

A Combination of Structural and *Cis*-Regulatory Factors Drives Biochemical Differences in *Drosophila melanogaster* Malic Enzyme

Teresa Z. Rzezniczak · Thomas E. Lum ·
Robert Harniman · Thomas J. S. Merritt

Received: 21 September 2011 / Accepted: 28 February 2012 / Published online: 26 June 2012
© Springer Science+Business Media, LLC 2012

Abstract The evolutionary significance of molecular variation is still contentious, with much current interest focusing on the relative contribution of structural changes in proteins versus regulatory variation in gene expression. We present a population genetic and biochemical study of molecular variation at the malic enzyme locus (*Men*) in *Drosophila melanogaster*. Two amino acid polymorphisms appear to affect substrate-binding kinetics, while only one appears to affect thermal stability. Interestingly, we find that enzyme activity differences previously assigned to one of the polymorphisms may, instead, be a function of linked regulatory differences. These results suggest that both regulatory and structural changes contribute to differences in protein function. Our examination of the *Men* coding sequences reveals no evidence for selection acting on the polymorphisms, but earlier work on this enzyme indicates that the biochemical variation observed has physiological repercussions and therefore could potentially be under natural selection.

Keywords *Malic enzyme (Men)* · Molecular evolution · Genetic variation · Enzyme function

Introduction

The functional and evolutionary impact of genetic variation is, surprisingly, still an open question. Great strides have been made in determining genotypes and quantifying phenotypes, but the connection between the two is often elusive. Similarly, the relative evolutionary importance of coding (structural) and noncoding (*cis*-regulatory) variation is still very much in question, with strong arguments being made in either camp (Hoekstra and Coyne 2007; Wray 2007). In the present study,

T. Z. Rzezniczak · T. E. Lum · R. Harniman · T. J. S. Merritt (✉)
Department of Chemistry and Biochemistry, Laurentian University, Sudbury, ON P3E 2C6, Canada
e-mail: tmerritt@laurentian.ca

we continue the examination of the population genetics and biochemistry of a genetic variation in *Drosophila melanogaster* at the *malic enzyme* locus (*Men*), a model system with interesting genetic, biochemical, and physiological variations that suggest the potential for natural selection (Merritt et al. 2005, 2009).

Malic enzyme (MEN) oxidizes malate to pyruvate with the concurrent reduction of NADP to NADPH, a major reductant involved in a number of cellular processes (Wise and Ball 1964; Pollak et al. 2007; Ying 2008). MEN appears to be a major player in the maintenance of the NADP/NADPH ratio in the cell (Geer et al. 1979a, b; Merritt et al. 2005, 2009). A single amino acid polymorphism was identified in a study of geographic variation in the genes of central metabolism across 10 *D. melanogaster* populations from the east coast of the United States (Sezgin et al. 2004). This guanine/cytosine (G/C) polymorphism is at base pair 338, resulting in a glycine (MEN^{113G})-to-alanine (MEN^{113A}) substitution at amino acid 113. Although Sezgin et al. (2004) did find clinal changes in allele frequency at some other loci suggesting selection, the *Men* alleles occur at approximately equal frequency in all populations surveyed. Further, a more thorough examination of the genetic variation in the *Men* coding region across these populations did not detect any evidence for selection between the alleles (Merritt et al. 2005). Merritt et al. (2005), however, did detect significant differences in the biochemistry between the two alleles and a significant effect of the alleles on the physiology of the fly. Coupling this examination with that of synthetic MEN activity variant alleles generated in the laboratory Merritt et al. (2005) suggested that the site could affect fitness and be under selection.

Specifically, Merritt et al. (2005) found differences between the two alleles in maximum enzyme velocity (V_{\max}) and concentration (K_m) for malate. Comparison with the three-dimensional structure of the pigeon cytosolic malic enzyme from Yang et al. (2002) suggests that the G/C polymorphism is in an α -helical region of the protein, structurally close to the active site of the enzyme. The site is near a tyrosine residue that is thought to interact with the substrate in the active site and is conserved across MEN proteins (Yang et al. 2002). This structural information suggests that the amino acid variation could affect enzyme activity, and in fact, the V_{\max} was found to be 46 % higher in MEN^{113G} than MEN^{113A}. This difference in activity could be a function of the amino acid polymorphism, or of linked regulatory variation, and the authors did report pronounced linkage disequilibrium between the amino acid polymorphism and other molecular variation in the gene region (Merritt et al. 2005). Merritt et al. (2005) also found significant variation in V_{\max} across MEN^{113G} and MEN^{113A} isochromosomal lines (i.e., lines of flies with different third chromosomes but the same allele at the polymorphic site). They attribute these line-specific differences in activity to regulatory differences either closely linked or perhaps distributed across the chromosomes examined. Such line-specific differences in enzyme activity are common and generally attributed to regulatory effects (Laurie-Ahlberg et al. 1982). Merritt et al. (2005) also reported a significant impact of the polymorphism on the K_m for malate (although not the K_m for NADP), with the MEN^{113G} allele having a higher observed K_m . The authors suggested that this difference in binding kinetics was explained by the location of the polymorphic site near the active site of the enzyme. It is interesting that the statistical analysis also

suggested significant, line-specific differences within the two genotypes. Since the entire coding region was sequenced and only the single polymorphic site identified, such differences would require that the binding kinetics were being modified by factors other than amino acid composition (i.e., that factors coded for elsewhere in the genome bind to the protein and modify its biochemistry). The authors found no significant differences in thermal stability between the two alleles. This result was somewhat surprising, given the location of the polymorphism in an α -helical region of the protein and the known impact of alanine/glycine variation on helical structures (Ganter and Plückthun 1990; Chakrabartty et al. 1991). The data did, however, show a trend toward the G allele, the ancestral allele, being more sensitive to thermal degradation.

