAPK mutations and *daf-16/FOXO* transcription factor mutations were also included in the double mutant analyses. However, we were not able not generate some double mutants among the set either because they were closely linked (*cpi-1;pmk-1*, *lys-10;pmk-1*, *lys-4;lys-10*, *lec-9;lec-8*) or were unable to be recovered in F2 generation, likely due to lethality of the double mutant (*lec-9:clec-50*, Table 3.2).

In our analysis of gene interaction using double mutants, we have only analyzed the double mutant lifespan of those single mutants that have lifespans that differ significantly from wild-type in each bacterial environment. We have not assigned epistatic relationships to combinations of mutants if their survivorship showed only subtle differences. Due to the complex and stochastic nature of the phenotype, we have assigned mutants to be in a pathway only if the individual mutants with opposing phenotypes were significantly different from each other and from wild-type ($p<0.05$) and the double mutant should also be significantly different from wild-type, but not from the epistatic gene.

Genetic regulation of survivorship in response to *E. coli*

Mutations in *lec-9*, *cpr-5* and *cpi-1* caused increased survivorship, while mutations in *lys-4*, *lys-10*, *clec-50*, *daf-16* and *pmk-1* caused decreased survivorship in response to *E. coli* (Table 3.3). Since the survivorship of *lec-8* mutants was not significantly different from wild-type, it might not be an important component in lifespan regulation in *E. coli* and was excluded from

double mutant analyses on *E. coli*. Analysis of epistatic relations among genes with opposing phenotypes, revealed that the pathways that regulate responses to *E. coli* could have at least two branches which act more or less in parallel (Figure 3.4a). Specifically, we found that the increased lifespan of *lec-9* mutants was suppressed by *lys-4*, *pmk-1* and *daf-16* in each of the double mutants. Furthermore, survivorship of each of the double mutants was significantly different from that of wild-type and *lec-9,* but not from *lys-4*, *daf-16* and *pmk-1* respectively. However, a mutation in *lys-10* was unable to suppress the increased survivorship of *lec-9* and the survivorship of the *lys-10; lec-9* double mutant was not significantly different from wild-type and only marginally significant from *lys-10* (p=0.0572), indicating the interaction could be additive and no relationship could be inferred between *lec-9* and *lys-10*. The daf-16 mutation completely suppressed the increased survivorship of both cpr-5 and cpi-1 mutants, indicating that daf-16 functions downstream of both cpi-1 and cpr-5. The survivorship of double mutant cpr-5; pmk-1 animals was significantly shorter than that of pmk-1 mutants. So, although the extended lifespan of cpr-5 is completely dependent on activity of pmk-1, pmk-1 cannot be considered to be epistatic to cpr-5. Finally, the decreased survivorship of clec-50 mutants is completely reversed by cpr-5 in cpr-5; clec-50 double mutants, thus cpr-5 is epistatic to clec-50 and they lie in the same pathway.

lec-9, cpr-5 and cpi-1 likely function as components of parallel branches in a survivorship regulation pathway

The survivorships of *cpr-5* and *lec-9* mutants on *E. coli* was not significantly different from each other and the survivorship of the double mutant, *cpr-5;lec-9* was intermediate to the single mutants. Even though the survivorship of the double mutant was different from *lec-9*, it was only marginally significant (p=0.0301), so their effect may not be epistatic. Furthermore, the difference is less than a day, thus it is difficult to conclude that there is a genetic interaction and more investigation is needed. The survivorship of *cpi-1:lec-9* double mutants, was significantly different from *lec-9* and similar to *cpi-1* (p=0.0843), suggesting that *cpi-1* is epistatic. However, although the survivorship of the *cpi-1* mutant was different from wild-type, the survivorship of the *cpi-1;lec-9* double mutant was not, thus there is not enough evidence to conclude that *cpi-1* is epistatic to *lec-9*. Interestingly, although both *cpi-1* and *cpr-5* mutants have increased survivorship, it was decreased in the *cpi-1;cpr-5* double mutant, demonstrating a synergistic effect, which suggests parallel pathways. However, because these lectin and protease regulated

pathways share an important downstream transcription factor, *daf-16*, a known major regulator of worm lifespan, they cannot be termed strictly parallel and the survivorship effects may not be expected to exhibit clear synergy.

We observed that survivorship of the *lys-4: clec-50* double mutant was significantly shorter than the wild-type or either single mutant, indicating a synergistic interaction that suggests they act in parallel. Since *clec-50* functions upstream of *cpr-5* and *lys-4* functions downstream of *lec-9,* this further suggests that *cpr-5* and *lec-9* pathways function in parallel to each other. Since the survivorships of *cpr-5; lys-4* and *cpi-1; lys-4* double mutants are intermediate to that of their respective single mutants, it appears that they are not in the same pathway. Survivorship of *lys-4; daf-16* and *lys-4; pmk*-1 double mutants, were not significantly different from their respective single mutants. Since each of these single mutants reduced survivorship and but did not show synergistic effects when combined, further investigation is needed to identify whether they function in the same or different pathways.

Survivorship of the *clec-50; pmk-1* double mutant was significantly less than each single mutant, indicating a synergistic interaction that suggest they act in parallel pathways. Survivorship of the *cpi-1;clec-50* was shorter than that of the *cpi-1* mutant and similar to that of the *clec-50* mutant. However, *clec-50* cannot be considered epistatic to *cpi-1* since survivorship of the double mutant was significantly different that of *clec-50*. Finally, survivorship of the *clec-50; daf-16* double mutant was not significantly different from either single mutant, indicating no interaction. However, since, *clec-50* appears to function upstream of *cpr-5* and *daf-16* functions downstream, suggesting that that they could function in the same pathway.

Finally, survivorship of *clec-50; lys-10*, *cpr-5; lys-10* and *cpi-1;lys-10* double mutants were not significantly different either from their single mutants or from wild-type, indicating that they do not interact. Although survivorship of the *lys10; daf-16* double mutant, is significantly lower than the *daf-16* mutant, *lys-10* is not be considered as epistatic since they both reduce survivorship and also because *lys-10* was not found to be epistatic to any of the pathway components upstream of *daf-16* (*cpr-5*, *lec-9* or *cpi-1*).

Genetic regulation of survivorship response to *Pseudomonas sp.*

Mutations in *lec-8*, *lec-9* and *cpr-5* extended the survivorship while those in *clec-50*, *cpi-1* , *daf-16* and *pmk-1* reduced survivorship in response to *Pseudomonas sp.* (Table 3.4).

Survivorships of *lys-4* and *lys-10* mutants were not significantly different from wild-type, suggesting they may not have significant role in lifespan regulation and were excluded from the double mutant analysis in response to *Pseudomonas sp*. Similar analysis of double mutants in response to *E. coli,* using mutants with opposing phenotypes allowed us to determine that the above genes appear to function parallel to each other in a branched pathway to regulate *C. elegans* lifespan in response to *Pseudomonas sp.* (Figure 3.4b). The extended survivorship of *lec-9* was suppressed by *cpi-1* and the survivorship of the *lec-9; cpi-1* double mutant was significantly different from *lec-9* and wild-type suggesting that *cpi-1* is epistatic to *lec-9* in response to *Pseudomonas sp*. Although the *daf-16* mutation suppressed the extended lifespan of *lec-9* mutant to some extent, their double mutant lifespan was not significantly different from both the single mutants as well as wild-type. So, the *daf-16* and *lec-9* interaction may be additive and no genetic relation could be inferred among them in response to *Pseudomonas sp*. Lifespan of *lec-9;pmk-1* is significantly lower than that of *pmk-,1* showing a similar effect as with *cpr-5;pmk-1* in response to *E. coli*. Thus in response to *Pseudomonas sp*., *pmk-1* cannot be considered to be epistatic to *lec-9*.

The extended survivorship of *cpr-5* and *lec-8* was completely suppressed by *pmk-1* suggesting that *pmk-1* functions downstream of *cpr-5* and *lec-8*. Mutation of *daf-16* also suppressed the extended survivorship of *cpr-5* and *lec-8*. However, survivorships of double mutants with *daf-16* are not significantly different from wild-type. However, the survivorship of the double mutants were not additive as compared to that of the single mutants, thus their genetic interaction appears to be epistatic in nature.

lec-9, cpr-5 and lec-8 likely function as components of parallel branches in a survivorship regulation pathway

The lifespan of the *cpr-5; lec-9* double mutant was not significantly different from either of the single mutants. Their interaction was neither found to be synergistic, nor epistatic. Mutations in *cpi-1*not only suppressed the extended lifespan of *lec-8* and *cpr-5*, but significantly reduced the survivorship of the double mutants below that of the *cpi-1* single mutant. Although the negative regulation of survivorship by *lec-8* and *cpr-5* mutations was dependent on *cpi-1* activity, their genetic interaction was not strictly epistatic. Mutation of *lec-8* could suppress extended longevity of *cpr-5*. However, survivorship of the *cpr-5;lec-8* double mutant was not significantly different from wild-type. Since they do not show synergistic effect or a strict

epistatic effect, they might belong to discrete pathways that negatively regulate *pmk-1* and *daf-16*. Further investigation is needed to find their exact genetic relationship as to whether or not they interact. Although the survivorship of the *cpi-1;daf-16* double mutant was lower than wild-type, it was not significantly less than either those of the single mutants. Thus they are not synergistic and their genetic interaction could not be determined.

