Genetic Background and Environmental Effects on Single Nucleotide Polymorphisms in the NADPH Pathway

by

Simran Baath

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science (MSc) in Chemical Sciences

Faculty of Graduate Studies
Laurentian University
Sudbury, Ontario, Canada

© Simran Baath, 2019

THESIS DEFENCE COMMITTEE/COMITÉ DE SOUTENANCE DE THÈSE
Laurentian Université/Université Laurentienne
Faculty of Graduate Studies/Faculté des études supérieures

Title of Thesis Titre de la thèse	Genetic Background and Environmental Effects on Single Nucleotide Polymorphisms in the NADPH Pathway	
Name of Candidate Nom du candidat	Baath, Simran	
Degree Diplôme	Master of Science	
Department/Program Département/Programme	Chemical Sciences	Date of Defence Date de la soutenance December 13, 2018

APPROVED/APPROUVÉ

Thesis Examiners/Examinateurs de thèse:

Dr. Thomas Merritt
(Supervisor/Directeur(trice) de thèse)

Dr. Eric Gauthier
(Committee member/Membre du comité)

Dr. Guandong Yang
(Committee member/Membre du comité)

Approved for the Faculty of Graduate Studies
Approuvé pour la Faculté des études supérieures
Dr. David Lesbarrères
Monsieur David Lesbarrères

Dr. Nadia Singh Dean, Faculty of Graduate Studies
(External Examiner/Examinateur externe)

Doyen, Faculté des études supérieures

ACCESSIBILITY CLAUSE AND PERMISSION TO USE

I, **Simran Baath**, hereby grant to Laurentian University and/or its agents the non-exclusive license to archive and make accessible my thesis, dissertation, or project report in whole or in part in all forms of media, now or for the duration of my copyright ownership. I retain all other ownership rights to the copyright of the thesis, dissertation or project report. I also reserve the right to use in future works (such as articles or books) all or part of this thesis, dissertation, or project report. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that this copy is being made available in this form by the authority of the copyright owner solely for the purpose of private study and research and may not be copied or reproduced except as permitted by the copyright laws without written authority from the copyright owner.

Abstract

A major focus in modern genomics is determining the connection between genotypes and quantifying phenotypes. In this connection, many factors come into play including different genetic backgrounds, genetic variation at a locus, and environmental conditions. Genetic variation in *Drosophila melanogaster*, and specifically the simple polymorphisms within *Malic enzyme* (*Men*), can provide insight into the pathways between genotypes and phenotypes. Globally, there are two polymorphic sites in the malic enzyme gene. On site is near the protein (MEN) active site and found at an allelic frequency of 50% glycine amino acid and 50% alanine amino acid. The second polymorphism is buried within the protein and found at an allelic frequency of 90% methionine amino acid and 10% leucine amino acid. To determine the complexity of the pathway between genotypes and phenotypes, multiple genetic backgrounds for each genotype, using multiple *D. melanogaster* lines, were included to explore and quantify genetic background effects, and paraquat was used to induce oxidative stress. The biochemical characteristics of the alleles varied significantly between the genotypes under benign conditions and both polymorphic sites effected some phenotypes. The first site played a role in the MEN V_{max} and K_m; the glycine allele had 14% higher V_{max} activity than the alanine allele and the glycine allele had 8% higher K_m than the alanine allele. The second site influenced the K_m and V_{max}/K_m ratio (relative activity); the methionine allele had 34% higher malate K_m than the leucine allele the leucine allele had 52% higher relative activity than the methionine allele. Interestingly, the protein product encoded by the rarer allele, leucine, had a higher relative activity and lower K_m concentration, having a large impact on the enzymatic phenotype. These extreme phenotypes of that allele may be an indication of the why the allele is maintained at 10% across populations. Different lines with the same genotype had different biochemical phenotypes, indicating the importance of backgrounds effects influencing the final phenotype. Further, the flies' phenotypes differed between benign and oxidative stress conditions. Flies exposed to paraquat had a decrease in MEN V_{max}, and the MEN alleles did not significantly differ from each other. Overall, the findings from this study suggest that the final phenotype are strongly influenced by the polymorphisms found in MEN, the interactions between genetic background and environmental conditions.

Acknowledgments

I would like to take this time to thank a few people in helping me with completing this thesis. First off, I would like to thank Thomas Merritt, my supervisor, for taking me on for a masters and supporting me throughout the years. I would also like to thank Thomas for pushing me forward to being a better researcher.

I would also like to thank Dr. Gauthier and Dr. Yang for taking the time to be a part of my committee and being available to answer any questions I have regarding my project. I truly believe that I had an amazing committee to support me through my research.

I'm grateful for my family and especially my brothers, Tej and Navjeet Baath for supporting me during my research phase specially the food trips they've done for me. I am also grateful to my parents for their advice, regardless of their non-science background, and support nonetheless. I would not have been able to reach here without them.

I am also thankful to Michelle Eng and Jessica Keegan, my best friends, for always being there and helping me with completing my thesis. Without them, I would not have been able to complete all my hundreds of crosses required for this project.

Lastly, I am grateful for Teresa Rzezniczak, my mentor, whom I meet on the first day in the lab and has taught me everything I know now in research. I'm grateful to have met a patient person, who has always been available to answer my questions and plan out my thesis.

Table of Contents

Abstract iii

Acknowledgements iv

List of Figures viii

List of Tables x

List of Abbreviations ii

CHAPTER 1: GENERAL INTRODUCTION 1

1.1) Overview of this thesis 2

1.2) Malic enzyme 3

1.3) NADPH Pathway 5

1.4) Genetic Background 6

1.5) The Environment 7

1.6) Oxidative Stress 8

1.7) Objectives & Hypothesis 8

1.8) References 10

CHAPTER 2: GENETIC BACKGROUND AND SINGLE NUCLEOTIDE POLYMORPHISM IN THE NADPH PATHWAY IN *DROSOPHILA MELANOGASTER* AT THE *MALIC ENZYME* LOCUS 14

2.1) Introduction 15

2.2) Materials and Methods 18

2.2.1) Fly stocks and lines 18
2.2.2) Fly wet weight 21
2.2.3) Fly homogenization 21
2.2.4) Enzyme activity measurement 21
2.2.5) Estimation of Michaelis-Menten Constant 22
2.2.6) Thermal Stability of Malic Enzyme 22
2.2.7) Soluble Triglyceride Content 22
2.2.8) Total Carbohydrate Content 23
2.2.9) Soluble Protein Content 23
2.2.10) RNA Extraction and Quantitative RT-PCR 23
2.2.11) Longevity 24
2.2.12) Data Analysis 24

2.3) Results 25

2.3.1) The G-A polymorphism has a much larger effect on maximum Malic Enzyme reaction velocity (V_{max}) than the M-L polymorphisms 25
2.3.2) *Men* expression and MEN activity are correlated 29
2.3.3) The M-L polymorphism has a larger effect on the Malate Michaelis-Menton constant 31
2.3.4) The M-L polymorphism has a larger effect on the V_{max}/K_m Ratio 34
2.3.5) Only the G-A polymorphism affects Thermal Stability 36
2.3.6) The polymorphic sites are not correlated with changes in G6PD and IDH activity 38
2.3.7) The polymorphic sites are also not correlated with changes in distal phenotypes .41

2.4) Discussion..........45

2.4.1) MEN Phenotypes are influenced by the G-A and M-L Polymorphisms..........45
2.4.2) Distal Phenotypes and the MEN polymorphism 49
2.4.3) Insight into the *Men* locus and Heterozygote Advantage 49

2.5) Reference..........51

CHAPTER 3: ENVIRONMENTAL EFFECTS ON SINGLE NUCLEOTIDE POLYMORPHISM IN THE NADPH PATHWAY AT THE *MALIC ENZYME* LOCUS 58

3.1) Introduction 59

3.2) Materials and Methods 63

3.2.1) Fly stocks and rearing conditions 63
3.2.2) Fly crosses 63
3.2.3) Oxidative stress treatment 64
3.2.4) Fly wet weight 64
3.2.5) Fly homogenization 64
3.2.6) Enzyme activity measurement..........64
3.2.7) Antioxidant enzyme activity measurements 65
3.2.8) Estimation of Michaelis-Menten Constant..........65
3.2.9) RNA Extraction and Quantitative RT-PCR..........66
3.2.10) Soluble Triglyceride Content 66
3.2.11) Total Carbohydrate Content 67
3.2.12) Soluble Protein Content..........67
3.2.13) Data Analysis..........67

3.3) Results 68

3.3.1) Both *Men* Polymorphisms Influence Maximum Velocity 68
3.3.2) Both Polymorphisms Influences *Men* Expression Levels 70

3.3.3) Both G-A and M-L Polymorphisms Influences the Binding Affinity 72
3.3.4) The M-L Polymorphism Influences the V_{max}/K_m Ratio 75
3.3.5) Responses to MEN Activity within the NADPH Pathway 76
3.3.5.1) The G-A Polymorphism affects IDH Activity 76
3.3.5.2) The G-A Polymorphism affects G6PD Activity 80
3.3.6) Responses of distal phenotypes to MEN activity 80
3.3.6.1) MEN Activity affected Carbohydrate Content 81
3.3.6.2) MEN Activity did not influence Triglyceride Storage 83
3.3.7) Response of NADPH Network to a Stressor 83
3.3.7.1) The G-A Polymorphism Influences MEN Maximum Velocity under Oxidative Stress 86
3.3.7.2) The MEN Genotype do not alter IDH Activity under Oxidative Stress ..86
3.3.7.3) The MEN genotypes do not alter G6PD Activity under Oxidative Stress 89
3.3.7.4) The MEN genotypes do not alter SOD Activity under Oxidative Stress 92
3.3.7.5) The MEN genotypes do not alter GST Activity under Oxidative Stress .94
3.3.7.6) The MEN genotypes do not alter CAT Activity under Oxidative Stress 94
3.4) Discussion 98
3.4.1) MEN Phenotypic Variation across the Different Genotypes..................................98
3.4.2) MEN Alleles and Interactions within the NADPH Pathway 103
3.4.3) MEN Alleles Effects Distal Phenotypes 104
3.4.4) *Men* genotypes Effects to Environmental Stressors 105
3.5) Reference 109

CHAPTER 4: CONCLUSION & FUTURE DIRECTION 115
4.1) Overview 116
4.2) MEN Alleles and the Biochemical Variation 116
4.3) Interactions with the NADPH Pathway & Genetic Background Effects 117
4.4) Characterizing Metabolic Response to Environmental Stressor 118
4.6) MEN Topology 119
4.7) The $NADP^+$: NADPH Ratio 119
4.8) 6-gluconatephosphate dehydrogenase 120
4.9) Environmental Conditions and Stressors 120
4.10) mRNA degradation 121
4.11) Conclusion 121
4.12) References 123

List of Figures

Figure 2.1. A. Schematic of crosses performed to generated desired alleles 20
Figure 2.2. Average maximum MEN enzyme activity was estimated for each genotype 26
Figure 2.3. Average maximum MEN enzyme activity (OD/time) was estimated for various crosses that contribute to the specific genotypes 28
Figure 2.4.Correlation between MEN protein activity and relative *Men* gene expression 30
Figure 2.5. Michaelis-Menton constant (K_m) for malate 32
Figure 2.6. Average MEN binding affinity was estimated for various crosses that contribute to the specific genotypes 33
Figure 2.7. Maximum MEN velocity to Michaelis-Menton constant for malate ratio 35
Figure 2.8. Malic enzyme thermal stability 37
Figure 2.9. Average maximum G6PD enzyme activity (OD/time) was estimated for each genotype 39
Figure 2.10. Average maximum IDH enzyme activity (OD/time) was estimated for each genotype 40
Figure 2.11. Average carbohydrate content was estimated for each genotype 42
Figure 2.12. Average Triglyceride content was estimated for each genotype 43
Figure 2.13. Average Longevity was estimated for each genotype 44
Figure 3.1. MEN V_{max} Activity across different alleles and backgrounds under saturating conditions. 69
Figure 3.2. *Men* relative expression levels 71
Figure 3.3. Malate K_m binding affinity across different alleles and backgrounds, using various malate concentrations 74
Figure 3.4. MEN V_{max}/K_m Activity, or relative *in vivo* activity across different alleles and backgrounds under saturating conditions...77
Figure 3.5. IDH V_{max} Activity across different *Men* alleles and backgrounds under saturating conditions 79
Figure 3.6. G6PD V_{max} Activity across different *Men* alleles and backgrounds under saturating conditions 82
Figure 3.7. Carbohydrate content across different *Men* alleles and backgrounds under saturating conditions 84
Figure 3.8. Triglycerides content across different *Men* alleles and backgrounds under saturating conditions 85
Figure 3.9. MEN V_{max} Activity across different alleles and backgrounds under paraquat condition 87
Figure 3.10. IDH V_{max} Activity across different alleles and backgrounds under paraquat condition 88
Figure 3.11. G6PD V_{max} Activity across different alleles and backgrounds under paraquat condition 91

Figure 3.12. SOD V_{max} Activity across different alleles and backgrounds under paraquat condition 93
Figure 3.13. GST V_{max} Activity across different alleles and backgrounds under paraquat condition 96
Figure 3.14. CAT Vmax Activity across different alleles and backgrounds under paraquat condition 97

List of Tables

Table 2.1: DGRP lines, and their genotypes, used to construct homozygous and heterozygous genotypes 19

Table 2.2. Predicted MEN activities for Heterozygote alleles base derived from the homozygous alleles 27

List of Abbreviations

μL: microlitre

6PGD: 6-phosphogluconate dehydrogenase

6pgd: 6-phsophogluconate dehydrogenase gene

ANCOVA: Analysis of co-variance

CAT: Catalase

g: grams

G6PD: Glucose-6-phosphate dehydrogenase

G6pd: *Glucose-6-phosphate dehydrogenase* gene

GST: Glutathione-s-transferase

H_2O_2: Hydrogen peroxide

IDH: Isocitrate dehydrogenase

Idh: *Isocitrate dehydrogenase* gene

MEN: Malic enzyme

Men: Malic enzyme gene

MenEx: Malic enzyme excision chromosome

mg: milligram

MgSO4: Magnesium sulfate

mL: millilitre

mM: millimolar

MnCl2: Manganese chloride

$NADP^+$: Nicotinamide adenine dinucleotide phosphate (oxidized)

NADPH: Nicotinamide adenine dinucleotide phosphate (reduced)

nmol: nanomoles

OD: Optical density

PPP: Pentose phosphate pathway

Para: Paraquat

ROS: Reactive oxygen species

SOD: Superoxide dismutase

Tris-HCl: Tris(hydroxymethyl) aminomethane –hydrochloride

Tukey's HSD Test: Tukey's Honestly significant difference test

Chapter 1: General Introduction

1.1 Overview of this thesis

With advances in technology used to study modern genetics, the main focus of research is determining genotypes and quantifying phenotypes. However, an essential mechanism in genetics has yet to be answered, a question that has existed for about 50 years (Merritt, Duvernell, & Eanes, 2005). To begin to understand the complex pathway from a genotype to phenotype, it is crucial to fully understand how one gene can lead to a variety of phenotypes. These phenotypes derived from one gene are responding to different influencers, such as other genes, the environment, and mutations (Boone, Bussey, & Andrews, 2007; Chandler, Chari, & Dworkin, 2013; Cordell, 2009; Dworkin et al., 2009; Rzezniczak & Merritt, 2012). Therefore, it is not the gene solely that drives the phenotype. In this thesis, *Drosophila melanogaster* is the model organism used to understand genetic inheritance, focusing on the *Malic enzyme* (*Men*) locus, an NADPH network enzyme. How the different alleles of *Men* can be affected by multiple genes, in this case other NADPH enzymes, and the environment is extensively studied. The examples throughout this introduction focus on *D. melanogaster* and the NADPH network.

A multiple genes and environment approach is used for this thesis, incorporating and measuring responses of multiple factors. The focus on the variation in this study is the nine different *Men* genotypes, their interactions with each other, other players in the NADPH network, and how MEN alleles acclimates to an environmental stressor, more specifically oxidative stress. As background for my work, I will go over the essential concepts of the central dogma of molecular genetics, heterozygote advantage, and complex traits. Furthermore, an in-depth discussion of the knowledge obtained from studying Malic enzyme, the NADPH pathway, and the oxidative stress condition is present in this chapter.

The central dogma is a great schematic that shows a mechanism of converting DNA to RNA to protein (Crick, 1970; Li & Xie, 2011; Schreiber, 2005). The limitation with the central dogma is that it is a simplified explanation. A gene, solely, does not determine the phenotype. Numerous factors have been uncovered, and are still being studied today, that can drive changes in the phenotype. Complex traits are a function of genes interacting

with other genes to drive changes in the phenotype (Chow, 2016; Dworkin et al., 2009; Lessel, Parkes, Dickinson, & Merritt, 2017). Furthermore, the biological system is dynamic; in response to variation from its surroundings, and phenotype will change (Bernard, Parkes, & Merritt, 2011; Griendling et al., 2016; Rzezniczak & Merritt, 2012). Therefore, one allele can display multiple phenotypes for a given circumstance (Bing et al., 2014; Dworkin et al., 2009; Lessel et al., 2017; Rzezniczak & Merritt, 2012).

1.2 Malic Enzyme

Malic enzyme (*Men*) in *Drosophila melanogaster* is a metabolic enzyme and a suitable model system to answer a fundamental question in genetics: What is the path between genetic diversity and biological complexity? With the advances in modern genomics, the focus is now on determining genotype and quantifying phenotypes. However, the connection between the two was not always clear. Within the Merritt Lab, the focus is attempting to answer that fundamental question, understanding the players that take us from DNA to the diverse phenotypes we have today (Merritt et al., 2005, 2009; Merritt, Sezgin, Zhu, & Eanes, 2006; Rzezniczak, Lum, Harniman, & Merritt, 2012).

MEN is a metabolic enzyme that oxidizes malate to pyruvate and reduces $NADP^+$ to NADPH (Merritt et al., 2005; Rzezniczak et al., 2012; Rzezniczak & Merritt, 2012; Ying, 2008). The *Men* locus is located in the third chromosome in *D. melanogaster*. Across the global population *D. melanogaster*, there are two known single nucleotide polymorphisms (SNPs) that alter the amino acid composition of the MEN protein. The first polymorphism occurs at base pair 338, which is a guanine to cytosine substitution that results in a glycine to alanine change at amino acid 113. Hereafter, we refer to this site as the G-A polymorphism. Between the two amino acids, the glycine amino acid is found in Africa *D. melanogaster* and *D. simulans* (a *Drosophila* species that shares a common ancestor with *D. melanogaster*), suggesting that glycine is the ancestral trait (Merritt et al., 2005; Rzezniczak et al., 2012; Sezgin et al., 2004). The glycine and alanine occur in equal frequencies across North American populations. The G-A polymorphism is in an alpha helix near the MEN active site. The second polymorphism occurs at base pair 1051, which is an adenine to thymine substitution that results in a

methionine to leucine change at amino acid 351. Hereafter, we refer to this as the M-L polymorphism. The leucine amino is present in two North American populations (Sudbury and Raleigh), however not in African populations or *D. simulans*, suggesting that the methionine amino is the ancestral trait. The M-L polymorphism is in a beta sheet buried in the protein, not close to functional components of the protein. The leucine amino acid is at an allelic frequency of 10%, whereas methionine is at an allelic frequency of 90% (Rzezniczak et al., 2012). Structure-wise, MEN is a tetra-homodimer, in which two MEN monomers form a dimer, and two dimers bind to form a tetramer (Merritt et al., 2005).

The biochemical characteristics of the polymorphism in MEN have been previously studied (Rzezniczak et al., 2012). The previous work used isothrid chromosomes, in which the third chromosomes were identical (homozygous). The phenotypes examined were V_{max}, which is the maximum velocity of the enzyme, K_m, the binding affinity of the enzyme and thermal stability, an indication of how stable the enzyme is under denaturing conditions. For V_{max}, the G-A polymorphism influenced the activity; the glycine amino acid had a 46% higher activity than alanine. No differences in MEN V_{max} were found at the M-L polymorphism. Differences in V_{max} were identified at the G-A polymorphism, but not the M-L polymorphism, likely because of the G-A polymorphism's location is near the active site of the enzyme, and the M-L polymorphism’s location within the protein away from the active site.

The *Men* gene expression for the different polymorphisms was also quantified. Gene expression was analyzed because the V_{max}'s phenotype is expected to be influenced by both the structure of the protein and the regulation of the gene. Regulatory effects consist of transcripts and regulatory elements, such as enhancers, insulators and promoters. In previous work, lines that had the guanine nucleotide allele has higher *Men* expression level (Rzezniczak et al., 2012). Interestingly, the M-L polymorphism (adenine and thymine) was also associated with significant differences in *Men* expression levels, even though their MEN V_{max} levels were not significantly different. The thymine nucleotide was associated with lower expression of *Men* than the adenine nucleotide. This difference in *Men* expression levels could possibly dictate the similarities in MEN V_{max}. The

polymorphisms on its own should not affect gene expression, since it is found within the gene. The differences in expression levels could be driven due to regulatory variations. The structural changes in the enzyme have yet to be analyzed and further quantification of the isolated protein will provide insight into structure variation and regulatory variation. The quantification of expression levels along with the MEN activity indicates that both structure and regulatory variations influence the differences and similarities found in MEN V_{max}.

The second phenotype quantified was the malate K_m of MEN. K_m is a measure of the binding affinity of an enzyme, and this phenotype is expected to be influenced solely by the structure of the enzyme, the amino acid sequence, and not by regulation of gene expression. Both polymorphisms, G-A and M-L, influenced the K_m of MEN; the glycine amino acid had a higher K_m than alanine, and the leucine had a higher K_m than methionine. The M-L polymorphism influence on the K_m phenotype is likely due to the alteration of the amino acid sequence, which leads to changes in the tertiary structure of the protein. Changes in the structure of a protein can change the function of the protein, ex. Hemoglobin and prions (Prusiner, 1990; Steinberg, 2008).

Thermal stability was another MEN phenotype quantified in this study. The authors placed the enzyme in denaturing conditions over time and compared the MEN activity with a control group. The G-A polymorphisms is in an alpha helix. The alanine amino acid at the G-A polymorphic site is assumed to be more stable than the glycine amino acid, since glycine amino acid tends to destabilize alpha helices (Chakrabartty, Schellman, & Baldwin, 1991; Ganter & Plückthun, 1990). However, the glycine is 15% more stable in denaturing condition than the alanine amino acid. With thermal stability, the protein as a whole is exposed to denaturating conditions, which may be the reason that glycine is more stable than alanine- the interactions between the R groups. The M-L polymorphism did not influence the stability of the enzyme, relating back to the proximity of the polymorphic sites.

Overall, previous studies were able to characterizations in the differences between the SNPs found in MEN. These biochemical differences indicate the possibility that MEN is under selection and can further provide insight into the maintenance of genetic variations.

1.3 The NADPH Pathway

We expect that some mutations, especially the ones that alter the amino acid sequence, can affect the immediate function of the enzyme. However, enzymes do not act in isolation; a single gene can affect multiple phenotypes through gene products and substrates (Rzezniczak et al., 2012; Rzezniczak & Merritt, 2012). To determine a gene's function often times the gene is quantified only in the gene's primary pathway. However, any changes in the direct pathway can disrupt other pathways as well, and vice versa, in which other pathways can alter the gene of interest (Bernard et al., 2011; Rzezniczak & Merritt, 2012). Therefore, when studying genes, other pathways should be quantified to grasps the effects and interactions of the gene; metabolism is a network, not a linear pathway.

MEN is a metabolic enzyme that reduces $NADP^+$ to NADPH while converting malate to pyruvate. Other metabolic enzymes also reduce $NADP^+$ to NADPH, including Isocitrate dehydrogenase (IDH), Glucose-6-phosphate dehydrogenase (G6PD) and 6-phosphogluconate dehydrogenase (6PGD; Rzezniczak et al. 2012; Ying 2008; Merritt, Duvernell, and Eanes 2005). These four enzymes interact to maintain the $NADP^+$: NADPH ratio within the cells (Merritt et al., 2005, 2009). These interactions are often compensatory, in which a reduction of one enzyme results in upregulation of the other), but can also be parallel and counterintuitive, in which a reduction in one enzyme leads to down-regulation of others (Rzezniczak & Merritt, 2012). Even though this network is small, it is a complex regulatory network. Previous studies found that an extreme reduction in MEN activity leads to an increase in G6PD and 6PGD activity, however a reduction in IDH activity. In adults, each enzyme contributes to the NADPH pathway differently. MEN produces about 30% of the available NADPH in *D. melanogaster*, IDH produces about 20%, and G6PD and 6PGD produces a combined total of 40% of NADPH in the cells. Other enzymes are currently being studied to understand this network further. The NADPH produced by the enzyme mentioned above have some downstream function, such as lipid synthesis and reactive oxygen species (ROS) clearance (Bernard et al., 2011; Geer, Lindel, & Lindel, 1979; Hosamani & Muralidhara, 2013; Merritt et al., 2005, 2006; Rzezniczak et al., 2011; Rzezniczak & Merritt, 2012).

1.4 Genetic Background

Complex traits are not dependent solely on their genotype, but their phenotype is also determined by interactions between multiple genes, the environment, and random chance. For example, IDH, G6PD, and 6PDG are all interacting components of the NADPH pathway whose function and activity modify that of MEN. Genetic background effects take into account the whole genome of an organism, and these genes interact with each other to regulate activities (Chandler et al., 2013; Chow, 2016). On a small scale, the genetic background can be thought out as IDH, G6PD, and 6PDG interacting with MEN. On a larger scale, it is the organism's whole genome interacting with MEN. Therefore, the phenotypes are a result of the interaction of the gene of interest with other genes, plus the environment. Nonidentical siblings can be used as an example to grasp the complexity of genetic background effects. When comparing siblings, they come from the same parents; however, the combination of the genes in their genome differs. This difference in their genome allows for different communication between alleles, overall leading to unique phenotypes from one another, hence not clones to each other.

A study within the Dworkin lab looked at the phenotype of a mutation for wing shape in different lines (also be known as different genetic backgrounds; (Dworkin et al., 2009). The mutation analyzed in this study was the scalloped allele. When the mutation is in two different *D. melanogaster* background lines, Samarkand and Oregon-R, the phenotype of the wings differed significantly. Therefore, one mutation will behave differently in different genetic background. Dworkin’s work with genetic background displays the importance of using multiple lines.

1.5 The Environment

Organisms' phenotypes can respond rapidly to changing environmental conditions, and this response is crucial for the survival of the organism. As mentioned before, this change in phenotype results from interactions with multiple genes plus the environment.

Most often, lab-derived mutations are studied in one condition- the lab condition. However, mutations' phenotypes can alter depending on the environment it is present. Furthermore, environmental factors can alter a phenotype differently across the different

genetic backgrounds. Going back to the previous example with the fly wings, a mutant allele phenotype of one line can behave differently compared to different environmental conditions. Genetic background effects and environmental changes further complicate a study, which is why many studies only include one condition and one gene to study mutation to reduce variability. However, it is essential to include this variability when understanding genetic variation maintenance within a natural population. This maintenance of genetic variation allows for organisms to acclimates to the changing environment, which is key to survival. In this study, paraquat was used to induce oxidative stress since NADPH is known to play a role in the clearance of ROS within a cell (Lessel et al., 2017; Rzezniczak & Merritt, 2012). Oxidative stress will be discussed further in the following section.

1.6 Oxidative Stress

One downstream function of NADPH is the clearance of reactive oxygen species, ROS, within a cell. When ROS begins to accumulate within the cells, it can damage DNA, protein, and lipids, sometimes leading to cell death (Bernard et al., 2011; Finkel & Holbrook, 2000; Hosamani & Muralidhara, 2013; Weber et al., 2012). The accumulation of these products is known as oxidative stress. Oxidative stress is a common environmental condition. To generate an oxidative stress environment for an organism, paraquat (1,1'-dimethyl-4-4'-bipyridynium dichloride) is one approach that is used often when working with *D. melanogaster* (Rzezniczak et al., 2011). Paraquat is an herbicide and a neurotoxin that can cause a severe hazard to animal and human health (Hosamani & Muralidhara, 2013; Rzezniczak et al., 2011). Ingested paraquat combines with an oxygen molecule to produce a superoxide anion radical along with a paraquat derivate radical, generating a ROS present in the cell. Interestingly, paraquat requires NADPH to generate the ROS mentioned above (Rzezniczak & Merritt, 2012). The superoxide anion is a ROS that can bind to macromolecules and activate apoptosis within a cell.

Under conditions of oxidative stress, the antioxidant enzymes superoxide dismutase (SOD), catalase (CAT) and glutathione- s-transferase (GST) are major consumers of NADPH (Abolaji, Olaiya, Oluwadahunsi, & Farombi, 2017; Aebi, 1984; Müller et al., 2017). SOD converts superoxide anion into oxygen molecules. CAT takes hydrogen

peroxide and converts it into oxygen and water molecules. GST is known to detoxify secondary ROS products using glutathione (GSH) as a substrate, it can cycle between its reduced (GSH) and oxidized (GSSG) forms, where GSH behaves as a reducing agent and ROS scavenger.

A previous study in the Merritt laboratory induced oxidative stress in *D. melanogaster* using paraquat and examined the interactions between the enzymes in the NADPH pathway (Rzezniczak & Merritt, 2012). In this study, the authors found that, compared to the control flies, the enzymatic activities decreased, the organism's response to a different environment. The reason why there was a decrease in the enzyme activities is to reduce the concentration of NADPH present within the cell.

1.7 Objectives & Hypothesis

Here, we are extending a combination of previous work to understand further the mechanism of genetic variation, biological complexity, and selection by using polymorphisms present within *Men*, a variety of genetic backgrounds and oxidative stress. Previous work on the system used isothrid chromosomes, quantifying one gene in one condition. In this study, the two *Men* polymorphisms are furthered biochemically characterized. Drosophila Genetic Reference Panel (DGRP) lines are used, with known alleles for the polymorphisms. Multiple lines for each allele (AM/AM, AL/AL, GM/GM and GL/GL) were crossed to one another to generate multiple heterozygous and homozygous genotypes (AM/AM, AM/AL, AL/AL, AM/GM, AM/GL, AL/GL, GM/GM, GM/GL and GL/GL). The biochemical characterization of *Men* will be first studied and compared to previous work, determining if as heterozygotes the trend observed remains consistent. The phenotypes that will quantify in Chapter 2 are MEN V_{max}, which is maximum velocity, K_m, the binding affinity and thermal stability, how stable the enzyme in denaturing condition.

By including multiple lines, we are generating flies with the same alleles at the *Men* locus that potentially differ at other loci, therefore studying genetic background effects. Furthermore, a phenotype is driven by multiple genes, the other NADPH pathway enzyme will be quantified as well. These biochemical phenotypes are immediate

functions of the allelic variation, and we refer to them as "proximal" phenotypes (Lessel et al. 2017). The distal phenotypes carbohydrate and triglyceride content, which are downstream functions of the enzymes and cofactors present here (Lessel et al. 2017), will be analyzed as well.

The final component of the study is the inclusion of variation in the environmental, Chapter 3. Flies will be exposed to oxidative stress, using paraquat-laced food to induce ROS and the various MEN phenotypes compared to oxidative stress and benign laboratory conditions.

In summary, the objective of this study is to quantify the biochemical characteristics of the *Men* homozygous and heterozygous alleles. The results obtained are compared to the biochemistry of these *Men* alleles with the trends previously established. Further, this study will explore across the NADPH network and see if interactions are occurring, and if there are, do the level of interactions between the *Men* alleles differ. Lastly, how the alleles respond to environmental conditions, in this case, oxidative stress will be quantified. If there are changes between the alleles, is there an advantage of an allele to another when under oxidative stress.

I hypothesized that the alleles would vary in their biochemistry at *Men* locus. These alleles will most likely show an additive response with their homozygous alleles. The differences between the alleles for the MEN phenotypes will not differ greatly; therefore, I do not expect there to be a high level of interactions with the other NADPH enzyme (previous work used excisions lines with larger-scale differences in MEN activity to understand the players in the pathway). For the environmental conditions, I hypothesized that the alleles would behave differently, compare to the control, where the trends will not be the same, and some alleles will potentially display a heterozygote advantage. This study can provide further insight into understanding the mechanisms between genetic variation and biological complexity.