Here we re-examine the biochemical characteristics of this amino acid polymorphism and a second, novel, amino acid polymorphism found in a Northern Ontario population, but absent from any of the 10 Atlantic coast sites previously surveyed (Sezgin et al. 2004; Merritt et al. 2005). The second site is an adenine/thymine (A/T) polymorphism at base pair 1051 in the coding region that results in a methionine (MEN^{351M})/leucine (MEN^{351L}) polymorphism at amino acid 351 in the protein. The ancestral nucleotide is adenine, found across all previously surveyed North American sites, African *D. melanogaster*, and *D. simulans*, and the ancestral amino acid is methionine. Our population genetic survey of the entire coding region of 35 third-chromosome isochromosomal lines indicates that the T allele has a frequency of ~14 %. Standard population genetic tests do not show any evidence of positive selection on either variable site, although these tests are notoriously weak. We find strong evidence for linkage disequilibrium between the nonsynonymous mutations and other, synonymous, mutations in the region as in the Merritt et al. (2005) study. Examining these isochromosomal lines, we find that both polymorphic sites significantly impact the biochemistry of the enzyme, although qPCR quantification of *Men* expression suggests that differences in V_{max} , including those previously attributed to the G/C polymorphism (Merritt et al. 2005), are best explained by *cis*-regulatory, not structural, variation. In sum, the results suggest that the nonsynonymous sites do impact the biochemistry of the enzyme and could therefore be under selection. Our results are most consistent, however, with the potential selective impact of molecular variation at this locus being a combination of structural and *cis*-regulatory variation.

Materials and Methods

Flies

One hundred isofemale lines were established from a wild population in Sudbury, Ontario, in August 2008 and again in August 2010. In 2008, 35 individual third chromosomes were made isogenic for the third chromosome using a dominantly marked balancer chromosome, 53 from the 2010 collection. The X chromosome was replaced with a w^{118} -marked chromosome. The second and fourth chromosomes were not replaced or made isogenic. Stocks and all crosses were maintained

on standard cornmeal food at 25 °C. All biochemical assays and the coding region sequence analysis were conducted on the 2008 lines. The 2010 lines were used only to screen for the continued presence and frequency of the leucine/methionine polymorphism.

DNA Sequence Analysis

The entire coding sequence of the *Men* gene was sequenced from cDNA for each isochromosomal line. Total RNA was isolated from 5 male flies for each of the isochromosomal lines using the RNeasy Mini Kit (Qiagen, Mississauga, ON). Complementary DNA (cDNA) was synthesized from total RNA using a commercially available kit (High Capacity cDNA Reverse Transcription Kit with RNase Inhibitor, Applied Biosystems, Carlsbad, CA) according to the manufacturer's suggested protocol. Amplifications were conducted according to the following conditions: 10 ng of cDNA was amplified using the TopTaq Master Mix (Qiagen) following the manufacturer's suggested protocol, with primers designed to span the coding region. Amplification products were purified and sequenced at the Génome Québec Innovation Centre (<http://www.genomequebec.com>). The primary sequence was determined for both strands and all ambiguous sites verified on both strands. Base calls from the chromatograms were checked and the coding regions assembled by means of CodonCode Aligner version 3.5.6 (CodonCode Corp., Dedham, MA). All sequences have been deposited in GenBank (accession nos. JQ690826–JQ690860).

DnaSP 4.0 (Rozas et al. 2003) was also used to calculate θ and π , in order to test for differences in θ using a coalescence model, to carry out the Tajima (1989) and Fu and Li (1993) tests, and to generate standardized estimates of linkage disequilibrium (R^2).

Malic Enzyme V_{\max} Measurements

Malic enzyme V_{\max} was approximated by the observed MEN activity under saturating substrate and cofactor conditions. Flies were homogenized in grinding buffer (100 mM Tris–HCl, 0.15 mM NADP, pH 7.4) at a concentration of 1 fly per 100 μ L of buffer, and the homogenate was spun at 13 K rpm for 5 min at 4 °C to pellet all solids. In general, assays were conducted using samples of four flies (four flies in 400 μ L grinding buffer). In a few cases, insufficient flies were available and fewer flies were assayed. The supernatant was collected and vortexed, and 250 μ L was added to one well of a 96-well plate. Aliquots were taken from this source plate for each of the analyses. Activity assays were conducted on a Molecular Designs SpectraMax 384 Plus 96-well plate spectrophotometer, using 10 μ L of extract and 100 μ L of assay buffer. Absorbance was measured every 9 s over 3 min at 25 °C. Samples were assayed twice and the means used in further analysis. The MEN assay buffer, optimized to give maximum activities (Merritt et al. 2005), contained 100 mM Tris–HCl, 0.34 mM NADP, 50 mM MnCl_2 , 50 mM malate, pH 7.4. Enzyme activity is expressed as nanomoles NADP^+ reduced per minute.