Although, the *clec-50* mutation suppressed the extended survivorship of *lec-8* and *cpr-5*, survivorship of the *lec-8; clec50* double mutant was not significantly different from wild-type, indicating an additive effect. Similarly, survivorship of the *cpr-5; clec-50* double mutant was significantly different from both single mutants, again, indicating an additive effect. Lifespans of remaining three double mutants with *clec-50*, *cpi-1; clec50*, *clec50; daf-16* and *clec50; pmk-1* were not significantly different from their single mutants. Thus, in response to *Pseudomonas sp*. the interaction of *clec-50* with other genes could not be determined.

Genetic regulation of survivorship response to *B. megaterium*

In response to *B. megaterium*, mutations in *lec-9* and *cpr-5* extended survivorship whereas *lys-4*, *clec-50*, *cpi-1*, *daf-16* and *pmk-1* reduced it (Table 3.5). Since survivorship of *lys-10 lec-8 and cpi-1* mutants did not differ significantly from wild-type, they were not included in the gene interaction analysis. Similar to *E. coli* and *Pseudomonas sp*. responses, survivorship appears to be regulated by a pathway where *lec-9* and *cpr-5* are in separate branches that exhibit cross talk between them (Figure 3.4c). In response to *B. megaterium*, mutations in *lys-4* completely suppressed extended survivorship of *lec-9* and *cpr-5* indicating that *lys-4* functions downstream of *lec-9* and *cpr-5*. Similarly mutations in *daf-16* also suppressed the extended survivorship due to the *cpr-5* mutation. Interestingly, a mutation in *cpr-5* completely reversed the reduced survivorship due to the *clec-50* mutation, indicating that *cpr-5* might function downstream of *clec-50*. Although mutations in *clec-50* and *daf-16* reduced survivorship, the survivorship of the *clec-50;daf-16* double mutant was not significantly different from *daf-16*. This supports the observation that *daf-16* could be in the pathway downstream of *cpr-5,* which in turn is downstream of *clec-50*. Mutations in *daf-16* and *pmk-1* could also suppress the extended longevity of *lec-9*. However, their effects were additive, since the survivorships of the double mutants were significantly different from respective single mutants. The interaction between *cpr-*

5 and *pmk-1* could not be determined since the survivorship of the *cpr-5; pmk-1* double mutant showed an additive effect.

lec-9, cpr-5 likely function as components of parallel branches in a survivorship regulation pathway

Although mutations in *cpr-5* and *lec-9* extended survivorship, the survivorship of the double mutant was less than wild-type. Although mutations in *lys-4* and *clec-50* showed reduced survivorship, the survivorship of the *lys-4; clec-50* double mutant, was significantly higher than either single mutant, indicating a synthetic effect. Similarly, survivorship of the *lys-4; daf-16* double mutant showed an additive genetic effect from the single mutants. That we did not detect synergy among the genes regulated by *lec-9* and *cpr-5* might be because *lys-4* appears to act downstream of both these genes and therefore they are not strictly parallel to each other.

Although the survivorship of the *lys-4; pmk-1* double mutant, was not significantly different from *lys-4*, the difference between their survivorships was less than a day. Moreover, since single mutants do not have opposite phenotypes, further investigation is needed to correctly assign their role in genetic pathways. Survivorship of the *clec-50; pmk-1* double mutant, *was* significantly higher than both single mutants, showing a synthetic phenotype. Thus in the *C. elegans* response to *B. megaterium*, unlike its response to *E. coli* and *Pseudomonas sp*., our analysis could not assign *pmk-1* to either *cpr-5*- or *lec-9*-regulated branches of the pathway.

Discussion

Biotic interactions with bacteria have been shown to dominate *C. elegans* ecology, where different bacteria can be perceived by the worm as both sources of food or potential pathogens (Felix and Braendle 2010; Felix and Duveau 2012). We chose to focus on genes most relevant to these interactions; those involved either *C. elegans* defense or metabolism functions in response to different bacterial environments. In order to get a better understanding of how these functions are integrated at the level of the whole organism, we assayed the effects of these interactions on survivorship, a phenotype that has been demonstrated to be a good reflection of the consequences of various bacterial interactions (Darby 2005). Survivorship is a complex trait measured in response to bacteria that is affected by many different genes acting in evolutionary

conserved pathways (Irazoqui and Ausubel 2010; Irazoqui, Urbach et al. 2010). We have found that the survivorship of wild-type *C. elegans* varied in response to grassland soil bacteria. From our previous transcriptomic study (Coolon, Jones et al. 2009) and current study involving loss of function of single genes in response to grassland soil bacteria, we saw bacterium-specific and bacterium-shared host responses in gene expression and function. For this study we focused on *lys*, *lec* and *protease* genes and the components of major evolutionarily conserved immune signaling pathways, functions of all of which have been previously implicated in bacterial responses. To further understand the functional basis of the contribution of gene interactions to the specificity of responses to bacterial environments, we generated double mutants and tested them in three bacterial environments*, E. coli, Pseudomonas sp.,* and *B. megetarium*. Based on these combined studies, our model of *C. elegans* gene interactions in response to grassland soil bacteria is depicted in Figure 3.4.

Conserved interaction patterns across environments

The interactions we observed were complex, involving components that function in parallel branches converging on common elements. We observed several interactions that were conserved across bacterial environments. Specifically, DAF-16 functioned downstream of CPR-5 in response to each bacterium tested and PMK-1 functioned downstream of CPR-5 in response to *Pseudomonas sp.* In addition, we observed that a lectin, either LEC-9 or LEC-8 functions upstream of either DAF-16 or PMK-1 or both in multiple bacterial environments. We also observed that the gene interaction cassettes *clec-50*-|*cpr-5*-|*daf-16* and *lec9*-|*lys-4* are conserved in response to *E. coli* and *B. megaterium*. Conserved genetic interactions to multiple bacteria could be due to similar bacterial challenges being dealt with by common genetic mechanisms. Similar signaling mechanisms might serve the purpose of metabolizing bacteria while it is perceived as food source as well as in defense responses when it triggers a defense response. Since we have found many genes from our transcriptomic study as well as components of known innate immune response pathways showed similar survivorships in response to multiple bacteria, it is not entirely surprising to find their interaction components conserved in response to multiple bacteria. In the evolution of complex genetic systems, functional epistasis has been suggested to be a likely outcome (Phillips 2008). This could represent an example where the signal transduction pathways involving components of p38 MAPK and DAF-16 pathways have been

used as a backbone to which new effectors involved in bacterial responses might be attached, as suggested by Phillips (Phillips 2008).

Environment-specific Gene interactions

Since no single pathway is activated by any given bacterium, Alper et al (2007) suggested that *C. elegans* can distinguish between different bacteria and the pathways that are activated can interact to turn on appropriate responses (Alper, McBride et al. 2007). Although there are conserved gene interactions across the multiple bacterial environments we also find that epistatic rewiring causing altered interaction components in response to a bacterium can provide additional specificity. Whereas *lec-9* regulates *pmk-1*, *daf-16* and *lys-4* in response to *E. coli*, it only regulates *lys-4* in response to *B. megaterium*. In response to *Pseudomonas sp*., instead of any of the above components *lec-9* regulates *cpi-1*. Similarly *lys-4* is negatively regulated by *lec-9* in response to *E. coli*, but both by *lec-9* and *cpr-5* in response to *B. megaterium*. This indicates that bacteria might be recognized differently and this could lead to bacterium-specific responses via differential gene interactions. We also found that *cpr-5* regulated both *pmk-1* and *daf-16* in response to *Pseudomonas sp*, but only *daf-16* in response *E. coli* and *B. megaterium*. Thus there is a shift in upstream regulators for major pathways such as p38 MAPK and DAF-16 pathways in response to specific bacteria. Similarly it appears that both of these pathways are regulated by more than one gene (*cpr-5* and *lec-8*) in response to *Pseudomonas sp*., but only *daf-16* is regulated by more than one gene (*cpr-5*, *cpi-1* and *lec-9*) in response to *E. coli*.

Correlation of gene expression and gene function

Among the *lys*, *lec* and *protease-related* genes in our model, we observed that the positive regulators of survivorship generally displayed increased expression and the negative regulators generally displayed decreased expression in the respective environments (Table 3.1). The only exception to this trend was increased expression of *cpr-5* in response to *B. megaterium* where it was found to function to negatively regulate survivorship. Survivorship may not reflect the entirety of functional responses in this case because when we examined a more comprehensive life history trait, fitness, we found that the *cpr-5* mutant was less fit on *B. megaterium* (see Chapter 2). In addition, there was at least one instance where the expression of *lys-4* was decreased in *E. coli* (where it was positive regulator of lifespan) as compared to *B. megaterium* and the expression of *cpi-1* was increased on *E. coli* (where it was negative

regulator) as compared to *B. megaterium*. In these cases, we cannot rule out the possibility that *lys-4* could be up regulated both in response to *E. coli* and *B. megaterium*, but is expressed at a much higher level in *B. megaterium* than *E. coli*. Similarly, *cpi-1* could be down regulated in response to both *E. coli* and *B. megaterium*, but much less so in *B. megaterium* as compared to *E. coli*.