1.8) References

Abolaji, A. O., Olaiya, C. O., Oluwadahunsi, O. J., & Farombi, E. O. (2017). Dietary consumption of monosodium L-glutamate induces adaptive response and reduction in the life span of Drosophila melanogaster. *Cell Biochemistry and Function*, *35*(3), 164–170. https://doi.org/10.1002/cbf.3259

Aebi, H. (1984). [13] Catalase in Vitro. *Methods in Enzymology*. https://doi.org/10.1016/S0076-6879(84)05016-3

Bernard, K. E., Parkes, T. L., & Merritt, T. J. S. (2011). A model of oxidative stress management: Moderation of carbohydrate metabolizing enzymes in SOD1-null drosophila melanogaster. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0024518

Bing, X., Rzezniczak, T. Z., Bateman, J. R., & Merritt, T. J. S. (2014). Transvection-based gene regulation in Drosophila is a complex and plastic trait. *G3 (Bethesda, Md.)*, *4*(11), 2175–2187. https://doi.org/10.1534/g3.114.012484

Boone, C., Bussey, H., & Andrews, B. J. (2007). Exploring genetic interactions and networks with yeast. *Nature Reviews Genetics*. https://doi.org/10.1038/nrg2085

Chakrabartty, A., Schellman, J. A., & Baldwin, R. L. (1991). Large differences in the helix propensities of alanine and glycine. *Nature*. https://doi.org/10.1038/351586a0

Chandler, C. H., Chari, S., & Dworkin, I. (2013). Does your gene need a background check? How genetic background impacts the analysis of mutations, genes, and evolution. *Trends in Genetics*. https://doi.org/10.1016/j.tig.2013.01.009

Chow, C. Y. (2016). Bringing genetic background into focus. *Nature Reviews Genetics*. https://doi.org/10.1038/nrg.2015.9

Cordell, H. J. (2009). Detecting gene-gene interactions that underlie human diseases. *Nature Reviews Genetics*. https://doi.org/10.1038/nrg2579

Crick, F. (1970). Central dogma of molecular biology. *Nature*. https://doi.org/10.1038/227561a0

Dworkin, I., Kennerly, E., Tack, D., Hutchinson, J., Brown, J., Mahaffey, J., & Gibson, G. (2009). Genomic consequences of background effects on scalloped mutant expressivity in the wing of drosophila melamogaster. *Genetics*. https://doi.org/10.1534/genetics.108.096453

Finkel, T., & Holbrook, N. J. (2000). Oxidants, oxidative stress and the biology of

ageing. *Nature*. https://doi.org/10.1038/35041687

Ganter, C., & Plückthun, A. (1990). Glycine to Alanine Substitutions in Helices of Glyceraldehyde-3-phosphate Dehydrogenase: Effects on Stability. *Biochemistry*. https://doi.org/10.1021/bi00492a013

Geer, B. W., Lindel, D. L., & Lindel, D. M. (1979). Relationship of the oxidative pentose shunt pathway to lipid synthesis in Drosophila melanogaster. *Biochemical Genetics*. https://doi.org/10.1007/BF00504310

Griendling, K. K., Touyz, R. M., Zweier, J. L., Dikalov, S., Chilian, W., Chen, Y. R., … Bhatnagar, A. (2016). *Measurement of Reactive Oxygen Species, Reactive Nitrogen Species, and Redox-Dependent Signaling in the Cardiovascular System: A Scientific Statement from the American Heart Association*. *Circulation Research* (Vol. 119). https://doi.org/10.1161/RES.0000000000000110

Hosamani, R., & Muralidhara. (2013). Acute exposure of drosophila melanogaster to paraquat causes oxidative stress and mitochondrial dysfunction. *Archives of Insect Biochemistry and Physiology*, *83*(1), 25–40. https://doi.org/10.1002/arch.21094

Lessel, C. E., Parkes, T. L., Dickinson, J., & Merritt, T. J. S. (2017). Sex and Genetic Background Influence Superoxide Dismutase (cSOD)-Related Phenotypic Variation in Drosophila melanogaster. *G3 (Bethesda, Md.)*. https://doi.org/10.1534/g3.117.043836

Li, G.-W., & Xie, X. S. (2011). Central dogma at the single-molecule level in living cells. *Nature*. https://doi.org/10.1038/nature10315

Merritt, T. J. S., Duvernell, D., & Eanes, W. F. (2005). Natural and synthetic alleles provide complementary insights into the nature of selection acting on the Men polymorphism of Drosophila melanogaster. *Genetics*, *171*(4), 1707–1718. https://doi.org/10.1534/genetics.105.048249

Merritt, T. J. S., Kuczynski, C., Sezgin, E., Zhu, C. T., Kumagai, S., & Eanes, W. F. (2009). Quantifying interactions within the NADP(H) enzyme network in Drosophila melanogaster. *Genetics*, *182*(2), 565–574. https://doi.org/10.1534/genetics.109.100677

Merritt, T. J. S., Sezgin, E., Zhu, C. T., & Eanes, W. F. (2006). Triglyceride pools, flight and activity variation at the Gpdh locus in Drosophila melanogaster. *Genetics*,

172(1), 293–304. https://doi.org/10.1534/genetics.105.047035

Müller, K. R., Martins, I. K., Rodrigues, N. R., da Cruz, L. C., Filho, V. M. B., Macedo, G. E., … Posser, T. (2017). Anacardium microcarpum extract and fractions protect against paraquat-induced toxicity in drosophila melanogaster. *EXCLI Journal*, *16*, 302–312. https://doi.org/10.17179/excli2016-684

Prusiner, S. B. (1990). Novel structure and genetics of prions causing neurodegeneration in humans and animals. *Biologicals*. https://doi.org/10.1016/1045-1056(90)90027-W

Rzezniczak, T. Z., Douglas, L. A., Watterson, J. H., & Merritt, T. J. S. (2011). Paraquat administration in Drosophila for use in metabolic studies of oxidative stress. *Analytical Biochemistry*. https://doi.org/10.1016/j.ab.2011.08.023

Rzezniczak, T. Z., Lum, T. E., Harniman, R., & Merritt, T. J. S. (2012). A combination of structural and cis-regulatory factors drives biochemical differences in drosophila melanogaster malic enzyme. *Biochemical Genetics*, *50*(11–12), 823–837. https://doi.org/10.1007/s10528-012-9523-3

Rzezniczak, T. Z., & Merritt, T. J. S. (2012). Interactions of NADP-Reducing Enzymes Across Varying Environmental Conditions: A Model of Biological Complexity. *G3*, *2*(12), 1613–1623. https://doi.org/10.1534/g3.112.003715

Schreiber, S. L. (2005). Small molecules: the missing link in the central dogma. *Nature Chemical Biology*. https://doi.org/10.1038/nchembio0705-64

Sezgin, E., Duvernell, D. D., Matzkin, L. M., Duan, Y., Zhu, C. T., Verrelli, B. C., & Eanes, W. F. (2004). Single-locus latitudinal clines and their relationship to temperate adaptation in metabolic genes and derived alleles in Drosophila melanogaster. *Genetics*, *168*(2), 923–931. https://doi.org/10.1534/genetics.104.027649

Steinberg, M. H. (2008). Sickle cell anemia, the first molecular disease: Overview of molecular etiology, pathophysiology, and therapeutic approaches. *TheScientificWorldJournal*. https://doi.org/10.1100/tsw.2008.157

Weber, A. L., Khan, G. F., Magwire, M. M., Tabor, C. L., Mackay, T. F., & Anholt, R. R. (2012). Genome-wide association analysis of oxidative stress resistance in Drosophila melanogaster. *PLoS One*. https://doi.org/10.1371/journal.pone.0034745

Ying, W. (2008). NAD $^+$ /NADH and NADP $^+$ /NADPH in Cellular Functions and Cell

Death: Regulation and Biological Consequences. *Antioxidants & Redox Signaling*, *10*(2), 179–206. https://doi.org/10.1089/ars.2007.1672

Chapter 2: Genetic Background and Single Nucleotide Polymorphism in the NADPH Pathway in *Drosophila melanogaster* at the *Malic Enzyme* Locus

2.1) Introduction

Given the huge amount of genetic data generated by modern genomics, our lack of understanding of biological molecular variation is striking. What is the effect of genetic variation? A fundamental task of modern genetics and genomics is still to uncover the path between genetic diversity and biological complexity. Part of this task is examining the biological effect of simple polymorphisms. The impact of these sites have been a central question since the earliest studies of allozymes and genetic molecular variation.

Genetic variation at the *Drosophila melanogaster Malic enzyme* (*Men*) locus is an interesting model system for the study of the biology of molecular variation in general (Merritt et al. 2005; Rzezniczak et al. 2012). Malic enzyme is a metabolic enzyme that oxidizes malate to pyruvate while reducing $NADP^+$ to NADPH(Merritt et al. 2009, 2005; Rzezniczak et al. 2012). In *D. melanogaster*, there are two known amino acid changing single nucleotide polymorphisms at the *Men* locus. The first polymorphism, at base pair 338, is a guanine to cytosine substitution that results in a glycine to alanine change at amino acid 113. This site is in an alpha helix close to the MEN active site. The two alleles occur in approximately equal frequency across North America, but only the guanine allele is found in Africa and *D. simulans*, suggesting that it is the ancestral condition (Rzezniczak et al. 2012; Sezgin et al. 2004). The second polymorphism, base pair 1051, is an adenine to thymine substitution that results in a methionine to leucine change at amino acid 351. The leucine allele is found at a frequency of 14% in at least two North American populations, but has not been found in African sites or *D. simulans*, suggesting methionine is the ancestral condition (Rzezniczak et al. 2012). This site is in a beta sheet buried in the protein, not in close proximity to the active site. Malic enzyme functions as a homotetramer. In homozygous individuals, at the *Men* locus, there will be a single form of the enzyme. However, in individuals heterozygous at one site, there will be five possible different enzymes (e.g. four alanines, three alanines and one glycine, two alanines and two glycines, etc.) For both sites, there is the possibility of 25 different enzymes (e.g. four alanine and four methionine, four alanine, three methionine and one leucine, etc.).

Both polymorphisms significantly affect the biochemistry of the MEN enzyme, suggesting that the sites could affect fitness and are under selection (Merritt et al. 2009, 2005; Rzezniczak et al. 2012). Using homozygous isochromosomal flies, we quantified that both polymorphic site influences the K_m for malate, with the glycine and leucine alleles having a higher K_m. Further, the first polymorphism, but not the second, influences the V_{max}; the glycine allele has 46% higher activity than the alanine. The glycine amino acid alleles are associated with higher levels of expression, but the differences in expression do not fully explain the differences in activity. This indicates that the differences are a function of both structural and regulatory variation (Rzezniczak et al. 2012). Similarly, only the first polymorphism influences MEN thermal stability with the glycine amino acid alleles being more thermostable.

The biochemical effects, widespread distribution, and the consistency of their frequencies, suggest that the alleles are under selection (Merritt et al. 2005). The approximately equal frequency of the alleles at the first site mean that most individuals will be heterozygous and is consistent with the possible heterozygote advantage. If heterozygote advantage is maintaining the alleles, then a prediction is that the biology of the heterozygote lies outside that of either homozygous genotypes (Charlesworth and Charlesworth 1987; Dudash and Fenster 2000). The implication of the rare allele at the second are less clear. It could be that this is a recent mutation, but its presence in at least two geographically distant population argues against this and suggest that it is being maintained at the population level. The biased frequency does mean that homozygotes of rare allele will be extremely rare and that it will predominantly occur in the heterozygote. We test this prediction for both these sites using a suite of biochemical and complex phenotypes. One of the first heterozygote advantage study completed was done in *Drosophila melanogaster*, examining the ebony allele (Kalmus 1945). In this study, 14 months of selection was carried out, and found that the heterozygotes were superior to both homozygotes at higher temperatures and at lower temperature increased the frequency of ebony flies emerging. An example of heterozygote advantage within humans is at the *Triosephosphate isomerase* (TPI) locus, where one mutation, coding for aspartic acid instead of glutamic acid, causes inactivation of the enzyme (Ralser et al.

2006). This null allele as homozygous causes a TPI deficiency, however, as heterozygote, it shows no effects to the organism. The frequency of the heterozygote allele (one null allele and one functional) is higher than excepted, also indicating a heterozygote advantage based on population genetics.

Most biological traits are complex, influenced by multiple genes and environmental factors and potential selection advantages or disadvantages of genetic variation is complicated by interactions across these multiple factors (Bernard et al. 2011; Mackay 2004; Wolf 2003). In previous studies, we showed that *Men* is an important contributor to the maintenance of the $NADP^+$/NADPH ratio within the cell (Merritt et al. 2009; Rzezniczak et al. 2012; Rzezniczak and Merritt 2012; Ying 2008). NADPH is a major reductant in lipid synthesis and oxidative stress resistance (Bernard et al. 2011; Lessel et al. 2017; Merritt et al. 2006; Rzezniczak and Merritt 2012; Wise and Ball 1964). MEN shares this role with at least three other enzymes that reduce $NADP^+$, Isocitrate dehydrogenase (IDH), Glucose-6-phosphate dehydrogenase (G6PD) and 6-phosphogluconate dehydrogenase (6PGD). These four enzymes interact and are co-regulated, likely through cellular monitoring of the $NADP^+$/NADPH ratio (Merritt et al. 2009; Rzezniczak and Merritt 2012). These interactions are often compensatory (reduction of one source of NADPH results in up regulation of others), but can also be parallel and counterintuitive (reduction of one source of NADPH results in down regulation of others), indicating a complex regulatory network (Rzezniczak and Merritt 2012). It has been estimated that MEN produces about 30% of available NADPH in *D. melanogaster*. The remaining portion of NADPH is produced from IDH (about 20%) and G6PD and 6PGD (combined 40%) (Geer et al. 1979). Previous studies have found that an extreme reduction in MEN activity can lead to an increase in G6PD and 6PGD activity, however a reduction in MEN can lead to a reduction in IDH (Rzezniczak and Merritt 2012). This work displays the importance of studying a gene with their respective influencers rather than in isolation (Chow 2016; Dixon et al. 2009; Huang et al. 2014).

Here we expand our investigation of the biochemical characteristics of the MEN amino acid polymorphisms to include heterozygote genotypes and a greater number of genetic

backgrounds. By including the heterozygotes, we are able to better understand the regulatory and structural variation between the alleles, which may explain the differences in allelic frequencies and overdominance at *Malic enzyme*. Furthermore, studying heterozygotes are of interest to us because they better represent a population, since heterozygotes tend to exist at a higher frequency than homozygotes. Examining the heterozygotes, we find that both polymorphic sites significantly impact the biochemistry of the enzyme. In addition, our results are consistent with previous work mentioned above.

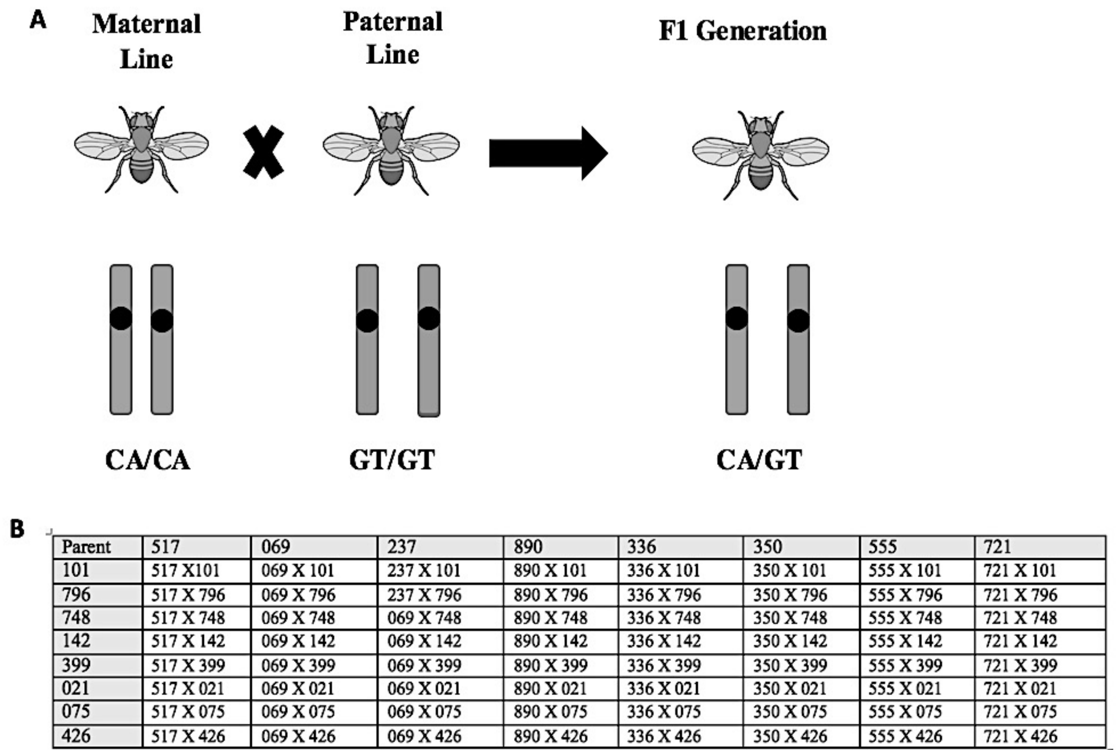
2.2) Materials and Methods

2.2.1) Fly Stocks and Lines

Fly lines are a subset of the DGRP lines obtained from the Bloomington Drosophila Stock Centre (Bloomington, USA). Fly lines were selected based on genotype at the two known polymorphism sites in the Malic enzyme gene: Position 351, a G-C polymorphism that results in a G-A substitution and position 1051, an A-T polymorphism that results in a M-L substitution. Flies were also selected based on parental V_{max} activity (mid-activity) and IDH and G6PD alleles to be consistent with one another. A lists the parental lines used and their respective genotypes can be found in Table 2.1. Parental lines were crossed with one another to create heterozygotes with the following genotypic combinations: C/C, A/A (CCAA), C/C, A/T (CCAT), C/C, T/T (CCTT), C/G, A/A (CGAA), C/G, A/T (CGAT), C/G, T/T (CGTT), G/G, A/A (GGAA), and G/G, T/T (GGTT). A representation of how the crosses were complete to obtain the desired genotypes is found in Figure 2.1A. All crosses included five adult male flies from one line paired with five virgin female flies from another line. Flies were maintained on a standard cornmeal medium with 12:12-hr light dark cycle at 25ºC. Emerging male flies were aged to 3-6 days, frozen, and stored at -80 ºC until further analysis.

Sex	DGRP Line #	Site 351	Site 1051
Maternal	517	C	A
Maternal	069	C	A
Maternal	237	G	A
Maternal	890	G	A
Maternal	336	C	T
Maternal	350	C	T
Maternal	555	G	T
Maternal	721	G	T
Parental	101	C	A
Parental	796	C	A
Parental	748	G	A
Parental	142	G	A
Parental	399	C	T
Parental	021	C	T
Parental	075	G	T
Parental	426	G	T

Table 2.1: DGRP lines, and their genotypes, used to construct homozygous and heterozygous genotypes. The sex column distinguishes with sex was used, maternal being females and parental being males. The DGRP Line # indicates which specific DGRP line was used. The last two columns state the genotypes at the first and second polymorphic sites.



Parent	517	069	237	890	336	350	555	721
101	517 X101	069 X 101	237 X 101	890 X 101	336 X 101	350 X 101	555 X 101	721 X 101
796	517 X 796	069 X 796	237 X 796	890 X 796	336 X 796	350 X 796	555 X 796	721 X 796
748	517 X 748	069 X 748	069 X 748	890 X 748	336 X 748	350 X 748	555 X 748	721 X 748
142	517 X 142	069 X 142	069 X 142	890 X 142	336 X 142	350 X 142	555 X 142	721 X 142
399	517 X 399	069 X 399	069 X 399	890 X 399	336 X 399	350 X 399	555 X 399	721 X 399
021	517 X 021	069 X 021	069 X 021	890 X 021	336 X 021	350 X 021	555 X 021	721 X 021
075	517 X 075	069 X 075	069 X 075	890 X 075	336 X 075	350 X 075	555 X 075	721 X 075
426	517 X 426	069 X 426	069 X 426	890 X 426	336 X 426	350 X 426	555 X 426	721 X 426

Figure 2.1. A. Schematic of crosses performed to generate desired alleles. A virgin maternal line from Table 2.1 is crossed to a parental line to generate the desired allele in the F1 generation. B. All the crosses performed to generate. The grey row represents the maternal lines used and the grey column represents the parental lines used. The white boxes are the crosses that were completed.

2.2.2) Fly Wet Weight

Flies were weighed to the nearest 0.01 mg using a Mettler Toledo microbalance MX5. Weight was used as a covariate in analyses of covariance (ANCOVA) to standardize enzyme activities, triglycerides and carbohydrate concentration for differences in fly size.

2.2.3) Fly Homogenization

Assays were conducted using whole fly homogenates created by homogenizing samples of five flies in grinding buffer (100 mM Tris-HCl, 0.15 mM $NADP^+$, pH 7.4) at a concentration of one fly per 100 μL of buffer, and centrifuged at 13 000 RPM for 10 min at 4 ºC to pellet all solids. Homogenate supernatant was collected and 300 μL of supernatant was transferred to a 96-well plate; aliquots were sampled from this master plate for each subsequent analysis.

2.2.4) Enzyme Activity Measurements

Enzyme activity was quantified by measuring the production of NADPH through time. Assays were conducted on a 96-well plate spectrophotometer (Molecular Devices SpectraMax 384 Plus), using 10 μL of fly homogenate and 100 μL of assay buffer (described below). Absorbance at 340 nm (the wavelength of absorbance of NADPH) was measured every 9 s for 3 minutes at 25 ºC, with the exception of the G6PD activity assays, which was measured for 5 min. 6PGD activity was not quantified since it directly follows G6PD in the pentose phosphate pathway and another study shown a correlation between G6PD and 6PGD activities (Wilton et al. 1982). Samples were assayed twice and the means were used for further analysis. The assay buffers were previously optimized to give maximum activities (Merritt *et al*. 2005, 2009) and were as follows:

G6PD: 100 mM Tris- HCl, 0.32 mM NADP, 3.5 mM D-glucose-6-phosphate (pH 7.4)

IDH: 100 mM Tris-HCl, 0.10 mM NADP, 0.84 mM $MgSO_4$, 1.37 mM DL-isocitrate (pH 8.6)

MEN: 100 mM Tris-HCl, 0.34 mM NADP, 50 mM $MnCl_2$, 50 mM malate (pH 7.4)

2.2.5) Estimation of Michaelis-Menten Constant

Genotype-specific Michaelis-Menten constants for malate were determined using a 10-point geometric design method following Rzezniczak et al (2012). Briefly, rates for each genotype were determined using the method described above, but using 10 substrate concentrations ranging from 0.1 mM-10 mM, each measurement was done in replicate. The average value for each of the 10 concentrations was then used to determine the Michaelis-Menten constant using the program GraFit 7.0 software.

2.2.6) Thermal Stability of Malic Enzyme

MEN thermal stability was estimated by following MEN activity decline over time (0-10 min) at 50 ºC, as described in (Hall 1985; Rzezniczak et al. 2012). Briefly, for each genotype, a control sample was kept on ice while matched samples were incubated at 50 ºC. Single heat-exposed aliquots were removed and placed on ice at interval of 1 minutes. At the end of 10 minutes, all aliquots were kept on ice until their MEN activities were measured. The activity of each sample was compared to the control sample to determine the proportion of activity remaining at each time point. The decline in enzyme activity with time was treated as a first-order exponential decay process. Denaturing constant (k_D) was determined by the relationship $\left(\frac{E}{E^o}\right)_t = e^{-k_D t}$ where $\left(\frac{E}{E^o}\right)_t$ is the proportion of initial enzyme activity remaining at the time t and k_D is the denaturation constant (Hall 1985). The slope of the line from linear regression of $ln\left(\frac{E}{E^o}\right)_t$ on the time is estimate of k_D. A mean k_D was calculated for each sample.

2.2.7) Soluble Triglyceride Content

Soluble triglyceride was measured using a commercially available kit (Triglyceride-SL Assay, Pointe Scientific, Canton, MI, Catalog No. T7531) following manufacturer's protocol. Briefly, each assay, 10 µL of homogenate and 100 µL of reagent, was incubated at 37 ºC for 5 min. Sample absorbance was measured at 500 nm and total soluble triglycerides concentrations were determined by comparison with commercially available standard (Pointe Scientific, Canton, MI, Catalog No. T7532). Each sample was assayed

twice, and the mean used in analysis. Results are reported as micrograms of triglycerides per sample.

2.2.8) Total Carbohydrate Content

Total carbohydrate content was measured as previously described (Merritt *et al*. 2006). Briefly, complex carbohydrates were converted to glucose using a digestion cocktail that contained 10 μL of fly homogenate sample and 2 μL of amyloglycosidase (Sigma Aldrich, St Louis, MO, A1602) at a concentration of 1 unit/sample in 2.0 M sodium acetate buffer (pH 5.7). Samples were digested at 55 ºC for 45 min. Following digestion, total glucose was measured using commercially available kit (Genzyme, Cambridge, MA, Catalog No. 23517) in which 10 μL of digested homogenate was combined with 200 μL of glucose reagent and incubated at 37 ºC for 10 min. Sample absorbance was measured at 340 nm and total carbohydrate concentration was determined by comparison to a glycogen standard (Sigma Aldrich, St Louis, MO, Catalog No. G0885). Results are reported as milligrams per liter.

2.2.9) Soluble protein content

Soluble protein was measured using the bicinchoninic acid (BCA) assay using a commercially available kit (Pierce, Thermo Scientific, Rockford, IL, Catalog No. 23225) following the manufacturer's protocol. In brief, assays contained 10 μL homogenate and 100 μL reagent and were incubated at 37 ºC for 30 min. Sample absorbance was measured at 562 nm, and total soluble protein concentrations were determined by comparison with bovine serum albumen standards (Sigma Aldrich, St Louis, MO, A4503). Each sample was assayed twice, and the mean was used in analysis. Soluble protein content was used as a covariate in analyses of covariance (ANCOVA) for enzyme activities, triglyceride and carbohydrate concentration to standardize for differences in fly size and homogenization.

2.2.10) RNA Extraction and Quantitative RT-PCR

RT-qPCR was used to test for differences in *Men* expression across the CCTT genotype set. Total RNA was isolated from three groups of five 3 to 5 day old male flies using the

RNeasy Kit (QIAGEN, Valencia, CA) according to manufacturer's instructions, and was stored at -80 ºC until needed for reverse transcription. For each sample, one microgram of total RNA was reverse-transcribed using random hexamers and High Capacity cDNA Reverse Transcription Kits with RNase Inhibitor (Applied Biosystem) The qPCR reaction consisted of 2 μL of undiluted cDNA template, 0.4 μM of each primer, and 0.2 μM probe, and Quantitect Probe PCR Master Mix (QIAGEN); a total volume of 20 μL. The primers and the probe flank the *Men* intron between exon 2 and exon 3 (5' GTATTGCCAACCTGTGCC, 3' AGCTTGTGTTCGGTGAGT and probe 56-FAM/ATGGTGGATAGCCGTGGTGTCA/3IABkFQ. cDNA synthesis of samples lacking reverse transcriptase were used as a negative control to ensure that there was no genomic DNA contamination and "no-template" blanks were used to ensure there was no contamination. Two reactions per template were performed using a Mastercycler Ep Realplex Thermal Cycler. Expression results were normalized to *RpL32* and are reported relative to *MenEx3* using ΔΔCT method (Livak and Schmittgen 2001).

2.2.11) Longevity

Adult male flies of from each cross were collected 48 hours post eclosion and transferred into fresh vials with a maximum of 20 flies per vial. Vials were maintained at 25ºC, 12hr light: dark cycle and mortality was recorded every two days. Flies were transferred to fresh media every seven days until no living flies remained.

2.2.12) Data Analysis

All crosses were replicated in two independent vials, with three samples taken from each vial. Assays were run in duplicates to account for technical error. ANCOVA and Tukey's honesty significant difference (HSD) multiple-comparison tests were performed using JMP 12.0 software (SAS Institute) to determine whether there were significant differences in enzyme activity, triglyceride or carbohydrate concentration using protein concentration and wet weight as covariates.

2.3) Results

The G-A polymorphism has a much larger effect on maximum Malic Enzyme reaction velocity (V_{max}) than the M-L polymorphisms

Both MEN polymorphisms affect the enzyme's V_{max} for malate, although to different degrees. The glycine-alanine (G-A) polymorphism at position 113 was associated with significant and substantial differences in V_{max} (Figure 2.2A, left-hand columns). Flies that are homozygous for the glycine amino acid, G/G (hereafter noted as GG, with similar nomenclature for each site and genotype, e.g. A/A is noted as AA, M/M as MM, A/G as AG, M/L as ML and A/G, M/L as AGML), had a 28% higher V_{max} than those homozygous for the alanine amino acids, A/A (hereafter noted as AA), consistent with previous work conducted using a distinct set of isothrid chromosome lines ($F_{8,372}$=15.8471, $P < 0.0001$; Rzezniczak et al. 2012). In contrast, the methionine-leucine (M-L) polymorphism at position 351 was not associated with any significant difference in the observed V_{max} (Figure 2.2A, right-hand columns), but there were significant interactions between the two sites (Figure 2.2B, white bars). The GGLL and GGMM had a 20% higher observed V_{max} than the alanine and leucine combination (AALL), and a 41% higher observed V_{max} than the alanine and methionine combination (AAMM).

A focus of this study was to compare the characteristics of the heterozygotes with those of the homozygous genotypes to identify a possible biochemical basis for any heterozygote advantage. In testing the effects of the two polymorphisms, we created a total of nine genotypes to explore all possible allelic combinations of the sites (Figure 2B). In all cases, the heterozygous genotypes, shown in black bars, are not significantly different from an average of their respective homozygous genotypes, shown in white bars. All of the heterozygote combinations fall within ± 11% of expected V_{max} based on the values observed for the homozygous flies (Table 2.2). If heterozygote advantage is maintaining either of these polymorphisms, it is not the function of non-linear biochemistry of the heterozygous combinations.

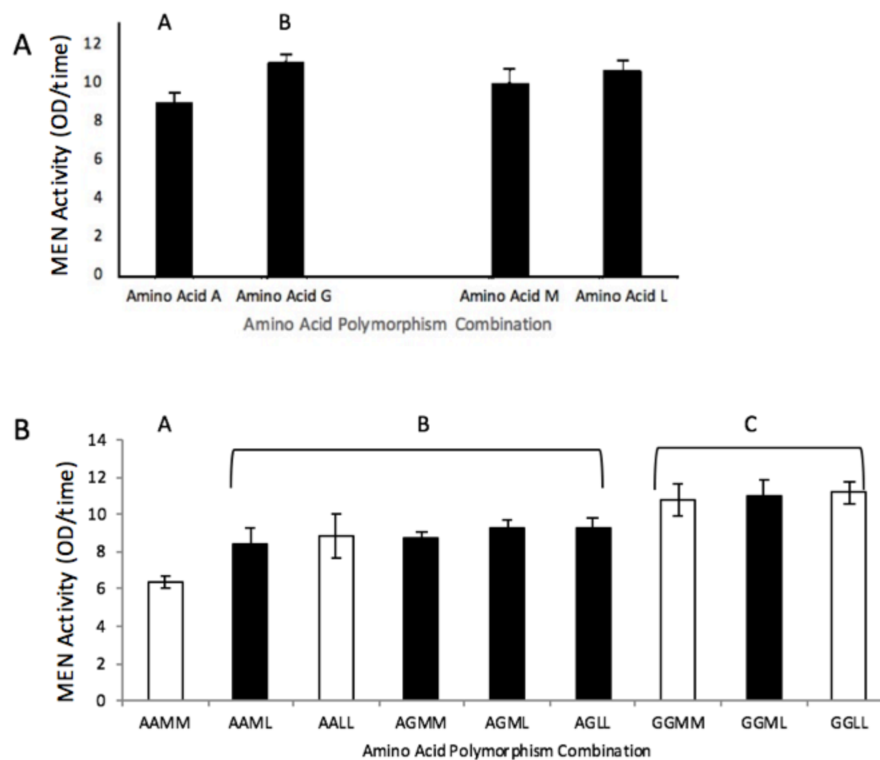


Figure 2.2. Average maximum MEN enzyme activity (OD/time) was estimated for each genotype as observed enzyme activity under saturating conditions. Bars mean ± standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different. (A) MEN enzyme activity for the first and second polymorphic site respectively. The first two set of bars represents the G/A polymorphic site, and the second set of bars represents the M/L polymorphic sites (B) MEN enzyme activity for all genotypes that were generated from the crosses. The white bars represent the alleles that are homozygous for the polymorphic sites and the black bars represents where the alleles are heterozygous for one or both polymorphic sites.

	AAML	AGMM	AGML	AGLL	GGML
V_{max}	7.5859 ± 0.489	8.5543 ± 0.455	9.2785 ± 0.339	10.0028 ± 0.4277	10.97125 ± 0.5502
K_m	0.5778 ± 0.0302	0.8096 ± 0.0179	0.6500 ± 0.5532	0.49041 ± 0.0156	0.7222 ± 0.01854
V_{max}/K_m	14.715 ± 0.487	10.39670 ± 0.435	15.4172 ± 0.461	20.43777± 0.376	16.11918 ± 0.4350

Table 2.2. Predicted MEN activities for Heterozygote alleles derived from the homozygous alleles. The homozygous alleles are AAMM, AALL, GGMM and GGLL. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different.

In addition to testing the polymorphisms' effects on V_{max}, and possible heterozygote difference, we tested the effect of genetic background within the genotype groups. Multiple crosses of independent parental lines were used to create each genotype (Figure 2.1) and the different line combinations has substantial differences in V_{max}. The crosses that were used to generate the AAMM, AALL and AAML genotypes are shown in Figure 2.3A-C (all alleles were analyzed, only AAMM, AALL and AAML are shown in Figure 2.3). The difference in V_{max} between the AALL lines and the AAMM is a function of one cross, 350 X 021, which has a much higher V_{max} than the rest. If this cross is removed, the two allelic combinations have similar V_{max}. This fact does not diminish the allelic effects so much as highlight the important of line effects; on average the alleles are significantly different but the differences between lines are as substantial as the differences between the alleles.

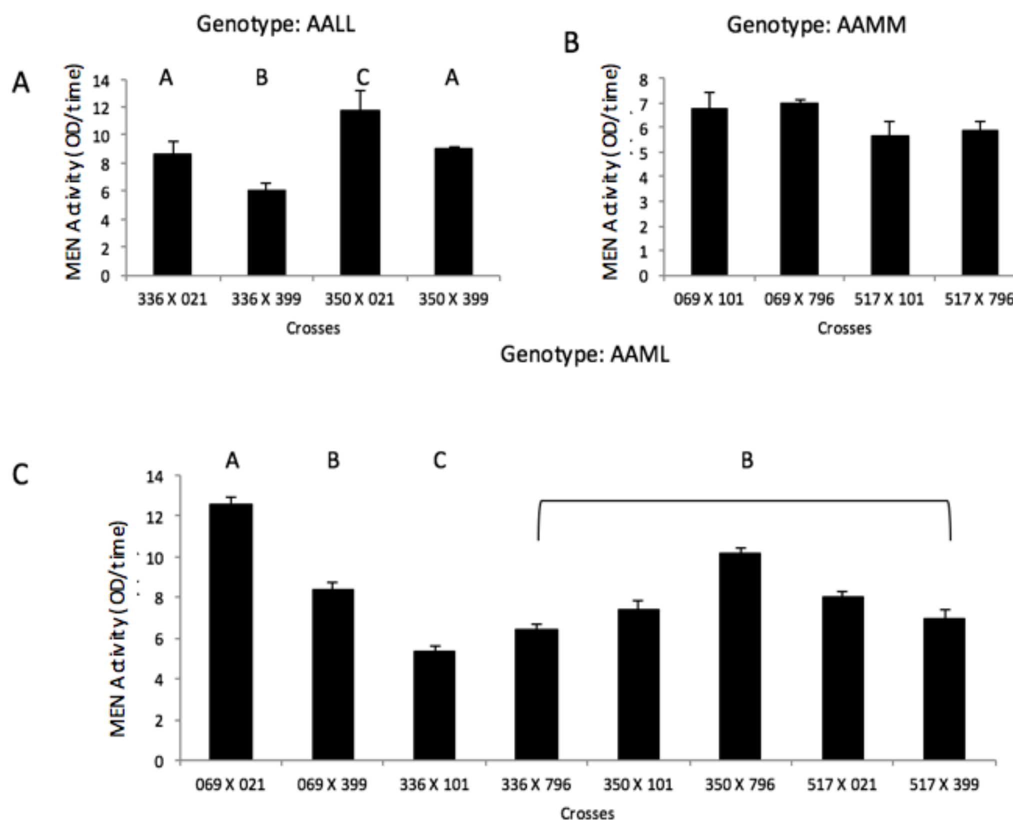


Figure 2.3. Average maximum MEN enzyme activity (OD/time) was estimated for various crosses that contribute to the specific genotypes. Bars mean ± standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different. (A). MEN activity is plotted against the crosses that contribute to the allele AALL. (B) MEN activity is plotted against the crosses that contribute to the allele AAMM. (C) MEN activity is plotted against the crosses that contribute to the allele AAML.