Thermal Stability

Genotype-specific enzyme thermostabilities were estimated by following MEN activity decline over time at 50 °C (Merritt et al. 2005; Hall 1985). Flies were homogenized, and a single aliquot of fly homogenate was immediately placed on ice for use as a reference. Ten aliquots were placed in a 50 °C heat block; at intervals of 1 min, a single aliquot was removed and placed on ice. All aliquots were subsequently kept on ice until their MEN activities were measured. The activity of each sample was compared to the reference 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 constants (k_D) were determined by the relationship $(E/E^0)_t = e^{-k_D t}$, where $(E/E^0)_t$ is the proportion of initial enzyme activity remaining at time t and k_D is the denaturation rate (Hall 1985). The slope of the line from the linear regression of $\ln(E/E^0)_t$ on time is an estimate of k_D . A mean k_D was calculated for each line.

Estimation of Michaelis–Menton constants (K_m)

Genotype-specific Michaelis–Menton constants for malate were determined using a 10-point geometric design method. Reaction rates were determined in parallel for all isochromosomal lines using the Molecular Designs SpectraMax 384 Plus 96-well plate spectrophotometer at 25 °C. The K_m of each line was determined by fitting two replicate initial velocity estimates made at each of 10 substrate concentrations, ranging from 0.2 K_M to 5 K_m , to the Michaelis–Menton equation, using the program GraFit 4.0 (Leatherbarrow 1998).

Soluble protein content

Soluble protein was measured by the bicinchoninic acid assay using a commercially available kit (Pierce Catalog No. 23225, Thermo Scientific, Rockford, IL) following the manufacturer's protocol. The assays contained 10 μ L homogenate and 100 μ L reagent and were incubated at 37 °C for 30 min. Reactions were measured at OD₅₆₂, and total soluble protein concentrations (μ g/mL) were determined by comparison with bovine serum albumen standards. Each sample was assayed twice for protein concentration and the mean used in analysis. Soluble protein was used as a covariate (along with weight) in ANCOVA analyses of MEN activity to account for differences in mass or body size between individual flies and possible differences in the degree of homogenization between samples.

Fly Weight

Flies were weighed to the nearest 0.01 mg using a Mettler Toledo MX5 microbalance. Weight was used as a covariate (along with soluble protein) in ANCOVA analyses to standardize MEN V_{max} for differences in fly size.

Quantitative RT-PCR

Total RNA was extracted and purified from four groups of 15 male flies for each genotype (only 2 samples from SON20 and 3 samples from SON50 were available) using the RNeasy kit (Qiagen, cat. no. 74106). Approximately 1 µg of total RNA was reverse transcribed using random hexamers and High Capacity cDNA Reverse Transcription kits with RNase Inhibitor (Applied Biosystems, cat. no. 4368814). The PCR reaction consisted of 2 µL undiluted cDNA template, 0.4 µM each primer, 0.2 µM probe, and Quantitect Probe PCR Master Mix (Qiagen, cat. no. 204363) in a total volume of 25 µL. The primers and probe flank the intron between exon 2 and exon 3 (5' GTATTGCCAACCTGTGCC, 3' AGCTTGTGTTCGGTGAGT, and probe 56-FAM/ATGGTGGATAGCCGTGGTGTCA/3IABkFQ). Two reactions per template were performed in parallel using a Mastercycler Ep Realplex Thermal Cycler (Eppendorf). All samples were normalized to *RpL32* and measured relative to an arbitrarily chosen biologic replicate of SON2 using the $\Delta\Delta C_T$ method (Livak and Schmittgen 2001).

Results

Sequence Variation

We amplified and sequenced all 1,737 bases of the *Men* coding region from 35 *D. melanogaster* isochromosomal lines. We observed 15 polymorphic sites across this population (Fig. 1), two of which were nonsynonymous changes: the G/C polymorphism at position 338, which results in a glycine/alanine amino acid polymorphism, and a novel A/T polymorphism at position 1051, which results in a methionine/leucine amino acid polymorphism (Sezgin et al. 2004; Merritt et al. 2005). The sequence and three-dimensional structure of the pigeon cytosolic malic enzyme and the human mitochondrial malic enzyme are known (Yang et al. 2000, 2002). The *D. melanogaster* MEN sequence is 55 % identical to the pigeon sequence. Given that the human and pigeon MEN sequences are 59 % identical, and all major structural features are conserved between them (Yang et al. 2002), we expect that the pigeon MEN structure is a reasonable predictor of the structure of the *D. melanogaster* molecule. Alignment of the *D. melanogaster* sequence with the human and pigeon sequences and comparison with the known pigeon (cytosolic MEN) structure indicates that the first polymorphic site is in an α -helical segment of the molecule, near the active site (Merritt et al. 2005), and the second site is in a β -sheet region buried within the molecule.