We have seen that the survivorship pathway in response to bacteria has examples of components that interact in an environmentally specific manner. If these multi-locus associations can confer differential fitness relative to the environments, they can contribute to the non-additive variance component in quantitative genetics parlance. This could eventually mean that host phenotypic differentiation could be explained based on non-additive variance components. Epistatic relationships due to such nonrandom associations of alleles are the basis of linkage disequilibrium. We have also seen that regulatory specificity of genes change in response to specific environments, for example *lec-9* interacts with *lys-4* in *B. megaterium* but with *cpi-1* in *Pseudomonas sp* and with *lys-4*, *pmk-1* and *daf-16* in *E. coli*. Such context specific regulation will affect the selective regime experienced by each allele and dampen the predictability of allele frequency as a response to selection (Fenster, Galloway et al. 1997). Other than in a few cases, we have also seen that the regulation of host transcription is adjusted to generate bacteria-specific interactions among various effectors that in turn regulate major signal transduction pathways. This suggests that there is a higher level of organization in which the environment dependent response is regulated at the transcriptional level, making the components and mechanism more complex and should be considered in further evolutionary analysis.

To our knowledge this is first demonstration that components of major evolutionary conserved innate immune pathways, p38 MAPK and DAF-16, might be under the control of a Lectin (*lec-9*) in response to *E. coli* OP50. Similarly, both pathways are negatively regulated by another Lectin (*lec-8*) in response to *Pseudomonas sp.* Since lectins function as pattern recognition receptors, it is possible that direct bacterial recognition at molecular level might trigger downstream responses. Troemel et al suggested that p38 MAPK and DAF-16 pathways might act in parallel in *C. elegans* innate immunity (Troemel, Chu et al. 2006). Although the genes up-regulated by these two pathways are distinct, there is considerable overlap between genes that are up-regulated by *pmk-1*, but down-regulated by *daf-16* (Troemel, Chu et al. 2006). Thus in

response to *E. coli*, the *pmk-1; cpr-5* double mutant might affect the same downstream targets to cause reduction in survivorship greater than that in *pmk-1*single mutants (Table 3.3).

The survivorship phenotype could be affected by bacterial environment both by its quality as well as quantity. *C. elegans* lifespan has been found to be extended with calorically restricted diets (Lakowski and Hekimi 1998). We found that wild-type *C. elegans* fed *B megaterium* have slower development and longer age to reproductive maturity (Table 2.11, see Chapter 2) suggesting that the extended survivorship in response to this bacterium could be due to caloric restriction. Thus the change in survivorship for wild-type as well as mutants could be due to a combined effect of food quantity, metabolic and/or pathogenic effects of the bacteria and we may not be able to tease apart their specific effects on survivorship. Lifespan extension due to different regimens of dietary restriction in *C. elegans* is controlled by different nutrient sensors and transcription factors which include DAF-16 (Kenyon 2010). Out of the two major immunity pathways (PMK-1 and DAF-16) tested, we indeed find DAF-16 to be involved in survivorship response pathway in response to *B megaterium*.

In a recent study involving natural *C. elegans* associated microbiota, Sirena et al (2013) found *Bacillus megaterium* and *Pseudomonas mendocina* as representative species belonging to two main genera (Montalvo-Katz, Huang et al. 2013). This environmental isolate is slightly different from our *B. megaterium* strain since it did not increase survivorship of wild-type *C. elegans* as compared to *E. coli* OP50. Interestingly, *C. elegans* exposed to both of the above bacteria provided better protection to subsequent challenge by pathogenic *Pseudomonas aeruginosa*. Since *P. mendocina* fed animals up-regulated *pmk*-regulated genes and a *pmk-1* mutant abolished this enhanced resistance to subsequent infection by *P. aeruginosa*, the p38 MAPK pathway was suggested to be involved in this interaction (Montalvo-Katz, Huang et al. 2013). We have also found that exposing to our *B. megaterium* and *Pseudomonas sp*. isolates, *lectins* that negatively regulate *pmk-1* and *daf-16* in these environments are reduced in expression. Since *C. elegans* is likely to confront a heterogeneous mix of bacteria in its natural setting, priming of such major pathways could be advantageous especially if these bacteria are relatively innocuous. Since we have found that lectins and proteases are modulators of above immune signaling pathways, acting as connecting links between these two environment responses, our study sheds light onto mechanisms of how this might occur.

C. elegans innate immune signaling in response to bacteria, especially against *P. aeruginosa,* has been shown to pass through multiple evolutionarily conserved pathways to finally affect downstream responses (Irazoqui and Ausubel 2010; Irazoqui, Urbach et al. 2010). Thus it is hard to imagine that all the downstream effectors have specific interactions with specific bacterial pathogens. As revealed by our study, a more realistic view would be that some effectors interact with each other and with conserved components to provide additional specificity that lead to an optimal response. Although we began by focusing on defense/metabolism related genes, these genes and pathways cannot be regarded strictly affecting only those functions. In fact it is likely they could be involved in different molecular or cellular process but eventually affecting the organismal level trait, survivorship. In addition, multiple pathways, including the p38 MAPK and DAF-16 pathways, have been found to control the expression of lectin, lysozymes and protease genes, which were then considered to be effectors of these signaling pathways (Troemel, Chu et al. 2006; Alper, McBride et al. 2007). However, since we found that *lec, lys* and *protease* genes were involved in interacting with components of these pathways, it raises the possibility that feedback regulation is occurring. Nicholas and Hodgkin (2004), suggested that up-regulation of a C-type lectin might increase production of its own protein by feedback regulation once it recognizes the pathogen (Nicholas and Hodgkin 2004). However the specific *lys*, *lec* and *protease* genes involved in our study have not been reported to be regulated by these pathways and there could be other transcriptional factors regulating their expression as well.

Multiple regulatory pathways might provide more efficient and fine-tuned responses. This might also provide flexibility for specific environmental responses by better adjusting to the stages of infection and severity of bacterial challenge. Since we used a system-level trait to uncover gene interactions and the candidate gene set was not specific to tissues or age of the host, the scope of our study remained on the whole organism level. So we might have missed any interactions that could be revealed by studying the phenotype sensitive to such gene functions. Nevertheless, we believe that our interaction model could provide a framework for future studies that might address questions relating to their role specific to conditions, tissues etc. It would also be interesting to study whether there is age-specific or tissue-specific rewiring of genetic interactions in response to various environments. A complex trait that is affected by multiple alleles will be influenced by genetic background. What would be most important allelic

interaction that governs a phenotype in a particular environment? It would be interesting to study how the gene interaction specificities might change with naturally occurring alleles among the interacting genes. Epistatic gene interactions have been found to be important component of genetic architecture of quantitative traits in *D. melanogaster* (Huang, Richards et al. 2012) and *Arabidopsis thaliana* (Rowe, Hansen et al. 2008). Thus non-additive allelic interactions are an important component of phenotypic variations within these species.

With a large repertoire of effectors likely to be involved in responses to bacterial challenges in *C. elegans*, for example, 15 Lysozymes, 11 Galectins, 265 C-type lectins, (Engelmann and Pujol 2010), there exist the possibility of forming extensive "interaction hubs" of effectors. This could provide and flexible responses especially since evolutionarily conserved pathway components like p38 MAPK and DAF-16 that interact with these effectors have pleiotropic functions and are likely to be highly constrained in the face of natural selection. In our study we found environment specific interaction among components of these pathway and some of the effectors setting the stage for this scenario. Thus structural variations in effector proteins or regulatory variations that alter their expression are important components which can provide selective advantage owing to their interaction specificity in environmental responses. Also, such gene interaction specificity could provide a functional basis for diversifying selection of gene duplications as suggested in the case of *C. elegans* lysozymes (Schulenburg and Boehnisch 2008). Our approach of finding environment-specific changes in gene interactions also give insights in selecting genes for studies that might involve a candidate-gene approach that looks at allelic variants associated with environmental specificity.

Figures and Tables

Table 3.1 Expression of candidate defense and/or metabolism genes

Out of 204 genes differentially expressed in the transcriptome analysis by Coolon, Jones et al. 2009, the defense and metabolism based category was enriched. Defense and metabolism functions are further categorized based on putative molecular functions. Genes differentially responding to all six pair-wise differential expression comparisons are tabulated. Genes with opposite patterns of expression in each bacterial comparison is shown in red. E, *E. coli* OP50; P, *Pseudomonas sp.*; B, *Bacillus megaterium* M, *Micrococcus luteus* (bacteria not used for further survivorship assays).

Expression Pattern	**Functional class of genes expressed**			
	Lysozyme	**Lectin**	**Protease**	**Others**
M>P	*lys-2* *lys-4* *lys-10*	*clec-50* *lec-6* *lec-8* *lec-9*	*cpr-5* *asp-6* *F57F5.1* *cpi-1*	*fat-7* *lbp-5*
B>P	*lys-4* *lys-10*	*clec-50* *lec-6* *lec-8*	*cpr-5* *asp-6* *F57F5.1* *cpi-1*	*hex-1* *mtl-1*
E<M	*lys-1* *lys-4* *lys-10*	*lec-6*	*cpi-1*	*fat-7* *lbp-5*
E>P	*lys-4* *lys-10*		*asp-6* *F57F5.1*	
B>E	*lys-4*		*cpr-5* *cpi-1*	*cyp-34A9* *mtl-1*
B<M		*lec-9*		*fat-7* *lbp-5* *mtl-1*

Figure 3.1 Wild-type and mutant survivorship exposed to grassland soil bacteria. Wild-type (*N2*) *C. elegans* and mutant strains were grown in the three bacterial environments and survivorship assays are conducted. Survivorship was determined as average lifespan of individual worms. Standard error is shown as error bars. Out of the 17 mutant strains tested only those with significantly different ($P<0.05$) survivorship compared to *N2* are shown.