***Men* expression and MEN activity are correlated**

Malic enzyme protein activity and gene expression are correlated, but the relationship is sensitive to genetic background (Chandler et al. 2013). MEN V_{max} is a function of both the amino acid composition of the enzyme and the level of gene expression (Rzezniczak et al. 2012) and our V_{max} data are consistent with the G-A polymorphism modifying enzyme activity with potentially additional variation due to differences in regulation of expression across the alanine amino acid containing lines. To test this possibility, we measured *Men* gene expression using quantitative RT-PCR (qRT- PCR), and compared it to MEN activity. We compared *Men* gene expression and MEN enzyme activity within the AALL genotype using eight genotypes, four heterozygotes lines and their respective homozygous parent lines: 336/021, 336/399, 350/021 and 350/ 399 and 336/336, 021/021, 399/399, 350/350. As expected, MEN enzyme activity and *Men* activity were correlated (Figure 2.4). In addition to the DGRP wild-derived alleles, we included a pair of lines with laboratory engineered *Men* alleles, *MenEx55*, a knockout allele, and *MenEx3*, a wild type allele (alleles are described in Lum and Merritt 2011), which allowed us to create "known" 100% and 50% *Men* expression and MEN activity genotype, *MenEx3/Sb* and *MenEx55⁻/Sb* (Bing et al. 2014; Lum and Merritt 2011). MEN activities are shown in Figure 2.4A and the relationship between activity and expression is shown in Figure 2.4B. Not surprisingly, there is a strong overall correlation between gene expression and protein activity within these flies. Also not surprisingly, there is more error associated with the gene expression data than the protein activity data, which is one reason that we generally quantify the later even in cases where phenotypes are driven by the former (Bing et al. 2014; Lum and Merritt 2011).There is, however, substantial variation in the relationship between the two values with some genotypes, strikingly the 021/021 flies, lying far from the average relationship; these flies have about 103% greater protein activity that their gene expression would predict. The correlation between protein activity and *Men* expression is a R^2 value of 0.69689, upon the removal of line 021/021, whereas with line 021/021 included, R^2 value is 0.11022). In addition, gene expression is not simply the average of the parent lines. The 350/021 flies, for

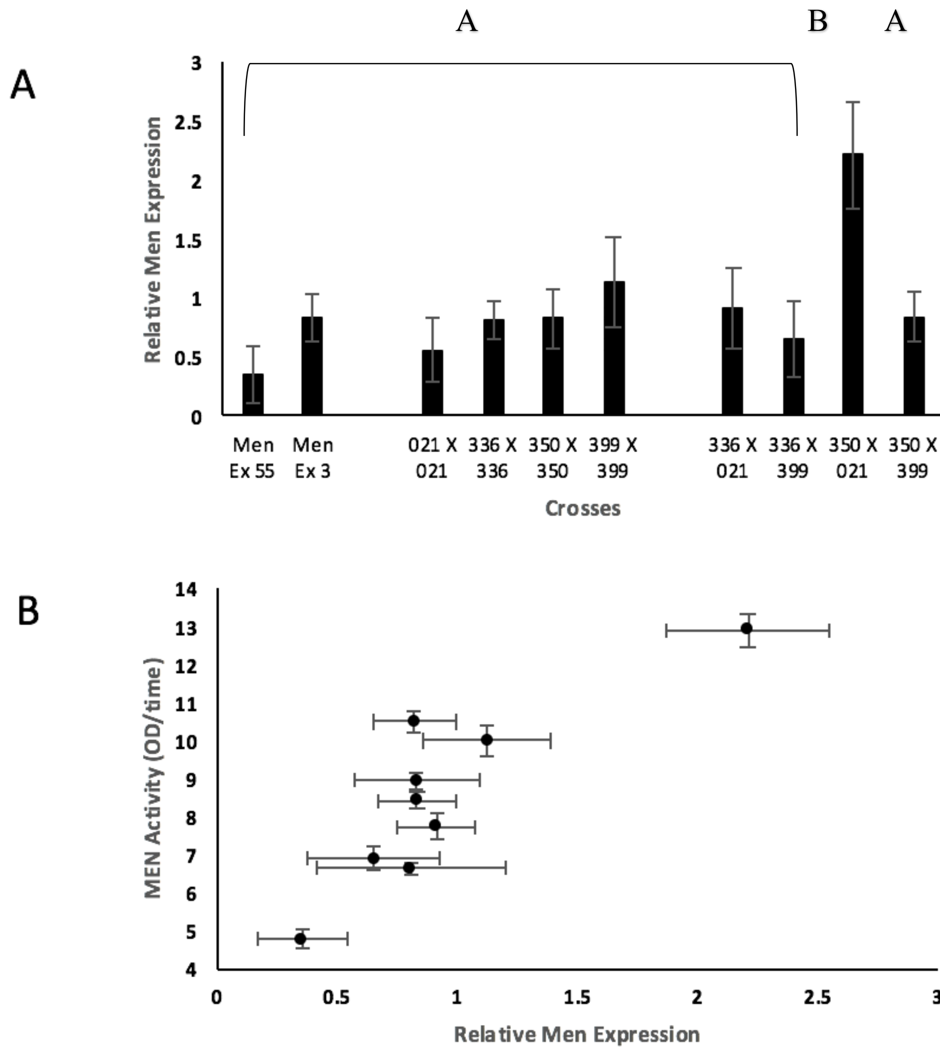


Figure 2.4. Correlation between MEN protein activity and relative *Men* gene expression. (A) Relative *Men* expression is plotted against AALL genotypes. MENEx55 and MenEx3 represent the synthetic lines used to quantify the assay. The crosses used generated the AALL genotype, plus the parent lines for that cross. Expression was normalized to MENEX3^{+}. Relative expression as calculated using the ΔΔCT method (Livak and Schmittgen 2001). (B). Correlation between MEN activity against Relative *Men* expression.

example, have twice the gene expression that would be predicted by the parent lines, suggesting broad interaction between regulatory regions. Combined, this variation indicates that gene expression of *Men* is a non-linear combination of its parent's lines, indicating that there is the potential of other factors leading to the expression levels, therefore influencing the V_{max} of MEN we seen above.

The M-L polymorphism has a larger effect on the Malate Michaelis-Menton constant

Both the G-A and M-L polymorphism significantly alter the MEN K_m for malate, but the M-L polymorphism has a much larger effect than the G-A polymorphism (Figure 2.5A-B). All four homozygous genotypes, AAMM, AALL, GGMM and GGLL, were significantly different (Figure 5B; $F_{8, 372} = 49.9519$, $P < 0.0001$). The flies homozygous for the GG polymorphism had a 19% higher K_m than flies homozygous for the AA polymorphism and the MM flies had a 41% higher value than that of the LL flies. Interestingly, flies homozygous for this rare allele had strikingly lower K_m values than any other genotype. Similar to the V_{max} results, all of the heterozygous combinations had K_m values intermediate between those of the homozygous genotypes (Figure 2.5B); all experimental values were not significantly different from the calculated values, falling within ± 10% of those values (Table 2.2). Previous work has been split on the question of genetic background effects on MEN malate K_m, with an early study finding effects (Merritt et al. 2005) and a later study filing to find any such effects (Rzezniczak et al. 2012). Comparison of three genotypes, AAMM, AAML and AALL, demonstrates the differences in the genetic background effects on K_m that we observed (Figure 2.6); some genotypes had large background effects (Figure 2.6A), while others showed little or only a few crosses that were significantly different from the rest (Figure 2.6B-C). Overall, our results indicate that genetic background can modify $K_{m,}$ e.g. that K_m is not strictly based on the amino acid composition, since variation was found in Figure 2.6A.

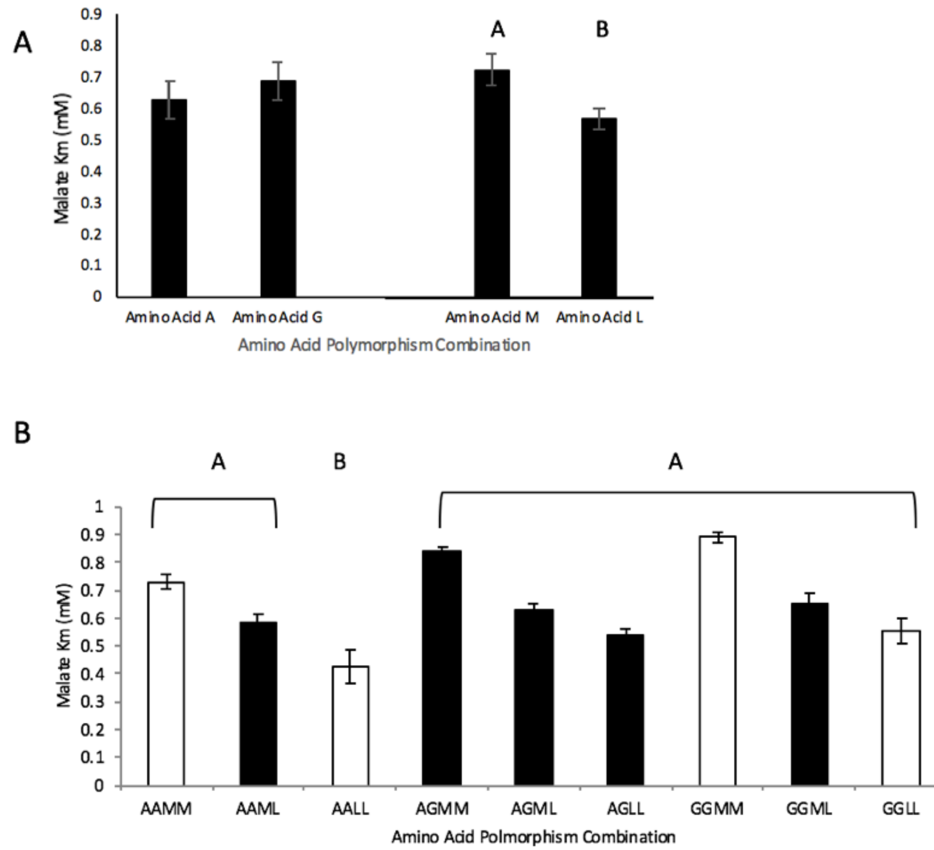


Figure 2.5. Michaelis-Menton constant (K_m) for malate. Substrate-binding kinetics were quantified as the K_m for malate for each genotype. Units are μM/L. Bars mean ± standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different. (A) MEN binding affinity for both polymorphic sites, the first two bars are the G/A polymorphic site and the second set of bars are the M/L polymorphic site. (B) MEN binding affinity for all genotypes, where the white bars represent the alleles that are homozygous for both polymorphic sites, and the black bars represent that alleles that are heterozygous for one or both polymorphic sites.

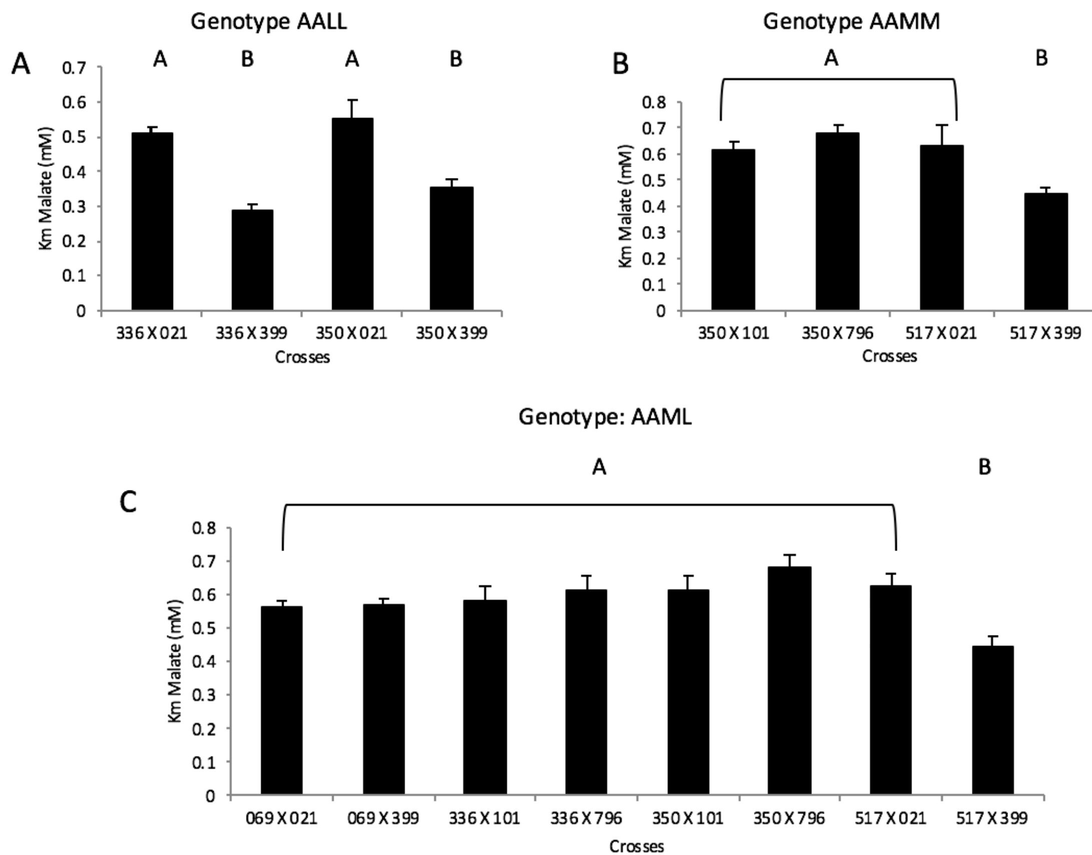
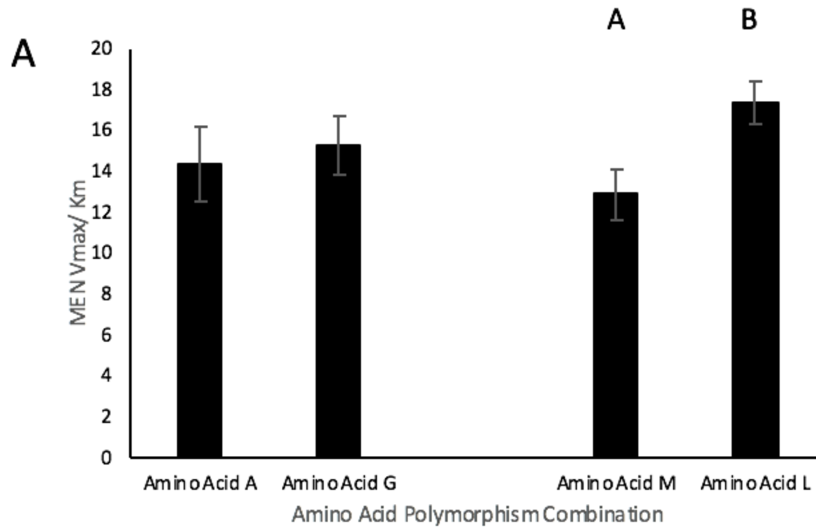


Figure 2.6 Average MEN binding affinity various crosses that contribute to the specific genotypes. Bars mean ± standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different. (A) K_m Malate is plotted against the crosses for the genotype AALL. (B) K_m Malate is plotted against genotype AAMM. (C) K_m Malate is plotted against genotype AAML.

The M-L polymorphism has a larger effect on the V_{max}/K_m Ratio

Previous work suggests that the relative *in vivo* enzyme activity is better estimated by the ratio of V_{max} to K_m than either phenotype alone (Hall and Koehn 1983; Merritt et al. 2005; Watt and Dean 2000). Overall, the ratio is very similar in the GG and AA alleles, but is significantly different between the MM and LL alleles (Figure 2.7A; $F_{8, 372} = 24.086$ $P < 0.0001$). Interestingly, because of the large differences in the denominator Km, the higher V_{max} genotypes have lower relative activities and the M-L polymorphism, which is not located in close proximity to the active site, has the largest influence on the relative enzymatic activity. The M-L polymorphism influencing the activity is interesting because the leucine amino acid has low K_m, has higher V_{max}, despite being the rarer allele. There is substantial variation in the relative activity across the heterozygotes genotypes (Figure 2.7B), although as with V_{max} and K_m, the values are generally intermediate between those of the homozygous genotypes. The predicted values and experimental values are not significantly different with almost all of the heterozygote combinations falling within ± 6% of predicted values. The AGLL flies have slightly lower than expected relative activities, 15% less of the predicted value (Table 2.2).



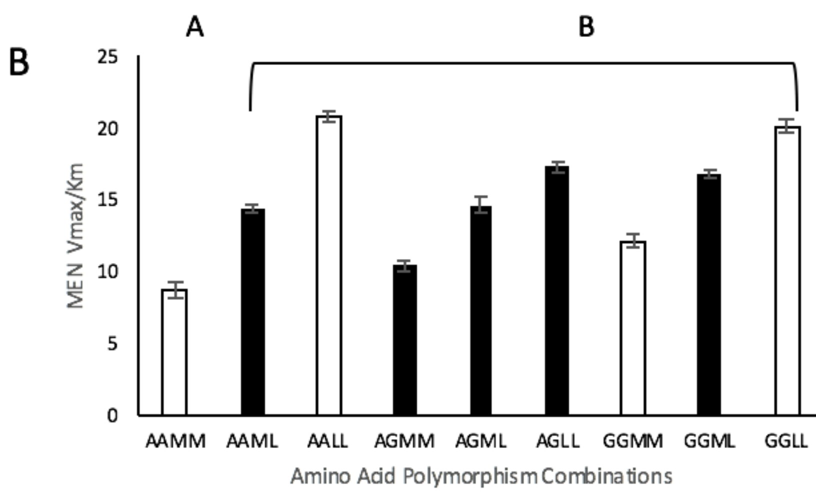


Figure 2.7. Maximum MEN velocity to Michaelis-Menton constant for malate ratio. Bars mean ± standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different. (A) MEN Vmax/Km ratio for both polymorphic sites, where the first 2 bars are the G/A polymorphic site and the second bard is the M/L polymorphic site. (B) MEN Vmax/Km ratio for all genotypes, where the white bars represent the alleles that are homozygous for the polymorphic sites and the black bars represents where the alleles are heterozygous for one or both polymorphic sites.

Only the G-A polymorphism affects Thermal Stability

We quantified thermal stability as a general indicator of the structural stability of Malic enzyme (Figure 2.8). Similar to previous studies, we found that the G-A polymorphism had a slight effect on stability, where GG allele has higher stability than that of the AA allele, but not significant (Figure 2.8A). Also similar to previous studies, we found that the M-L polymorphism did not influence thermal stability (Figure 2.8A, $F_{8, 372} = 1.9574$, $P=0.1458$). Comparing the genotypes, the AGML and AGLL combination resembles the thermal stability constant of the observed values for all the alleles that are homozygous for the GG alleles; the alleles do not significantly differ from one another (Figure 2.8B; $F_{8, 372} = 0.9213$, $P = 0.5167$). Even though past studies found differences in thermal stability between the polymorphisms, here, this was not the case. The lack of consistency between studies in terms of thermal stability likely reflect the small absolute differences in stability between the alleles and the different genetic backgrounds used in each study.

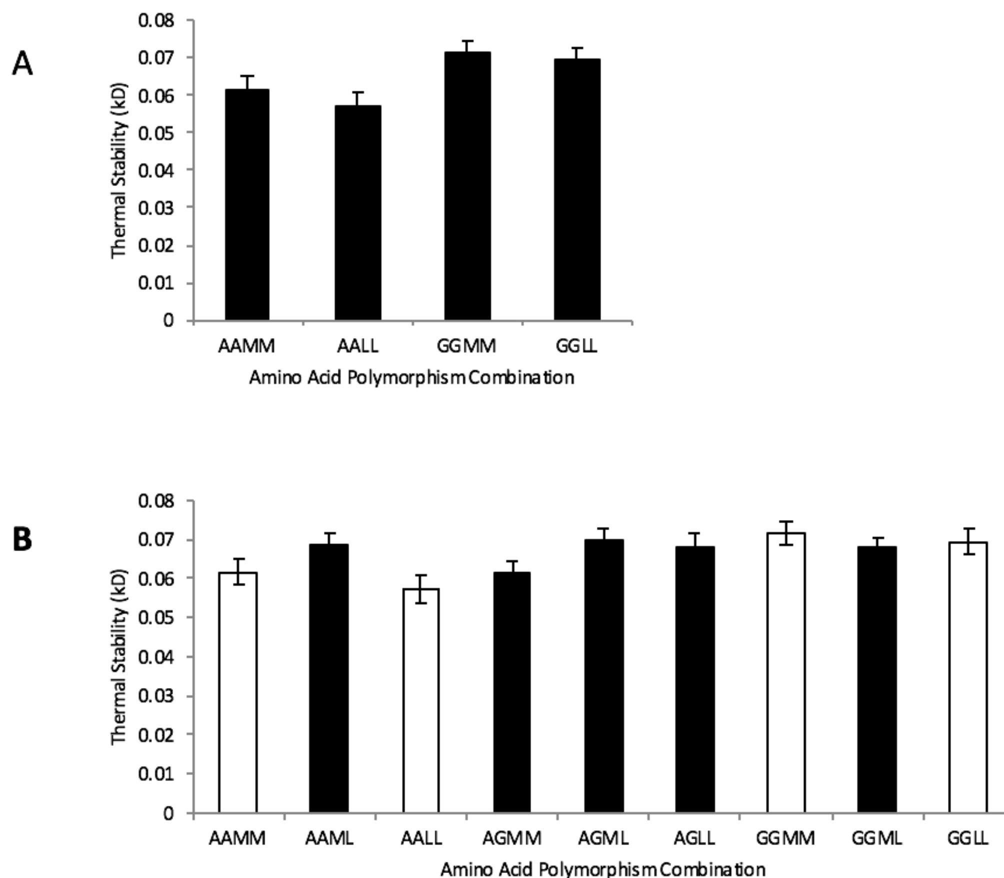


Figure 2.8. Malic enzyme thermal stability. Thermal stability was measured for each genotype as a proxy for overall structural stability of the enzyme and was quantified as absolute value for K_D of the enzyme after denaturing at 50 ºC. Bars mean ± standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different. (A) MEN K_D for both polymorphic sites, where the first 2 bars are the G/A polymorphic site and the second bard is the M/L polymorphic site. (B) MEN K_D for all genotypes, where white bars represent the alleles that are homozygous for the polymorphic sites and the black bars represents where the alleles are heterozygous for one or both polymorphic sites.

The polymorphic sites are not correlated with changes in G6PD and IDH activity

Malic enzyme activity is only one of the major biological sources of NADPH, the reduced cofactor, in the fly and previous work has shown that it interacts with other sources, specifically IDH and G6PD. In earlier studies, we found interactions between MEN activity and the activities of both of these enzymes, however over large-scale (engineered) differences in V_{max} (Merritt et al 2005). Overall, the MEN amino acid polymorphisms are not associated with any significant differences in G6PD activity (Figure 2.9A, $F_{8, 372} = 1.9574$, $P=0.1458$), but there were significant differences in G6PD V_{max} across all nine genotypes, although this was limited to one genotype being significantly higher than the rest (Figure 2.9B; $F_{8,372} = 4.3606$, $P < 0.0001$). We also found no significant overall correlation between G6PD and MEN relative activities (Figure 2.9C; $R^2=0.01549$), in contrast to some earlier studies (Merritt et al 2005). We find very similar patterns when comparing IDH and MEN; there was no significant difference in IDH V_{max} across the MEN polymorphism (Figure 2.10A, $F_{8, 372} = 1.9574$, $P=0.1458$) only limited, and here not significant, variation in IDH Vmax across the nine genotypes (Figure 2.10B; $F_{8, 372} = 0.967$, $P = 0.461$) and no significant correlation between IDH activity and MEN activity (Figure 2.10C; $R^2=0.17374$). MEN, G6PD, and IDH are key players in the NADPH pathway, however, their phenotypes display trends not found previously, indicating the complexity of this pathway and its activation upon larger differences in MEN enzymatic activity.

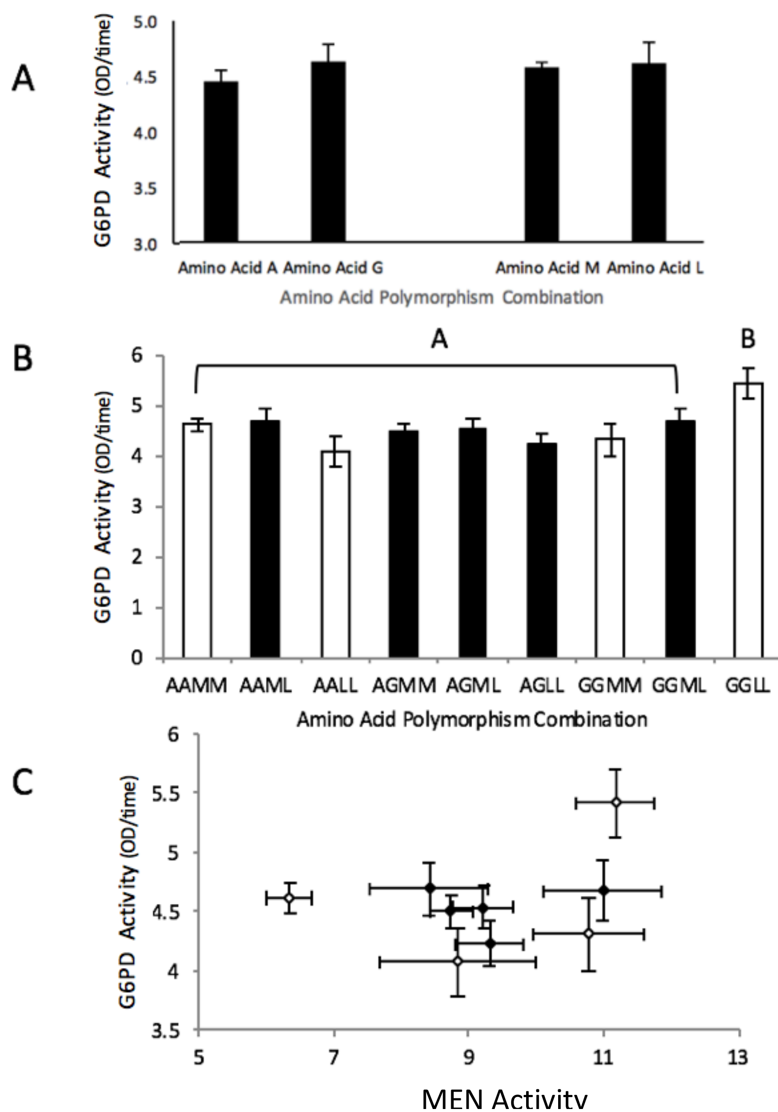


Figure 2.9. Average maximum G6PD enzyme activity (OD/time) was estimated for each genotype as observed enzyme activity under saturating conditions. Bars mean ± standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different (A) Average G6PD activity for MEN polymorphic sites (B) Average G6PD activity for all genotype. White bars represent the homozygous alleles and black bars represent the heterozygote alleles. (C) Average G6PD activity against MEN V_{max}/K_m, also known as Men Activity. White points represent the homozygous alleles and black points represent the heterozygous alleles.

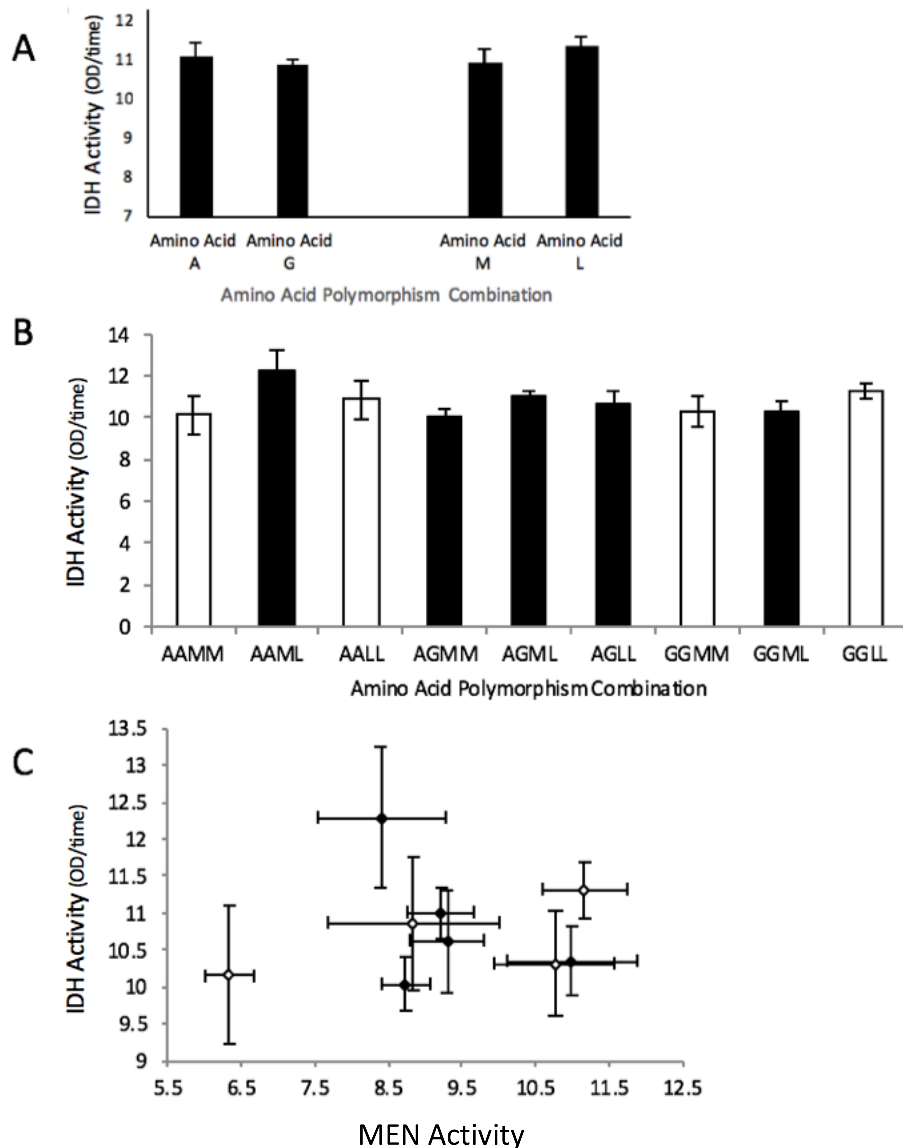


Figure 2.10. Average maximum IDH enzyme activity (OD/time) was estimated for each genotype as observed enzyme activity under saturating conditions. Bars mean ± standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different. (A) Average IDH activity for MEN polymorphic sites (B) Average IDH activity for all genotype. White bars represent the homozygous alleles and the black bars represent the heterozygous alleles. (C) Average IDH activity against MEN V_{max}/K_m. White points represent the homozygous alleles and black points represent the heterozygous alleles.

The polymorphic sites are also not correlated with changes in distal phenotypes

In addition to these relatively proximal phenotypes of enzyme activity, we also quantified the effects of the polymorphisms on three more distal phenotypes that are also sensitive to levels of NADPH: triglyceride concentrations, carbohydrate concentrations, and longevity. NADPH is a major source of reducing power in triglyceride synthesis and other studies have found that triglyceride concentration is sensitive to activities of $NADP^+$ or NAD^+ reducing enzymes (Merritt et al. 2006, 2009, 2005). Previous work in the MEN system has found mixed results (Merritt et al. 2009, 2005) largely differences in triglyceride concentration driven by large-scale differences in MEN activity (Merritt et al. 2005). Neither carbohydrate concentrations nor longevity have been quantified across the MEN alleles although carbohydrate concentrations do change with metabolic condition and enzyme activities and the evidence that the MEN polymorphisms are under selection (Merritt et al. 2005) suggest that they may affect longevity. The amino acid polymorphisms were associated with significant differences in carbohydrate concentrations (Figure 2.11A; $F_{8,372} = 8.5579$, $P < 0.0001$), but no significant differences across the *Men* alleles (Figure 2.11B; $F_{8,372} = 0.245$, $P = 0.873$). Overall, there is a slight, but not statistically significant, trend toward increased carbohydrate concentration with higher MEN activity (Figure 2.11C; $R^2 = 0.29107$). Similarly, the amino acid polymorphisms were not associated with significant differences in triglycerides concentration (Figure 2.12A, $F_{8,372} = 0.6609$, $P = 0.327$), and across the alleles (Figure 2.12A-B, $F_{8,372} = 1.6609$, $P = 1.839$). There was no overall trend between MEN activity and triglyceride concentration (Figure 2.12C; $R^2 = 0.06845$). Similarly, there were no significant differences between the polymorphisms and alleles in longevity (Figure 13A-B; $F_{8,372} = 3.6609$, $P = 0.115$) and no overall correlation (Figure 2.13C; $R^2 = 0.24988$). Overall, the three distal phenotypes were not significantly influenced by the *Men* polymorphisms or differences in relative MEN activity, likely reflecting the complex nature of these phenotypes and the multiple factors that regulate them.

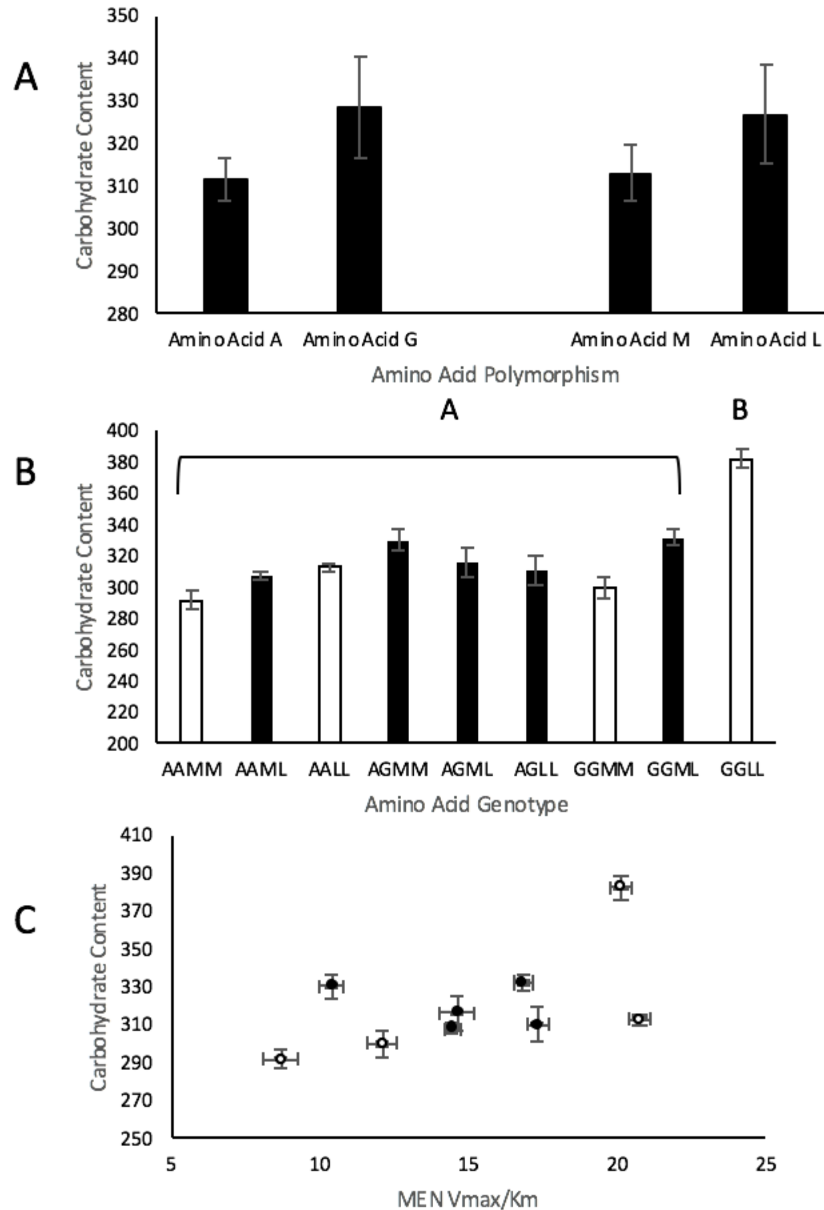


Figure 2.11. **Average carbohydrate content was estimated for each genotype.** Bars mean ± standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different. (A) Average Carbohydrate content for MEN polymorphic sites (B) Average Carbohydrate content for all genotype. White bars represent the homozygous alleles and black bars represent the heterozygous alleles. (C) Average carbohydrate content against MEN V_{max}/K_m. White points represent the homozygous alleles and black points represent the heterozygous alleles.