Our sample of the *Men* gene in Northern Ontario shows normal levels of synonymous polymorphism for *D. melanogaster*, similar to our earlier findings from sampling along the eastern coast of the United States (Table 1; Merritt et al. 2005; Moriyama and Powell 1996). Genetic diversity was slightly higher in the Canadian sample than the USA sample, but not significantly different at the $P < 0.05$ level using a coalescence simulation (Rozas et al. 2003). African *D. melanogaster* and the *D. simulans* lines were all the *Men*^{CA} haplotype (Merritt et al. 2005), suggesting that

Fig. 1 Nucleotide polymorphism in the amino acid coding region of the *Men* gene from Canadian population of *Drosophila melanogaster*. **Bold** variable sites responsible for the glycine/alanine (338) and methionine/leucine (1051) amino acid polymorphisms. The first 19 sequences have the *Men^{GA}* (glycine, methionine coding) genotype. The remaining sequences all have the *Men^{CA}* (alanine, methionine) genotype. The final 5 sequences have the *Men^{CT}* (alanine, leucine) genotype

	111111	11111
	1238001223	34556
	4337252352	72497
	1186618343	18891
SON44	CCGCGAGTCC	GATTG
SON34
SON28
SON19	.T.....	...C.
SON15A...	...C.
SON08A...	...C.
SON37C.
SON30C.
SON05C.
SON46C.
SON31C.
SON25C.
SON17C.
SON11C.
SON21C.
SON09	...G...TT	TTCC.
SON20	...G...TT	TTCC.
SON18	...G...TT	TTCC.
SON12	...G...TT	TTCC.
SON13	...G...TT	TTCC.
SON50	A.C...C..	.T.C.
SON02	..CG...TT	TTCC.
SON03	..CG...TT	TTCCCT
SON23	..CG...TT	TTCC.
SON24	..CGA...TT	TTCC.
SON45	..CGA...TT	TTCC.
SON10	..CGA...TT	TTCC.
SON22	..CG...TT	TTCC.
SON40	..CG...TT	TTCC.
SON48	..CG...TT	TTCC.
SON27	..CG.T...TT	TTCC.
SON47	..CG.T...TT	TTCC.
SON29	..CG.T...TT	TTCC.
SON33	..CG.T...TT	TTCC.
SON39	..CG.T...TT	TTCC.

this is the ancestral state. In their 2005 study, Merritt et al. found that the G/C polymorphism is at roughly 50/50 allele frequency in the USA samples, as did Sezgin et al. (2004) in their study across North American populations. Similarly, in our Canadian sample, 20 of the 35 lines we sequenced had a glycine at base pair 338 (*Men^G*), and 15 had a cytosine at base pair 1051 (*Men^C*). All of the *Men^G* lines had an adenine at the second nonsynonymous site (*Men^A*), as did 10 of the *Men^C* lines. Five of the *Men^C* lines had a thymine at the second nonsynonymous site (*Men^T*), ~14 % of the population. There are no fixed synonymous differences between the *Men^A* and *Men^G* alleles. Three positions, 1323, 1428, and 1548, that were fixed in almost all of the USA *Men^A* lines, show a similar nonrandom distribution in the Canadian sample, although not as pronounced as in the USA populations. Overall, we found significant linkage disequilibrium between the two amino acid-altering polymorphisms and synonymous sites in this data set ($P < 0.05$) by a Bonferroni corrected chi-square

test. The level of polymorphism within the *Men*^G alleles is not significantly different from that within the North American *D. melanogaster* *Men*^A alleles (Table 1), and all of the Tajima (1989) and Fu and Li tests (1993) conducted on the *D. melanogaster* sequences, Canadian only or combined North American sequences, were not statistically significant. There was no genetic variation across the entire coding region of the five *Men*^T lines, nor did we find any variation across 1,282 bases from the large first intron of the *Men* gene (data not shown). This lack of genetic variation suggests that although the novel mutation was found in over 14 % of the isofemale lines, it is a relatively new mutation. We sequenced the two polymorphic sites from the 53 isochromosomal third-chromosome lines established in 2010 and found the novel T allele in six lines, or ~11 % of the population; these allele frequencies are not significantly different by Fisher's exact test. These results from the 2010 lines suggest that though likely a relatively recent mutation, the novel allele does appear to be persisting in the Sudbury population and possibly at a stable frequency.

MEN Thermal Stability

Thermal stability was quantified as a proxy for overall enzyme stability. We found that the glycine/alanine polymorphism, but not the methionine/leucine polymorphism, affected thermal stability (Fig. 2). An ANOVA indicated that genotype significantly affected the observed K_D ($F_{2,137} = 18.99$, $P < 0.0001$), and Tukey's HSD test indicated that the *MEN*^{GM} (*MEN*^{113G} at the first polymorphic site and *MEN*^{351M} at the second polymorphic site; $n = 20$) genotype k_D was significantly higher than the k_D of either the *MEN*^{AM} (*MEN*^{113A} at the first polymorphic site and *MEN*^{351M} at the second polymorphic site; $n = 11$) or *MEN*^{AL} (*MEN*^{113A} at the first polymorphic site and *MEN*^{351L} at the second polymorphic site; $n = 4$) genotypes, but the *MEN*^{AM} and *MEN*^{AL} genotypes did not significantly differ in k_D . The *MEN*^{GM} genotype had a 15 % higher k_D than the *MEN*^{AM} or *MEN*^{AL} genotypes. A larger k_D reflects more thermal degradation (i.e., a less stable enzyme). This result is similar to earlier results (Merritt et al. 2005), although the differences observed in the earlier work were not found to be statistically significant, possibly because of the smaller sample size. Significant line (third chromosome) effects were also found

Table 1 Levels of silent polymorphism in the *Men* coding region across sampled populations of *Drosophila melanogaster*