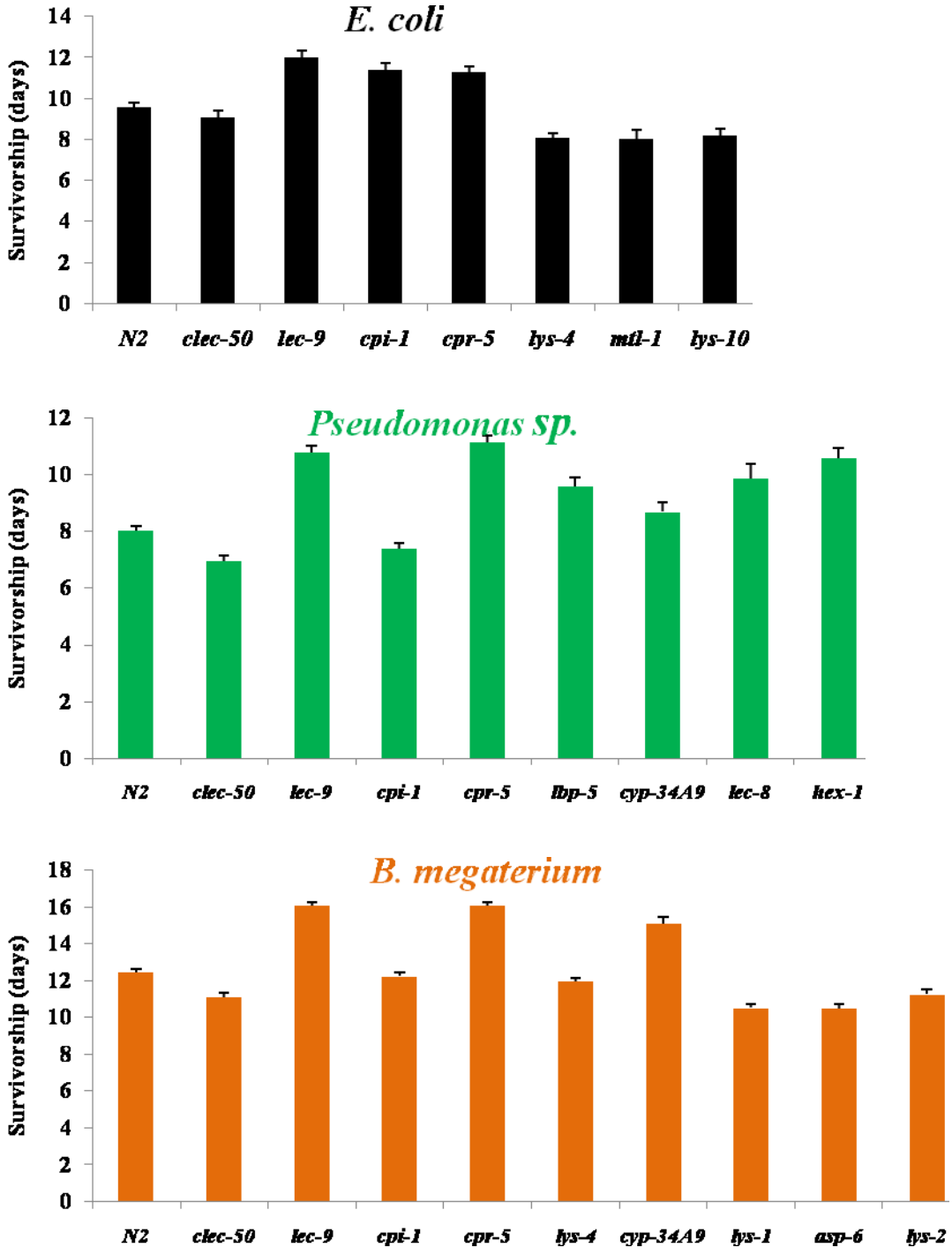


Figure 3.2 Evolutionary conserved innate immunity pathway components modulate *C. elegans* responses to grassland soil bacteria.

Wild-type *C. elegans* and mutant strains were grown in the three bacterial environments and survivorship assays are conducted. Survivorship was determined as average lifespan of individual worms. Standard error is shown as error bars. Mutant strains that showed significantly different survivorship ($P<0.05$) compared to *N2* are shown in asterisk on their histograms.

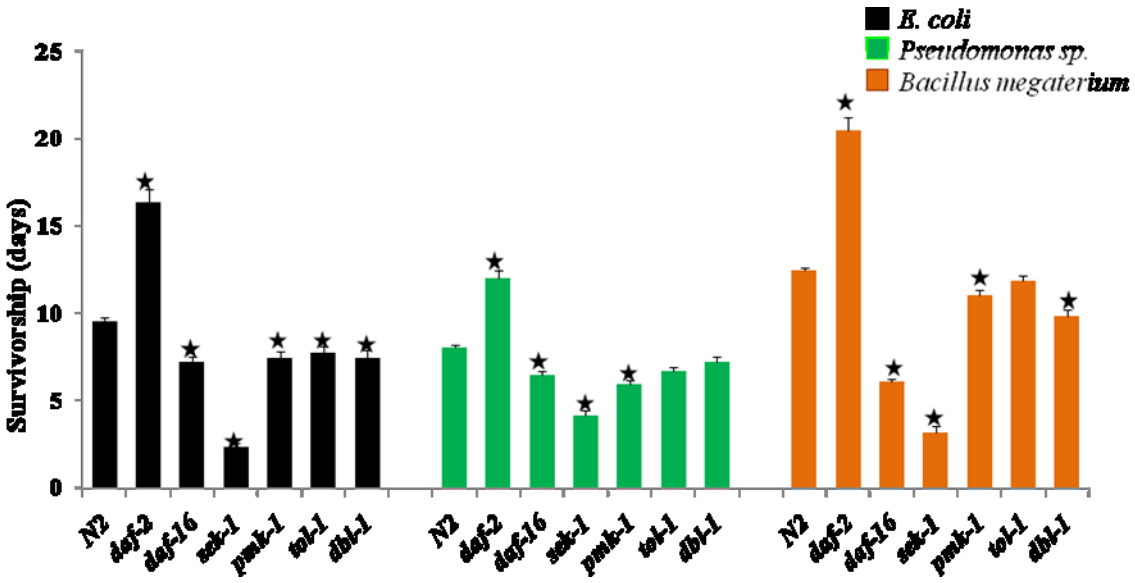


Table 3.2 Component genes chosen for double mutant analysis in soil bacteria.

Genes with putative molecular functions of lysozymes, lectin and proteases forms a major part of defense/metabolism group. Genes selected for double mutant analysis based on their expression and functional characteristics are shown with their respective chromosome number in parenthesis. Doubles were made in all possible pairs except for the genes that belonged close to each other in a chromosome or those which was not able to be recovered as double mutants probably due to lethality.

	Lysozyme		**Lectins**			**Proteases**	
	lys-4 (IV)	*lys-10* (IV)	*clec-50* (V)	*lec-8* (X)	*lec-9* (X)	*cpi-1* (IV)	*cpr-5* (V)
lys-4 (IV)							
lys-10 (IV)	-						
clec-50 (V)	X	X					
lec-8 (X)	X	X	X				
lec-9 (X)	X	X	-	-			
cpi-1 (IV)	X	X	X	X	X		
cpr-5 (V)	X	X	X	X	X	X	
pmk-1 (IV)	X	-	X	X	X	-	X
daf-16 (I)	X	X	X	X	X	X	X

Figure 3.3 Model for epistasis analysis in soil bacteria

Single mutants that showed opposing (higher or lower survivorships significantly different (P<0.05) from Wild-type) phenotypes were picked for epistasis analysis. For example, *x,* mutant allele of gene X causes increased survivorship in bacteria (negative regulator) and *y*, mutant allele of gene Y has opposing phenotype, reduced survivorship (positive regulator). If survivorship of the double mutant *xy*, is same as *y,* there is an epistatic relation among them X and Y where Y is downstream of X. An underlying assumption is that the two genes regulating lifespan could be in a regulatory pathway. Also, direct interaction among the genes may or may not occur.