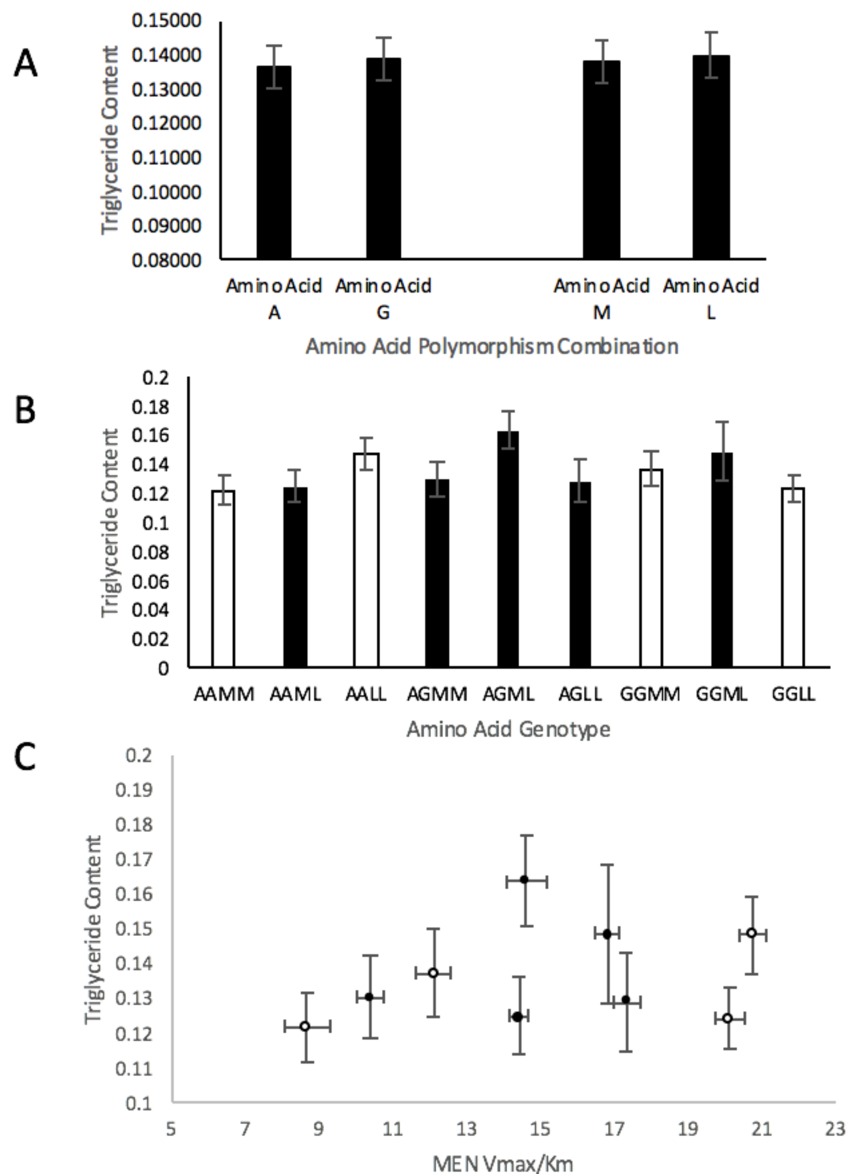


Figure 2.12. Average Triglyceride content was estimated for each genotype. Bars mean ± standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different. (A) Average Triglyceride content for MEN polymorphic sites (B) Average Triglyceride content for all genotype. The white bars represent the homozygous alleles and the black bars represent the heterozygous alleles. (C) Average Triglyceride content against MEN V_{max}/K_m. The white points represent the homozygous alleles and the black points represent the heterozygous alleles.

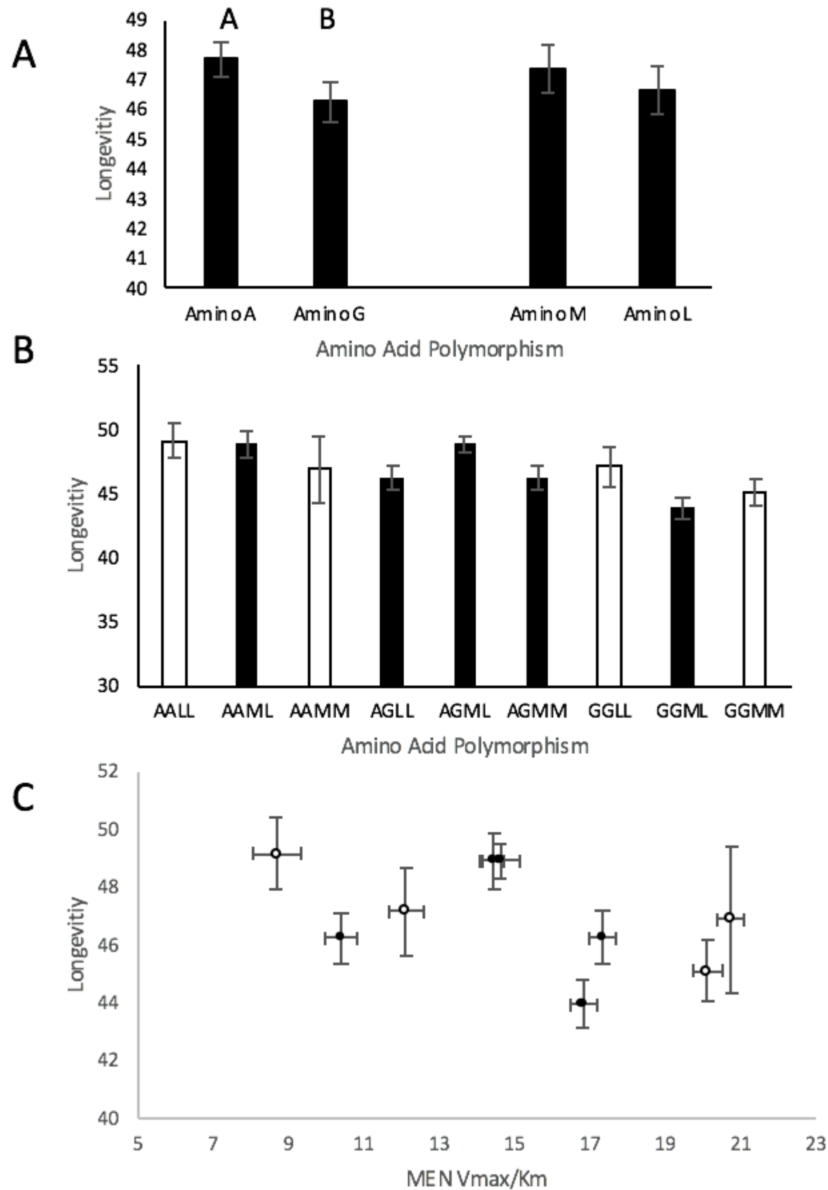


Figure 2.13. Average Longevity was estimated for each genotype. Bars mean ± standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different (A) Average longevity for MEN polymorphic sites (B) Longevity for all genotype. The white bars represent the homozygous alleles and the black bars represents the heterozygous alleles. (C) Longevity against MEN V_{max}/K_m. The white points represent the homozygous alleles and the black points represents the heterozygous alleles.

2.4) Discussion

As biology makes great strides in determining both genotypes and phenotypes, connecting the two continues to challenges us; what are the mechanisms through which genetic diversity leads to biological complexity? Using the simple polymorphisms within *Men*, and studying the biological impact caused by these polymorphisms, provides insight into these pathways. Furthermore, the difference in allele frequencies at the two sites, essentially 50/50 at the G-A site and 90/10 at the M-L site, suggest that different mechanisms maintain the variability at the sites. Can we identify biochemical differences that are the basis of these differences?

MEN Phenotypes are influenced by the G-A and M-L Polymorphisms

In this study, we measured the biochemical characteristics of MEN alleles in *Drosophila melanogaster* to further understand the effects of genetic variation in both homozygous and heterozygous individuals. The allelic frequencies of the polymorphisms are relatively constant across different populations, suggesting that the alleles may be under selection. Presumably, the effects that the polymorphisms cause are a function of many factors, including the location of the polymorphic sites within the protein. As mentioned, the G-A polymorphism is located within an alpha helix in close proximity to the active site and the higher V_{max} of the glycine allele than the alanine allele, may be due to changes in stability or flexibility in this region of the enzyme. The glycine amino acid likely shortens this end of the alpha helix and this change in structure, may alter flexibility at the active site affecting substrate binding (Chakrabartty et al. 1991). Previous work found that the glycine allele was associated with higher thermal stability in enzyme activity, a general proxy for enzyme stability (Rzezniczak et al. 2012). Our data (Figure 2.8) show the same trend, although the results are not statistically significant, but combined with the earlier findings suggest that the G-A polymorphism changes conformation of the active site or the region around this site. The M-L polymorphism does not significantly affect V_{max}, likely reflecting its location buried within the protein far from the active site. In addition to differences driven by the G-A polymorphism, we also found significant line effects; different lines with the same allelic genotype had significantly different V_{max}, indicating that V_{max} is a function of both local and global modifiers. Current work is using genome

wide association study (GWAS) to identify these more distal modifiers. Metabolic pathways are complex and interact with multiple loci to produce a phenotype, the genetic background effects at MEN likely reflect interactions between variable sites across such loci. The location of the G-A polymorphism within the tertiary structure of the enzyme suggests a structural explanation for the difference in the G-A genotypes' V_{max}, but these differences, and certainly the genetic background effects across all genotypes, could also be driven by differences in gene expression driven by either local regulatory variation or variation potentially spread across the genome.

Previous work indicates that MEN V_{max} is a function of both structural and regulatory variation; higher V_{max} is also associated with higher *Men* expression level (Bing et al. 2014; Lum and Merritt 2011; Rzezniczak et al. 2012). Overall, we find a similar pattern, high MEN V_{max} is associated with higher levels of MEN activity (Figure 2.4). Interestingly, however, while parental values generally predict the large-scale differences in V_{max} that we attribute to structure, there is strong non-additively in the levels of gene expression; heterozygous offspring are not simply the average of the homozygous parental values (Figure 2.4A). For example, cross 350 X 021 has higher gene expression level than either 350/350 or 021/021 parental lines. This non-additively suggests complex interactions between the regulatory variation in each parental genotype. In *D. melanogaster*, the homologous chromosomes are physically paired and this pairing modifies gene expression through a process known as transvection, where trans-interactions occur between paired homologous chromosomes driving misregulation (Bing et al. 2014; Duncan 2002; Lum and Merritt 2011; Mellert and Truman 2012; Morris et al. 1998; Noble, Dolph, and Supattapone 2016; Wu and Morris 1999). The *Men* locus is known to be a hotspot for transvection and such inter-chromosomal gene regulation may be driving this non-additively (Bing et al. 2014; Lum and Merritt 2011; Merritt et al. 2005). The cross-specific differences that we see in V_{max} are likely functions of complex interactions between regulatory elements in both cis (along a chromosome) and trans (between homologous chromosomes). Further, the 021 line itself is unusual in that its Men expression is much higher than would be predicted giving its MEN V_{max} (Figure 2.4B) suggesting unique regulatory differences.

While the M-L polymorphism does not significantly modify MEN V_{max}, both it and the G-A polymorphism significantly modify the binding affinity for malate of the enzyme, K_m. K_m is dependent on the structure of the protein, and the differences in K_m of the genotypes likely reflects structural changes driven by the physio-chemical differences between methionine and leucine at this polymorphism buried within the protein. The leucine allele has a substantially lower K_m for malate than the methione allele, indicating a higher binding affinity for the substrate, consistent with previous work. Methionine residues are commonly found buried within the protein, and are unique since they can have both hydrophobic and hydrophilic interactions, hydrogen bonding are involved with polar oxygen atom (Biswal et al. 2012; Pal and Chakrabarti 2001). In contrast, leucine has the potential for fewer interactions with other residue since the R group is simpler and smaller, possibly leading to structural changes that modify substrate biding. The fact that the M-L polymorphism is not near the active site suggests that its influence is not directly on substrate binding, but more likely through broader-scale changes in enzyme structure. The site could possibly affect formation of the MEN homotetramer, but we don't find any indication that the site influences stability of the enzyme, at least as indicated by thermal stability. Furthermore, the intermediate binding affinity values we see for K_m between the heterozygotes can be a result of the different enzyme possibility for one genotype (Figure 2.5B). The intermediate value could be an resultant of the mean binding affinity of the different enzymes for one genotype (e.g. the genotype AGLL has the possibility of all alanines and all leucines, three alanines, one glycine and all leucines, two alanines, two glycine and all leucines, etc.).

Although the effects are smaller, the G-A polymorphism is associated with differences in K_m, with the alanine allele associated with a lower K_m indicating higher substrate binding. As with V_{max}, this difference likely reflects changes in the alpha helix at the active site. Unlike V_{max}, which is a function of both expression levels and structure, K_m is independent of expression and expected to be only a function of structure. Based on this structure-only model we would not expect to find differences across genetic background within an allele class as these lines all have identical amino acid composition and,

presumably, structure, but line-specific differences were apparent, nonetheless (Figure 2.3C). These genetic background-specific differences indicate that overall structure is likely modified by factors coded elsewhere in the genome. Previous work on MEN has been split on this issue with studies both finding, and failing to find background effects (Merritt et al. 2005; Rzezniczak et al. 2012).The background effects are smaller than those seen in V_{max}, suggesting that while background has a modifying effect on this structural phenotype it is not as pronounced as in a phenotype driven by both structure and expression.

A more biologically relevant phenotype than either V_{max} or K_m alone is the V_{max}/K_m ratio, an indication of the relative of the *in vivo* enzyme activity. Comparison of the V_{max} (Figure 2.2) and K_m (Figure 2.5) results with this estimate of relative activity (Figure 2.7) clearly shows the different implication of each value. In our results, the G-A polymorphism has little effect on the relative activity, while the M-L polymorphism has substantial effect. Further, because of the large difference in K_m associated with the leucine allele, the rare allele, lines with this variant have higher relative enzyme activity (Figure 2.7B). Comparison of V_{max} alone would suggest that the M-L polymorphism has little effect and the glycine allele was associated with the highest activity. Relative activity, then, suggests that the largest effect on the biochemistry of this system is from the rare allele, frequency of the leucine allele is ~10%, and that the more equal frequency polymorphism has limit effect on the performance of the enzyme We had speculated that the polymorphisms, or at least the approximate equal frequency of the G-A polymorphism could be maintained by some heterozygote advantage possibly the result of non-linear combinatorial effects on biochemistry. Instead, the relative activities of the heterozygotes were all essentially intermediate to those of the homozygous genotypes (AAMM, AALL, GGMM and GGLL). Perhaps the frequencies of variants are maintained by selection for intermediate, not extreme, activities.

Distal Phenotypes and the MEN polymorphism

Metabolic enzymes function as components of complex pathways and networks, interacting with other genes and regulatory components. Previous research using natural and laboratory-derived Men alleles found correlations between MEN activity and increasingly distal phenotypes including activities of the other NADPH network enzymes IDH and G6PD and carbohydrate and triglyceride concentrations two downstream metabolites whose production is dependent on the reducing power of NADPH. These correlations were, however, sensitive to both genetic background and environmental conditions (Merritt et al. 2009, 2005; Rzezniczak et al. 2012; Rzezniczak and Merritt 2012). While we did find variation in these phenotypes, and the even more distal phenotype of longevity, we did not find a consistent pattern of correlation between MEN activity (either V_{max} or relative activity) and any of these phenotypes. Our inability to replicate the broad-scale correlations that other studies have demonstrated may reflect the genetic backgrounds we used, the relatively small scale of natural variation compared to laboratory-derived alleles, or the complex nature of these traits, especially metabolite concentrations and longevity, or a combination of all of these factors. The striking consistency of the allele frequencies, however, strongly suggests that these sites are under selection. The fundamental question remains – what are the selective forces acting on the Men locus and across the NADPH network as a whole?

Insight into the *Men* locus and Heterozygote Advantage

The roughly equal allele frequency at the G-A site means that most individuals will be heterozygosis and the consistency of these frequencies across all North American populations surveyed suggests heterozygote advantage (Cormack, Hartl, and Clark 1990). Population genetics based expectations for the M-L site are less clear, but the rare allele is present, and rare, at the two sites in which it has been looked for and the rare nature means that individuals will be either homozygous for the common allele or heterozygous for the rare allele. Both sites clearly affect the fundamental biochemistry of the MEN enzyme. Interestingly, phenotypes of the heterozygous individuals are consistently intermediate between those of the homozygotes (with the noted exception of gene expression). While we suspected that heterozygote advantage would manifest as non-

linear combinations of biochemical phenotypes, perhaps the message from our results is that the genetic variation results in relatively consistent biochemistry. But, if the consistency is maintained by selection, why haven't populations simply fixed for one allele or the other? The answer to this question may lie in phenotypes we have not, yet, discovered, or in responses to changes, such as environmental variation, that we have not yet explored.

This study is just a start in understanding the connection between genotype to phenotype and the impact of genetic molecular variation. Future work needs to be completed in understanding the phenotype changes in environmental stressors to understand the impact of the allelic frequencies using larger-scale. With the different environmental stressors, the flies will learn to acclimate to those condition, which will provide more information on heterozygous advantage at the *Men* locus.

2.7) References

Abolaji, Amos O., Charles O. Olaiya, Oluwagbenga J. Oluwadahunsi, and Ebenezer O. Farombi. 2017. "Dietary Consumption of Monosodium L-Glutamate Induces Adaptive Response and Reduction in the Life Span of Drosophila Melanogaster." *Cell Biochemistry and Function* 35(3):164–70.

Aebi, Hugo. 1984. "Catalase in Vitro." *Methods in Enzymology*.

Apel, Klaus and Heribert Hirt. 2004. "REACTIVE OXYGEN SPECIES: Metabolism, Oxidative Stress, and Signal Transduction." *Annual Review of Plant Biology*.

Baker, Monya. 2012. "De Novo Genome Assembly: What Every Biologist Should Know." *Nature Methods*.

Beelman, Clare A. and Roy Parker. 1995. "Degradation of MRNA in Eukaryotes." *Cell*.

Bernard, Kristine E., Tony L. Parkes, and Thomas J. S. Merritt. 2011. "A Model of Oxidative Stress Management: Moderation of Carbohydrate Metabolizing Enzymes in SOD1-Null Drosophila Melanogaster." *PLoS ONE*.

Bing, Xinyang, Teresa Z. Rzezniczak, Jack R. Bateman, and Thomas J. S. Merritt. 2014. "Transvection-Based Gene Regulation in Drosophila Is a Complex and Plastic Trait." *G3 (Bethesda, Md.)* 4(11):2175–87.

Biswal, Himansu S., Eric Gloaguen, Yohan Loquais, Benjamin Tardivel, and Michel Mons. 2012. "Strength of NH···S Hydrogen Bonds in Methionine Residues Revealed by Gas-Phase IR/UV Spectroscopy." *Journal of Physical Chemistry Letters* 3(6):755–59.

Boone, Charles, Howard Bussey, and Brenda J. Andrews. 2007. "Exploring Genetic Interactions and Networks with Yeast." *Nature Reviews Genetics*.

Chakrabartty, Avijit, John A. Schellman, and Robert L. Baldwin. 1991. "Large Differences in the Helix Propensities of Alanine and Glycine." *Nature*.

Chandler, Christopher H., Sudarshan Chari, and Ian Dworkin. 2013. "Does Your Gene Need a Background Check? How Genetic Background Impacts the Analysis of Mutations, Genes, and Evolution." *Trends in Genetics*.

Charlesworth, D. and B. Charlesworth. 1987. "Inbreeding Depression and Its Evolutionary Consequences." *Annual Review of Ecology and Systematics*.

Chow, Clement Y. 2016. "Bringing Genetic Background into Focus." *Nature Reviews*

Genetics.

Cordell, Heather J. 2009. "Detecting Gene-Gene Interactions That Underlie Human Diseases." *Nature Reviews Genetics*.

Cormack, R. M., D. L. Hartl, and A. G. Clark. 1990. "Principles of Population Genetics." *Biometrics*.

Crick, F. 1970. "Central Dogma of Molecular Biology." *Nature*.

Dixon, Scott J., Michael Costanzo, Anastasia Baryshnikova, Brenda Andrews, and Charles Boone. 2009. "Systematic Mapping of Genetic Interaction Networks." *Annual Review of Genetics*.

Dudash, Mr and Cb Fenster. 2000. "Inbreeding and Outbreeding Depression in Fragmented Populations." *Genetics, Demography and Viability of Fragmented Populations*.

Duncan, Ian W. 2002. "Transvection Effects in Drosophila." *Annual Review of Genetics*.

Dworkin, Ian et al. 2009. "Genomic Consequences of Background Effects on Scalloped Mutant Expressivity in the Wing of Drosophila Melamogaster." *Genetics*.

Finkel, Toren and Nikki J. Holbrook. 2000. "Oxidants, Oxidative Stress and the Biology of Ageing." *Nature*.

Fujioka, Miki, Hemlata Mistry, Paul Schedl, and James B. Jaynes. 2016. "Determinants of Chromosome Architecture: Insulator Pairing in Cis and in Trans." *PLoS Genetics* 12(2):1–25.

Ganter, Christiane and Andreas Plückthun. 1990. "Glycine to Alanine Substitutions in Helices of Glyceraldehyde-3-Phosphate Dehydrogenase: Effects on Stability." *Biochemistry*.

Geer, B. W., D. L. Lindel, and D. M. Lindel. 1979. "Relationship of the Oxidative Pentose Shunt Pathway to Lipid Synthesis in Drosophila Melanogaster." *Biochemical Genetics*.

Griendling, Kathy K. et al. 2016. *Measurement of Reactive Oxygen Species, Reactive Nitrogen Species, and Redox-Dependent Signaling in the Cardiovascular System: A Scientific Statement from the American Heart Association*. Vol. 119.

Habig WH, Pabst MJ, Jakoby WB. 1974. "Glutathione S-Transferases." *J Biol Chem*.

Hall, J. G. and R. K. Koehn. 1983. "The Evolution Of Enzyme Catalytic Efficiency and

Adaptive Inference From Steady-State Kinetic Data." *Evolutionary Biology*.

Hall, John G. 1985. "Temperature-Related Kinetic Differentiation of Glucosephosphate Isomerase Alleloenzymes Isolated from the Blue Mussel, Mytilus Edulis." *Biochemical Genetics*.

Hosamani, Ravikumar and Muralidhara. 2013. "Acute Exposure of Drosophila Melanogaster to Paraquat Causes Oxidative Stress and Mitochondrial Dysfunction." *Archives of Insect Biochemistry and Physiology* 83(1):25–40.

Huang, Wen et al. 2014. "Natural Variation in Genome Architecture among 205 Drosophila Melanogaster Genetic Reference Panel Lines." *Genome Research*.

Kalmus, H. 1945. "Adaptative and Selective Responses of a Population of Drosophila Melanogaster Containing e and E+ to Differences in Temperature, Humidity and to Selection for Developmental Speed." *Journal of Genetics* 47(1):58–63.

Kimura, M., Benjamin J. Callahan, Dmitri A. Petrov, and Philipp W. Messer. 2011. "RULES FOR TESTING STABILITY OF A SELECTIVE POLYMORPHISM." *PNAS* 42(6):336–40.

Kostyuk, V. A. and A. I. Potapovich. 1989. "Superoxide--Driven Oxidation of Quercetin and a Simple Sensitive Assay for Determination of Superoxide Dismutase." *Biochemistry International* 19(5):1117–24.

Labate, J. and W. F. Eanes. 1992. "Direct Measurement of in Vivo Flux Differences between Electrophoretic Variants of G6PD from Drosophila Melanogaster." *Genetics*.

Lessel, Courtney E., Tony L. Parkes, Joel Dickinson, and Thomas J. S. Merritt. 2017. "Sex and Genetic Background Influence Superoxide Dismutase (CSOD)-Related Phenotypic Variation in Drosophila Melanogaster." *G3 (Bethesda, Md.)*.

Li, Gene-Wei and X. Sunney Xie. 2011. "Central Dogma at the Single-Molecule Level in Living Cells." *Nature*.

Livak, Kenneth J. and Thomas D. Schmittgen. 2001. "Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2-ΔΔCT Method." *Methods* 25(4):402–8.

Lum, Thomas E. and Thomas J. S. Merritt. 2011. "Nonclassical Regulation of Transcription: Interchromosomal Interactions at the Malic Enzyme Locus of

Drosophila Melanogaster.” *Genetics* 189(3):837–49.

Mackay, Trudy F. C. 2004. “The Genetic Architecture of Quantitative Traits: Lessons from Drosophila.” *Current Opinion in Genetics and Development*.

Mellert, David J. and James W. Truman. 2012. “Transvection Is Common throughout the Drosophila Genome.” *Genetics*.

Merritt, Thomas J. S. et al. 2009. “Quantifying Interactions within the NADP(H) Enzyme Network in Drosophila Melanogaster.” *Genetics* 182(2):565–74.

Merritt, Thomas J. S., David Duvernell, and Walter F. Eanes. 2005. “Natural and Synthetic Alleles Provide Complementary Insights into the Nature of Selection Acting on the Men Polymorphism of Drosophila Melanogaster.” *Genetics* 171(4):1707–18.

Merritt, Thomas J. S., Efe Sezgin, Chen Tseh Zhu, and Walter F. Eanes. 2006. “Triglyceride Pools, Flight and Activity Variation at the Gpdh Locus in Drosophila Melanogaster.” *Genetics* 172(1):293–304.

Migula, P. et al. 2004. “Antioxidative Defence Enzymes, in Beetles, from a Metal Pollution, Gradient.” *Biologia* 59(5):645–54.

Morris, J. R., J. L. Chen, P. K. Geyer, and C. T. Wu. 1998. “Two Modes of Transvection: Enhancer Action in Trans and Bypass of a Chromatin Insulator in Cis.” *Proceedings of the National Academy of Sciences of the United States of America*.

Müller, Katiane Raquel et al. 2017. “Anacardium Microcarpum Extract and Fractions Protect against Paraquat-Induced Toxicity in Drosophila Melanogaster.” *EXCLI Journal* 16:302–12.

Noble, G. P., P. J. Dolph, and S. Supattapone. 2016. “Interallelic Transcriptional Enhancement as an In Vivo Measure of Transvection in Drosophila Melanogaster.” *G3: Genes|Genomes|Genetics* 6(October):3139–48.

Pal, Debnath and Pinak Chakrabarti. 2001. “Non-Hydrogen Bond Interactions Involving the Methionine Sulfur Atom.” *Journal of Biomolecular Structure and Dynamics*.

Prusiner, Stanley B. 1990. “Novel Structure and Genetics of Prions Causing Neurodegeneration in Humans and Animals.” *Biologicals*.

Ralser, Markus, Gino Heeren, Michael Breitenbach, Hans Lehrach, and Sylvia Krobitsch. 2006. “Triose Phosphate Isomerase Deficiency Is Caused by Altered Dimerization-

Not Catalytic Inactivity-of the Mutant Enzymes." *PLoS ONE* 1(1).

Reczek, Colleen R. and Navdeep S. Chandel. 2015. "ROS-Dependent Signal Transduction." *Current Opinion in Cell Biology*.

Rzezniczak, T. Z., L. A. Douglas, J. H. Watterson, and T. J. S. Merritt. 2011. "Paraquat Administration in Drosophila for Use in Metabolic Studies of Oxidative Stress." *Analytical Biochemistry*.

Rzezniczak, Teresa Z., Thomas E. Lum, Robert Harniman, and Thomas J. S. Merritt. 2012. "A Combination of Structural and Cis-Regulatory Factors Drives Biochemical Differences in Drosophila Melanogaster Malic Enzyme." *Biochemical Genetics* 50(11–12):823–37.

Rzezniczak, Teresa Z. and Thomas J. S. Merritt. 2012. "Interactions of NADP-Reducing Enzymes Across Varying Environmental Conditions: A Model of Biological Complexity." *G3* 2(12):1613–23.

Scherz-Shouval, Ruth and Zvulun Elazar. 2007. "ROS, Mitochondria and the Regulation of Autophagy." *Trends in Cell Biology*.

Schreiber, Stuart L. 2005. "Small Molecules: The Missing Link in the Central Dogma." *Nature Chemical Biology*.

Sellis, D., B. J. Callahan, D. A. Petrov, and P. W. Messer. 2011. "Heterozygote Advantage as a Natural Consequence of Adaptation in Diploids." *Proceedings of the National Academy of Sciences* 108(51):20666–71.

Sellis, Diamantis, Benjamin J. Callahan, Dmitri A. Petrov, and Philipp W. Messer. 2011. "Heterozygote Advantage as a Natural Consequence of Adaptation in Diploids." *Proceedings of the National Academy of Sciences* 108(51):20666–71.

Sezgin, Efe et al. 2004. "Single-Locus Latitudinal Clines and Their Relationship to Temperate Adaptation in Metabolic Genes and Derived Alleles in Drosophila Melanogaster." *Genetics* 168(2):923–31.

Steinberg, Martin H. 2008. "Sickle Cell Anemia, the First Molecular Disease: Overview of Molecular Etiology, Pathophysiology, and Therapeutic Approaches." *TheScientificWorldJournal*.

Valencia-Sanchez, Marco Antonio, Jidong Liu, Gregory J. Hannon, and Roy Parker. 2006. "Control of Translation and MRNA Degradation by MiRNAs and SiRNAs."

Genes and Development.

Valko, Marian et al. 2007. "Free Radicals and Antioxidants in Normal Physiological Functions and Human Disease." *The International Journal of Biochemistry & Cell Biology*.

Wang, K. et al. 2009. "Genome-Wide Identification of Post-Translational Modulators of Transcription Factor Activity in Human B Cells." *Nature Biotechnology*.

Watt, W. B. and a M. Dean. 2000. "Molecular-Functional Studies of Adaptive Genetic Variation in Prokaryotes and Eukaryotes." *Annual Review of Genetics*.

Weber, A. L. et al. 2012. "Genome-Wide Association Analysis of Oxidative Stress Resistance in Drosophila Melanogaster." *PLoS One*.

Wilton, A. N., C. C. Laurie-Ahlberg, T. H. Emigh, and J. W. Curtsinger. 1982. "Naturally Occurring Enzyme Activity Variation in Drosophila Melanogaster. II. Relationships among Enzymes." *Genetics*.

Wise, E. M. and E. G. Ball. 1964. "Malic Enzyme and Lipogenesis." *Proceedings of the National Academy of Sciences of the United States of America* 52(1933):1255–63.

Wolf, J. B. 2003. "Genetic Architecture and Evolutionary Constraint When the Environment Contains Genes." *Proceedings of the National Academy of Sciences*.

Wu, C. ting and James R. Morris. 1999. "Transvection and Other Homology Effects." *Current Opinion in Genetics and Development*.

Ying, Weihai. 2008. "NAD $^{+}$ /NADH and NADP $^{+}$ /NADPH in Cellular Functions and Cell Death: Regulation and Biological Consequences." *Antioxidants & Redox Signaling* 10(2):179–206.

Zhang, Jixiang et al. 2016. "ROS and ROS-Mediated Cellular Signaling." *Oxidative Medicine and Cellular Longevity*.

Zhu, Chen Tseh and David M. Rand. 2012. "A Hydrazine Coupled Cycling Assay Validates the Decrease in Redox Ratio under Starvation in Drosophila." *PLoS ONE* 7(10).

Chapter 3 of this manuscript will have some overlapping information found in the introduction and discussion compared to Chapter 2. Chapter 3 is a second, independent manuscript in preparation for publication. Chapter 2 is also an independent manuscript in preparation for publication.

Chapter 3: Environmental Effects on Single Nucleotide Polymorphism in the NADPH Pathway at the *Malic Enzyme* Locus

3.1) Introduction

Biological systems are dynamic, taking input from the surroundings and responds accordingly for the survival of that organism. These biological systems consist of numerous networks that interact with one another to drive changes in their phenotypes. Therefore, it is essential to consider that complex traits can be influenced directly or indirectly by multiple genes and the environment (and the interaction between the two), rather than acting in isolation (Boone, Bussey, & Andrews, 2007; Chandler, Chari, & Dworkin, 2013; Dixon, Costanzo, Baryshnikova, Andrews, & Boone, 2009). When studying a gene within a network, a common theme amongst most studies is that these interactions are quantified in only one condition (e.g. include an example); however, changes in the environment can lead to important differences that should be explored (e.g. given an example).

Malic enzyme (*Men)* codes for a metabolic enzyme that oxidizes malate to pyruvate, concurrently reducing $NADP^+$ to NADPH (Merritt, Duvernell, & Eanes, 2005; Merritt et al., 2009; Rzezniczak, Lum, Harniman, & Merritt, 2012; Wise & Ball, 1964). In *D. melanogaster*, there are two known single nucleotide polymorphisms (SNPs), which each lead to amino acid substitutions in this gene. The first polymorphism is a guanine to cytosine substitution, that causes an alanine being coded for instead of the ancestral glycine, noted as the G-A polymorphism. This first polymorphism occurs at base pair 338; or amino acid 113, found within an α-helix near the enzyme's active site (Merritt et al., 2005; Rzezniczak et al., 2012). The second polymorphism occurs at base pair 1051; at amino acid 351, in which an adenine is substituted for the ancestral thymine, coding for a leucine amino acid rather than a methionine, at amino acid 351, noted as the M-L polymorphism. The M-L polymorphism is buried in a β-sheet away from the active site (Rzezniczak et al., 2012). The allelic frequencies were found to be approximately 50:50 for the G-A polymorphism and 90:10 for the M-L polymorphism, throughout different populations (Rzezniczak et al., 2012).

The biochemistry of these amino acid substitutions has been previously studied, and both polymorphisms play a role in altering the phenotype of the enzyme. *D. melanogaster* with isothrid chromosomes was used to study the polymorphism, with the MEN alleles of GM/GM (glycine at the G-A polymorphism and methionine and the M-L polymorphism), AM/AM and AL/AL. The leucine is the rarer allele, and fly lines with GL/ chromosomes were not available for this early study. The alanine allele was associated with a 20% higher MEN maximum velocity (V_{max}), than the glycine allele and there were no significant differences in V_{max} between the methione and leucine alleles (Rzezniczak et al., 2012).

Interestingly, in Chapter 2, there were significant line effects between the GM/GM and AM/AM alleles. The variation between the genotypes' V_{max} may also be driven by differences in gene expression driven by either local regulatory variation or non-regulatory variation potentially spread across the genome: genetic background effects. For the *Men* expression levels, the guanine nucleotide (which codes for glycine) had 51% higher expression than cytosine (which codes for alanine), which suggests that the differences in V_{max} could be due to a difference in gene expression levels. Strikingly, the M-L polymorphism did not display any differences in its V_{max} of MEN; however, there were differences in *Men* expression levels; alleles with the adenine nucleotide (which codes of methionine) were associated with higher Men expression levels than the thymine (which codes for leucine) alleles. The difference in expression levels between the alleles further confirms that V_{max} is a combination of structural and regulatory variation. The G-A polymorphism shows regulatory variations and the M-L polymorphism shows both. Enzyme thermal stability was also modified by the G-A polymorphism. Strikingly, in this study, the glycine allele was more stable, by 15% than the alanine allele. The more stable glycine allele was surprising since it is known that glycine amino acid tends to destabilize α-helices, which was not the case here. Thermal stability was quantified as a proxy for the overall stability of the protein, as opposed to helical stability; however, it was proposed that with the destabilization of the helix, that would affect the overall protein. Another possibility for the increase stability with the glycine allele is that a shorten helix would stabilize the protein. Lastly, the binding affinity, K_m, was influenced by both polymorphisms, where at the G-A polymorphism, the glycine amino acid containing

alleles were associated with 10% higher K_m values than the alanine alleles, and the methionine alleles were associated with 42% higher K_m values than the leucine alleles. It is interesting to note that leucine alleles, the rare allele, had the lowest K_m values. The consistency of the allele frequencies between populations and the differences in the biochemistry of MEN suggest that the alleles are under selection. We study the polymorphisms in *Men* as a model system for the biological complexity effects derived from the genetic variation of on complex traits.