Population	<i>n</i>	Sites	<i>S</i>	π	θ
<i>D. melanogaster</i>	68	433.87	19	0.00895	0.00921
African	9	433.76	10	0.00967	0.00854
North American	59	433.89	15	0.00858	0.00749
USA	24	433.83	10	0.00728	0.00622
Canadian	35	433.93	15	0.00924	0.00845
<i>Men</i> ^{GA}	31	433.83	9	0.00649	0.00523
<i>Men</i> ^{CA}	32	433.81	17	0.00708	0.00908
<i>Men</i> ^{CT}	5	434.5	0	0.00000	0.00000

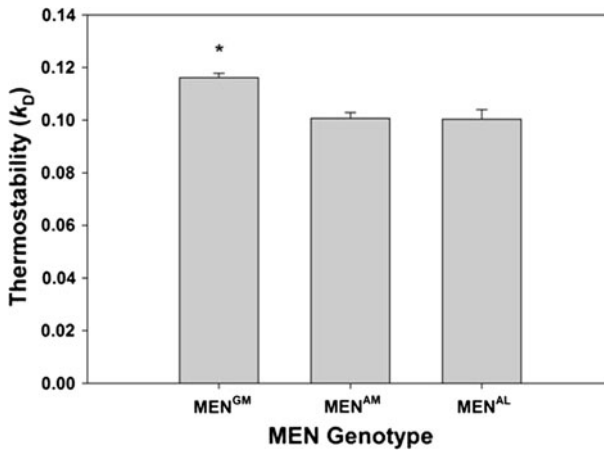


Fig. 2 Malic enzyme thermal stability (k_D). Thermal stability was measured for each genotype as a proxy for overall structural stability of the enzyme and was quantified as the absolute value of the K_D of the enzyme after denaturing at 50 °C. Bars mean \pm standard error across all isochromosomal lines for each genotype

across the MEN^{GM} ($F_{19,77} = 4.93$, $P < 0.0001$), MEN^{AM} ($F_{10,43} = 6.13$, $P < 0.0001$), and MEN^{AL} lines ($F_{3,15} = 7.18$, $P < 0.005$), suggesting that factors outside the amino acid polymorphism (perhaps coded by other loci on the third chromosome) also impact enzyme stability (data not shown).

Malate Michaelis–Menton Constant (K_m)

Because of the role of binding kinetics in in vivo enzyme activity, we quantified the effect on the K_m for malate of both polymorphisms and found that both the glycine/alanine and methionine/leucine polymorphisms significantly alter the K_m (Fig. 3). An ANOVA indicated that genotype significantly affected the observed K_m ($F_{2,133} = 92.26$, $P < 0.0001$), and Tukey's HSD test indicated that all three genotypes were significantly different. The MEN^{GM} genotype K_m was 10 % higher than the MEN^{AM} genotype. This result is similar to that of Merritt et al. (2005), although the earlier work found a larger difference in K_m between the MEN113G and MEN113A alleles. In the results presented here, the most striking difference was between the MEN^{AL} ($n = 4$) and MEN^{AM} ($n = 11$) genotypes. The MEN^{AL} genotype K_m was 42 % lower than the MEN^{AM} genotype. No significant line (third chromosome) effects were found across any of the genotypes, suggesting that the amino acid composition of the proteins is solely responsible for the observed K_m (i.e., K_m was not modified by factors coded elsewhere on the third chromosome; data not shown).

Reaction Velocity (V_{max})

We also quantified the effect of each polymorphic site on the MEN reaction velocity under saturating substrate conditions, an empirical approximation of V_{max} . We

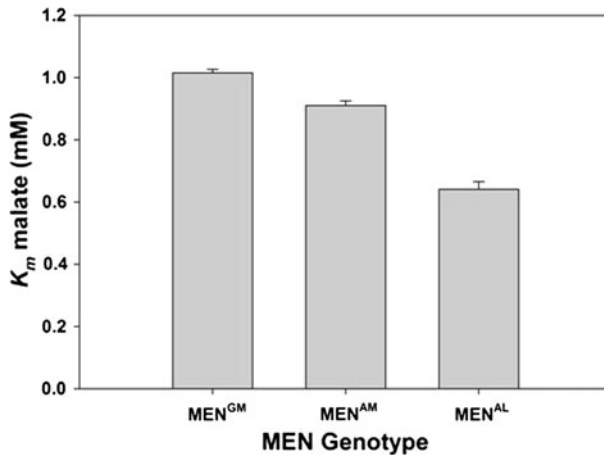


Fig. 3 Michaelis-Menton constant (K_m) for malate. Substrate-binding kinetics were quantified as the K_m for malate for each of the three genotypes. Units are $\mu\text{M/L}$. Bars mean \pm standard error across all isochromosomal lines for each genotype

found that the glycine/alanine polymorphism impacted the observed V_{\max} , but the M/L polymorphism did not (Fig. 4). An ANCOVA with weight and soluble protein as covariates indicated that genotype significantly affected the observed V_{\max} ($F_{2,134} = 46.71$, $P < 0.0001$), and Tukey's HSD test indicated that the MEN^{GM} ($n = 21$) genotype V_{\max} was significantly higher than the V_{\max} of either the MEN^{AM} ($n = 11$) or MEN^{AL} ($n = 4$) genotypes, but the MEN^{AM} and MEN^{AL} genotypes did not significantly differ in V_{\max} . The MEN^{GM} genotype had a 20 % higher observed