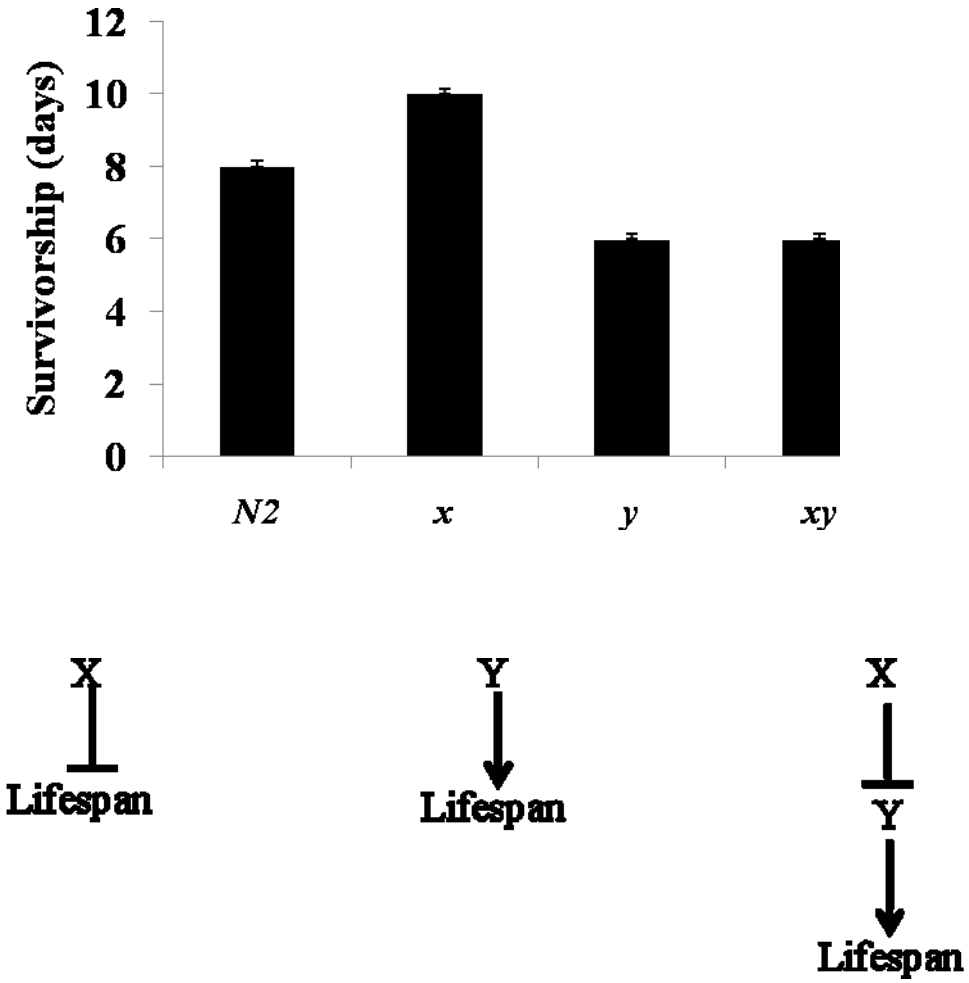


Table 3.3 Gene interactions that affect *C. elegans* lifespan in response to *E. coli*

Wild-type and mutant strains were grown in *E. coli* and survivorship (days) is determined. Survivorship of wild-type and single mutants are shown in first column/rows. Survivorships of the double mutants are shown on cells that correspond to single mutants. Survivorships significantly different (P<0.05) from wild-type is shown as yellow shades. Alphabet, 'a' inside the parenthesis represents those survivorships significantly different from corresponding single mutants in first row and 'b' represents those significantly different from corresponding mutants in first column.

	wild-type	***lys-4***	***lys-10***	***clec-50***	***lec-9***	***cpi-1***	***cpr-5***	***daf-16***	***pmk-1***
wild - type	9.58	8.10	8.21	9.08	12.01	11.41	11.31	8.38	7.49
lys-4	8.10			6.61 (a, b)	8.67 (a)	9.52 (a,b)	10.23 (a,b)	8.03	6.23
lys-10	8.21			9.3	11.7	10.03	9.62	6.30 (a)	
clec-50	9.08	(a)				8.32 (a,b)	11.94 (b)	8.93	6.29 (a, b)
lec-9	12.01	(a)	(a)			10.40 (b)	11.79 (b)	7.20 (b)	8.46 (b)
cpi-1	11.41	(a)	(a)	(a)	(a)		7.39 (a,b)	8.50 (b)	
cpr-5	11.31	(a)	(a)	(a)				7.97 (b)	5.95 (a,b)
daf-16	8.38				(a)	(a)	(a)		
pmk-1	7.49				(a)	(a)	(a)		

Table 3.4 Gene interactions that affect *C. elegans* lifespan in response to *Pseudomonas sp.*
Wild-type and mutant strains were grown in *Pseudomonas sp.* and survivorship (days) is determined. Survivorship of wild-type and single mutants are shown in first column/rows. Survivorships of the double mutants are shown on cells that correspond to single mutants. Survivorships significantly different ($P<0.05$) from wild-type is shown as yellow shades. Alphabet, 'a' inside the parenthesis represents those survivorships significantly different from corresponding single mutants in first row and 'b' represents those significantly different from corresponding mutants in first column.

	wild-type	***lec-8***	***clec-50***	***lec-9***	***cpi-1***	***cpr-5***	***daf-16***	***pmk-1***
wild-type	8.05	9.89	6.98	10.80	7.42	11.15	7.33	5.94
lec-8	9.89		7.75 (b)		6.19 (a,b)	8.81 (a)	6.67 (b)	5.90 (b)
clec-50	6.98	(a)			7.48	10.64 (a,b)	6.93	6.32
lec-9	10.80				7.39 (b)	10.78	7.96	4.98 (a,b)
cpi-1	7.42	(a)		(a)		5.37 (a,b)	6.58	
cpr-5	11.15	(a)	(a)		(a)		6.80 (b)	6.33 (b)
daf-16	7.33	(a)		(a)		(a)		
pmk-1	5.94	(a)		(a)		(a)		

Table 3.5 Gene interactions that affect *C. elegans* lifespan in response to *Bacillus megaterium*

Wild-type and mutant strains were grown in *Bacillus megaterium* and survivorship (days) is determined. Survivorship of wild-type and single mutants are shown in first column/rows. Survivorships of the double mutants are shown on cells that correspond to single mutants. Survivorships significantly different ($P<0.05$) from wild-type is shown as yellow shades. Alphabet, 'a' inside the parenthesis represents those survivorships significantly different from corresponding single mutants in first row and 'b' represents those significantly different from corresponding mutants in first column.

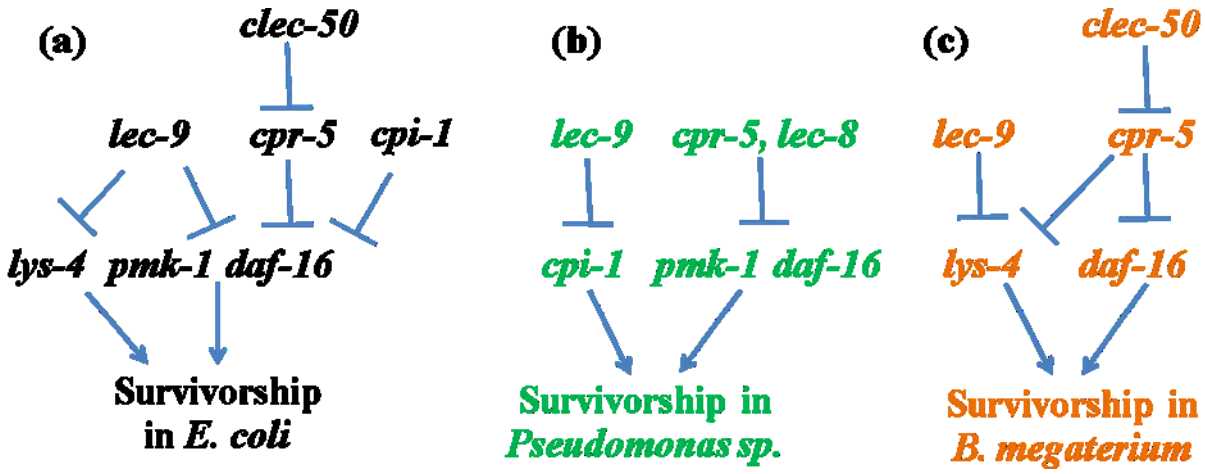
	wild-type	***lys-4***	***clec-50***	***lec-9***	***cpr-5***	***daf-16***	***pmk-1***
wild-type	12.48	11.95	11.11	16.09	16.08	8.19	11.07
lys-4	11.95		12.69 (a,b)	11.65 (a)	11.59 (a)	9.9 (a,b)	11.34 (a)
clec-50	11.11	(a)			16.61 (b)	7.86 (b)	12.41 (a,b)
lec-9	16.09	(a)			11.02 (a,b)	11.59 (a,b)	12.15 (a,b)
cpr-5	16.08	(a)	(a)			8.29 (b)	11.41 (a,b)
daf-16	8.19	(a)	(a)	(a)	(a)		
pmk-1	11.07	(a)		(a)		(a)	

Table 3.6 PCR primers used for genes involved in generation of double mutants

Gene	Forward primer	Reverse primer
lys-4	GGGACCATTAGCTTGTAGAA	CTAGACAGTCAGAGGGGACA
lys-10	CCATCGTATGTTGCCCACCA	ACCAGCATTGTTTCCGGTAC
clec-50	ACACAACCCGGACACCTTA	AGTTGTTCGCATCCTTTTG
lec-8	TACTCGACCTTAGTCATCGT	TGCGATTGGCATATTGGTAC
lec-9	TCCCTATCCCACCTGTACTA	AACCAGCGGATTCATGGCAT
cpi-1	ATACGGTGTCTATCGCGGAC	AAGAACGTAGCGCGAGTGAT
cpr-5	TTGTGACACCCCGAAATTCT	GGTTTTTCACCTCGAATGGA
daf-16	CAAGACAGGCGGTATCCAAT	GAGCCCATCAATGCTCTCTC
N2 (daf-16)	CAAGACAGGCGGTATCCAAT	AAGCCATTTGTCGTGGAAAC
pmk-1	GTTGCCATGACCTCAGAGCCTC	ATGTGGTCATCGTTGAGTCGCTG

Figure 3.4 Genetic architecture underlying *C. elegans* - bacteria interaction

From our gene interaction analysis in three bacteria, we propose a model of plasticity in gene interactions that could in turn confer specificity in bacterial response. Although we find conserved components and interactions involved in *C. elegans* - bacterial responses, their regulation is found to have different interaction partners and altered interaction specificity. Interaction components, *lec-9*, *cpr-5* and *daf-16* are found to be common in responses to all bacteria. Also, *cpr-5* negative regulating *daf-16* is a cassette that is common against all bacteria. Similarly, for *E. coli* and *B. megaterium clec-50* ⊣ *cpr-5* ⊣ *daf-16* and *lec-9* ⊣ *lys-4* is common and *pmk-1* is common for *E. coli* and *Pseudomonas sp*. However, responding to each bacterium there are unique interactions, like *cpi-1* and *lec-9* negatively regulating *daf-16* in *E. coli*, *lec-9* regulating *cpi-1* and *lec-8* regulating *daf-16* in *Pseudomonas sp*. and *cpr-5* negatively regulating *lys-4* in *B. megaterium.* Conserved components are speculated to be involved in general metabolism/defense functions and specific interactions for bacteria-specific challenges.