Mutations occurring in genes lead to effects in a primary pathway; however, mutations in one gene can also induce changes in connected pathways as well (Bernard, Parkes, & Merritt, 2011; Mackay, 2004; Rzezniczak & Merritt, 2012; Wolf, 2003). MEN is known to be part of the well-studied, small, network known as the nicotinamide adenine dinucleotide phosphate (NADPH) enzyme pathway, in *D. melanogaster* (Merritt et al., 2009; Rzezniczak et al., 2012; Ying, 2008). In this pathway, there are four key players: cytosolic MEN, cytosolic Isocitrate dehydrogenase (IDH), Glucose-6-phosphate dehydrogenase (G6PD) and 6-phosphogluconate dehydrogenase (6PGD). These enzymes have been shown to interact and are co-regulated through monitoring of the $NADP^+$: NADPH ratio within a cell. In a previous study, engineered excision lines were used to generate approximately 50% and 100% enzymatic activity for MEN, IDH and G6PD. This study quantified the interactions between the enzymes using these large-scale differences in activities. These interactions were often compensatory (reduction of one source of NADPH results in up-regulation of others), but some were parallel and counterintuitive (reduction of one source of NADPH results in down-regulation of others; (Rzezniczak & Merritt, 2012). Therefore, these enzymes should be taken into consideration as an interacting set when quantifying the biochemical characteristics of MEN to further account for interactions driving changes in the phenotype. Also, studying the other enzymes may provide insight into the maintenance of the alleles in *Men* responsible for the known amino acid substitutions across different populations.

NADPH is known to play a role in reactive oxygen species (ROS) clearance and involvement in the immune response (Bernard et al., 2011; Hosamani & Muralidhara,

2013; Lessel, Parkes, Dickinson, & Merritt, 2017; Rzezniczak & Merritt, 2012). ROS are a byproduct of aerobic metabolism and can include the superoxide anion, hydroxyl radical and hydrogen peroxide. ROS is at low concentrations within the cells, serving as intermediates in cell signaling. However, at high concentrations, ROS can bind to and damage DNA, protein and lipids, leading to cell death (Apel & Hirt, 2004; Bernard et al., 2011; Reczek & Chandel, 2015; Scherz-Shouval & Elazar, 2007; Valko et al., 2007; Zhang et al., 2016). The accumulation of ROS and these damaged macromolecules is known as oxidative stress. NADPH is used directly, as a cofactor, or indirectly for antioxidant enzymes that aid in clearing out ROS (Bernard et al., 2011; Merritt et al., 2005; Rzezniczak et al., 2012; Rzezniczak & Merritt, 2012). In another study, Rzezniczak and Merritt (2012) used the NADP(H) enzyme knockout and control alleles and environmental stress to further understand the interactions within the network (Rzezniczak & Merritt, 2012). Paraquat, an herbicide, was used to induce an oxidative stress environment (Rzezniczak, Douglas, Watterson, & Merritt, 2011; Rzezniczak & Merritt, 2012). Paraquat treatment led to an overall reduction in enzyme activity, but the effects varied and paraquat treatment also led to differences in the interactions between enzymes. For example, under control conditions a reduction in MEN activity led to a reduction in IDH activity, but an increase in IDH activity under paraquat stress. These changes in different environmental conditions underscore the importance of studying genes under multiple conditions. The focus of previous work on the NADPH network has mainly been on the enzymes upstream of NADPH, namely those involved in its production (MEN, IDH, and G6PD). The interactions of these enzymes with enzymes further downstream, or those who consume NADPH, has not been studied as extensively. Under conditions of oxidative stress, the antioxidant enzymes superoxide dismutase (SOD), catalase (CAT) and glutathione- s-transferase (GST) are significant consumers of NADPH (Abolaji, Olaiya, Oluwadahunsi, & Farombi, 2017; Bernard et al., 2011; Lessel et al., 2017; Migula et al., 2004; Ralser, Heeren, Breitenbach, Lehrach, & Krobitsch, 2006).

Here, we extend our investigation of the *Men* polymorphisms, the forces that are driving their genetic variation and the further understanding the allelic frequency maintenance,

by studying the biochemical characteristics and communication in the NADPH pathway with environmental stressors. Previous work focused on the polymorphisms and genetic background effects, here we add environmental variations in addition to a wider range of lines. Nine different *Men* genotypes were created using lines from the Drosophila Genetic Reference Panel (DGRP) collection, to study line effects on the *Men* genotypes (Dworkin et al., 2009). Oxidative stress was induced to further quantify the complex interactions within the NADPH network. We find that the biochemistry of the MEN alleles reasonably consistent with that found in previous studies, but also find interactions that were not observed previously. Interestingly, we find that the M-L polymorphism impacts the V_{max}, whereas it was silent in previous studies. We also see interactions between proximal phenotypes for the G-A polymorphism. Under oxidative stress, enzyme activities responded to the paraquat administration, similarly to other work completed on the network.

3.2 Materials and Methods

3.2.1 Fly Stocks and Rearing Conditions

Fly lines used are a subset of the DGRP lines obtain from the Bloomington Drosophila Stock Centre (Bloomington, USA). The subset of lines was selected for their genotype at the two known polymorphism sites in Malic enzyme gene. The first polymorphism is at position 351, with a G/C nucleotide polymorphism that results in a G/A amino acid substitution. The second is at position 1051, an A/T nucleotide polymorphism that results in an M/L amino acid substitution. Table 1 lists the parental lines used and their respective genotypes. We included more and new lines for this study compared to Chapter 2 of this manuscript. Flies were maintained on a standard cornmeal medium with 12:12-hr light: dark cycle at 25 ºC.

3.2.2 Fly Crosses

The parental lines were crossed with one another to create heterozygotes with the following genotype combinations: C/C, A/A (CCAA), C/C, A/T (CCAT), C/C, T/T (CCTT), C/G, A/A (CGAA), C/G, A/T (CGAT), C/G, T/T (CGTT), G/G, A/A (GGAA),

and G/G, T/T (GGTT). All crosses included five adult male flies from one line paired with five virgin female flies from another line. A sample cross is shown in Figure 3.0. Flies were maintained on a standard cornmeal medium with 12:12-hr light: dark cycle at 25 ºC. Emerging male flies were aged to 3-6 days, frozen and stored at -80 ºC until further analysis.

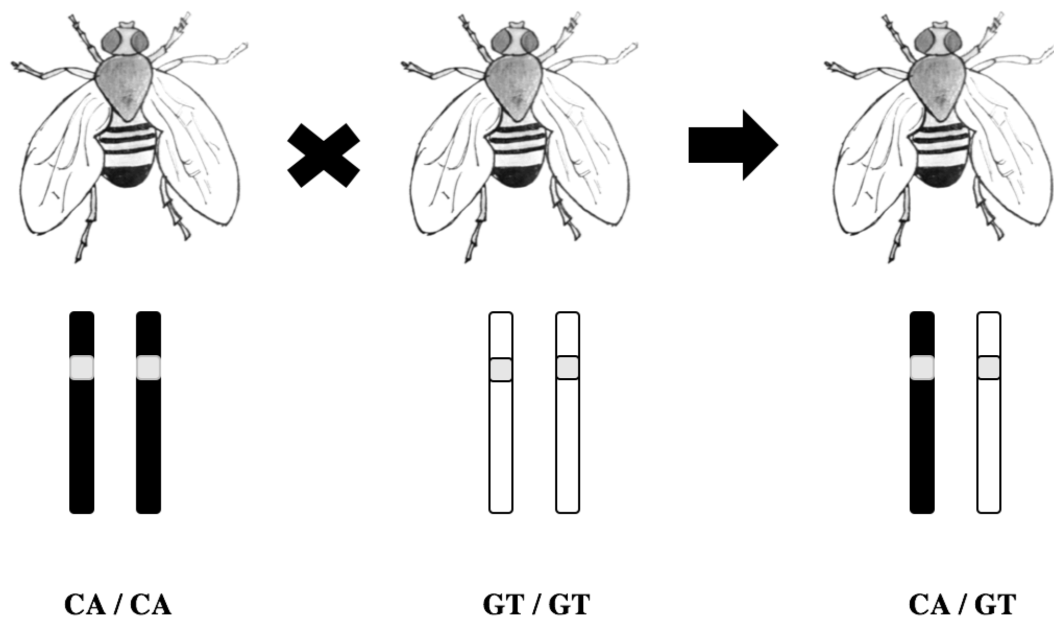


Figure 3.0. Schematic of crosses performed to generate desired alleles. The crosses were performed as followed: Five isogenic males of one line is crossed to five virgin isogenic females of another line to generate the respective heterozygote combination.

3.2.3 Oxidative stress treatment

A subset of flies from the above crosses were selected to place on oxidative stress treatment. From the 35 parental lines, 21 lines were chosen to cross and expose the offspring to the treatment. Flies were fed 20 mM paraquat incorporated into cornmeal-yeast-agar-corn syrup diet to induce oxidative stress. Replicates of 20 male flies aged to

3-5 days were maintained on the food for 24 hr. Stress-treated flies were compared to control flies that were maintained on standard cornmeal-yeast-agar-corn syrup diet (Rzezniczak et al., 2011).

3.2.4 Fly wet weight

Flies were weighed to the nearest 0.01 mg using a Mettler Toledo microbalance MX5. Weight was used as a covariate in analyses of covariance (ANCOVA) to standardize enzyme activity, triglycerides and carbohydrate concentration for differences in fly size.

3.2.5 Fly homogenization

Assays were conducted using whole fly homogenates created by homogenizing samples of five flies in grinding buffer (100 mM Tris-HCl, 0.15 mM $NADP^+$, pH 7.4 or 0.1 M Phosphate buffer, 0.2 M EDTA, pH 7.0) at a concentration of one fly per 100 μL, and centrifuged at 13 000 RPM for 10 min at 4 ºC to pellet all solids. Tris- Buffer was used for the determination of the following enzymatic assays, MEN, IDH and G6PD, as well as soluble protein concentration, triglycerides content, and carbohydrate content. Phosphate buffer was used for the determination of the following antioxidant activity, CAT, GST, and SOD, in addition to total ROS generation. Homogenate supernatant was collected, and 300 μL of supernatant was transferred to a 96-well plate; aliquots were sampled from this master plate for each subsequent analysis.

3.2.6 Enzyme activity measurement

Enzyme activity was quantified by measuring the production of NADPH through time. For this set of assays, a Tris-HCl homogenizing buffer was used. Assays were conducted on a 96-well plate spectrophotometer (Molecular Devices SpectraMax 384 Plus), using ten μL of fly homogenate and 100 μL of assay buffer (described below). Absorbance at 340 nm was measured every 9 s for 3 min at 25 ºC, except for G6PD activity assays, which was measured for 5 min. Samples were assayed twice, and the means were used for further analysis. The assay buffers were previously optimized to give maximum activities and were as follow:

G6PD: 100 mM Tris- HCl, 0.32 mM NADP, 3.5 mM D-glucose-6-phosphate (pH 7.4)

IDH: 100 mM Tris-HCl, 0.10 mM NADP, 0.84 mM $MgSO_4$, 1.37 mM DL-isocitrate (pH 8.6)

MEN: 100 mM Tris-HCl, 0.34 mM NADP, 50 mM $MnCl_2$, 50 mM malate (pH 7.4)

3.2.7 Antioxidant Enzyme activity measurement

The following antioxidant enzymes were analyzed: Catalase (CAT), Glutathione-S-transferase (GST) and Superoxide dismutase (SOD). This set of assays, phosphate homogenizing buffer, was used, except for CAT, phosphate buffer with the addition of 0.1% Triton-X was used. Assays were conducted on a 96-well plate spectrophotometer (Molecular Devices SpectraMax 384 Plus), using ten μL of fly homogenate and 100 μL of assay buffer (described below). For CAT, samples were diluted 1:4 before the addition of assay buffer.

CAT activity was quantified by observing the clearance of H_2O_2 at 240 nm, which was measured every 9 s for 5 min. The assay buffer contained 0.05 M Phosphate buffer and 15 mM H_2O_2 (stock 30% H_2O_2) (Aebi, 1984; Müller et al., 2017).

GST activity was quantified by observing the formation of the conjugate complex of the substrate 1-chloro-2,4-dinitrobenzene (CDNB) and GSH at 340 nm, measured every 9 s for 5 min. The assay buffer contained 100 mM phosphate buffer (pH 7.0), 1 mM EDTA, 1 mM GSH and 2.5 mM CDNB (Habig WH, Pabst MJ, 1974).

SOD activity was quantified by observing the inhibition of superoxide driven oxidation of quercetin by SOD at 406 nm, measured every 9 s for 10 min. The assay buffer contained 25 mM phosphate (pH 10), 0.25 mM EDTA, 0.8 mM TEMED and 0.05 μM quercetin (Kostyuk & Potapovich, 1989).

3.2.9 Estimation of Michaelis-Menten Constant

Genotype-specific Michaelis-Menten constants for malate were determined using a 10-point geometric design method. Rates for each genotype was determined using the method described above, but using 10 substrate concentration ranging from 0.1 mM- 10 mM, each measurement was done in replicate. The average value for each of the ten

concentrations was then used to determine the Michaelis-Menten constant using the program GraFit 7.0 software.

3.2.10 RNA Extraction and quantitative RT-PCR

RT-qPCR was used to test for differences in *Men* expression across the CCTT genotype set. Total RNA was isolated from three groups of fifteen 3 to 5-day old male flies using the RNeasy Kit (QIAGEN, Valencia, CA) according to manufacturer's instructions, and was stored at -80 ºC until needed for reverse transcription. For each sample, one microgram of total RNA was reverse-transcribed using random hexamers and High Capacity cDNA Reverse Transcription Kits with RNase Inhibitor (Applied Biosystem) The qPCR reaction consisted of 2 μL of undiluted cDNA template, 0.4 μM of each primer, and 0.2 μM probe, and Quantitect Probe PCR Master Mix (QIAGEN); a total volume of 20 μL. The primers and the probe flank the *Men* intron between exon 2 and exon 3 (5' GTATTGCCAACCTGTGCC, 3' AGCTTGTGTTCGGTGAGT and probe 56-FAM/ATGGTGGATAGCCGTGGTGTCA/3IABkFQ. cDNA synthesis of samples lacking reverse transcriptase were used as a negative control to ensure that there was no genomic DNA contamination and "no-template" blanks were used to ensure there was no contamination. Two reactions per template were performed using a Mastercycler Ep Realplex Thermal Cycler. Expression results were normalized to *RpL32* and are reported relative to *MenEx3* using ΔΔCT method (Livak & Schmittgen, 2001).

3.2.11 Soluble Triglycerides Content

Soluble triglyceride content was measured using a commercially available kit (Triglyceride-SL Assay, Pointe Scientific, Canton, MI, Catalog No. T7531) following manufacturer's protocol. Each assay, 10 μL of homogenate and 100 μL of reagent, was incubated at 37 ºC for 5 min. Sample absorbance was measured at 500 nm, and total soluble triglycerides concentration were determined by comparison with a commercially available standard (Pointe Scientific, Canton, MI, Catalog No. T7532). Each sample was assayed twice, and the mean used in the analysis. Results are reported as micrograms of triglycerides per sample.

3.2.12 Total Carbohydrate Content

Total carbohydrate content was measured as previously described. Complex carbohydrates were converted to glucose using a digestion solution that contained 10 μL of fly homogenate sample and 2 μL of amyloglucosidase (Sigma Aldrich, St Louis, MO, A1602) at a concentration of 1 unit/sample in 2.0 M sodium acetate buffer, pH 5.7. Samples and glycogen standards (Sigma Aldrich, St. Louis, MO, Catalog No. G0885) were digested at 55 oC for 45 min. Following digestion, total glucose was measured using the commercially available kit (Genzyme, Cambridge, MA, Catalog No. 23517) in which 10 μL of digested homogenate was combined with 200 μL of glucose reagent and incubated at 37 oC for 10 min. Each sample was assayed twice, and the mean was used in the analysis. Sample absorbance was measured at 340 nm, and comparison to digested glycogen standards determined total carbohydrate concentration. Results are reported as milligrams per liter.

3.2.13 Soluble Protein Content

Soluble protein was measured using the bicinchoninic acid (BCA) assay using a commercially available kit (Pierce, Thermo Scientific, Rockford, IL, Catalog No. 23225) following manufacturer's protocol. In brief, assays contained 10 μL homogenate and 100 μL reagent and were incubated at 37 ºC for 30 min. Sample absorbance was measured at 562 nm, and total soluble protein concentrations were determined by comparison with bovine serum albumin standards (Sigma Aldrich, St Louis, MO, A4503). Each sample was assayed twice, and the mean was used in the analysis. Soluble protein content was used as a covariate in analyses of covariance (ANCOVA) for enzyme activities, triglyceride, and carbohydrate concentration to standardize for differences in fly size and homogenization.

3.2.14 Data analysis

All crosses were replicated in two independent vials, with three samples taken from each vial. Assays were run in duplicates to account for technical error. ANCOVA and Tukey's honesty significant difference (HSD) multiple-comparison tests were performed using JMP 12.0 software (SAS Institute) to determine whether there were significant

differences in enzyme activity, triglyceride or carbohydrate concentration using protein concentration and wet weight as covariates.

3.2.15 Intermediate phenotype calculations

The heterozygote's phenotypes were predicted based off of the homozygous phenotypes. For example, the V_{max} of AALL and AAMM alleles crosses were averaged and the standard error was determined. The mean V_{max} from AALL and AAMM alleles is the theoretical value of V_{max} for AAML. An ANCOVA determined if the theoretical value and experimental value were significantly different. If the values were not significantly different, the heterozygote allele is said to be an intermediate of its respective homozygous allele.

3.3 Results

3.3.1 Both *Men* Polymorphisms Influence Maximum Velocity

Effect of polymorphism: Both the G-A and M-L polymorphisms in *Men* significantly alter the V_{max} of MEN, with the G-A polymorphism having a more significant influence on the MEN activity than the M-L polymorphism (Figure 3.1A). Flies homozygous for the glycine amino acid allele, $Men^{G/G}$ (hereafter noted as GG, with similar nomenclature for each site and genotype, e.g. A/A is noted as AA, M/M as MM, A/G as AG, M/L as ML and A/G, M/L as AGML) had 14% higher activity than the flies with the AA allele ($F_{2,1875}$=17.1253, $p < 0.0001$). At the M-L polymorphism, flies with the LL allele had 12% higher activity than flies with the MM allele ($F_{2,1875}$=6.0558, $p = 0.0024$). The intermediate activities of the heterozygous genotypes were calculated based on the homozygous combinations (AAMM, AALL, GGMM, and GGLL). For example, the homozygous combinations for the genotype AAML would be AAMM and AALL. The average activity of AAMM and AALL genotype would be the expected (or the calculated) value for AAML. If the experimental value of AAML falls within the range of the expected value, it would be noted as an intermediate of the homozygous genotypes. The GA allele is an intermediate between the GG and AA allele activity, not significantly different from the expected value ($F_{8,1875}$= 4.8514, p= 0.1148). The ML allele also had an

intermediate activity for the homozygous amino acid combination, not significantly different from the calculated value. Since the M-L polymorphism influences the enzymatic activity of MEN, this suggests that structural changes can play a role in the function of the enzyme, where the rare allele influences the activity. In Chapter 2, the significant difference was observed at the G-A polymorphism. With Chapter 3, the experiments increased the number of lines used for each polymorphism to introduce more genetic backgrounds into the study. Here, the variation is present at both polymorphic site, likely due to the increase in lines used.

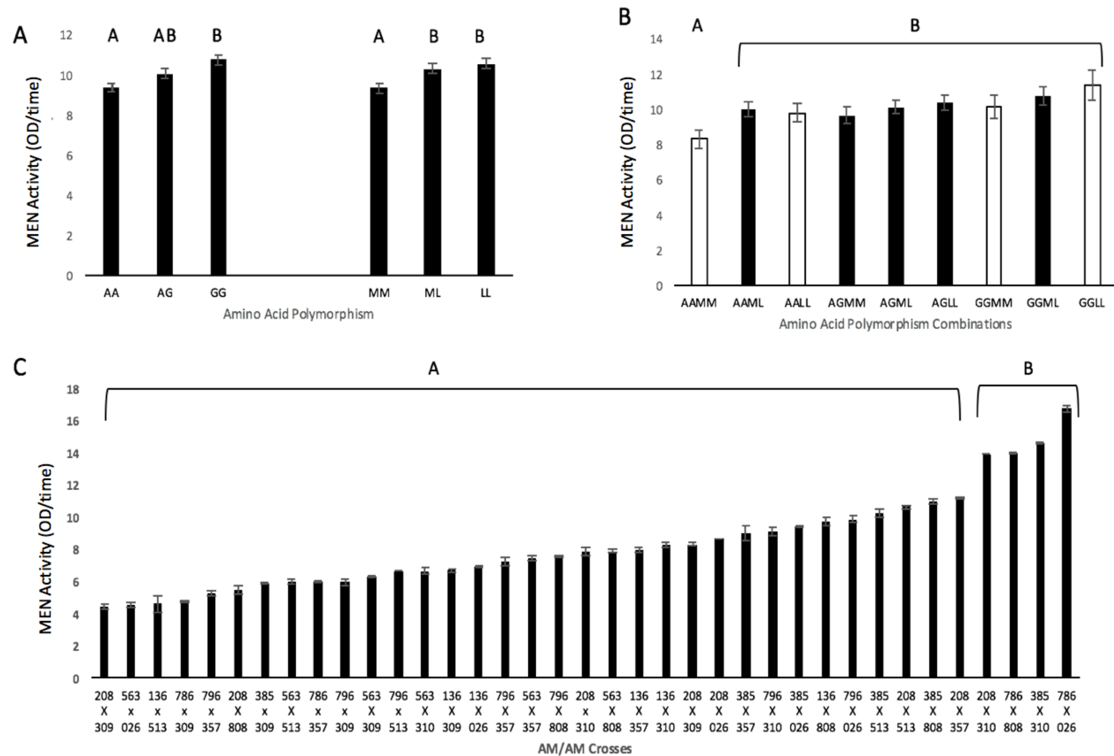


Figure 3.1. MEN V_{max} Activity (OD/time) across different alleles and backgrounds under saturating conditions. (A) Displays the genotype at the G-A polymorphisms (leftmost three bars) and the M-L polymorphisms (rightmost three bars) mean MEN V_{max} activity. (B) The nine different alleles with both polymorphic site in the genotype. White bars represent the alleles that are homozygous at both polymorphic sites, and black bars represents where the allele differs at one or both polymorphic site. (C) The experimental flies from various crosses that contribute to the allele AM/AM. The letters above the bars indicate which genotypes are different, where letters that differ are significantly different. Bars indicate ± one standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different

Effect of genotypes: We compared the biochemical characteristics of the *Men* genotypes to test for possible heterozygote advantage (Figure 3.1B). Overall, MEN V_{max} were significantly different across the genotypes ($F_{8,1875}$=6.8582, $p < 0.0001$). The heterozygous combinations were not significantly different from the expected values, indicating that there is an additive effect towards the V_{max} phenotype ($F_{8,1875}$=2.4382, $p < 0.2360$). A similar trend in genotypes were also observed in Chapter 2.

Genetic background effects: Genetic background effects were analyzed across all the crosses; a subset of the variation for genotype AAMM is shown in Figure 3.1C. Multiple lines were used for each polymorphism, introducing a range of genetic background to analyze per genotype. Despite the identical Men allelic genotype, Figure 3.1C, there is variation within the activity of the enzyme, indicating that MEN V_{max} is sensitive to genetic background effects, the same trend observed in Chapter 2 as well. (Chandler et al., 2013; Dworkin et al., 2009). In addition to a sensitivity to background effects, this also suggests that other factors can influence the V_{max}, not solely the genotype, hinting that genotype does not necessarily dictate the phenotype (Rzezniczak et al., 2012).

3.3.2. Both Polymorphisms Influences *Men* Expression Levels

Effect of polymorphism: The variation of MEN activity between the different polymorphism and the alleles within each polymorphism could be a function of variation of *Men* expression levels, as opposed to actual variation in the maximum velocity of the enzyme itself (i.e expression effects, not structural effects alone). Expression levels of *Men* were quantified using a qPCR. At the G-A polymorphism, alleles with the cytosine nucleotide (which codes for alanine) had a 40% higher transcript level than guanine alleles (which codes for glycine), a significant difference (Figure 3.2A; $F_{2,18}$=14.192 $p < 0.0001$). At the M-L polymorphism, the adenine nucleotide (which codes for methionine) has about 20% higher relative *Men* transcript than the thymine nucleotide (which codes for the leucine), also significantly different.

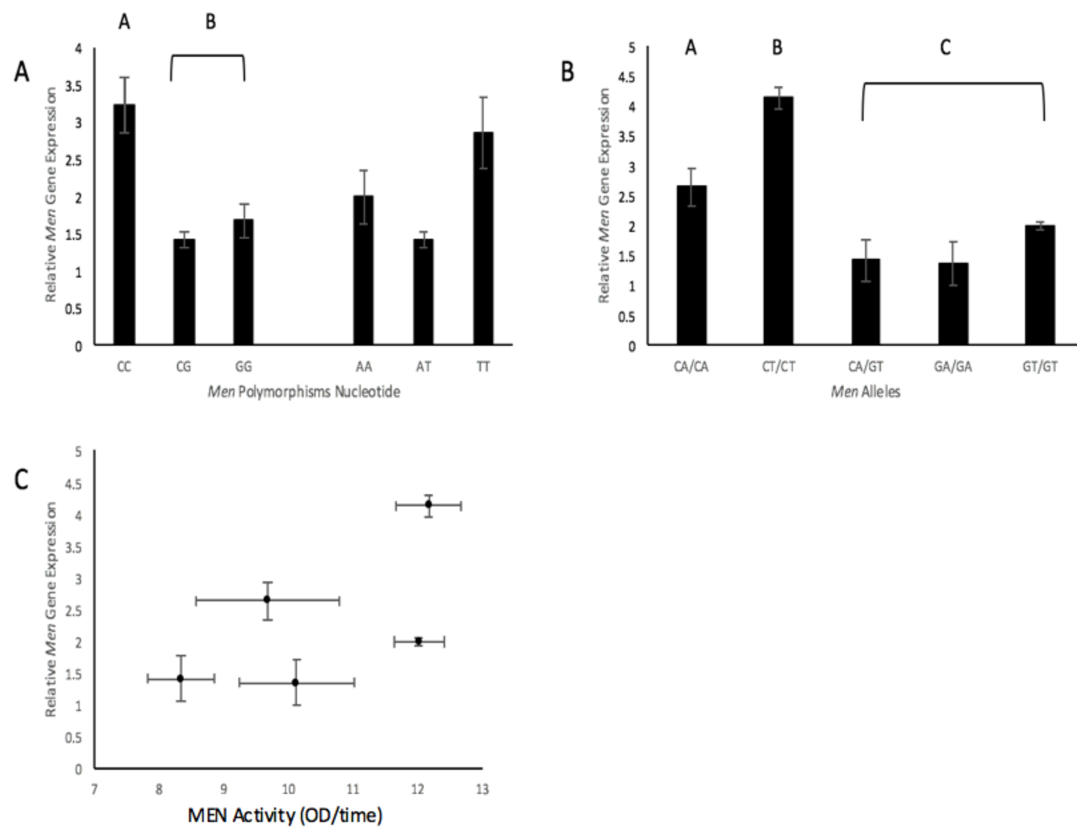


Figure 3.2. *Men* relative expression levels. (A) Relative *Men* expression for the single nucleotide polymorphism in *Men*. The leftmost three bars represent the genotypes found at G-A polymorphism, and the rightmost three bars represent the genotypes found at the M-L polymorphism. (B) *Men* alleles against relative *Men* expression. (C) Correlation between *Men* relative expression and MEN activity. Expression was normalized to MENEX3^{+}. Relative expression as calculated using the ΔΔCT method (Livak and Schmittgen 2001). Bars indicate ± one standard error.

Effect of genotypes: Interestingly, the double heterozygous genotype, CGAT, has lower than expected expression levels, more similar to the expression of the GG alleles, GA/GA and GT/GT, than an average of the GG and CC alleles (Figure 3.2B). There were also differences between the second polymorphism (adenine and thymine) when combined with cytosine (CA/CA and CT/CT), where the CT/CT genotype has higher *Men* expression levels than CA/CA ($F_{2,18}$=9.925, p = 0.0004; Figure 3.2B), however not with the guanine alleles (GA/GA and GT/GT).

Men gene expression levels and MEN protein activity for the genotypes CA/CA, CT/CT/ CA/GT, GA/GA and GT/GT are plotted in Figure 3.2C. Not surprisingly, as expression levels of *Men* increase, there is a general increase in the activity of MEN, indicating a positive correlation. This correlation suggests that the variation in MEN V_{max} activity across the polymorphisms, where higher MEN activity has higher *Men* expression levels. Therefore, there is a regulatory effect that can influence the differences in the MEN activity.

The most striking result is that the trends that are observed in Figure 3.2A, where lines with the cytosine alleles have higher expression levels than those with guanine allele. In previous studies, guanine allele expression levels were higher than cytosine (Rzezniczak et al., 2012). The variation seen in *Men* expression is highly dependent on line effects. The different trend that we observe may be a function of the different lines, and different backgrounds, in our study than in the earlier study. Even with the different trends observed, the same correlation exists between *Men* expression and MEN activity, Figure 3.2B, indicating that MEN activity is manipulated by regulatory expression (Bing et al., 2014; Lum & Merritt, 2011; Rzezniczak et al., 2012). In Chapter 2, one allele expression level was analyzed, AALL. In the analysis, a positive correlation between *Men* expression levels and MEN protein activity was found.

3.3.3 Both G-A and M-L Polymorphisms Influences the Binding Affinity

Effect of polymorphism: To determine if the polymorphisms affected the binding affinity of MEN, the K_m for malate was calculated using various malate concentrations. The G-A polymorphism has a small, but significant effect on K_m; the GG genotype is 8% higher than the AA genotype ($F_{2,625}$=7.7925, p = 0.0004, Figure 3.3A). The M-L polymorphism has a greater effect on the binding affinity of MEN, which has been consistent with previous studies. The MM genotype is 34% higher in malate concentration compared to LL genotype ($F_{2,625}$=54.993, p <0.0001). This is interesting that the rare allele, leucine, has a larger effect on the binding affinity. Furthermore, the ML allele is intermediate for the homozygous allele; not significantly different from the calculated value. A similar trend was also observed in Chapter 2, where the G-A polymorphism had a small significant effect on K_m and the M-L polymorphism had a larger effect.

Effect of genotypes: The different genotypes were analyzed, and a decrease in malate concentration can be observed when going from when going from methionine to leucine at the M-L polymorphism, ignoring the G-A polymorphism (AAMM to AALL and GGMM to GGLL) (white bars; Figure 3.3B), which was also present in previous studies ($F_{8,1875}$=17.22442, p < 0.0001) (Rzezniczak et al., 2012). The experimental values of the heterozygous alleles (black bars; Figure 3.3B) were quantified and not significantly different from the expected value, therefore noted as intermediates values. There is a linear effect in the binding affinity of the MEN polymorphisms, similar to MEN $V_{max;}$ also holds true for the data observed in Chapter 2 ($F_{8,1875}$= 0.086, p = 3.4282).

Genetic background effects: K_m is a function of the structure, not the amount, of a protein and should, therefore, be sensitive to the amino acid sequence of a protein, but not the expression level of a gene. Given this structural dependence and expression independence, we expected K_m to be potentially sensitive to the polymorphic sites, but insensitive to background effects (which can change expression, but not sequence). In other words, the K_m phenotype is not expected to be influenced by the line effects. However, previous studies have shown that K_m is sensitive to genetic background effects, e.g. lines within one genotype have significantly different K_m values, suggesting that

elements coded elsewhere in the genome are influencing the proteins substrate binding. Upon analyzing the different genetic background used, genetic backgrounds impacted the K_m values, Figure 3.3C, indicating that K_m is sensitive to genetic background effects. The same genetic background effects were found across lines used in Chapter 2.

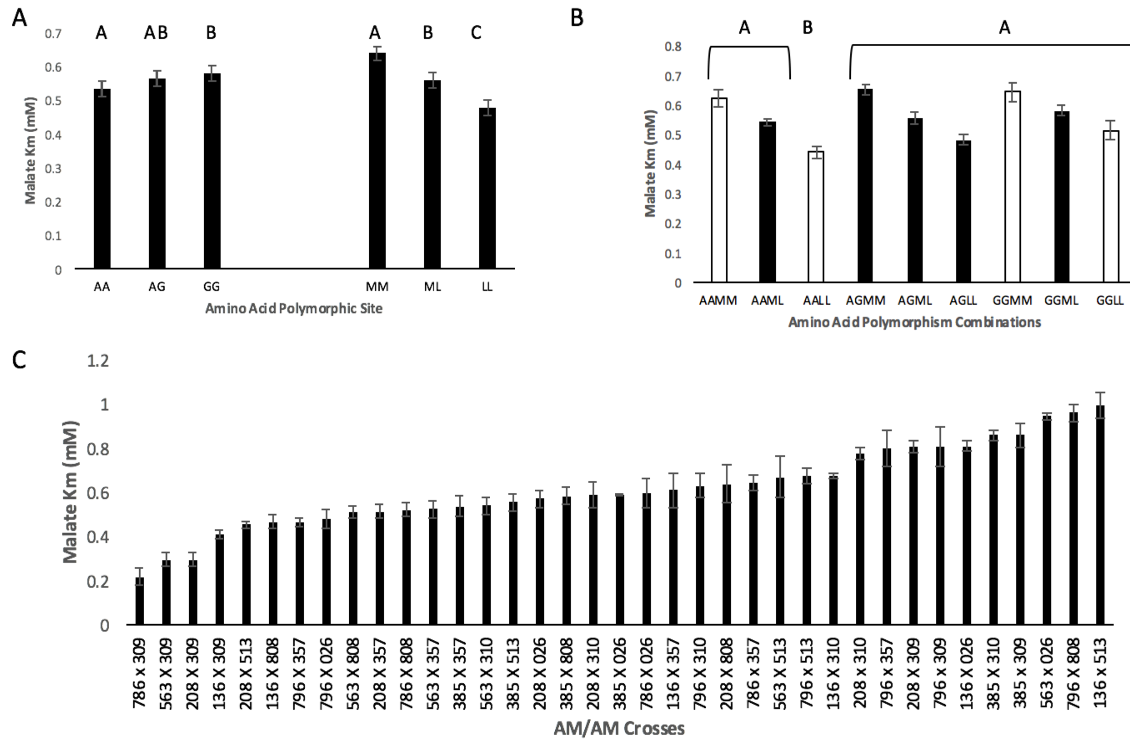


Figure 3.3. Malate K_m binding affinity across different alleles and backgrounds, using various malate concentrations. (A) the G-A polymorphisms' genotype (leftmost three bars) and the M-L polymorphisms' genotype (rightmost three bars) mean malate K_m. (B) The nine different alleles, taking into account both polymorphisms. White bars represent the alleles that are homozygous at both polymorphic sites, and black bars represents where the allele differs at one or both polymorphic site. (C) The experimental flies from various crosses that contribute to the allele AM/AM. The letters above the bars indicate which genotypes are different, where letters that differ are significantly different. Bars indicate ± one standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different.

3.3.4 The M-L Polymorphism Influences the V_{max}/K_m Ratio

Effect of polymorphism: The G-A polymorphism has very similar relative MEN activities, not significantly different (Figure 3.4A). The G-A polymorphism allelic frequency is at 50:50, and the similarity in activity quantified here may provide insight into that. The M-L polymorphism has an evident trend, whereas going from MM to LL, there is an increase in the relative activity, where the LL genotype has a 52% greater relative activity than MM genotype ($F_{2,625}$=64.5599, $p < 0.0001$). Furthermore, the intermediate value for ML genotype was calculated from the MM and LL genotypes, and the experimental value falls within the range, therefore not significantly different from the calculated value ($F_{8,1875}$= 0.2004, $p = 0.2360$). Similar results are found in Chapter 2, where the leucine allele had higher relative activity.