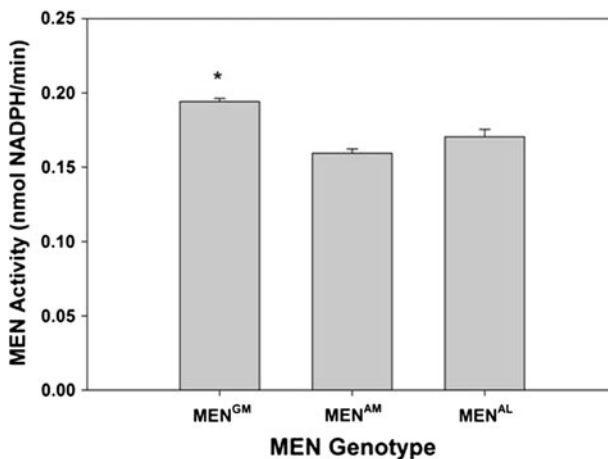


Fig. 4 MEN maximum activity (V_{\max}). Maximum MEN activity was estimated for each genotype as the observed enzyme activity under saturating substrate and cofactor (Merritt et al. 2005). Enzyme activity is expressed as nanomoles NADP^+ reduced per minute. Bars mean \pm standard error across all isochromosomal lines for each genotype

V_{\max} than the other two genotypes. Significant line (third chromosome) effects were also found across the MEN^{GM} ($F_{20,328} = 13.04$, $P < 0.0001$) and MEN^{AM} lines ($F_{10,137} = 19.33$, $P < 0.0001$), but not the MEN^{AL} lines (data not shown). Such chromosome effects are common for enzyme activities and likely reflect differences in gene expression or enzymatic interactions with factors coded elsewhere on the chromosome. The lack of chromosome effects across the MEN^{AL} lines likely reflects the lack of genetic variation across these lines.

Men Expression

We used qPCR to quantify differences in relative *Men* expression between the genotypes and found that the G/A, but not M/L, polymorphism was associated with differences in *Men* expression (Fig. 5), presumably through linked regulatory variation. An ANOVA indicated that genotype significantly affected the observed gene expression ($F_{2,47} = 11.00$, $P < 0.0001$). Tukey's HSD test indicated that *Men* expression was higher in the Men^{GA} ($n = 4$) genotype (codes for MEN^{GM}) lines than *Men* expression in either the Men^{CA} ($n = 4$) genotype (codes for MEN^{AM}) or Men^{CT} ($n = 4$) genotype (codes for MEN^{AL}) lines, but *Men* expression levels did not differ between the Men^{CA} and Men^{CT} genotype lines. Relative *Men* expression was 51 % higher in the Men^{GA} genotype lines than in the Men^{CA} or Men^{CT} genotype lines. Significant line (chromosome) effects were also found across the Men^{CA} lines ($F_{3,15} = 4.48$, $P < 0.025$), but not the Men^{GA} and Men^{CT} lines. The lack of chromosome effects on *Men* expression in the Men^{GA} genotype lines was surprising, given the significant variation found in *MEN* activity in these lines; it may reflect the relatively higher error associated with the qPCR experiments than the enzyme activity experiments, a phenomenon we have noted elsewhere (Lum and Merritt 2011).

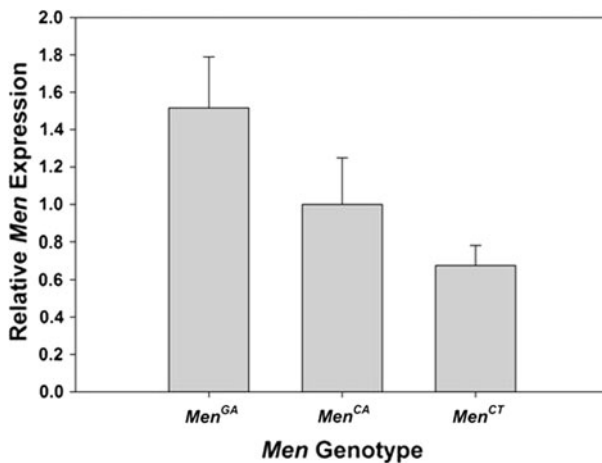


Fig. 5 Relative *Men* expression, quantified for each genotype using qPCR. Activity was standardized by expression of a housekeeping gene. Units are presented relative to an arbitrarily chosen biological replicate

Discussion

The evolutionary significance of molecular variation across enzymes within and between populations and species has been debated since the recognition of this variation in the 1950s (Lewontin 1974; Laurie-Ahlberg et al. 1982; Eanes 1999; Hoekstra and Coyne 2007; Wray 2007). Although the last half century has seen impressive advances in genotyping and our ability to identify and quantify diverse phenotypes, connecting genotype to phenotype remains difficult, and establishing this connection is a central challenge in modern genetics. The research presented here contributes to our understanding of this elusive connection by quantifying the impact of two simple nonsynonymous changes that result in a pair of amino acid substitutions in a metabolic enzyme. We report that both amino acid polymorphisms appear to affect substrate-binding kinetics, while only one appears to affect thermal stability. It is interesting to find that the activity differences previously assigned to one of the polymorphisms (Merritt et al. 2005) may, instead, be a function of linked regulatory differences. This last point begins to address the question of the contribution of gene expression, as well as amino acid variation, to functional variation. Our results support the conclusions of Hoekstra and Coyne (2007) that both regulatory and structural change impact gene product function. Our examination of the *Men*-coding sequences reveals no evidence for selection acting on the polymorphisms, but earlier work on this enzyme (Merritt et al. 2005, 2009) indicates that the biochemical variation observed could have physiological repercussions and therefore could potentially be under selection.