References

Agrawal, A. F. and M. C. Whitlock (2010). "Environmental duress and epistasis: how does stress affect the strength of selection on new mutations?" Trends in Ecology & Evolution **25**(8): 450-458.

Agrawal, A. F. and M. C. Whitlock (2012). "Mutation Load: The Fitness of Individuals in Populations Where Deleterious Alleles Are Abundant." Annual Review of Ecology, Evolution, and Systematics, Vol 43 **43**: 115-+.

Anderson, J. L., R. M. Reynolds, et al. (2011). "Experimental evolution reveals antagonistic pleiotropy in reproductive timing but not life span in Caenorhabditis elegans." J Gerontol A Biol Sci Med Sci **66**(12): 1300-1308.

Alper, S., S. J. McBride, et al. (2007). "Specificity and complexity of the Caenorhabditis elegans innate immune response." Mol Cell Biol **27**(15): 5544-5553.

Avery, L. and S. Wasserman (1992). "Ordering gene function: the interpretation of epistasis in regulatory hierarchies." Trends Genet **8**(9): 312-316.

Breen, M. S., C. Kemena, et al. (2012). "Epistasis as the primary factor in molecular evolution." Nature **490**(7421): 535-+.

Brenner, S. (1974). "The genetics of Caenorhabditis elegans." Genetics **77**(1): 71-94.

Bulmer, M. (1991). "The Selection-Mutation-Drift Theory of Synonymous Codon Usage." Genetics **129**(3): 897-907.

Burga, A., M. O. Casanueva, et al. (2011). "Predicting mutation outcome from early stochastic variation in genetic interaction partners." Nature **480**(7376): 250-U133.

Carlborg, O. and C. S. Haley (2004). "Epistasis: too often neglected in complex trait studies?" Nat Rev Genet **5**(8): 618-625.

Casanueva, M. O., A. Burga, et al. (2012). "Fitness Trade-Offs and Environmentally Induced Mutation Buffering in Isogenic C. elegans." Science **335**(6064): 82-85.

Charlesworth, B. and D. Charlesworth (1998). "Some evolutionary consequences of deleterious mutations." Genetica **102-3**: 3-19.

Charlesworth, B. and K. A. Hughes (1996). "Age-specific inbreeding depression and components of genetic variance in relation to the evolution of senescence." Proc Natl Acad Sci U S A **93**(12): 6140-6145.

Chen, J., D. Senturk, et al. (2007). "A demographic analysis of the fitness cost of extended longevity in Caenorhabditis elegans." J Gerontol A Biol Sci Med Sci **62**(2): 126-135.

Cheverud, J. M. and E. J. Routman (1995). "Epistasis and Its Contribution to Genetic Variance-Components." Genetics **139**(3): 1455-1461.

Chiu, H. C., C. J. Marx, et al. (2012). "Epistasis from functional dependence of fitness on underlying traits." Proceedings of the Royal Society B-Biological Sciences **279**(1745): 4156-4164.

Ciosk, R., M. DePalma, et al. (2004). "ATX-2, the C. elegans ortholog of ataxin 2, functions in translational regulation in the germline." Development **131**(19): 4831-4841.

Coolon, J. D., K. L. Jones, et al. (2009). "Caenorhabditis elegans genomic response to soil bacteria predicts environment-specific genetic effects on life history traits." PLoS Genet **5**(6): e1000503.

Coolon, J. D., K. L. Jones, et al. (2013). "Long-term nitrogen amendment alters the diversity and assemblage of soil bacterial communities in tallgrass prairie." Plos One **8**(6): e67884
Cooper, T. F., R. E. Lenski, et al. (2005). "Parasites and mutational load: an experimental test of a pluralistic theory for the evolution of sex." Proc Biol Sci **272**(1560): 311-317.
Costanzo, M., A. Baryshnikova, et al. (2010). "The genetic landscape of a cell." Science **327**(5964): 425-431.
Costello, J. C., M. M. Dalkilic, et al. (2009). "Gene networks in Drosophila melanogaster: integrating experimental data to predict gene function." Genome Biol **10**(9): R97.
Coyne, J. A. (1992). "Genetics and Speciation." Nature **355**(6360): 511-515.
Dam, T. K. and C. F. Brewer (2010). "Lectins as pattern recognition molecules: the effects of epitope density in innate immunity." Glycobiology **20**(3): 270-279.
Darby, C. (2005). "Interactions with microbial pathogens." WormBook: 1-15.
de Visser, J. A. G. M. and S. F. Elena (2007). "The evolution of sex: empirical insights into the roles of epistasis and drift." Nature Reviews Genetics **8**(2): 139-149.
Demeeus, T., Y. Michalakis, et al. (1993). "Polymorphism in Heterogeneous Environments, Evolution of Habitat Selection and Sympatric Speciation - Soft and Hard Selection Models." Evolutionary Ecology **7**(2): 175-198.
Desai, M. M., D. Weissman, et al. (2007). "Evolution can favor antagonistic epistasis." Genetics **177**(2): 1001-1010.
Drake, J. W., B. Charlesworth, et al. (1998). "Rates of spontaneous mutation." Genetics **148**(4): 1667-1686.
Elena, S. F. and J. A. de Visser (2003). "Environmental stress and the effects of mutation." J Biol **2**(2): 12.
Elena, S. F. and R. E. Lenski (1997). "Test of synergistic interactions among deleterious mutations in bacteria." Nature **390**(6658): 395-398.
Elena, S. F. and R. E. Lenski (2003). "Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation." Nature Reviews Genetics **4**(6): 457-469.
Elena, S. F. and R. Sanjuan (2003). "Climb every mountain?" Science **302**(5653): 2074-2075.
Engelmann, I. and N. Pujol (2010). "Innate immunity in C. elegans." Adv Exp Med Biol **708**: 105-121.
Ewbank, J. J. (2006). "Signaling in the immune response." WormBook: 1-12.
Eyre-Walker, A. and P. D. Keightley (2007). "The distribution of fitness effects of new mutations." Nature Reviews Genetics **8**(8): 610-618.
Felix, M. A. and C. Braendle (2010). "The natural history of Caenorhabditis elegans." Curr Biol **20**(22): R965-969.
Felix, M. A. and F. Duveau (2012). "Population dynamics and habitat sharing of natural populations of Caenorhabditis elegans and C. briggsae." BMC Biol **10**: 59.
Fenster, C. B., L. F. Galloway, et al. (1997). "Epistasis and its consequences for the evolution of natural populations." Trends in Ecology & Evolution **12**(7): 282-286.
Flynn, K. M., T. F. Cooper, et al. (2013). "The Environment Affects Epistatic Interactions to Alter the Topology of an Empirical Fitness Landscape." PLoS Genet **9**(4)
Force, A., M. Lynch, et al. (1999). "Preservation of duplicate genes by complementary, degenerative mutations." Genetics **151**(4): 1531-1545.
Fry, J. D. (1996). "The evolution of host specialization: Are trade-offs overrated?" American Naturalist **148**: S84-S107.