Effect of genotype: When analyzing the genotypes, observing both polymorphisms together, the opposite trend of K_m (Figure 3.3B) was observed. Interestingly, as seen with previous work (Chapter 2 of this manuscript), the rarer amino acid, leucine, had the highest relative MEN activity, compared to the other amino acids, glycine, alanine, and methionine. The heterozygous genotypes, blue bars in Figure 3.3B, relative activities were not significantly different from the calculated values, therefore being intermediate values. Between the genotypes, significant differences exist ($F_{8,1875}$=17.3150, $p < 0.0001$). Once again, we see the additive effects between the heterozygotes ($F_{8,1875}$= 0.0351, $p = 0.6761$).

Genetic background effect: Since there were variations between the MEN V_{max} and K_m and line effects, it is expected to see variation between the ratio of those phenotypes, which was the case (Figure 4C). In Chapter 2, the genetic background effects across the lines were present.

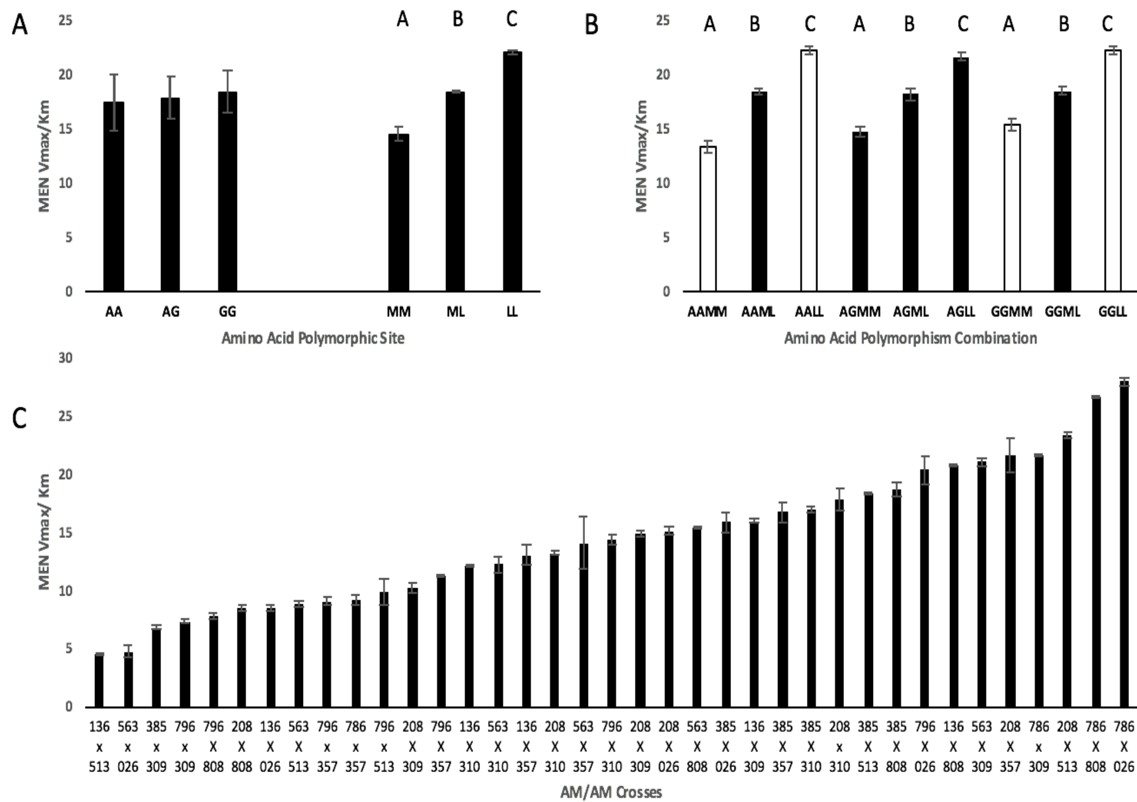


Figure 3.4. MEN V_{max}/K_m Activity, or relative *in vivo* activity across different alleles and backgrounds under saturating conditions. (A) the G-A polymorphisms' genotype (leftmost three bars) and the M-L polymorphisms' genotype (rightmost three bars) mean MEN relative activity. (B) The nine different alleles, taking into account both polymorphisms. White bars represent the alleles that are homozygous at both polymorphic sites, and black bars represents where the allele differs at one or both polymorphic site. (C) The experimental flies from various crosses that contribute to the allele AM/AM. Bars indicate ± one standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different.

3.3.5 Responses to MEN Activity within the NADPH Pathway

When studying an enzyme within a biological system, it is important to consider that enzymes do not act in isolation. For MEN, three other enzymes contribute to the NADPH pathway: IDH, G6PD, and 6PGD. The analysis of IDH and G6PD were included to quantify the potential interaction between the genotypes.

3.3.5.1 The G-A Polymorphism affects IDH Activity

Effect of polymorphism: To determine if the *Men* polymorphisms were associated with variation in IDH activity for, we quantified the maximum velocity of IDH was determined under saturating conditions for all of the Men genotypes. At the MEN G-A polymorphism, the AA genotype had approximately 12% higher IDH activity than the GG genotype ($F_{2,625}$=11.4158, $p < 0.0001$; Figure 3.5A). The GA genotype was not significantly different from the calculated value, from the homozygous genotypes, suggesting it to be an intermediate of the homozygous alleles GG and AA, regarding IDH activity. The *Men* M-L polymorphism did not influence IDH activity ($F_{2,625}$=0.1425, p= 0.8672; Figure 3.5A). With the larger sample size used in this study, we can quantify differences at the G-A polymorphism in IDH activity, whereas previous studies could not (Chapter 2).

Effect of genotypes: Once the polymorphisms were analyzed individually, the suite of combined genotypes were interpreted (Figure 3.5B). IDH activity generally decreased when going from the alanine to the glycine allele. However, two genotypes expressed higher IDH activity than expected. AAML had 9% higher IDH activity compared to the calculated value (Figure 3.5B), and this activity surpasses both respective homozygous genotypes (AAMM and AALL). The second genotype, AGLL, had 4% higher than expected activity and also surpassing the respective homozygous genotypes (AALL and GGLL; $F_{8,1875}$=5.8951, $p < 0.0001$).

MEN Correlation: MEN relative activity (V_{max}/K_m) and IDH activity were plotted against each other in order to determine if a correlation existed between the activities (Figure 3.5C; R^2= $2x10^{-5}$); however, no correlation was observed. The lack of a trend in

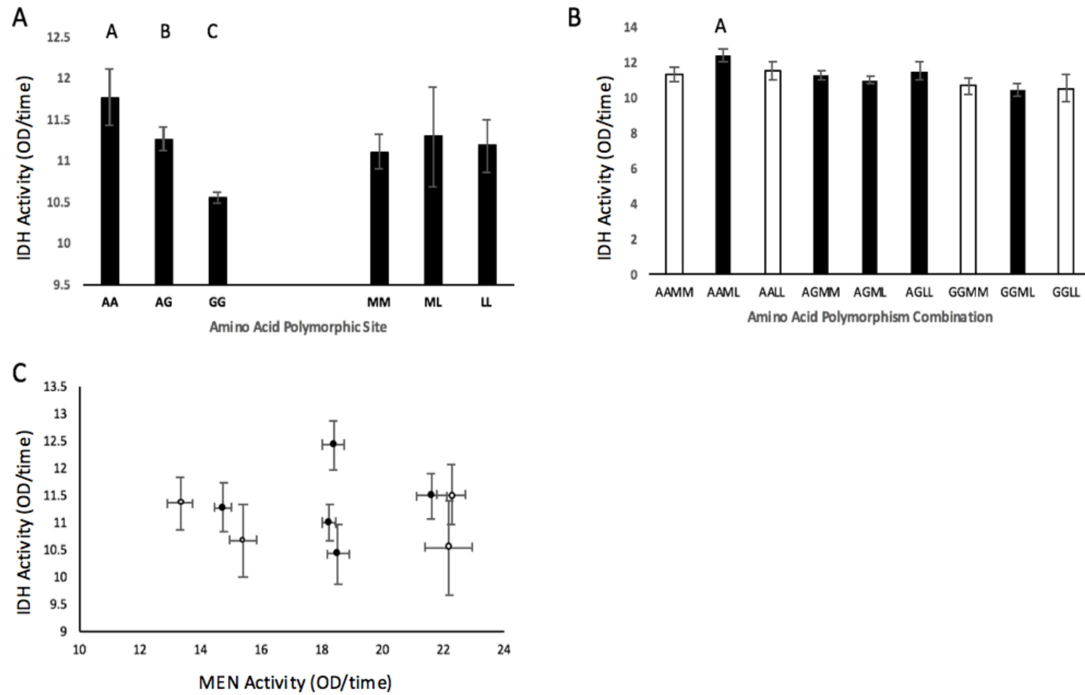


Figure 3.5. IDH V_{max} Activity (OD/time) across different *Men* alleles and backgrounds under saturating conditions. (A) the G-A polymorphisms' genotype (leftmost three bars) and the M-L polymorphisms' genotype (rightmost three bars) mean IDH V_{max} activity. (B) The nine different alleles, taking into account both polymorphisms. White bars represent the alleles that are homozygous at both polymorphic sites, and black bars represents where the allele differs at one or both polymorphic site. (C) IDH activity plotted against MEN activity. White points represent the alleles that are homozygous, and the black points represent the alleles that are heterozygous. Bars indicate ± one standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different

Figures 3.5B-C suggest that IDH activity is independent from the *Men* polymorphisms, which could be because the differences in MEN activity between the genotypes were not substantial to drive changes in IDH activity. More substantial differences drive more significant interaction between loci.

3.3.5.2 The G-A Polymorphism affects G6PD Activity

Effect of polymorphism: The third NADPH enzyme analyzed in this study was G6PD to study the possible interactions between MEN and the network. The G-A polymorphism significantly affected the activity of G6PD, where the GG genotype had 15% higher G6PD activity than the AA genotype ($F_{2,625}$=18.2082, $p < 0.0001$; Figure 3.6A). Furthermore, the AG genotype was had intermediate levels of activity, since it is value was not significantly different from the calculated value. The *Men* M-L polymorphism did not significantly influence the activity of G6PD ($F_{2,625}$=0.8391, p =0.4324; Figure 3.6A). With the larger sample size used in this study, we can quantify differences at the G-A polymorphism in G6PD activity, whereas previous studies could not (Chapter 2).

Effect of genotypes: The genotypes of *Men* were compared against the G6PD activity, overall having similar G6PD activities to one another and no significant differences between the genotype (Figure 3.6B; $F_{8,1875}$=0.3395, p=0.7140). To determine if there was a correlation between MEN activity and G6PD activity, both activities were plotted against each other (Figure 3.6C) and no correlation was quantified (R^2= 0.08134).

MEN Correlation: Similar to IDH activity, the G-A polymorphism influenced G6PD activity, but no overall trend was apparent when comparing the MEN genotype and activity, suggesting the G6PD activity is independent of the *Men* polymorphisms or the differences in MEN activity was not sufficient to drive changes in G6PD, Figure 3.6 C.

3.3.6. Responses of distal phenotypes to MEN activity

As mentioned before, MEN is a metabolic enzyme and a contributor to the NADPH pathway. Therefore, variation in *Men* expression and MEN activity can influence distal phenotypes, such as carbohydrate content, which is related to metabolism, and

triglycerides stores, which are related to metabolism and NADPH concentration. To determine if variation in MEN affected distal phenotypes, carbohydrate content and triglyceride concentration were quantified.

3.3.6.1 MEN Activity affected Carbohydrate Content

Effect of polymorphism: Carbohydrate content was measured across the different *Men* polymorphisms and genotypes in order to determine if differences in the MEN genotype and activity lead to changes in the amount of carbohydrate present in the samples. At the G-A polymorphism, as the genotype went from AA to GG, there was a decrease in carbohydrate content, significantly different ($F_{2,625}$=3.5078, p=0.0386; Figure 3.7A). At the M-L polymorphism, going from MM to LL there was an increase in carbohydrate content, significantly different ($F_{2,625}$= 3.5514, p = 0.0372; Figure 3.7A). With the larger sample size used in this study, we can quantify differences at the G-A polymorphism in carbohydrate content, whereas previous studies could not (Chapter 2).

Effect of genotypes: A similar trend was observed across the different genotype, especially at the M-L polymorphism (Figure 3.7B) where the genotypes homozygous with the leucine amino acid alleles (AALL and AGLL) had the highest carbohydrate content ($F_{8,1875}$=0.9124, p=0.5167).

MEN Correlation: Carbohydrate content was compared to MEN activity to determine if a correlation existed between the phenotypes, and as MEN activity increases, a general increase in carbohydrate content occurred (Figure 3.7C; R^2= 0.4818). MEN and carbohydrate were correlated, and this likely stems from the fact that MEN cycles between glycolysis and the Krebs cycle, which are pathways of respiration.

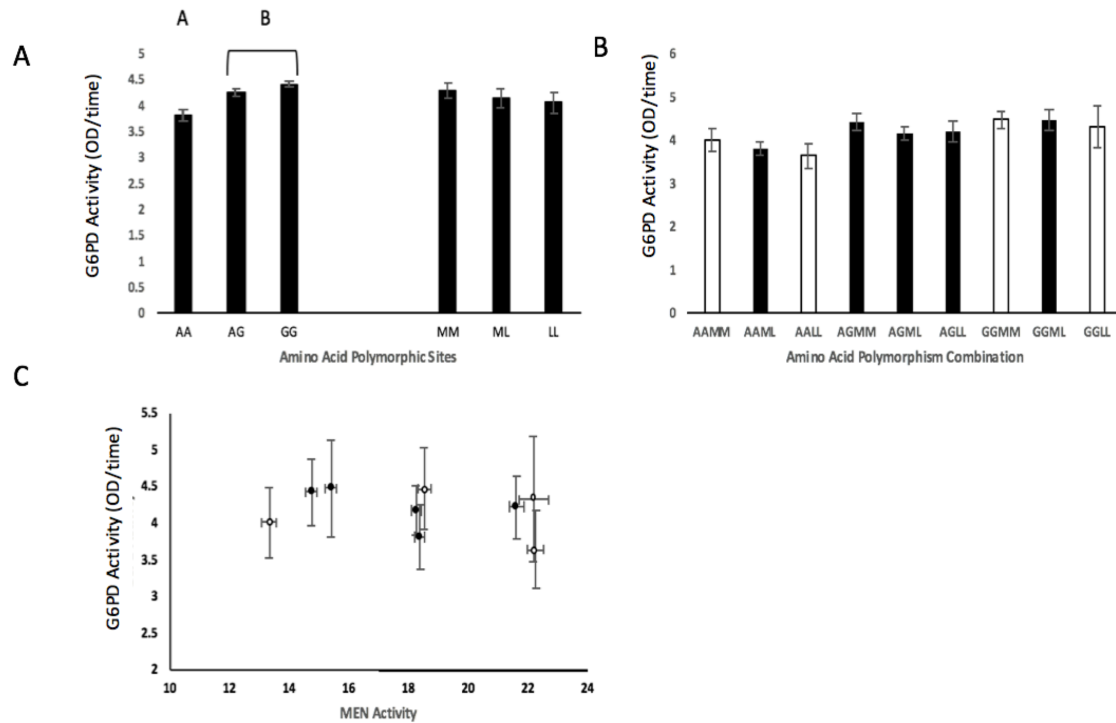


Figure 3.6 G6PD V_{max} Activity across different *Men* alleles and backgrounds under saturating conditions. (A) the G-A polymorphisms' genotype (leftmost three bars) and the M-L polymorphisms' genotype (rightmost three bars) mean G6PD V_{max} activity. (B) The nine different alleles, taking into account both polymorphisms. White bars represent the alleles that are homozygous at both polymorphic sites, and black bars represents where the allele differs at one or both polymorphic site. (C) Displays G6PD activity plotted against MEN activity. The white points represent the alleles that are homozygous, and the black points represent the alleles that are heterozygous. Bars indicate ± one standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different

3.3.6.2 MEN Activity did not influence Triglyceride Storage

Effect of polymorphism & genotype: Lipid synthesis is known to be a downstream target of NADPH, which is why triglyceride stores were also included in this study (Geer, Lindel, & Lindel, 1979). Triglyceride content remained fairly consistent across the MEN polymorphisms and genotypes, suggesting that the *Men* genotype did not impact this phenotype (Figure 3.8A-B; $F_{2,625}$=1.5953, p=0.2144; $F_{8,1875}$=5.8951, $p < 0.0001$). No differences in the polymorphism and genotypes were found in Chapter 2 as well.

MEN Correlation: Triglyceride content and MEN activity were compared to determine if a correlation existed between the two phenotypes (Figure 3.8C; R^2= 0.01137). Similar to the analysis of the polymorphism and genotype, there was no correlation between triglycerides and MEN activity. Multiple metabolites influence triglyceride content, NADPH being one of them. Since multiple loci contribute to the NADPH network, triglyceride storage is not likely to be significantly affected by the variation in MEN activity. A sufficient impact, such as 50% reduced activity is likely to lead to changes in lipid stores, as seen in previous work.

3.3.7 Response of NADPH Network to a Stressor

To fully understand one player's role in a metabolic system, two factors are essential to consider when analyzing its effects. As mentioned previously, systems do not act in isolation, which is why pathways should be studied together. The second is that metabolic systems are continually acclimating to the different conditions they are exposed to, in order to fulfill the needs of an organism. Since *Men* is a part of the NADPH pathway, and NADPH is an essential metabolite for the clearance of ROS, the *Men* genotype were further studied under oxidative stress, using paraquat as a source.

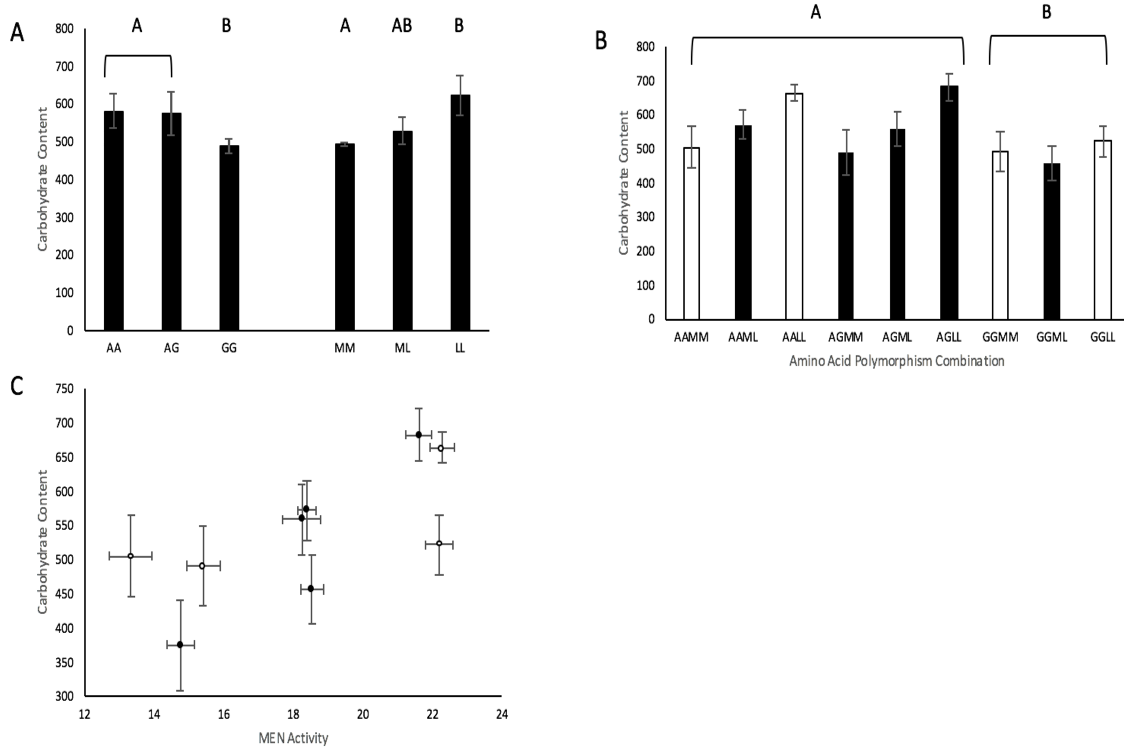


Figure 3.7. Carbohydrate content across different *Men* alleles and backgrounds under saturating conditions. (A) the G-A polymorphisms' genotype (leftmost three bars) and the M-L polymorphisms' genotype (rightmost three bars) mean carbohydrate content. (B) The nine different alleles, taking into account both polymorphisms. White bars represent the alleles that are homozygous at both polymorphic (C) Carbohydrate content plotted against MEN activity. The white points represent the alleles that are homozygous, and the black points represent the alleles that are heterozygous. Bars indicate ± one standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different.

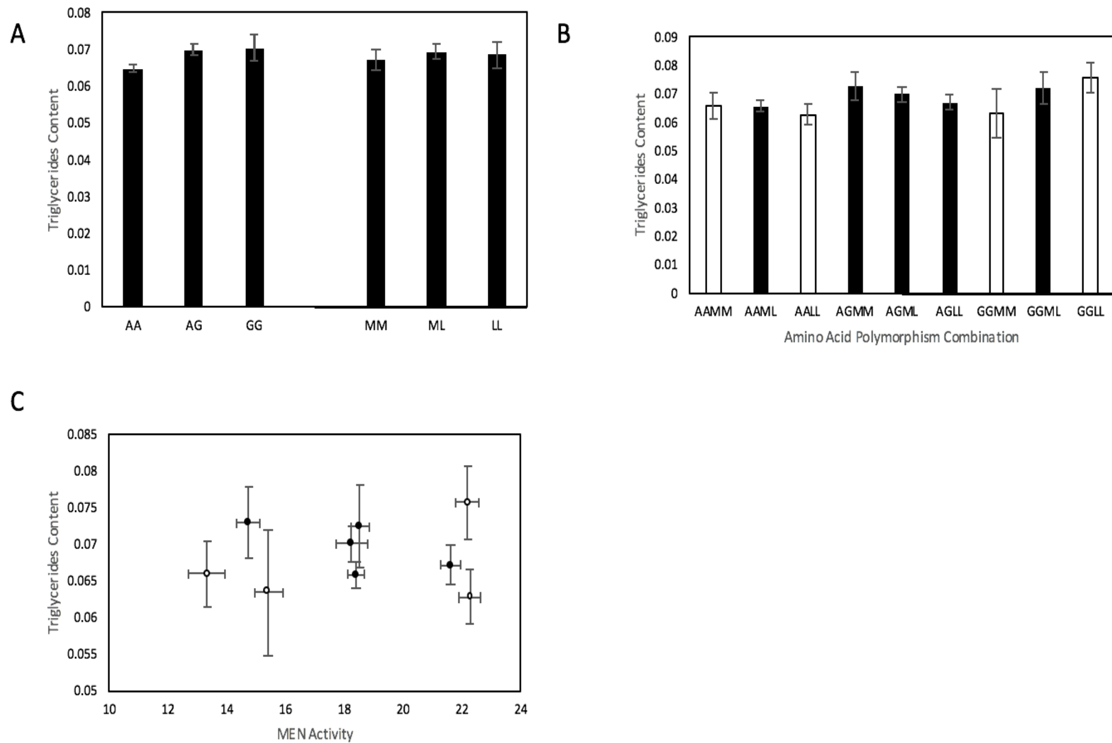


Figure 3.8. Triglycerides content across different *Men* alleles and backgrounds under saturating conditions. (A) the triglyceride content of the G-A genotypes (leftmost three bars) and the M-L genotypes (rightmost three bars) mean triglycerides content. (B) The nine different alleles, taking into account both polymorphisms. White bars represent the alleles that are homozygous at both polymorphic sites, and black bars represents where the allele differs at one or both polymorphic site. (C) Triglycerides content plotted against MEN activity. The white points represent the alleles that are homozygous, and the black points represent the alleles that are heterozygous. Bars indicate ± one standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different

3.3.7.1 The G-A Polymorphism Influences MEN Maximum Velocity under Oxidative Stress

To determine if MEN activity changed under an environmental stressor, the flies were placed on paraquat food for 24 hours. Paraquat exposure led to a significant decrease in MEN activity of 57% (Figure 3.9A; $F_{2,625}$=64.2272 $p < 0.0001$), similar to other studies analyzing the NADPH enzymes under paraquat. A decrease in MEN activity is also found in previous work with paraquat studying interactions within the NADPH pathway (Rzezniczak & Merritt, 2012).

Effect of polymorphism: The G-A polymorphism displayed significant differences, where the GG genotype had 27% higher MEN activity than the AA genotype ($F_{2,625}$=4.5364 $p = 0.0162$). The M-L polymorphism did not show significant differences when placed on paraquat food from each other ($F_{2,625}$=0.9939 $p = 0.3783$).

Effect of genotype: Paraquat treatment did not change the trend of effect of genotype on MEN activity. The genotype for MEN, on paraquat food, overall had the same trend as seen in Figure 3.1B, whereas you switch from an alanine to a glycine allele, there was an increase in MEN activity. The genotypes were not significantly different ($F_{8,1875}$=1.3934 $p = 0.2308$). Interestingly the heterozygous combinations did not fall in the range of the calculated values, or for them to be considered intermediates of their respected homozygous genotype. All the heterozygous genotype, which the exception of GGML had lower than expected activities. This lower activity may be advantageous due to the decreased concentration of NADPH available for paraquat. The lack of trend viewed for MEN activity is largely due to the cells responding to an environmental condition, allowing them to behave similarly.

3.3.7.2 The MEN Genotype does not alter IDH Activity under Oxidative Stress

To understand if IDH activity changed for the MEN alleles under oxidative stress condition, the same protocol was followed as mentioned previously. The IDH activity significantly decreased once exposed to paraquat (Figure 3.10A; $F_{2,625}$=67.7448, $p < 0.0001$), however, not to the same extent as MEN activity, decreasing by 27%. This slight decrease in IDH is consistent with previous work (Rzezniczak & Merritt, 2012).

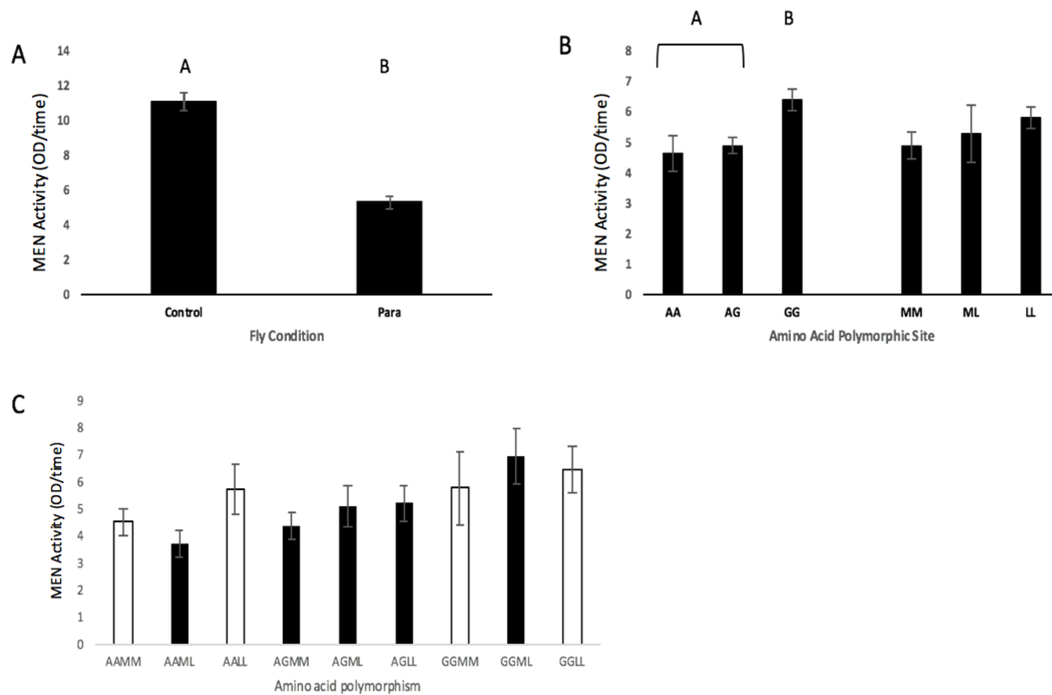


Figure 3.9. MEN V_{max} Activity (OD/time) across different alleles and backgrounds under paraquat condition. (A) MEN activity plotted against Fly condition, where control is standard cornmeal fly food and para is where paraquat is administered into the food. (B) the G-A polymorphisms' genotype (leftmost three bars) and the M-L polymorphisms' genotype (rightmost three bars) mean MEN V_{max} activity. (C) The nine different alleles, taking into account both polymorphisms. White bars represent the alleles that are homozygous at both polymorphic sites, and Black bars represent where the allele differs at one or both polymorphic site. Bars indicate ± one standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different

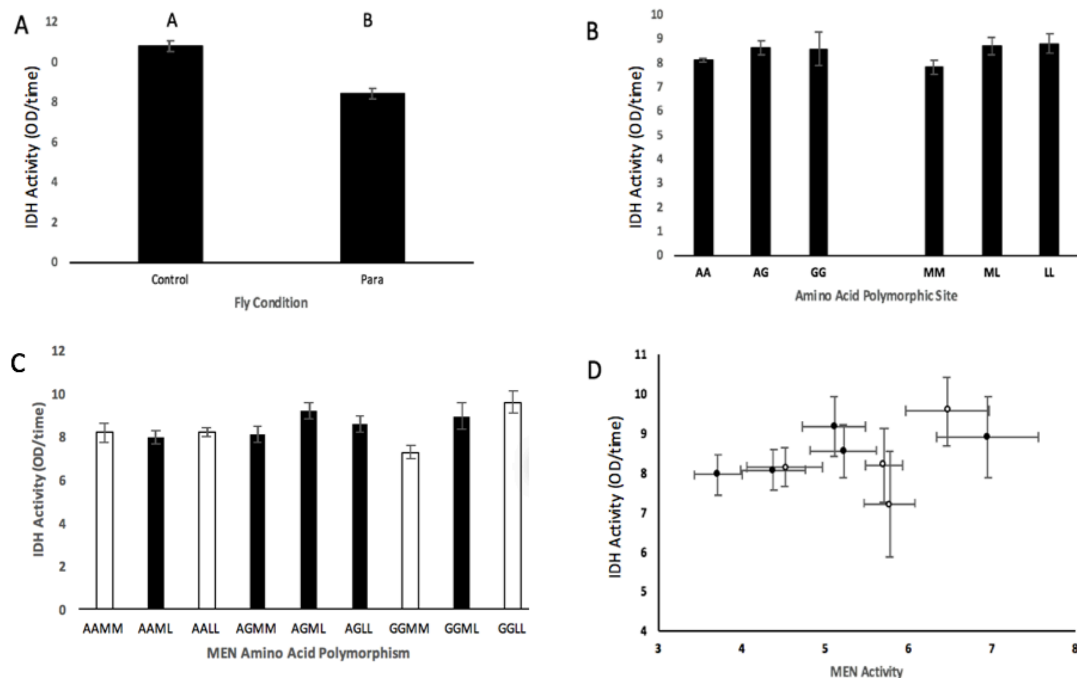


Figure 3.10. IDH V_{max} Activity across different alleles and backgrounds under paraquat condition. (A) IDH activity plotted against Fly condition, where control is standard cornmeal fly food and para is where paraquat is administered into the food. (B) the G-A polymorphisms' genotype (leftmost three bars) and the M-L polymorphisms' genotype (rightmost three bars) mean IDH V_{max} activity. (C) The nine different alleles, taking into account both polymorphisms. White bars represent the alleles that are homozygous at both polymorphic sites, and black bars represent where the allele differs at one or both polymorphic site. (D) IDH activity is plotted against MEN activity. White points represent the homozygous alleles, and black points represent the heterozygous alleles. Bars indicate ± one standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different.

Effect of polymorphism: At the G-A polymorphism, the IDH activities are relatively similar to one another, and do not significantly differ, (Figure 3.10B; $F_{2,625}$=0.5090 p = 0.6046). The MEN M-L genotypes also displayed similar IDH activities; however, the LL genotype had a 12% higher mean activity than the MM genotype, but these activities were not significantly different ($F_{2,625}$=2.2457 p = 0.1179).

Effect of genotype: The IDH activities are consistent across the genotypes of MEN activity, no significant trend observed (Figure 3.10C; $F_{8,1875}$=1.0260, p = 0.4338)

MEN Correlation: As MEN activity increases, a slight increase in IDH activity was observed, however, this difference was not significant (Figure 3.10D; R^2= 0.19641). The lack of difference in IDH activity across the MEN genotypes could be a function of a lack of difference initially, under control conditions or due to the system focusing on the clearance of ROS.

3.3.7.3 The *Men* genotypes do not alter G6PD Activity under Oxidative Stress

To determine whether G6PD activity changed for the MEN genotypes under oxidative stress, the same protocol was followed as mentioned previously. The G6PD activity remained consistent when exposed to paraquat compared to control samples, likely because G6PD is not as high a contributor to the NADPH ratio as MEN or IDH (Figure 3.11A). Previous work did not find significant differences between the paraquat and control flies for G6PD activity (Rzezniczak & Merritt, 2012).

Effect of polymorphism: At the MEN G-A polymorphism, the mean G6PD activity for the GG genotype was approximately 16% higher than the AA genotype; however, this difference was not significant (Figure 3.11B; $F_{2,625}$=2.4722 p = 0.0960). The MEN M-L polymorphism also displayed similar activity across genotypes, however, the LL genotype had 17% higher mean activity than the MM genotype, although the difference was not significant ($F_{2,625}$=2.4038 p = 0.1022).

Effect of genotype: The G6PD activities were consistent across the different MEN genotypes (Figure 3.11C; $F_{8,1875}$=1.1048, p = 0.3817). To determine if there was a correlation between MEN and G6PD activity, G6PD activity was compared to MEN activity under paraquat condition.

MEN Correlation: As MEN activity increased, an increase in G6PD activity was also observed, but due to the increase in variation (the error bars) the trend was not significant (Figure 3.11D; R^2= 4.1743). Overall, the *Men* genotypes do not alter the activity of G6PD under conditions of oxidative stress.

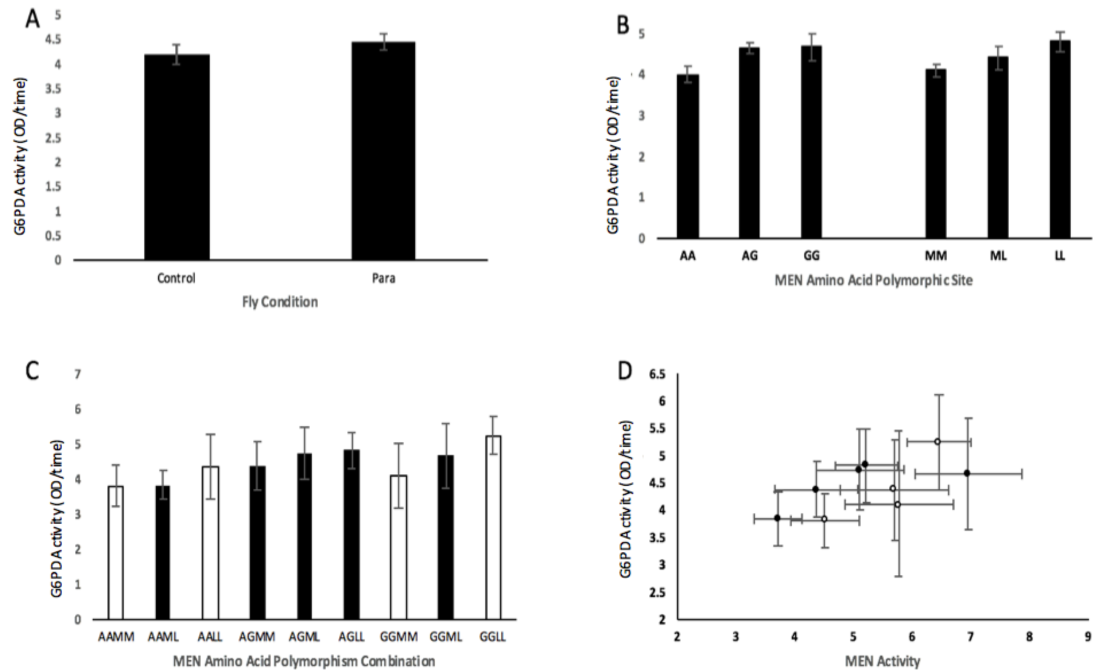


Figure 3.11. G6PD V_{max} Activity across different alleles and backgrounds under paraquat condition. (A) G6PD activity plotted against Fly condition, where control is standard cornmeal fly food and para is where paraquat is administered into the food. (B) the G-A polymorphisms' genotype (leftmost three bars) and the M-L polymorphisms' genotype (rightmost three bars) mean G6PD V_{max} activity. (C) The nine different alleles, taking into account both polymorphisms. White bars represent the alleles that are homozygous at both polymorphic sites, and black bars represent where the allele differs at one or both polymorphic site. (D) G6PD activity plotted against MEN activity. White points represent the homozygous alleles, and black points represent the heterozygous alleles. Bars indicate ± one standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different.