Changes in amino acid composition can result in changes in enzyme stability or binding coefficients, presumably through alteration of secondary, tertiary, or quaternary structure. The glycine/alanine polymorphism appears to be in an α -helical region of the protein, structurally close to the active site of the enzyme (Yang et al. 2000, 2002). Structural comparison also suggests that the methionine/leucine polymorphism is in a region of β -sheet, buried within the protein. Although assigning functional consequences to amino acid variation within a protein is complicated at best, the location of these polymorphic sites in regions of known secondary structure suggests possible mechanisms for the impact of this variation. Glycine/alanine polymorphisms have been shown to alter helical stability in other systems (Ganter and Plückthun 1990; Chakrabartty et al. 1991). Glycine, with the smallest R group of any amino acid, generally destabilizes α helices (Ganter and Plückthun 1990; Chakrabartty et al. 1991). In the case studied here, the change from glycine to alanine increases stability by approximately 15 %. It is interesting to speculate that the increased stability of the alanine allele may result from extension or stabilization of this region of α -helical secondary structure in the enzyme or interaction of this region of secondary structure with other structural units within the enzyme. Further, methionine-to-leucine substitutions can, but do not always, increase protein stability in other systems (Lipscomb et al. 1998). We find no effect of the methionine/leucine variation on stability in our system, supporting the conclusion of Lipscomb et al. (1998) that the effect of this substitution is context dependent.

The biologic implications of the difference in stability are more difficult to determine. In some systems, northern alleles have been shown to be more

thermolabile (i.e., less stable; Hall 1985; Hochachka and Somero 2002), which may translate into more active enzymes under colder conditions. The lower thermal stability could also reflect the absence of selection for heat tolerance in colder climates. *Drosophila* may, however, select very similar microclimates across geographic distances, and selection may be more for life history changes associated with seasonal differences. In any case, the northern allele in this system is the less thermolabile allele and has the greater stability, at least in vitro, the opposite of the pattern observed in some other systems. This result, contrary to a priori expectations, and the significant variation between lines with the same allele suggest that future research with purified enzyme is justified to determine more completely the impact of these variable amino acid sites on enzyme stability.

Both amino acid polymorphisms appear to affect the substrate-binding kinetics of the enzyme. The glycine/alanine polymorphism is near the active site, and modification of the structure of the active site pocket may result in changes in binding kinetics. The methionine/leucine site is in a region of β -sheet secondary structure. Although the methionine/leucine substitution is not particularly close to any known binding sites, it may affect overall protein structure (Lipscomb et al. 1998). At any rate, the significant difference in binding kinetics between the methionine and leucine genotypes suggests that this amino acid substitution alters the protein structure sufficiently to affect substrate binding. The low K_m of the MEN^{AL} genotype indicates a high binding affinity for this protein and may reflect strong enzyme-substrate loading under low substrate concentrations. As with the glycine/alanine polymorphism, these results suggest the need for complete characterization of purified protein.

In the initial characterization of the glycine/alanine polymorphism (Merritt et al. 2005), the authors suggested that the observed difference in V_{max} could be a function of the amino acid polymorphism or linked regulatory variation. Our examination of expression of the three *Men* genotypes suggests that the latter is the case; the variation appears to be a regulatory phenomenon. Variation in observed V_{max} , independent of amino acid variation, is common for *Men* and other metabolic enzymes (Laurie-Ahlberg et al. 1982; Merritt et al. 2005, 2006, 2009) and is generally attributed to regulatory variation. Further, gene expression has been documented to drive differences in enzyme activity with significant life history effects in other naturally occurring insect systems (e.g., Schilder et al. 2011). Linkage disequilibrium is very apparent in our data set, and linked, but unknown, variation is always a concern in attributing observations to known polymorphisms. In fact, concerns over such linkage were a driving force behind the authors' generating synthetic *Men* alleles in trying to determine the physiological impact of variation in malic enzyme activity (Merritt et al. 2005). Other studies in the Merritt lab (Lum and Merritt 2011) are currently annotating the *Men* regulatory region, and future research will combine these projects to attempt to identify the regulatory elements that are driving the differences in expression between the genotype classes.

In results presented here, variation in thermal stability and the K_m appear to be driven by the observed amino acid polymorphisms, while MEN V_{max} is more likely a function of linked regulatory variation. Interestingly, we find significant variation

in thermal stability but not K_m across isoallele lines, lines that should have identical enzymes (no amino acid variation) but differ in the third chromosome. The lack of line-specific differences in K_m suggests that the substrate-binding kinetics are a function of the proteins themselves. The presence of line-specific differences in k_D , however, suggests that protein stability is a function of the protein and other, interacting, factors outside this locus. As other authors have pointed out (e.g., Schilder et al. 2011), complete understanding of the impact of molecular variation requires purification of protein. Future research will examine these biochemical characteristics in purified protein to determine the amino acid-based control of this variation versus the impact of outside elements. Finally, no MEN^{GL} genotypes have been isolated and characterized, possibly reflecting the recent age of the MEN^{ML} polymorphism. Large-scale collections are ongoing to isolate MEN^{GL} genotypes to allow a complete characterization of the interactions between the two polymorphisms.