Fry, J. D., S. L. Heinsohn, et al. (1996). "The contribution of new mutations to genotype-environment interaction for fitness in Drosophila melanogaster." Evolution **50**(6): 2316-2327.
Garfield, D. A. and G. A. Wray (2010). "The Evolution of Gene Regulatory Interactions." Bioscience **60**(1): 15-23.Greenspan, R. J. (2001). "The flexible genome." Nat Rev Genet **2**(5): 383-387.
Gros, P. A., H. Le Nagard, et al. (2009). "The Evolution of Epistasis and Its Links With Genetic Robustness, Complexity and Drift in a Phenotypic Model of Adaptation." Genetics **182**(1): 277-293.
Gross, J. B., R. Borowsky, et al. (2009). "A Novel Role for Mc1r in the Parallel Evolution of Depigmentation in Independent Populations of the Cavefish Astyanax mexicanus." PLoS Genet **5**(1).
Hahm, J. H., S. Kim, et al. (2011). "GPA-9 is a novel regulator of innate immunity against Escherichia coli foods in adult Caenorhabditis elegans." Aging Cell **10**(2): 208-219.
Haldane, J. B. S. (1937). "The Effect of Variation of Fitness". The American Naturalist , 71 (735): 337-349
Halligan, D. L. and P. D. Keightley (2009). "Spontaneous Mutation Accumulation Studies in Evolutionary Genetics." Annual Review of Ecology Evolution and Systematics **40**: 151-172.
Harrison, R., B. Papp, et al. (2007). "Plasticity of genetic interactions in metabolic networks of yeast." Proc Natl Acad Sci U S A **104**(7): 2307-2312.
Hayden, E. J. and A. Wagner (2012). "Environmental change exposes beneficial epistatic interactions in a catalytic RNA." Proceedings of the Royal Society B-Biological Sciences **279**(1742): 3418-3425.
Hodgkin, J. and T. M. Barnes (1991). "More Is Not Better - Brood Size and Population-Growth in a Self-Fertilizing Nematode." Proceedings of the Royal Society B-Biological Sciences **246**(1315): 19-24.
Hoffmann, A. A. and J. Merila (1999). "Heritable variation and evolution under favourable and unfavourable conditions." Trends in Ecology & Evolution **14**(3): 96-101.
Huang, W., S. Richards, et al. (2012). "Epistasis dominates the genetic architecture of Drosophila quantitative traits." Proc Natl Acad Sci U S A **109**(39): 15553-15559.
Ideo, H., K. Fukushima, et al. (2009). "A Caenorhabditis elegans glycolipid-binding galectin functions in host defense against bacterial infection." J Biol Chem **284**(39): 26493-26501.
Irazoqui, J. E. and F. M. Ausubel (2010). "99th Dahlem conference on infection, inflammation and chronic inflammatory disorders: Caenorhabditis elegans as a model to study tissues involved in host immunity and microbial pathogenesis." Clin Exp Immunol **160**(1): 48-57.
Irazoqui, J. E., E. R. Troemel, et al. (2010). "Distinct pathogenesis and host responses during infection of C. elegans by P. aeruginosa and S. aureus." PLoS Pathog **6**: e1000982.
Irazoqui, J. E., J. M. Urbach, et al. (2010). "Evolution of host innate defence: insights from Caenorhabditis elegans and primitive invertebrates." Nat Rev Immunol **10**(1): 47-58
Jackson, R. B., C. R. Linder, et al. (2002). "Linking molecular insight and ecological research." Trends in Ecology & Evolution **17**(9): 409-414.
Jasnos, L. and R. Korona (2007). "Epistatic buffering of fitness loss in yeast double deletion strains." Nat Genet **39**(4): 550-554.

Jasnos, L., K. Tomala, et al. (2008). "Interactions between stressful environment and gene deletions alleviate the expected average loss of fitness in yeast." Genetics **178**(4): 2105-2111.
JebaMercy, G., S. K. Pandian, et al. (2011). "Changes in Caenorhabditis elegans life span and selective innate immune genes during Staphylococcus aureus infection." Folia Microbiol (Praha) **56**(5): 373-380.
Jeffery, W. R. (2009). "Regressive Evolution in Astyanax Cavefish." Annual Review of Genetics **43**: 25-47.
Jenkins, N. L., G. McColl, et al. (2004). "Fitness cost of extended lifespan in Caenorhabditis elegans." Proceedings of the Royal Society B-Biological Sciences **271**(1556): 2523-2526.
Jones, K. L., T. C. Todd, et al. (2006). "Molecular approach for assessing responses of microbial-feeding nematodes to burning and chronic nitrogen enrichment in a native grassland." Molecular Ecology **15**(9): 2601-2609.
Kawecki, T. J., N. H. Barton, et al. (1997). "Mutational collapse of fitness in marginal habitats and the evolution of ecological specialisation." Journal of Evolutionary Biology **10**(3): 407-429.
Keightley, P. D. and M. Lynch (2003). "Toward a realistic model of mutations affecting fitness." Evolution **57**(3): 683-685.
Kenyon, C. J. (2010). "The genetics of ageing." Nature **464**(7288): 504-512.
Kim, D. H., R. Feinbaum, et al. (2002). "A conserved p38 MAP kinase pathway in Caenorhabditis elegans innate immunity." Science **297**(5581): 623-626.
Kimura, M. and T. Maruyama (1966). "The mutational load with epistatic gene interactions in fitness." Genetics **54**(6): 1337-1351.
Kishony, R. and S. Leibler (2003). "Environmental stresses can alleviate the average deleterious effect of mutations." J Biol **2**(2): 14.
Kniazeva, M., Q. T. Crawford, et al. (2004). "Monomethyl branched-chain fatty acids play an essential role in Caenorhabditis elegans development." Plos Biology **2**(9): E257.
Koch, E. N., M. Costanzo, et al. (2012). "Conserved rules govern genetic interaction degree across species." Genome Biol **13**(7): R57.
Kondrashov, A. S. (1988). "Deleterious Mutations and the Evolution of Sexual Reproduction." Nature **336**(6198): 435-440.
Kondrashov, A. S. and J. F. Crow (1991). "Haploidy or Diploidy - Which Is Better." Nature **351**(6324): 314-315.
Kondrashov, A. S. and D. Houle (1994). "Genotype-Environment Interactions and the Estimation of the Genomic Mutation-Rate in Drosophila-Melanogaster." Proceedings of the Royal Society B-Biological Sciences **258**(1353): 221-227.
Kopitar-Jerala, N. (2012). "The role of cysteine proteinases and their inhibitors in the host-pathogen cross talk." Curr Protein Pept Sci **13**(8): 767-775.
Kover, P. X. and A. L. Caicedo (2001). "The genetic architecture of disease resistance in plants and the maintenance of recombination by parasites." Mol Ecol **10**(1): 1-16.
Kurz, C. L. and J. J. Ewbank (2003). "Caenorhabditis elegans: an emerging genetic model for the study of innate immunity." Nat Rev Genet **4**(5): 380-390.
Kwon, E. S., S. D. Narasimhan, et al. (2010). "A new DAF-16 isoform regulates longevity." Nature **466**(7305): 498-502.
Laffafian, A., J. D. King, et al. (2010). "Variation in the Strength and Softness of Selection on Deleterious Mutations." Evolution **64**(11): 3232-3241.

Lakowski, B. and S. Hekimi (1998). "The genetics of caloric restriction in Caenorhabditis elegans." Proc Natl Acad Sci U S A **95**(22): 13091-13096.
Lalic, J. and S. F. Elena (2013). "Epistasis between mutations is host-dependent for an RNA virus." Biology Letters **9**(1).
Lande, R. and D. W. Schemske (1985). "The Evolution of Self-Fertilization and Inbreeding Depression in Plants .1. Genetic Models." Evolution **39**(1): 24-40.
Lee, I., B. Lehner, et al. (2008). "A single gene network accurately predicts phenotypic effects of gene perturbation in Caenorhabditis elegans." Nat Genet **40**(2): 181-188.
Lynch, M., J. Blanchard, et al. (1999). "Perspective: Spontaneous deleterious mutation." Evolution **53**(3): 645-663.
Lynch, M. and J. L. Blanchard (1998). "Deleterious mutation accumulation in organelle genomes." Genetica **102-3**: 29-39.
Lynch, M., J. Conery, et al. (1995). "Mutational meltdowns in sexual populations." Evolution **49**(6): 1067-1080.
Lynch, M. and J. S. Conery (2003). "The origins of genome complexity." Science **302**(5649): 1401-1404.
Lynch, M., R. Burger, et al. (1993). "The Mutational Meltdown in Asexual Populations." Journal of Heredity **84**(5): 339-344.
Lynch, M., L. Latta, et al. (1998). "Mutation, selection, and the maintenance of life-history variation in a natural population." Evolution **52**(3): 727-733.
MacLellan, K., L. Kwan, et al. (2012). "Dietary stress does not strengthen selection against single deleterious mutations in Drosophila melanogaster." Heredity (Edinb) **108**(3): 203-210.
Mallo, G. V., C. L. Kurz, et al. (2002). "Inducible antibacterial defense system in C. elegans." Curr Biol **12**(14): 1209-1214.
Martin, G., S. F. Elena, et al. (2007). "Distributions of epistasis in microbes fit predictions from a fitness landscape model." Nat Genet **39**(4): 555-560.
Martin, G. and T. Lenormand (2006). "The fitness effect of mutations across environments: A survey in light of fitness landscape models." Evolution **60**(12): 2413-2427.
Montalvo-Katz, S., H. Huang, et al. (2013). "Association with soil bacteria enhances p38-dependent infection resistance in Caenorhabditis elegans." Infect Immun **81**(2): 514-520.
Moore, J. H. (2005). "A global view of epistasis." Nat Genet **37**(1): 13-14.
Murphy, C. T., S. A. McCarroll, et al. (2003). "Genes that act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans." Nature **424**(6946): 277-283.
Nicholas, H. R. and J. Hodgkin (2004). "Responses to infection and possible recognition strategies in the innate immune system of Caenorhabditis elegans." Mol Immunol **41**(5): 479-493.
Ohta, T. (1992). "The Nearly Neutral Theory of Molecular Evolution." Annual Review of Ecology and Systematics **23**: 263-286.
O'Rourke, D., D. Baban, et al. (2006). "Genomic clusters, putative pathogen recognition molecules, and antimicrobial genes are induced by infection of C. elegans with M. nematophilum." Genome Res **16**(8): 1005-1016.
Orr, H. A. (1998). "The population genetics of adaptation: The distribution of factors fixed during adaptive evolution." Evolution **52**(4): 935-949.
Otto, S. P. and A. C. Gerstein (2006). "Why have sex? The population genetics of sex and recombination." Biochemical Society Transactions **34**: 519-522.