3.3.7.4 The MEN genotypes do not alter SOD Activity under Oxidative Stress

SOD is an antioxidant enzyme that plays a role in clearing up superoxide molecules present within an organism. SOD is a crucial antioxidant to study since superoxide molecules are generated with paraquat. To determine if SOD activity changes for the MEN genotypes under oxidative stress condition, a simple enzymatic assay was performed under saturating conditions. SOD activity was significantly higher under paraquat conditions compared to the control sample (Figure 3.12A, $F_{2,625}=34.6686$, $p < 0.0001$).

Effect of polymorphism: At the MEN G-A polymorphism, the mean SOD activity for the AA genotype was approximately 19% higher than the GG genotype; however, the difference was not significant (Figure 3.12B; $F_{2,625}=0.9617$ $p = 0.3903$). The MEN M-L polymorphism also displayed variation in activity, where the LL genotype had 4% higher meaner SOD activity than the MM genotype, although the difference was not significant ($F_{2,625}=0.1875$ $p = 0.8297$).

Effect of genotype: There was a lot of variation between the genotypes; however, no overall trend was observed between the genotypes and SOD activity (Figure 3.12C; $F_{8,1875}=0.4925$, $p = 0.8536$).

MEN Correlation: As MEN activity increases, there is a considerable variation between SOD activity, but no overall trend was observed (Figure 3.12D; $R^2= 0.0053$). SOD is a line of defense against ROS that accumulates within the cell, SOD indirectly uses NADPH, which might explain the lack of communication between SOD and MEN.

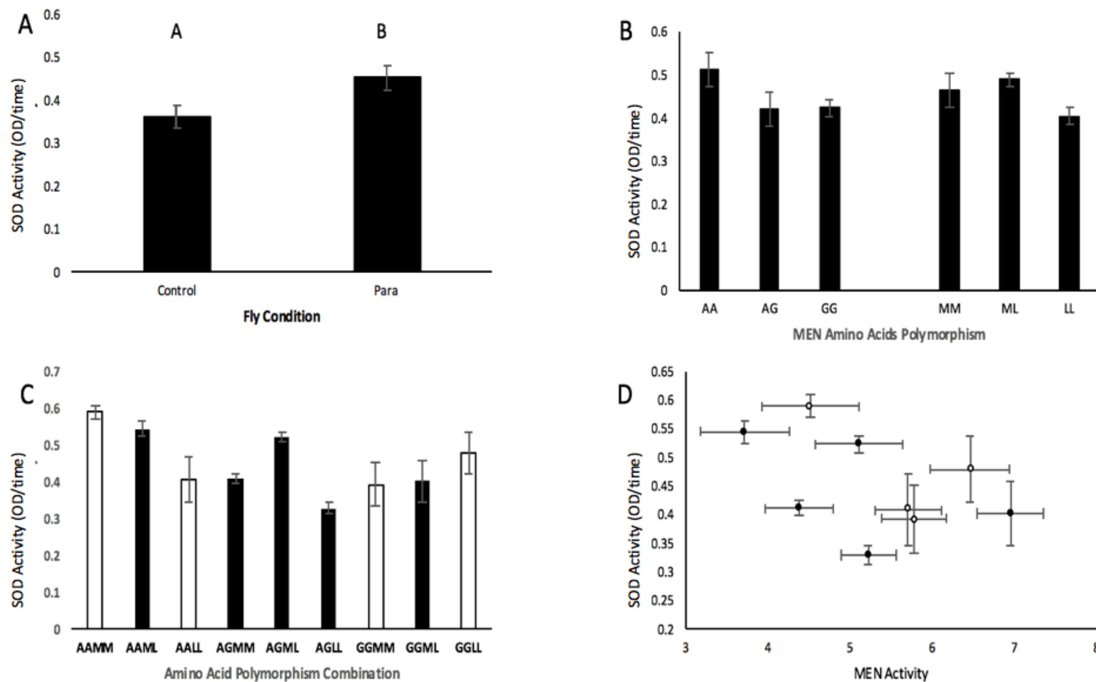


Figure 3.12. SOD V_{max} Activity across different alleles and backgrounds under paraquat condition. (A) is SOD activity plotted against Fly condition, where control is standard cornmeal fly food, and para is where paraquat is administered into the food. (B) the G-A polymorphisms' genotype (leftmost three bars) and the M-L polymorphisms' genotype (rightmost three bars) mean SOD V_{max} activity. (C) The nine different alleles, taking into account both polymorphisms. White bars represent the alleles that are homozygous at both polymorphic sites, and black bars represent where the allele differs at one or both polymorphic site. (D) SOD activity is plotted against MEN activity. White points represent the homozygous alleles, and black points represent the heterozygous alleles. Bars indicate ± one standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different.

3.3.7.5 The MEN genotypes do not alter GST Activity under Oxidative Stress

To determine if the activity of GST, an antioxidant enzyme, changed for the MEN genotypes under oxidative stress conditions, an enzymatic assay was performed under saturating conditions. Unexpectedly, paraquat treatment did not lead to elevated GST activity (Figure 3.13A). The lack of increase in activity under high ROS stress may mean that GST takes longer than the 24 hours exposure used here to become more active.

Effect of polymorphism: At the MEN G-A polymorphism, the mean GST activity for the AA genotype was approximately 17% higher than the GG genotype, but this difference was not significant (Figure 3.13 B; $F_{2,625}$=1.5391, p = 0.2262). GST activity remained consistent between the different allele groupings for the M-L polymorphism ($F_{2,625}$=0.7729, p = 0.4680).

Effect of genotype: Similar to SOD activity, there was variation between the MEN genotypes for GST activity, but no overall trend was observed (Figure 3.13C; $F_{8,1875}$=0.8243, p = 0.5866).

MEN Correlation: When MEN activity was compared to GST activity, there was a positive correlation, but it was not significant (Figure 3.13D; R^2= 0.02229). The lack of trend observed is most likely due to GST being further down the chain in the antioxidant defense system.

3.3.7.6 The MEN genotypes do not alter CAT Activity under Oxidative Stress

To determine if CAT activity changed for the MEN genotypes under oxidative stress conditions, an enzymatic assay was performed under saturating conditions. CAT activity was significantly higher in the samples exposed to paraquat compared to the control samples (Figure 3.14A; $F_{2,625}$=28.936, $p < 0.0001$).

Effect of polymorphism: At the MEN G-A polymorphism, the mean CAT activity for the AG genotype was approximately 24% higher than GG and AA homozygous genotypes, however, the difference was not significant (Figure 3.14B; $F_{2,625}$=0.8081, p = 0.4530). By being 24% higher, this suggests that CAT activity is not following an additive trend, and may point towards heterozygote advantage. For the MEN M-L

polymorphism, the mean CAT activity with the MM genotype was 22% higher than the LL genotype, but the difference was not significant (Figure 3.14B; $F_{2,625}$=0.6892, p = 0.5080).

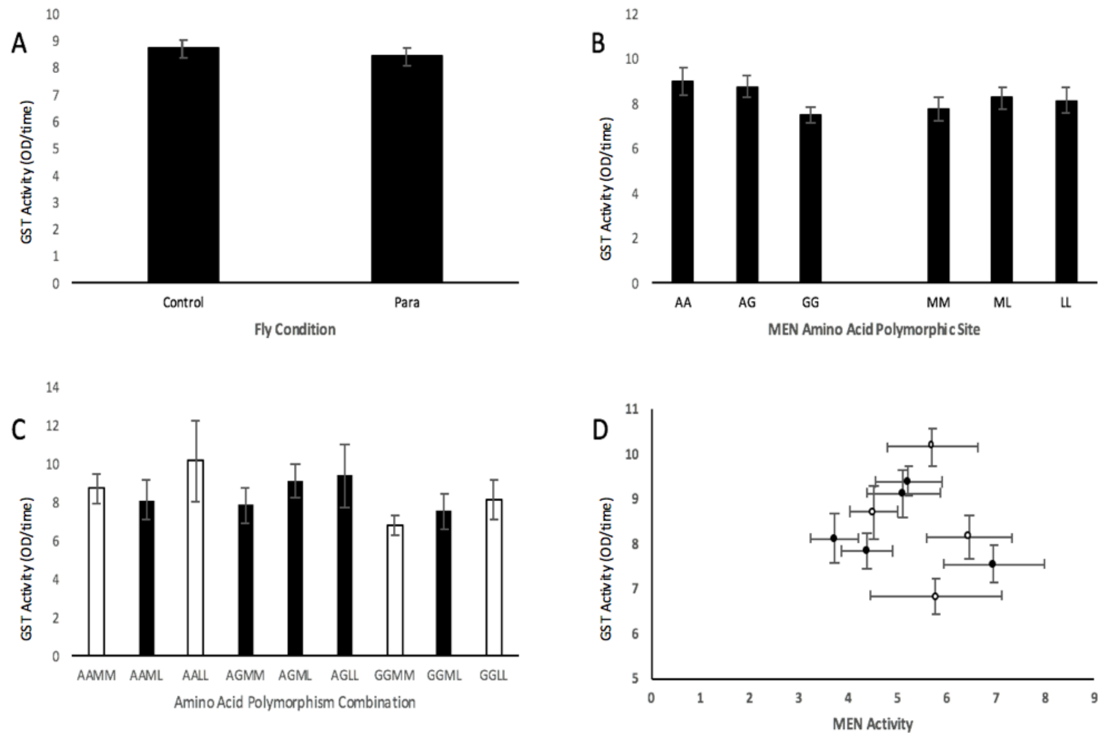


Figure 3.13. GST V_{max} Activity across different alleles and backgrounds under paraquat condition. (A) GST activity plotted against Fly condition, where control is standard cornmeal fly food and para is where paraquat is administered into the food. Figure (B) the G-A polymorphisms' genotype (leftmost three bars) and the M-L polymorphisms' genotype (rightmost three bars) mean GST V_{max} activity. (C) The nine different alleles, taking into account both polymorphisms. White bars represent the alleles that are homozygous at both polymorphic sites, and black bars represent where the allele differs at one or both polymorphic site. (D) GST activity is plotted against MEN activity. White points represent the homozygous alleles, and black points represent the heterozygous alleles. Bars indicate ± one standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different.

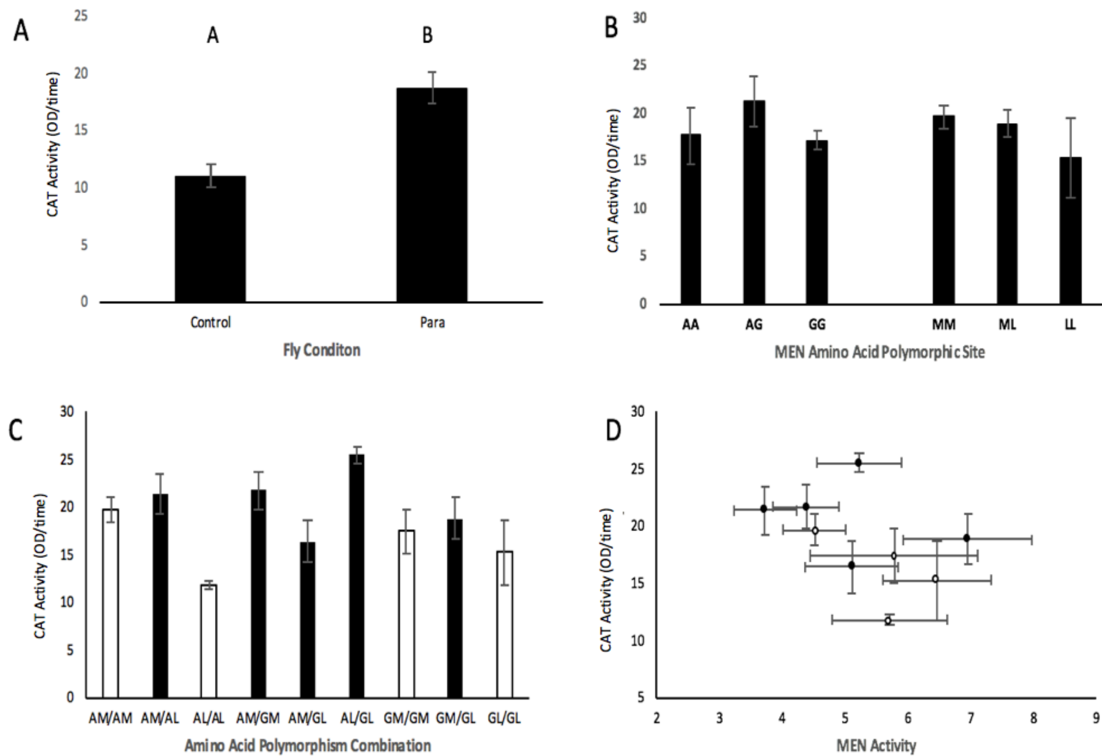


Figure 3.14. CAT V_{max} Activity across different alleles and backgrounds under paraquat condition. (A) CAT activity plotted against Fly condition, where control is standard cornmeal fly food and para is where paraquat is administered into the food. Figure (B) the G-A polymorphisms' genotype (leftmost three bars) and the M-L polymorphisms' genotype (rightmost three bars) mean CAT V_{max} activity. (C) The nine different alleles, taking into account both polymorphisms. White bars represent the alleles that are homozygous at both polymorphic sites, and black bars represents where the allele differs at one or both polymorphic site. Figure 3.14 D is where CAT activity is plotted against MEN activity. White points represent the homozygous alleles, and black points represent the heterozygous alleles. Bars indicate ± one standard error. An ANCOVA determined significant differences within the data set, and a Tukey Honesty Test provided the alleles/crosses that were significantly different.

Effect of genotype: Interestingly, between the genotypes, there are large variations, where all the heterozygous genotypes were superior in CAT activity compared to their respective homozygous lines (Figure 3.14C; $F_{8,1875}$=1.8281, p = 0.1069). The higher activities suggest the possibility of heterozygote advantage.

MEN Correlation: When comparing CAT activity against MEN activity, a negative correlation exists, whereas MEN activity increases there was a decrease in CAT activity, however, this was not significant (Figure 3.14D; R^2= 0.20646).

3.4 Discussion

3.4.1 MEN Phenotypic Variation across the Different Genotypes

Understanding the biochemical characteristics of alleles is a crucial first step in unraveling the biology of genetic variation and the maintenance of genotypes. In the case of single nucleotide polymorphisms (SNPs), two crucial factors play a role in phenotypic variation: the type of base pair change, and the base pair location. In *D. melanogaster* MEN, two SNPs lead to nucleotide base substitutions that change the amino acid sequence. The location of one of these polymorphisms is near the active site of the enzyme, whereas the other is buried within the protein; therefore, we expect to observe variation in MEN activity for the polymorphism near the active site (Merritt et al., 2005; Rzezniczak et al., 2012). Previous studies completed within the Merritt lab have examined these polymorphisms and their biochemical characteristics. Multiple lines are incorporated in this study to introduce genetic background effects that can drive changes in phenotypes. In Chapter 2, 16 lines were used, whereas Chapter 3 has 35 lines were used. The lines chosen for Chapter 2 had similar MEN V_{max} activities when compared to one another, whereas the lines in Chapter 3 had various MEN V_{max} activities ranging from low, mid and high activity.

For MEN V_{max}, the glycine allele has higher activity than the alanine allele, which also holds true with previous studies (Figure 3.1A). In previous studies, the M-L

polymorphism did not influence MEN activity, likely because the M-L polymorphism is buried within the protein, and not near functional components of the enzyme (Rzezniczak et al., 2012). Interestingly, the leucine allele has higher activity than the methionine allele (Figure 3.1A). Although the M-L polymorphism is known to be distant from an active site, it could be affected MEN activity in two ways. The first is the overall conformation of the protein is altered due to that amino acid substitution. Methionine and leucine are both nonpolar and bulky amino acids. However, methionine contains a sulfur atom in it is R group, which can form disulfide bonds with other sulfur-containing amino acids. By replacing methionine with a leucine, those potential disulfide bonds can be disrupted, where there are fewer interactions within the protein with the other R groups (Biswal et al., 2012; Pal & Chakrabarti, 2001). This loss of potential interactions could lead to a less compact protein, allowing greater flexibility overall and better interaction with substrates, hence the higher activity with the leucine allele. Furthermore, MEN is a homodimer tetramer, where two MEN monomers form dimers, and these dimers form tetramers. The M-L polymorphism could also potentially be located at the site where these dimers interact with one another. It is possible that the leucine allele leads to stronger positive cooperativity for MEN, through interactions between the dimers, corresponding to the higher activity at this polymorphic site. The reason why the M-L site did not appear to affect the MEN activity in previous studies is potential since the sample size, and genetic background effects were not studied as in-depth compared to this study. Here multiple lines were included for each genotype to study line effects. The variation between lines, in addition to the larger sample size, could lead to the quantification we see at the M-L polymorphism. Another interesting point is that leucine, the rarer allele, contributes to an overall higher V_{max} activity, compared to the other allele. This could potentially provide insight into why the leucine genotype is rare, where it might not be favorable to have high MEN activity.

The heterozygous genotype at the MEN locus was analyzed throughout, to understand the allelic frequency across the population further and provide insight into heterozygote advantage at the *Men* locus. The two *Men* polymorphisms exist in populations with a 50:50 ratios at for G-A polymorphism and a 90:10 ratio for the M-L polymorphism,

potentially under selection (Sezgin et al., 2004). The consistency of these frequencies across populations suggests a factor present within the biochemistry that may lead to this maintenance (Cormack, Hartl, & Clark, 1990). In Figure 1A, also viewed in Figure 1B, the heterozygous genotypes were quantified to be intermediates with their respected homozygous genotype. For example, AAML heterozygous genotypes would have the respective homozygous genotypes of AAMM and AALL. This is interesting because the heterozygous advantage is defined as the heterozygotes being superior to either homozygotes genotype (Kimura, Callahan, Petrov, & Messer, 2011; D. Sellis, Callahan, Petrov, & Messer, 2011; Diamantis Sellis, Callahan, Petrov, & Messer, 2011). In this scenario, it is likely that being at an intermediate value is the advantage.

As mentioned previously, it is important to study multiple lines, to include accounts for genetic backgrounds effects, since metabolic pathways interact with other genes, another source of phenotypic variation (Dworkin et al., 2009). The different crosses that contribute to the mean activity of the allele AAMM are shown in Figure 3.1C. Here, the genotypes at the *Men* locus are identical to one another; however, there is variation in the phenotypes between the lines, where 208 X 309 has the lowest activity, and 786 X 026 has the highest. This variation is known as the genetic background effect. Complex traits are dynamic such that phenotypes vary dependent on conditions. *Men* locus interacts with other loci that drives the changes within its phenotype, which was quantified across the different lines.

In previous studies, MEN V_{max} activity is a combination of structural changes and regulatory variation (Rzezniczak et al., 2012). Interestingly, the cytosine genotype has higher *Men* expression levels compared to the guanine genotype, Figure 3.2A. Previous studies within this network have shown that the guanine genotype had higher *Men* expression levels than the cytosine genotype, which correlated with the glycine allele having higher V_{max} activity than the alanine allele. Genetic background can explain why we see this variation in gene expression. Another potential reason is that novel factors outside of the known network are interacting with the locus to drive relative expression levels of *Men*. As these polymorphisms are located within the actual protein itself and not

upstream of the gene, the changes at the regulatory level is probably a consequence of another molecule, or set of molecules, interacting with this gene, leading to the observed variation in activity and expression levels (Bing et al., 2014; Rzezniczak et al., 2012). The second polymorphism is an adenine nucleotide substitution to a thymine nucleotide. Although the difference in V_{max} at the M-L polymorphism was not observed until now, a difference in *Men* expression level was quantified at the M-L polymorphism, where the adenine nucleotide had higher expression levels. However, in this study, the opposite was observed, where thymine had relatively higher *Men* expression levels, therefore a positive correlation between expression levels and activity levels was observed. In this study, it was also noted that there was a slight difference in V_{max} activity at the M-L polymorphism, where leucine (coded for from the thymine nucleotide) had higher V_{max} activity in comparison to methionine (adenine nucleotide). The higher V_{max} activity with the leucine allele could correspond to higher gene expression with the thymine nucleotide. The heterozygous genotypes, CG and AT, were also observed and both genotypes had relatively lower expression levels compared to the homozygous genotypes. Within the V_{max} phenotype, the heterozygotes appeared to have a linear additive effect, where the heterozygous genotypes are a mean of the homozygous genotypes. However, expression levels do not have the same linear additive effect as seen with the V_{max}. One common type of gene regulation within *D. melanogaster* is transvection, which is a form of *trans-* interaction between paired homologous chromosomes leading to misregulation (Duncan, 2002; Mellert & Truman, 2012; Morris, Chen, Geyer, & Wu, 1998; Wu & Morris, 1999). Transvection has been studied at the *Men* locus, indicating the possibility of trans-interactions with our results. When comparing *Men* expression levels and MEN activity, there is a correlation between the two phenotypes, where higher expression levels have higher MEN activity (Bing et al., 2014; Lum & Merritt, 2011). Regardless of the opposite trend viewed based on the previous study, the correlation still holds, indicating that structural and regulatory variations influence the V_{max} phenotype.

Changes to the amino acid sequence can affect the binding affinity of an enzyme due to the alteration in the structure. The structure is an essential factor that plays a role in the function of a protein. Both polymorphism (G-A and M-L) significantly influence the

binding affinity of the enzyme, although the M-L polymorphism had a more significant effect, Figure 3.3A, consistent with previous studies (Rzezniczak et al., 2012). Furthermore, the leucine allele has a lower K_m value when compared to the other allele, which indicates that leucine requires a lower concentration of substrate to effectively function. The leucine allele phenotype of a higher V_{max} activity and lower K_m is interesting, despite being the rarer genotype. The K_m phenotype may be an indicator as to why the leucine allele is being maintained within the population at low frequencies. It might not be beneficial to the organism to have such a low K_m value and high V_{max}, essentially being at extremes for these phenotypes. Previously, it was suggested that going from a methionine allele to a leucine allows for greater flexibility within the enzyme, since some intermolecular bonds are being lost, possibly aiding in the stronger binding affinity observed with leucine (Chapter 2).

Interestingly, the lines that contribute to the AAMM homozygous allele show much variation between one another. It is known that K_m is only dependent on the amino acid sequence; however, these variations suggest that K_m is not solely dependent on the protein's primary structure. The first factor that can be distinguished is that genetic background influences the binding affinity, seeing as each cross has a different K_m value. The second factor could be interaction at the MEN V_{max} level since these two phenotypes work together to drive the function of the enzyme. Differences in V_{max} can be based on the differences in how well the enzyme binds to its substrate. Therefore, the next phenotype observed was the ratio between V_{max} and K_m.

Looking at V_{max} and K_m as individual factors are essential to study the biochemistry of the enzyme. In order to understand the biological implications of MEN. These two phenotypes are examined together as they both drive the function of the enzyme, and that ratio provides a more relative *in vivo* activity of the enzyme. At the G-A polymorphism, there was no difference between either allele for activity level, Figure 3.4A. This could explain why the G-A polymorphism is found at a 50:50 allelic frequencies, as there is no distinguishable activity between the alleles. This indistinguishable activity also suggests that this polymorphism might not lead to a heterozygote advantage. However, the M-L polymorphism has variation between activities. Here the LL genotype has higher MEN

activity in comparison to the MM genotype, and the ML genotype has moderate MEN activity. This is an indicator of the allelic frequency, suggesting the M-L polymorphism is under selection. The heterozygotes are intermediates, consistent with previous studies, indicating that the heterozygous alleles for MEN phenotypes display a linear effect throughout its phenotypes (V_{max}, K_m, and V_{max}/K_m). The MEN activity was plotted against AAMM crosses (Figure 3.4C) to better understand the effects of genetic background on the phenotype. Here, there is a variation between the activities across the different lines, suggesting that genetic background does play a role in the phenotype of activity. However, it is also possible that there are interactions that are occurring in or outside of the NADPH pathway that is driving the observed changes (Figure 3.4C), which is a factor of complex traits (Rzezniczak et al., 2012).

3.4.2 MEN Alleles and Interactions within the NADPH Pathway

With complex traits, communication exists in and between networks, leading to diverse phenotypes; where one allele can display the result in multiple phenotypes depending on interactions between other genes and the environment. This dynamic system is crucial for an organism because it allows them to acclimate to changes over time (Kalmus, 1945; Merritt et al., 2005). The enzymes IDH and G6PD were analyzed to see if their activities varied for the MEN genotypes.

The MEN G-A polymorphism was with a significant difference in IDH activity, where the AA allele had higher activity than the GG (Figure 3.5A). For MEN activity, the GG genotype had higher activity than the AA genotype, Figure 3.1A. Therefore, to aid in maintaining the NADPH ratio in cells, it is possible that there is a decrease in IDH activity to compensate for the higher MEN activity. However, this is not consistent with a previous study, where a decrease in MEN activity leads to a decrease in IDH activity. In that paper, they used lab-engineered excision lines, where 100% and 50% enzymatic flies were generated (Merritt et al., 2005; Rzezniczak et al., 2012). In this study, we are using wild-type flies, and the differences between the alleles are not as pronounced.

Interestingly, only the G-A polymorphisms had significant differences in IDH activity, suggesting that the communication between loci is mainly dependent on the functional components of the enzymes, as opposed to structure. There was a slight decrease in

activity from the alanine allele towards glycine when comparing the different genotype (Figure 3.5B). The lack of interactions between IDH and MEN activity could be because the MEN alleles do not differ significantly compared to previous studies, as mentioned before, or that the interaction is not as strong because the NADPH ratio is maintained.

The G6PD activity also was found to significantly differed across the G-A polymorphism (Figure 3.6A). Specifically, G6PD activity followed a similar trend to MEN activity, where the GG genotype has a higher G6PD activity than AA genotype. As MEN activity increases, so does G6PD activity. With only the G-A polymorphism leading to differences in G6PD activity, similar to IDH activity, it further suggests that the communication between loci may be dependent on the functional components of the enzymes, rather than the structural components. G6PD activities did not differ across the genotypes, and when compared to MEN activity, there was no overall trend (Figure 3.6B-C).

The lack of trend observed in G6PD and IDH activity is likely due to the lack of substantial differences in MEN activity, consistent with previous studies (Chapter 2). Furthermore, the activities of the NADPH enzymes are dependent on the $NADP^+$: NADPH ratio. Therefore, if the ratio does not differ significantly between the MEN genotypes, it may not require an acclimatize response in IDH or G6PD (Bernard et al., 2011; Ying, 2008).

3.4.3 MEN Alleles Effects Distal Phenotypes

The main focus of this study is to understand the *Men* polymorphisms for their allelic frequencies. NADPH is a central component when studying *Men* since it is involved with numerous loci. In previous studies, variation in lipid content was observed for the MEN engineered flies (Merritt, Sezgin, Zhu, & Eanes, 2006). Therefore, lipid content was quantified here, in addition to carbohydrate concentration, since MEN is a metabolic enzyme.

Carbohydrate content varied across the different MEN genotypes (Figure 3.7A-B). The results were interesting, as a consistent difference in carbohydrate content was observed at the M-L polymorphism, where the leucine allele had higher carbohydrate content than

methionine, the rare allele had the more extreme phenotype. A positive correlation between carbohydrate content and MEN activity was observed, where an increase in MEN activity lead to an increase in carbohydrate content (Figure 3.7C). Overall, seeing a difference in carbohydrate content was striking as multiple lines were used, suggesting the potential for line effects masking a trend. Furthermore, carbohydrate content is a phenotype that is dependent on multiple loci, so it was expected that some of the variations would be masked due to this. However, The M-L polymorphism had an impact, suggesting that changes in the overall structure of the enzyme had more of an effect on this distal phenotype. From these observations, *Men* is potentially a key player for carbohydrate content.

Triglyceride storage is a vital phenotype to include for this investigation since NADPH is involved in lipid synthesis (Geer et al., 1979; Merritt et al., 2005; Rzezniczak & Merritt, 2012). Surprisingly, the variation in MEN activity did not alter the triglyceride content across the different polymorphism or genotypes (Figure 3.8A-C). This lack of interaction could be due to that MEN activity does not differ significantly from driving changes in triglyceride concentration, keeping an ideal ratio between $NADP^+$ and NADPH. Furthermore, triglyceride content is a distal phenotype of MEN, where MEN is an indirect contributor to the synthesis of triglycerides rather than a direct contributor. The polymorphisms may not drive significant changes in triglyceride storage as compared to a phenotype that is in its direct pathway.

3.4.4 *Men* genotypes Effects to Environmental Stressors

Biological systems are known to be dynamic, acclimatizing to changes that occur in the environment, therefore, essential to study metabolic systems in multiple environmental conditions. In this study, oxidative stress was chosen to be the environmental stressor, since NADPH is known to play a role in ROS clearance (Hosamani & Muralidhara, 2013; Scherz-Shouval & Elazar, 2007; Valko et al., 2007; Weber et al., 2012). Paraquat, a herbicide, induces oxidative stress, using NADPH to generate ROS (Bernard et al., 2011; Rzezniczak & Merritt, 2012; Ying, 2008). Here NADPH enzymes and antioxidant enzymes were examined, comparing paraquat fed flies to control flies.

For MEN activity, the control flies had significantly higher MEN activity than the paraquat fed flies (Figure 3.9A). The decrease in MEN activity is due to paraquat using NADPH to produce ROS in the cells. By reducing MEN activity, the NADPH concentration in cells is reduced, inhibiting the effect of paraquat on the flies. Concerning the SNPs, the G-A polymorphism was significantly different, where the GG genotype had higher MEN activity, compared to the AA genotype, consistent with the control conditions (Figure 3.9B). Therefore, similar trends were observed in the paraquat conditions compared to the control, and the difference is lower activity levels overall. The AA and AG genotypes had lower MEN activity, potentially suggesting an advantage for an organism to have either of those genotypes since they acclimatize well to paraquat stress. This lower MEN activity hints towards the potential of heterozygote advantage. For the genotypes, the heterozygous alleles (in the black bars of Figure 3.9C) had lower activity in comparison to the homozygous alleles (in the white bars), which the exception of GGML. This lower MEN activity suggests the possibility of an advantage as a heterozygote at the *Men* locus since lower MEN activity will lead to a low concentration of NADPH available for paraquat to use, reducing ROS present.

NADPH enzyme IDH activity has similar results as MEN activity (Figure 3.10A) where IDH activity was lower in paraquat exposed flies. The extent that IDH activity is driven down by paraquat was not the same as observed with MEN activity, since MEN contributes a higher percentage to the NADPH ratio in the cell, as opposed to the other enzymes. The reduction in MEN activity could be sufficient to lower the NADPH concentrations such that IDH activity was not required to be reduced to the same degree. In Figure 3.10B, the polymorphisms appear to be different from one another, but not to the level of difference observed in control. IDH activity was similar across MEN genotypes, such that when IDH activity was compared to MEN activity, no overall trend was observed (Figure 3.10C-D). One reason why IDH activity seems similar across the polymorphisms in the paraquat condition could be since MEN played a superior role in reducing the available NADPH for paraquat (Merritt et al., 2005). Therefore, IDH's response to paraquat is not as large. Another reason that the IDH activity does not differ across the polymorphism is that its function is halted, and the flies first response is to clear ROS.

G6PD activities were similar in the paraquat and control conditions, Figure 3.11A. This is since G6PD is not a large contributor to the NADPH ratio as compared to MEN and IDH; its activity is not required to reduce NADPH concentration. The AA genotype at the G-A polymorphism and the MM genotype at the M-L polymorphism is advantageous for being in a paraquat environment since the G6PD activity was the most reduced, Figure 3.11B. Nonetheless, this is a similar trend to that observed with the control sample.

To further understand the difference between the MEN genotypes, antioxidant enzyme activities were measured to quantify interactions in the NADPH network. SOD is an essential antioxidant enzyme since superoxide anion is the ROS that is generated through paraquat stress (Abolaji et al., 2017; Bernard et al., 2011; Hosamani & Muralidhara, 2013; Lessel et al., 2017). Overall, under the paraquat condition, there was higher SOD activity (Figure 3.12A). However, the SOD activity did not differ at the individual polymorphisms or the genotype in general (Figure 3.12B-C). The lack of difference in SOD activity between the genotype was interesting since the MEN activities differed. Potentially SOD activity is not interacting with the MEN genotype, possibly due to their indirect connection via the NADPH pathway (Bernard et al., 2011; Lessel et al., 2017). Variation in SOD activity did not correlate to differences in MEN activity. The absence of an observable trend between SOD activity and MEN activity could be due to the decrease in NADPH production as the first line of defense against paraquat. The low concentrations of NADPH available could be a factor of SOD activity since SOD indirectly requires NADPH to function.

The second antioxidant enzyme that was analyzed is GST. The activity of GST did not significantly differ between the paraquat and the control flies (Figure 3.13A). When comparing the GST activity to the polymorphism and genotype, there were no significant differences (Figure 3.13B-C). Since there were no differences between the MEN polymorphisms and genotypes, there was also no trend observed when comparing GST activity against MEN activity (Figure 3.13D). The lack of variation for GST activity across MEN genotypes is since GST is activated under secondary ROS responses. GST was not required to modify to acute paraquat conditioning, in this case, since the flies were placed on paraquat for 24 hours.

The last antioxidant enzyme that was studied was CAT. Under the paraquat condition, CAT activity was higher than in the control condition (Figure 3.14A). Similar to SOD activity, when compared across the polymorphisms and genotypes, CAT activity did not differ significantly (Figure 3.14 B-C). This could be due to the lack of communication between MEN and CAT. One overall trend is that the heterozygous genotypes (Figure 3.14C) had higher CAT activity. SOD is known to produce hydrogen peroxide as one of its by-products, which is a reactant for CAT (Lessel et al., 2017). Therefore, the increasing activity in SOD will increase the hydrogen peroxide concentration, leading to an increase in CAT activity.

In this investigation, we further studied the SNPs present in *Men* and how the biochemistry of the enzyme may lead to the allelic frequency we see across populations. Interestingly, by including more genetic backgrounds in this study, differences between the polymorphisms were amplified. The leucine allele, which is the rare polymorphisms, had a higher relative activity, which is an indicator as to why it is maintained at a lower allelic frequency- an extreme phenotype. By including an environmental stressor, we can test the acclimatization of the enzyme. When the flies were placed on paraquat to generate oxidative stress, the enzymes, MEN, IDH, and G6PD, the response was to prevent the accumulation and generation of ROS, thereby decreasing its activity. This decrease in activity becomes advantageous to the organisms, being able to acclimatize quickly to a stressor. Furthermore, the antioxidant enzymes do not have an overall trend for the polymorphism of MEN, but in the flies used, we do see an overall change in response, indicating some form of alteration.