Acknowledgments This work was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (3414-07) and Canada Research Chair (950-215763) to TJSM. TZR was supported by a NSERC Undergraduate Student Research Award and a Laurentian University Research Fund award.

References

- Chakrabarty A, Schellman JA, Baldwin RL (1991) Large differences in the helix propensities of alanine and glycine. *Nature* 351:586–588
- Eanes WF (1999) Analysis of selection on enzyme polymorphisms. *Ann Rev Ecol Syst* 30:301–326
- Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics* 133:693–709
- Ganter C, Plückthun A (1990) Glycine to alanine substitutions in helices of glyceraldehyde-3-phosphate dehydrogenase: effects on stability. *Biochemistry* 29:9395–9402
- Geer BW, Krochko D, Williamson JH (1979a) Ontogeny, cell distribution, and the physiological role of NADP-malic enzyme in *Drosophila melanogaster*. *Biochem Genet* 17:867–879
- Geer BW, Lindel DL, Lindel DM (1979b) Relationship of the oxidative pentose shunt pathway to lipid synthesis in *Drosophila melanogaster*. *Biochem Genet* 17:881–895
- Hall JG (1985) Temperature-related kinetic differentiation of glucosephosphate isomerase alleloenzymes isolated from the blue mussel, *Mytilus edulis*. *Biochem Genet* 23:705–728
- Hochachka PW, Somero GN (2002) Biochemical adaptation: mechanism and process in physiological evolution, 1st edn. Oxford University Press, New York
- Hoekstra HE, Coyne JA (2007) The locus of evolution: evo devo and the genetics of adaptation. *Evolution* 61:995–1016
- Laurie-Ahlberg CC, Wilton AN, Curtsinger JW, Emigh TH (1982) Naturally occurring enzyme activity variation in *Drosophila melanogaster*, I: sources of variation for 23 enzymes. *Genetics* 102:191–206
- Leatherbarrow R (1998) GraFit version 4. Erithacus Software Ltd, Staines
- Lewontin RC (1974) The genetic basis of evolutionary change. Columbia University Press, New York
- Lipscomb LA, Gassner NC, Snow SD, Eldridge AM, Baase WA, Drew CL, Matthews BW (1998) Context-dependent protein stabilization by methionine-to-leucine substitution shown in T4 lysozyme. *Protein Sci* 7:765–773
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods* 25:402–408
- Lum TE, Merritt TJS (2011) Non-classical regulation of transcription: interchromosomal-interactions at the malic enzyme locus of *Drosophila melanogaster*. *Genetics* 189:837–849
- Merritt TJS, Duvernell D, Eanes WF (2005) Natural and synthetic alleles provide complementary insights into the nature of selection acting on the *Men* polymorphism of *Drosophila melanogaster*. *Genetics* 171:1707–1718

- Merritt TJS, Sezgin E, Zhu C-T, Eanes WF (2006) Triglyceride pools, flight and activity variation at the *Gpdh* locus in *Drosophila melanogaster*. *Genetics* 172:293–304
- Merritt TJS, Kuczynski C, Sezgin E, Zhu C-T, Kumagai S, Eanes WF (2009) Quantifying interactions within the NADP(H) enzyme network in *Drosophila melanogaster*. *Genetics* 182:565–574
- Moriyama EN, Powell JR (1996) Intraspecific nuclear DNA variation in *Drosophila*. *Mol Biol Evol* 13:261–277
- Pollak N, Dölle C, Ziegler M (2007) The power to reduce: pyridine nucleotides, small molecules with a multitude of functions. *Biochem J* 402:205–218
- Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19:2496–2497
- Schilder R, Zera AJ, Black C, Hoidal M, Wehrkamp C (2011) The biochemical basis of life history adaptation: molecular and enzymological causes of NADP+–isocitrate dehydrogenase activity differences between morphs of *Gryllus firmus* that differ in lipid biosynthesis and life history. *Mol Biol Evol* 28:3381–3393
- Sezgin E, Duvernell DD, Matzkin LM, Duan Y, Zhu C-T, Verrelli BC, Eanes WF (2004) Single-locus latitudinal clines and their relationship to temperate adaptation in metabolic genes and derived alleles in *Drosophila melanogaster*. *Genetics* 168:923–931
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595
- Wise EM, Ball EG (1964) Malic enzyme and lipogenesis. *Proc Natl Acad Sci USA* 52:1255–1263
- Wray GA (2007) The evolutionary significance of *cis*-regulatory mutations. *Nat Rev Genet* 8:206–216
- Yang Z, Floyd DL, Loeber G, Tong L (2000) Structure of a closed form of human malic enzyme and implications for catalytic mechanism. *Nat Struct Biol* 7:251–257
- Yang Z, Zhang H, Hung H-C, Kuo C-C, Tsai L-C, Yuan HS, Chou W-Y, Chang G-G, Tong L (2002) Structural studies of the pigeon cytosolic NADP(+)-dependent malic enzyme. *Protein Sci* 11:332–341
- Ying W (2008) NAD+/NADH and NADP+/NADPH in cellular functions and cell death: regulation and biological consequences. *Antioxid Redox Signal* 10:179–206