Otto, S. P. and D. B. Goldstein (1992). "Recombination and the Evolution of Diploidy." Genetics **131**(3): 745-751.
Partridge, F. A., M. J. Gravato-Nobre, et al. (2010). "Signal transduction pathways that function in both development and innate immunity." Dev Dyn **239**(5): 1330-1336.
Partridge, L. and N. H. Barton (1993). "Optimality, Mutation and the Evolution of Aging." Nature **362**(6418): 305-311.
Pujol, N., E. M. Link, et al. (2001). "A reverse genetic analysis of components of the Toll signaling pathway in Caenorhabditis elegans." Curr Biol **11**(11): 809-821.
Peters, A.D., Lively C.M. (2000). Beyond the average: Epistasis and the maintenance of sex. Oxford [England];New York : Oxford University Press
Peters, A. D. and P. D. Keightley (2000). "A test for epistasis among induced mutations in Caenorhabditis elegans." Genetics **156**(4): 1635-1647.
Phillips, P. C. (2008). "Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems." Nat Rev Genet **9**(11): 855-867.
Phillips, P. C., Otto, S. P., et al. (2000). Beyond the average: the evolutionary importance of gene interactions and variability of epistatic effects. Oxford [England]; New York : Oxford University Press
Remold, S. K. and R. E. Lenski (2001). "Contribution of individual random mutations to genotype-by-environment interactions in Escherichia coli." Proc Natl Acad Sci U S A **98**(20): 11388-11393.
Rice, W. R. (1998). "Requisite mutational load, pathway epistasis, and deterministic mutation accumulation in sexual versus asexual populations." Genetica **102-3**: 71-81.
Rowe, H. C., B. G. Hansen, et al. (2008). "Biochemical networks and epistasis shape the Arabidopsis thaliana metabolome." Plant Cell **20**(5): 1199-1216.
Ryan, C. J., A. Roguev, et al. (2012). "Hierarchical modularity and the evolution of genetic interactomes across species." Mol Cell **46**(5): 691-704.
Sanjuan, R. and S. F. Elena (2006). "Epistasis correlates to genomic complexity." Proc Natl Acad Sci U S A **103**(39): 14402-14405.
Sanjuan, R., A. Moya, et al. (2004). "The contribution of epistasis to the architecture of fitness in an RNA virus." Proc Natl Acad Sci U S A **101**(43): 15376-15379.
Sanjuan, R. and M. R. Nebot (2008). "A Network Model for the Correlation between Epistasis and Genomic Complexity." Plos One **3**(7).
Scheiner, S. M. (1993). "Genetics and Evolution of Phenotypic Plasticity." Annual Review of Ecology and Systematics **24**: 35-68.
Schulenburg, H. and C. Boehnisch (2008). "Diversification and adaptive sequence evolution of Caenorhabditis lysozymes (Nematoda: Rhabditidae)." BMC Evol Biol **8**: 114.
Schulenburg, H., M. P. Hoeppner, et al. (2008). "Specificity of the innate immune system and diversity of C-type lectin domain (CTLD) proteins in the nematode Caenorhabditis elegans." Immunobiology **213**(3-4): 237-250.
Segre, D., A. DeLuna, et al. (2005). "Modular epistasis in yeast metabolism." Nat Genet **37**(1): 77-83.
Shaw, F. H., C. J. Geyer, et al. (2002). "A comprehensive model of mutations affecting fitness and inferences for Arabidopsis thaliana." Evolution **56**(3): 453-463.
Sniegowski, P. D., P. J. Gerrish, et al. (2000). "The evolution of mutation rates: separating causes from consequences." Bioessays **22**(12): 1057-1066.
Stearns, S. C. (1989). "Trade-Offs in Life-History Evolution." Functional Ecology **3**(3): 259-268.

Steiner, C. C., J. N. Weber, et al. (2007). "Adaptive variation in beach mice produced by two interacting pigmentation genes." Plos Biology **5**(9): 1880-1889.
Sulston J. and Hodgkin J (1988). The nematode *Caenorhabditis elegans* .WBWe, ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory press.pp 587-606
Szathmary, E. (1993). "Do Deleterious Mutations Act Synergistically - Metabolic Control-Theory Provides a Partial Answer." Genetics **133**(1): 127-132.
Tan, M. W. and F. M. Ausubel (2000). "Caenorhabditis elegans: a model genetic host to study Pseudomonas aeruginosa pathogenesis." Current Opinion in Microbiology **3**(1): 29-34.
Tan, M. W., S. Mahajan-Miklos, et al. (1999). "Killing of Caenorhabditis elegans by Pseudomonas aeruginosa used to model mammalian bacterial pathogenesis." Proc Natl Acad Sci U S A **96**(2): 715-720.
Tenor, J. L. and A. Aballay (2008). "A conserved Toll-like receptor is required for Caenorhabditis elegans innate immunity." EMBO Rep **9**(1): 103-109.
Troemel, E. R., S. W. Chu, et al. (2006). "p38 MAPK regulates expression of immune response genes and contributes to longevity in C. elegans." PLoS Genet **2**(11): e183.
Ungerer, M. C., L. C. Johnson, et al. (2008). "Ecological genomics: understanding gene and genome function in the natural environment." Heredity (Edinb) **100**(2): 178-183.
Vassilieva, L. L., A. M. Hook, et al. (2000). "The fitness effects of spontaneous mutations in Caenorhabditis elegans." Evolution **54**(4): 1234-1246.
Vassilieva, L. L. and M. Lynch (1999). "The rate of spontaneous mutation for life-history traits in Caenorhabditis elegans." Genetics **151**(1): 119-129.
Wade, M. J., R. G. Winther, et al. (2001). "Alternative definitions of epistasis: dependence and interaction." Trends in Ecology & Evolution **16**(9): 498-504.
Walker, D. W., G. McColl, et al. (2000). "Natural selection - Evolution of lifespan in C-elegans." Nature **405**(6784): 296-297.
Wang, A. D., N. P. Sharp, et al. (2009). "Selection, epistasis, and parent-of-origin effects on deleterious mutations across environments in Drosophila melanogaster." Am Nat **174**(6): 863-874.
Wang, C., A. Marshall, et al. (2012). "ANAP: an integrated knowledge base for Arabidopsis protein interaction network analysis." Plant Physiology **158**(4): 1523-1533.
Wang, Z. and D. R. Sherwood (2011). "Dissection of genetic pathways in C. elegans." Methods Cell Biol **106**: 113-157.
Wegner, K. M., C. Berenos, et al. (2008). "Nonadditive genetic components in resistance of the red flour beetle Tribolium castanaeum against parasite infection." Evolution **62**(9): 2381-2392.
Whitlock, M. C. and D. Bourguet (2000). "Factors affecting the genetic load in Drosophila: Synergistic epistasis and correlations among fitness components." Evolution **54**(5): 1654-1660.
Whitlock, M. C., P. C. Phillips, et al. (1995). "Multiple Fitness Peaks and Epistasis." Annual Review of Ecology and Systematics **26**: 601-629.
Williams, G. C. (1957). "Pleiotropy, Natural-Selection, and the Evolution of Senescence." Evolution **11**(4): 398-411.Agrawal, A. F. and M. C. Whitlock (2010). "Environmental duress and epistasis: how does stress affect the strength of selection on new mutations?" Trends in Ecology & Evolution **25**(8): 450-458.
Williams, G. C. (1957). "Pleiotropy, Natural-Selection, and the Evolution of Senescence." Evolution **11**(4): 398-411.

Wong, D., D. Bazopoulou, et al. (2007). "Genome-wide investigation reveals pathogen-specific and shared signatures in the response of Caenorhabditis elegans to infection." Genome Biol **8**(9): R194.
Wu, C. I. and C. T. Ting (2004). "Genes and speciation." Nature Reviews Genetics **5**(2): 114-122.
Yang, H. P., A. Y. Tanikawa, et al. (2001). "Whole-genome effects of ethyl methanesulfonate-induced mutation on nine quantitative traits in outbred Drosophila melanogaster." Genetics **157**(3): 1257-1265.
You, L. C. and J. Yin (2002). "Dependence of epistasis on environment and mutation severity as revealed by in silico mutagenesis of phage T7." Genetics **160**(4): 1273-1281.
Young, J. A., C. P. Yourth, et al. (2009). "The effect of pathogens on selection against deleterious mutations in Drosophila melanogaster." Journal of Evolutionary Biology **22**(10): 2125-2129.
Yang, H. P., A. Y. Tanikawa, et al. (2001). "Whole-genome effects of ethyl methanesulfonate-induced mutation on nine quantitative traits in outbred Drosophila melanogaster." Genetics **157**(3): 1257-1265.
You, L. C. and J. Yin (2002). "Dependence of epistasis on environment and mutation severity as revealed by in silico mutagenesis of phage T7." Genetics **160**(4): 1273-1281.
Young, J. A., C. P. Yourth, et al. (2009). "The effect of pathogens on selection against deleterious mutations in Drosophila melanogaster." Journal of Evolutionary Biology **22**(10): 2125-2129.