Future work will be focused on teasing apart the structure of the enzyme, to understand how the polymorphism affects the overall structure of MEN. Another component that requires focus is the difference in expression levels of *Men* between polymorphisms. The polymorphisms are found within the gene, whereas regulatory elements can be found upstream and sometimes downstream of a gene. Therefore, the differences in regulation between the polymorphism may be indicated in other location than the gene itself. Studying these locations can provide insight into the maintenance of the alleles across the population.

3.5) References

Abolaji, A. O., Olaiya, C. O., Oluwadahunsi, O. J., & Farombi, E. O. (2017). Dietary consumption of monosodium L-glutamate induces adaptive response and reduction in the life span of Drosophila melanogaster. *Cell Biochemistry and Function*, *35*(3), 164–170. https://doi.org/10.1002/cbf.3259

Aebi, H. (1984). [13] Catalase in Vitro. *Methods in Enzymology*. https://doi.org/10.1016/S0076-6879(84)05016-3

Apel, K., & Hirt, H. (2004). Reactive Oxygen Species: Metabolism, Oxidative Stress, and Signal Transduction. *Annual Review of Plant Biology*. https://doi.org/10.1146/annurev.arplant.55.031903.141701

Bernard, K. E., Parkes, T. L., & Merritt, T. J. S. (2011). A model of oxidative stress management: Moderation of carbohydrate metabolizing enzymes in SOD1-null drosophila melanogaster. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0024518

Bing, X., Rzezniczak, T. Z., Bateman, J. R., & Merritt, T. J. S. (2014). Transvection-based gene regulation in Drosophila is a complex and plastic trait. *G3 (Bethesda, Md.)*, *4*(11), 2175–2187. https://doi.org/10.1534/g3.114.012484

Biswal, H. S., Gloaguen, E., Loquais, Y., Tardivel, B., & Mons, M. (2012). Strength of NH···S hydrogen bonds in methionine residues revealed by gas-phase IR/UV spectroscopy. *Journal of Physical Chemistry Letters*, *3*(6), 755–759. https://doi.org/10.1021/jz300207k

Boone, C., Bussey, H., & Andrews, B. J. (2007). Exploring genetic interactions and networks with yeast. *Nature Reviews Genetics*. https://doi.org/10.1038/nrg2085

Chandler, C. H., Chari, S., & Dworkin, I. (2013). Does your gene need a background check? How genetic background impacts the analysis of mutations, genes, and evolution. *Trends in Genetics*. https://doi.org/10.1016/j.tig.2013.01.009

Cormack, R. M., Hartl, D. L., & Clark, A. G. (1990). Principles of Population Genetics. *Biometrics*. https://doi.org/10.2307/2531471

Dixon, S. J., Costanzo, M., Baryshnikova, A., Andrews, B., & Boone, C. (2009). Systematic Mapping of Genetic Interaction Networks. *Annual Review of Genetics*. https://doi.org/10.1146/annurev.genet.39.073003.114751

Duncan, I. W. (2002). Transvection Effects in Drosophila. *Annual Review of Genetics*.

https://doi.org/10.1146/annurev.genet.36.060402.100441

Dworkin, I., Kennerly, E., Tack, D., Hutchinson, J., Brown, J., Mahaffey, J., & Gibson, G. (2009). Genomic consequences of background effects on scalloped mutant expressivity in the wing of drosophila melamogaster. *Genetics*. https://doi.org/10.1534/genetics.108.096453

Geer, B. W., Lindel, D. L., & Lindel, D. M. (1979). Relationship of the oxidative pentose shunt pathway to lipid synthesis in Drosophila melanogaster. *Biochemical Genetics*. https://doi.org/10.1007/BF00504310

Habig WH, Pabst MJ, J. W. (1974). Glutathione S-transferases. *J Biol Chem*.

Hosamani, R., & Muralidhara. (2013). Acute exposure of drosophila melanogaster to paraquat causes oxidative stress and mitochondrial dysfunction. *Archives of Insect Biochemistry and Physiology*, *83*(1), 25–40. https://doi.org/10.1002/arch.21094

Kalmus, H. (1945). Adaptative and selective responses of a population of Drosophila melanogaster containing e and e+ to differences in temperature, humidity and to selection for developmental speed. *Journal of Genetics*, *47*(1), 58–63. https://doi.org/10.1007/BF02989038

Kimura, M., Callahan, B. J., Petrov, D. A., & Messer, P. W. (2011). RULES FOR TESTING STABILITY OF A SELECTIVE POLYMORPHISM. *PNAS*, *42*(6), 336–340. https://doi.org/10.1073/pnas.42.6.336

Kostyuk, V. A., & Potapovich, A. I. (1989). Superoxide--driven oxidation of quercetin and a simple sensitive assay for determination of superoxide dismutase. *Biochemistry International*, *19*(5), 1117–1124.

Lessel, C. E., Parkes, T. L., Dickinson, J., & Merritt, T. J. S. (2017). Sex and Genetic Background Influence Superoxide Dismutase (cSOD)-Related Phenotypic Variation in Drosophila melanogaster. *G3 (Bethesda, Md.)*. https://doi.org/10.1534/g3.117.043836

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods*, *25*(4), 402–408. https://doi.org/10.1006/meth.2001.1262

Lum, T. E., & Merritt, T. J. S. (2011). Nonclassical regulation of transcription: Interchromosomal interactions at the Malic enzyme locus of Drosophila

melanogaster. *Genetics*, *189*(3), 837–849. https://doi.org/10.1534/genetics.111.133231

Mackay, T. F. C. (2004). The genetic architecture of quantitative traits: Lessons from Drosophila. *Current Opinion in Genetics and Development*. https://doi.org/10.1016/j.gde.2004.04.003

Mellert, D. J., & Truman, J. W. (2012). Transvection is common throughout the Drosophila genome. *Genetics*. https://doi.org/10.1534/genetics.112.140475

Merritt, T. J. S., Duvernell, D., & Eanes, W. F. (2005). Natural and synthetic alleles provide complementary insights into the nature of selection acting on the Men polymorphism of Drosophila melanogaster. *Genetics*, *171*(4), 1707–1718. https://doi.org/10.1534/genetics.105.048249

Merritt, T. J. S., Kuczynski, C., Sezgin, E., Zhu, C. T., Kumagai, S., & Eanes, W. F. (2009). Quantifying interactions within the NADP(H) enzyme network in Drosophila melanogaster. *Genetics*, *182*(2), 565–574. https://doi.org/10.1534/genetics.109.100677

Merritt, T. J. S., Sezgin, E., Zhu, C. T., & Eanes, W. F. (2006). Triglyceride pools, flight and activity variation at the Gpdh locus in Drosophila melanogaster. *Genetics*, *172*(1), 293–304. https://doi.org/10.1534/genetics.105.047035

Migula, P., Laszczyca, P., Augustyniak, M., Wilczek, G., Rozpedek, K., Kafel, A., & Woloszyn, M. (2004). Antioxidative defence enzymes, in beetles, from a metal pollution, gradient. *Biologia*, *59*(5), 645–654.

Morris, J. R., Chen, J.-L., Geyer, P. K., & Wu, C.-T. (1998). Two modes of transvection: enhancer action in trans and bypass of a chromatin insulator in cis. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.95.18.10740

Müller, K. R., Martins, I. K., Rodrigues, N. R., da Cruz, L. C., Filho, V. M. B., Macedo, G. E., … Posser, T. (2017). Anacardium microcarpum extract and fractions protect against paraquat-induced toxicity in drosophila melanogaster. *EXCLI Journal*, *16*, 302–312. https://doi.org/10.17179/excli2016-684

Pal, D., & Chakrabarti, P. (2001). Non-hydrogen bond interactions involving the methionine sulfur atom. *Journal of Biomolecular Structure and Dynamics*.

https://doi.org/10.1080/07391102.2001.10506725

Ralser, M., Heeren, G., Breitenbach, M., Lehrach, H., & Krobitsch, S. (2006). Triose phosphate isomerase deficiency is caused by altered dimerization-not catalytic inactivity-of the mutant enzymes. *PLoS ONE*, *1*(1). https://doi.org/10.1371/journal.pone.0000030

Reczek, C. R., & Chandel, N. S. (2015). ROS-dependent signal transduction. *Current Opinion in Cell Biology*. https://doi.org/10.1016/j.ceb.2014.09.010

Rzezniczak, T. Z., Douglas, L. A., Watterson, J. H., & Merritt, T. J. S. (2011). Paraquat administration in Drosophila for use in metabolic studies of oxidative stress. *Analytical Biochemistry*. https://doi.org/10.1016/j.ab.2011.08.023

Rzezniczak, T. Z., Lum, T. E., Harniman, R., & Merritt, T. J. S. (2012). A combination of structural and cis-regulatory factors drives biochemical differences in drosophila melanogaster malic enzyme. *Biochemical Genetics*, *50*(11–12), 823–837. https://doi.org/10.1007/s10528-012-9523-3

Rzezniczak, T. Z., & Merritt, T. J. S. (2012). Interactions of NADP-Reducing Enzymes Across Varying Environmental Conditions: A Model of Biological Complexity. *G3*, *2*(12), 1613–1623. https://doi.org/10.1534/g3.112.003715

Scherz-Shouval, R., & Elazar, Z. (2007). ROS, mitochondria and the regulation of autophagy. *Trends in Cell Biology*. https://doi.org/10.1016/j.tcb.2007.07.009

Sellis, D., Callahan, B. J., Petrov, D. A., & Messer, P. W. (2011). Heterozygote advantage as a natural consequence of adaptation in diploids. *Proceedings of the National Academy of Sciences*, *108*(51), 20666–20671. https://doi.org/10.1073/pnas.1114573108

Sellis, D., Callahan, B. J., Petrov, D. A., & Messer, P. W. (2011). Heterozygote advantage as a natural consequence of adaptation in diploids. *Proceedings of the National Academy of Sciences*, *108*(51), 20666–20671. https://doi.org/10.1073/PNAS.1114573108

Sezgin, E., Duvernell, D. D., Matzkin, L. M., Duan, Y., Zhu, C. T., Verrelli, B. C., & Eanes, W. F. (2004). Single-locus latitudinal clines and their relationship to temperate adaptation in metabolic genes and derived alleles in Drosophila melanogaster. *Genetics*, *168*(2), 923–931.

https://doi.org/10.1534/genetics.104.027649

Valko, M., Leibfritz, D., Moncol, J., Cronin, M. T. D., Mazur, M., & Telser, J. (2007). Free radicals and antioxidants in normal physiological functions and human disease. *The International Journal of Biochemistry & Cell Biology*. https://doi.org/10.1016/j.biocel.2006.07.001

Weber, A. L., Khan, G. F., Magwire, M. M., Tabor, C. L., Mackay, T. F., & Anholt, R. R. (2012). Genome-wide association analysis of oxidative stress resistance in Drosophila melanogaster. *PLoS One*. https://doi.org/10.1371/journal.pone.0034745

Wise, E. M., & Ball, E. G. (1964). Malic Enzyme and Lipogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *52*(1933), 1255–1263. https://doi.org/10.1073/pnas.52.5.1255

Wolf, J. B. (2003). Genetic architecture and evolutionary constraint when the environment contains genes. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.0635741100

Wu, C. ting, & Morris, J. R. (1999). Transvection and other homology effects. *Current Opinion in Genetics and Development*. https://doi.org/10.1016/S0959-437X(99)80035-5

Ying, W. (2008). NAD $^+$ /NADH and NADP $^+$ /NADPH in Cellular Functions and Cell Death: Regulation and Biological Consequences. *Antioxidants & Redox Signaling*, *10*(2), 179–206. https://doi.org/10.1089/ars.2007.1672

Zhang, J., Wang, X., Vikash, V., Ye, Q., Wu, D., Liu, Y., & Dong, W. (2016). ROS and ROS-Mediated Cellular Signaling. *Oxidative Medicine and Cellular Longevity*. https://doi.org/10.1155/2016/4350965

Chapter 4: Conclusion & Future Direction

4.1 Overview

In this current study, we discovered interesting biochemical and biological variations between the Men alleles and interactions between NADPH enzymes. My hypothesis about variation across the Men alleles was supported; the phenotypes of the heterozygous alleles were mainly a linear combination of their respected homozygous alleles. Furthermore, the level of interaction between the Men alleles and the known NADPH network enzymes was minimal. Interestingly, the flies' phenotypes' were altered by the environmental stressor, paraquat. The previously observed trend of decrease NADPH enzyme MEN and IDH activity is also present here when the flies were exposed to paraquat, while G6PD activity did not change. Future work will focus on understanding the current players in the NADPH network to the Men polymorphism. In this chapter, I discuss the conclusions obtained from the study and provide suggestions for future work. The future work would provide further insight into the results obtained here and understanding of the connection between genetic variation and biological complexity. The goal of understanding the mechanism between a genotype and its phenotypes is a project that will extend for many years. My master's work is the first few steps to answering those questions.

4.2 Men Alleles and the Biochemical Variation

Overall, the heterozygotes followed the same trend found in previous work using isothrid chromosomes. Surprisingly, in Chapter 3, with larger sample size, we found that the second site does impact the maximum velocity of MEN, although only the first site was located to impact the velocity in Chapter 2 and previous studies (Rzezniczak et al., 2012). We also found that the heterozygous combinations do not differ significantly from the average of the respective homozygous parental lines. This pattern tells us that the phenotype for Men has a linear relationship. This linear relationship was also true for the Km of MEN, which is the binding affinity of the enzyme. Similar to Vmax, the two polymorphisms were found to play a role in the Km phenotypes observed, which is consistent with previous work. Another more biologically relevant phenotype included in this study is the ratio between Vmax and Km, also known as the relative activity.

Interestingly, the rarer allele of the second polymorphism, leucine,, had a more significant influence on the ratio than methionine. This difference in phenotype could explain why the second polymorphism, M-L, is being maintained across the different population. Thermal stability did not show any significant differences between the Men alleles, which is most probably due to the fact of genetic background effects, where the different background can control the amount of protein produced (Dworkin et al., 2009).

In Chapter 2 and 3, we quantified a set of flies for Men relative expression levels. For Chapter 2, one allelic combination was quantified, which was AALL. Chapter 3, five allelic combinations were quantified. In Chapter 2, we found that the gene expressions are not linear like the Vmax activities were found to be. There is the potential of transvection as the reason for nonlinear expression levels, in which the pairing of homologous chromosomes and sharing of regulatory elements leads to misregulation (Duncan 2002; Fujioka et al. 2016; Mellert and Truman 2012; Morris et al. 1998; Wu and Morris 1999). It is also known that transvection occurs at the Men locus (Bing et al. 2014; Lum and Merritt 2011).

Furthermore, Vmax is a factor of structural and regulatory effects (Rzezniczak et al. 2012). We know that the alleles' overall Vmax shows a linear fashion, but the expression does not, therefore the combinations of the two are likely leading to this additive result. In Chapter 3, the Men expression results found were not consistent with previous work (Rzezniczak et al. 2012). At the C-G polymorphism, the cytosine allele was associated with higher expression and at the A-T polymorphism, the thymine with higher expression. Regardless of the differences, the one similar trend observed in the positive correlation between Men expression and MEN activity levels.

Future work for the biochemical characterization of the Men alleles could isolate the protein and analyze the Vmax under known protein concentration, as opposed to using the crude homogenate. This type of experiment with purified protein could provide insight on the how the protein concentrations differ for each allele since the expression levels differ, and how much the Vmax changes under isolation.

4.3 Interactions with the NADPH Pathway & Genetic Background Effects

For interactions within the NADPH Pathway, as expected, there were minimal interactions between the Men alleles and the other enzymes and downstream functions. In the analysis with a large sample set, with multiple lines, the interactions between other enzymes were expected to be small for a few reasons. The first is that the Men alleles do not differ substantially in their Vmax phenotype, as pointed out in previous studies (Rzezniczak et al. 2012). Lines used in this study are from wild-caught flies, with different MEN activities. The lines used in the previous study that did show more strong interactions were lab-engineered excision levels, generating 100% and 50% MEN activity (Merritt et al. 2005; Rzezniczak et al. 2012). Therefore, the differences between alleles were not as substantial as those in the previous study with the engineered lines. Secondly, with using multiple lines, any outliers would be averaged out. For the downstream phenotype, such as carbohydrate content, triglyceride content, and longevity, many genes contribute to driving those phenotypes. Once again, the differences in Men alleles are not significant to induce changes in these distal phenotypes.

Consistent with previous studies of Men, we found that, regardless of the allele, genetic background is a crucial player in determining the phenotype. Across all of the phenotypes studied, genetic background effect impacted the phenotype observed. Interestingly, Km is influenced by genetic background effects. The biochemistry of Km is expected to be only a function of the structure of the enzyme, specifically the amino acid sequence. However, our significant background effect results indicate is that there is another factor, or factors, that influence the Km, possibly suggesting bind of factors coded elsewhere in the genome to the MEN protein and modifying the binding kinetics of this enzyme.

Future work can be completed on the interactions with the NADPH pathway, discussed below. With regards to genetic background and Km, isolating for the protein is crucial to

study the background effects quantified in Chapter 2 and 3. Since Km is supposedly a factor of the amino acid sequence, understanding how the structure changes with the polymorphisms are essential next step, further discussed below.

4.4 Characterizing Metabolic Response to Environmental Stressors

For metabolic responses within oxidative stress, we found similar results to previous studies using excision lines (Rzezniczak and Merritt 2012). Overall, when flies were fed paraquat food, the NADPH enzymatic activities were decreased compared to the control group. This decrease in activity is how the flies' metabolism changed, likely because paraquat consumes NADPH in producing ROS. The same trends that occurred in the control group occurred in the paraquat flies, but with overall reduced activities value. Another way to induce oxidative stress within the flies is to control the oxygen levels that are present (Finkel and Holbrook 2000). This control on oxygen levels could potentially be another way of studying oxidative stress, rather than using NADPH requiring paraquat. The issue of using paraquat here is that MEN is an NADPH network enzyme and paraquat uses NADPH as a cofactor. The overlap between the two, paraquat and the NADPH network, can inhibit difference between the Men alleles. There are other ways to generate oxidative stress that would provide insight into a response between the Men alleles. In addition to oxidative stress, other stressors can be used to see the acclimatization of the enzyme to the inducing environment. Temperature stress, starvation, and desiccator stress are known to alter MEN activity within D. melanogaster (Bing et al. 2014; Rzezniczak and Merritt 2012). A more in-depth understanding of the polymorphisms and the maintenance of the alleles can be further studied using different stressors.

4.5 MEN Structure

One crucial factor discussed briefly is the structure of MEN and how the polymorphisms could potentially alter it. When the amino acid sequence changes, the interactions with the background and the R groups alter too. We assume that the changes in MEN

biochemistry are a function of changes in structure (Rzezniczak et al. 2012). However, the structure of the protein exclusively has yet to be studied. Future work can look into using software to understand how the structure can potentially change with the different alleles studied here. De novo protein structure prediction exists and can take the amino acid sequence and predict what the tertiary protein structure will be (Baker 2012). A complication is that MEN is a homo-tetrameric quaternary structure, with 25 different quaternary combinations from the different alleles. Being able to predict the changes in the structure of the protein will provide insight into the different phenotypes quantified in this study.

Once a protein is translated, there are post-translational modifications that can occur as well, such as phosphorylation, acetylation and so on (Wang et al. 2009). Theses addition groups added on the protein can alter the structure and function of the protein. Therefore, future work should be completed to study the post-translational modification as well as the structure.

4.6 The NADP+: NADPH Ratio

The NADPH network and the enzymes found in the network, work with one another to keep the NADP+ and NADPH ratio constant within cells since they are frequently required. NADPH acts as a cofactor for numerous enzymes and is in high demand within the cell. One assumption made throughout studying the NADPH network is that the NADP+ / NADPH ratio must remain constant, which is why the enzymatic activities are adjusting (Merritt et al. 2005; Rzezniczak et al. 2012; Ying 2008). However, it would be informative to study this phenotype directly. Future work regarding the NADP+: NADPH ratio should be considered. If the ratio is not being kept constant, that can explain why the phenotypes of the same alleles differ from one, on top of genetic background effects. Two methods can be completed: find the concentration with HPLC or through coupled enzymatic reaction assays, where UV-Vis spectrophotometry or Fluorescent spectrophotometry can be used to measure the concentration of the cofactors (Bernard et al. 2011; Zhu and Rand 2012).

4.7 6-phosphogluconate dehydrogenase

The final enzyme of the NADPH network is 6-phosphogluconate dehydrogenase, which follows glucose-6-phosphate dehydrogenase in the pentose phosphate pathway. Previous studies showed that 6-phosphogluconate dehydrogenase activities did not differ from glucose-6-phosphate since it follows immediately after. This result is the reason why most studies examine G6PD, but not 6PGD. Nonetheless, there can be different alleles for this enzyme, leading to variable phenotype, as we seen with genetic background effects. Therefore, there is a possibility that the enzymatic activities are not similar to glucose-6-phosphate (Wilton et al. 1982). Future work studying this network could include 6-phosphogluconate dehydrogenase to grasp the bigger picture. It is possible that the variation expected in the enzymes of the network can be found within in this enzyme. A simple protocol can be designed within the Merritt Lab for this enzyme and can be applied to the same methodology that is currently being used (Labate and Eanes 1992).

4.8 Environmental Conditions and Stressors

In this study, there was only one environmental stressor administered, which was paraquat incorporated into the fly food. Paraquat was a good stressor to include since previous work completed in the lab used paraquat to generated oxidative stress (Rzezniczak et al. 2011; Rzezniczak and Merritt 2012). However, since we are studying the NADPH network, and paraquat uses NADPH to generate the ROS, some of the differences between the alleles could have been masked. We wanted to study the heterozygote advantage, and the next step to understanding the system was to include another environmental condition. By using paraquat, which uses NADPH, we can restrict the acclimatization the organism needs to make since they are focusing on the cofactor rather than the ROS generated. Other possible stressors to test are temperature sensitivity, starvation, desiccator food, and oxidative stress through changes in atmospheric pressure (Bing et al. 2014; Finkel and Holbrook 2000; Rzezniczak and Merritt 2012). By studying each condition, it will provide more insight into how the maintenance of the Men alleles, since they will have to acclimatize to circumstances outside of their network.

4.9 mRNA degradation

The central dogma of molecular genetics is a simplistic overview of how we go from genotype to phenotype. However, the process is more complex. Once DNA is transcribed into RNA, numerous processes occur before it is translated into protein, and as mentioned before, there are post-translation processes that also occur. With regards to mRNA, by switching the nucleotide base, it can affect the half-life and stability of the mRNA (Beelman and Parker 1995; Valencia-Sanchez et al. 2006). Therefore, the transcript levels can vary between alleles, leading to different enzyme activities. The general expression levels obtained from the qPCR experiment is reverse transcriptase, where RNA is converted in cDNA. Therefore, the altered expression levels could be dependent on the mRNA half-life, where one allele could have a smaller half-life leading to less protein present in that sample. Since one of the focus is attempting to understand what is maintaining these alleles in across the different population, and another is understanding how we go from genotype to phenotype, including the effects of RNA would be crucial in answering these questions.

4.10 Conclusion

My research is a step towards understanding the pathway from genetic variation and biological complexity. It is expected that the mechanism between genotype and phenotype is quite complex, however, here, we used SNPs in the Men locus in Drosophila melanogaster to take the first few steps to answer that question. We quantified important biochemical characteristics within MEN and interactions that occur within the NADPH network. Future work on understanding the connection between genotype and phenotype, described in this chapter, could be potentially completed on a larger scale, allowing for differences between alleles to be teased out. It is important to isolate for the protein from the homogenate so that we can separate protein structure and regulatory effects from one another. As of now, we see a linear effect of activity and non-linear effect for regulation. There is a disconnect between the two that remains unsolved. Furthermore, isolating the NADPH enzymes, it will allow us to quantify interactions between the pathway better. If isolation of protein is required, the number of

genetic background effects will be reduced since a larger fly sample is needed. Future work should also look at different environmental stressors exposures since paraquat uses NADPH to generate ROS. The response between the alleles to another environmental stressor will lead to different quantifications since it will not impede in the direct pathway, as paraquat did. Furthermore, I also suggest understanding the exact location of the polymorphism in MEN's structure, which can explain some of the results we obtained in this study. These main approaches, based on my results presented in Chapter 2 and Chapter 3, will provide more insight into the quantification here and how genotypes overall lead to phenotype in an organism.

4.11 References

Baker, Monya. 2012. "De Novo Genome Assembly: What Every Biologist Should Know." *Nature Methods*.

Beelman, Clare A. and Roy Parker. 1995. "Degradation of MRNA in Eukaryotes." *Cell*.

Bernard, Kristine E., Tony L. Parkes, and Thomas J. S. Merritt. 2011. "A Model of Oxidative Stress Management: Moderation of Carbohydrate Metabolizing Enzymes in SOD1-Null Drosophila Melanogaster." *PLoS ONE*.

Bing, Xinyang, Teresa Z. Rzezniczak, Jack R. Bateman, and Thomas J. S. Merritt. 2014. "Transvection-Based Gene Regulation in Drosophila Is a Complex and Plastic Trait." *G3 (Bethesda, Md.)* 4(11):2175–87.

Boone, Charles, Howard Bussey, and Brenda J. Andrews. 2007. "Exploring Genetic Interactions and Networks with Yeast." *Nature Reviews Genetics*.

Chandler, Christopher H., Sudarshan Chari, and Ian Dworkin. 2013. "Does Your Gene Need a Background Check? How Genetic Background Impacts the Analysis of Mutations, Genes, and Evolution." *Trends in Genetics*.

Charlesworth, D. and B. Charlesworth. 1987. "Inbreeding Depression and Its Evolutionary Consequences." *Annual Review of Ecology and Systematics*.

Chow, Clement Y. 2016. "Bringing Genetic Background into Focus." *Nature Reviews Genetics*.

Cordell, Heather J. 2009. "Detecting Gene-Gene Interactions That Underlie Human Diseases." *Nature Reviews Genetics*.

Cormack, R. M., D. L. Hartl, and A. G. Clark. 1990. “Principles of Population Genetics.” *Biometrics*.

Dixon, Scott J., Michael Costanzo, Anastasia Baryshnikova, Brenda Andrews, and Charles Boone. 2009. “Systematic Mapping of Genetic Interaction Networks.” *Annual Review of Genetics*.

Dudash, Mr and Cb Fenster. 2000. “Inbreeding and Outbreeding Depression in Fragmented Populations.” *Genetics, Demography and Viability of Fragmented Populations*.

Duncan, Ian W. 2002. “Transvection Effects in Drosophila.” *Annual Review of Genetics*.

Dworkin, Ian et al. 2009. “Genomic Consequences of Background Effects on Scalloped Mutant Expressivity in the Wing of Drosophila Melamogaster.” *Genetics*.

Finkel, Toren and Nikki J. Holbrook. 2000. “Oxidants, Oxidative Stress and the Biology of Ageing.” *Nature*.

Fujioka, Miki, Hemlata Mistry, Paul Schedl, and James B. Jaynes. 2016. “Determinants of Chromosome Architecture: Insulator Pairing in Cis and in Trans.” *PLoS Genetics* 12(2):1–25.

Geer, B. W., D. L. Lindel, and D. M. Lindel. 1979. “Relationship of the Oxidative Pentose Shunt Pathway to Lipid Synthesis in Drosophila Melanogaster.” *Biochemical Genetics*.

Griendling, Kathy K. et al. 2016. *Measurement of Reactive Oxygen Species, Reactive Nitrogen Species, and Redox-Dependent Signaling in the Cardiovascular System: A Scientific Statement from the American Heart Association*. Vol. 119.

Habig WH, Pabst MJ, Jakoby WB. 1974. “Glutathione S-Transferases.” *J Biol Chem*.

Hall, J. G. and R. K. Koehn. 1983. “The Evolution Of Enzyme Catalytic Efficiency and Adaptive Inference From Steady-State Kinetic Data.” *Evolutionary Biology*.

Hall, John G. 1985. “Temperature-Related Kinetic Differentiation of Glucosephosphate Isomerase Alleloenzymes Isolated from the Blue Mussel, Mytilus Edulis.” *Biochemical Genetics*.

Hosamani, Ravikumar and Muralidhara. 2013. “Acute Exposure of Drosophila Melanogaster to Paraquat Causes Oxidative Stress and Mitochondrial Dysfunction.” *Archives of Insect Biochemistry and Physiology* 83(1):25–40.

Huang, Wen et al. 2014. "Natural Variation in Genome Architecture among 205 Drosophila Melanogaster Genetic Reference Panel Lines." *Genome Research*.

Kalmus, H. 1945. "Adaptative and Selective Responses of a Population of Drosophila Melanogaster Containing e and E+ to Differences in Temperature, Humidity and to Selection for Developmental Speed." *Journal of Genetics* 47(1):58–63.

Kimura, M., Benjamin J. Callahan, Dmitri A. Petrov, and Philipp W. Messer. 2011. "RULES FOR TESTING STABILITY OF A SELECTIVE POLYMORPHISM." *PNAS* 42(6):336–40.

Kostyuk, V. A. and A. I. Potapovich. 1989. "Superoxide--Driven Oxidation of Quercetin and a Simple Sensitive Assay for Determination of Superoxide Dismutase." *Biochemistry International* 19(5):1117–24.

Labate, J. and W. F. Eanes. 1992. "Direct Measurement of in Vivo Flux Differences between Electrophoretic Variants of G6PD from Drosophila Melanogaster." *Genetics*.

Lessel, Courtney E., Tony L. Parkes, Joel Dickinson, and Thomas J. S. Merritt. 2017. "Sex and Genetic Background Influence Superoxide Dismutase (CSOD)-Related Phenotypic Variation in Drosophila Melanogaster." *G3 (Bethesda, Md.)*.

Real-Time Quantitative PCR and the 2-ΔΔCT Method." *Methods* 25(4):402–8.

Lum, Thomas E. and Thomas J. S. Merritt. 2011. "Nonclassical Regulation of Transcription: Interchromosomal Interactions at the Malic Enzyme Locus of Drosophila Melanogaster." *Genetics* 189(3):837–49.

Mackay, Trudy F. C. 2004. "The Genetic Architecture of Quantitative Traits: Lessons from Drosophila." *Current Opinion in Genetics and Development*.

Mellert, David J. and James W. Truman. 2012. "Transvection Is Common throughout the Drosophila Genome." *Genetics*.

Merritt, Thomas J. S. et al. 2009. "Quantifying Interactions within the NADP(H) Enzyme Network in Drosophila Melanogaster." *Genetics* 182(2):565–74.

Merritt, Thomas J. S., David Duvernell, and Walter F. Eanes. 2005. "Natural and Synthetic Alleles Provide Complementary Insights into the Nature of Selection Acting on the Men Polymorphism of Drosophila Melanogaster." *Genetics* 171(4):1707–18.

Merritt, Thomas J. S., Efe Sezgin, Chen Tseh Zhu, and Walter F. Eanes. 2006. "Triglyceride Pools, Flight and Activity Variation at the Gpdh Locus in Drosophila Melanogaster." *Genetics* 172(1):293–304.

Morris, J. R., J. L. Chen, P. K. Geyer, and C. T. Wu. 1998. "Two Modes of Transvection: Enhancer Action in Trans and Bypass of a Chromatin Insulator in Cis." *Proceedings of the National Academy of Sciences of the United States of America*.

Müller, Katiane Raquel et al. 2017. "Anacardium Microcarpum Extract and Fractions Protect against Paraquat-Induced Toxicity in Drosophila Melanogaster." *EXCLI Journal* 16:302–12.

Rzezniczak, T. Z., L. A. Douglas, J. H. Watterson, and T. J. S. Merritt. 2011. "Paraquat Administration in Drosophila for Use in Metabolic Studies of Oxidative Stress." *Analytical Biochemistry*.

Rzezniczak, Teresa Z., Thomas E. Lum, Robert Harniman, and Thomas J. S. Merritt. 2012. "A Combination of Structural and Cis-Regulatory Factors Drives Biochemical Differences in Drosophila Melanogaster Malic Enzyme." *Biochemical Genetics* 50(11–12):823–37.

Rzezniczak, Teresa Z. and Thomas J. S. Merritt. 2012. "Interactions of NADP-Reducing Enzymes Across Varying Environmental Conditions: A Model of Biological Complexity." *G3* 2(12):1613–23.

Scherz-Shouval, Ruth and Zvulun Elazar. 2007. "ROS, Mitochondria and the Regulation of Autophagy." *Trends in Cell Biology*.

Schreiber, Stuart L. 2005. "Small Molecules: The Missing Link in the Central Dogma." *Nature Chemical Biology*.

Sellis, D., B. J. Callahan, D. A. Petrov, and P. W. Messer. 2011. "Heterozygote Advantage as a Natural Consequence of Adaptation in Diploids." *Proceedings of the National Academy of Sciences* 108(51):20666–71.

Sellis, Diamantis, Benjamin J. Callahan, Dmitri A. Petrov, and Philipp W. Messer. 2011. "Heterozygote Advantage as a Natural Consequence of Adaptation in Diploids." *Proceedings of the National Academy of Sciences* 108(51):20666–71.

Steinberg, Martin H. 2008. "Sickle Cell Anemia, the First Molecular Disease: Overview

of Molecular Etiology, Pathophysiology, and Therapeutic Approaches." *TheScientificWorldJournal*.

Valencia-Sanchez, Marco Antonio, Jidong Liu, Gregory J. Hannon, and Roy Parker. 2006. "Control of Translation and MRNA Degradation by MiRNAs and SiRNAs." *Genes and Development*.

Valko, Marian et al. 2007. "Free Radicals and Antioxidants in Normal Physiological Functions and Human Disease." *The International Journal of Biochemistry & Cell Biology*.

Wang, K. et al. 2009. "Genome-Wide Identification of Post-Translational Modulators of Transcription Factor Activity in Human B Cells." *Nature Biotechnology*.

Watt, W. B. and a M. Dean. 2000. "Molecular-Functional Studies of Adaptive Genetic Variation in Prokaryotes and Eukaryotes." *Annual Review of Genetics*.

Weber, A. L. et al. 2012. "Genome-Wide Association Analysis of Oxidative Stress Resistance in Drosophila Melanogaster." *PLoS One*.

Wilton, A. N., C. C. Laurie-Ahlberg, T. H. Emigh, and J. W. Curtsinger. 1982. "Naturally Occurring Enzyme Activity Variation in Drosophila Melanogaster. II. Relationships among Enzymes." *Genetics*.

Wise, E. M. and E. G. Ball. 1964. "Malic Enzyme and Lipogenesis." *Proceedings of the National Academy of Sciences of the United States of America* 52(1933):1255–63.

Wolf, J. B. 2003. "Genetic Architecture and Evolutionary Constraint When the Environment Contains Genes." *Proceedings of the National Academy of Sciences*.

Wu, C. ting and James R. Morris. 1999. "Transvection and Other Homology Effects." *Current Opinion in Genetics and Development*.

Ying, Weihai. 2008. "NAD $^{+}$ /NADH and NADP $^{+}$ /NADPH in Cellular Functions and Cell Death: Regulation and Biological Consequences." *Antioxidants & Redox Signaling* 10(2):179–206.

Zhang, Jixiang et al. 2016. "ROS and ROS-Mediated Cellular Signaling." *Oxidative Medicine and Cellular Longevity*.

Zhu, Chen Tseh and David M. Rand. 2012. "A Hydrazine Coupled Cycling Assay Validates the Decrease in Redox Ratio under Starvation in Drosophila." *PLoS ONE* 7(10).