Published in final edited form as:

Evolution. 2016 October ; 70(10): 2239–2246. doi:10.1111/evo.13020.

THE MUTATIONAL STRUCTURE OF METABOLISM IN *CAENORHABDITIS ELEGANS*

Sarah K. Davies^{1,2,†}, Armand Leroi^{1,†}, Austin Burt¹, Jacob G. Bundy², and Charles F. Baer³

¹Department of Life Sciences, Imperial College London

²Department of Surgery and Cancer, Imperial College London

³Department of Biology, University of Florida and University of Florida Genetics Institute

Abstract

A properly functioning organism must maintain metabolic homeostasis. Deleterious mutations degrade organismal function, presumably at least in part via effects on metabolic function. Here we present an initial investigation into the mutational structure of the *Caenorhabditis elegans* metabolome by means of a mutation accumulation experiment. We find that pool sizes of 29 metabolites vary greatly in their vulnerability to mutation, both in terms of the rate of accumulation of genetic variance (the mutational variance, VM) and the rate of change of the trait mean (the mutational bias, M). Strikingly, some metabolites are much more vulnerable to mutation than any other trait previously studied in the same way. Although we cannot statistically assess the strength of mutational correlations between individual metabolites, principal component analysis provides strong evidence that some metabolite pools are genetically correlated, but also that there is substantial scope for independent evolution of different groups of metabolites. Averaged over MA lines, PC3 is positively correlated with relative fitness, but a model in which metabolites are uncorrelated with fitness is nearly as good by Akaike's Information Criterion (AIC).

Keywords

Fitness; Metabolome; Mutation accumulation; Mutational Bias; Mutational variance

INTRODUCTION

In the long run, the course of phenotypic evolution depends on the quantity and qualities of new mutations. The quantity and quality of available mutations must, in turn, depend on the functional architectures – the networks of matter, energy and signals that regulate the development of living things and their workings as adults. These thoughts, put together, lead to the notion that functional architectures *constrain* the evolution of living things, making some paths more likely than others, quite independently of the adaptive demands of the environment (Maynard Smith et al. 1985; Houle 1991; Arnold 1992; Brakefield 2006;

Correspondence to: Charles F. Baer.

[†]Joint first authors.

Stoltzfus 2006; Hendrikse et al. 2007; Psujek and Beer 2008; Braendle et al. 2010). They also lead to another thought: that if we can learn enough about the structure of functional architectures we may be able to predict the quality and quantity of mutations available for natural selection and hence the ability of living things to respond to new adaptive challenges (Ibarra et al. 2002; Dean and Thornton 2007; Lewis et al. 2010; Papp et al. 2011).

This, in brief, is the structuralist programme in evolutionary biology (Gould 2002) – or, at least, one version of it. Its execution faces two difficulties. The first is that it is difficult to estimate the quantity and quality of variation that mutation produces each generation. The second difficulty is that, for all that we have learned about how organisms build themselves and work, our causal models remain fragmentary.

The first of these difficulties - assessing the input of genetic variation by mutation - can be addressed by means of evolution experiments in which the effects of natural selection are minimized, leaving only those of mutation. This is the principle upon which mutation accumulation (MA) experiments depend. In a MA experiment, many replicate lines derived from a genetically homogeneous common ancestor are allowed to evolve in the near absence of natural selection by minimizing effective population size. Over the course of many generations, the lines accumulate mutations and so diverge from each other. The among-line variance reflects the input of new genetic variance by mutation - the mutational variance, VM, or more generally, the mutational variance-covariance matrix, **M** (Lynch and Walsh 1998). The change in the mean of a trait over the course of the experiment represents the phenotypic bias of new mutations, **M**.

Mutation accumulation experiments have been carried out in a number of species – all of them small and possessed of rapid generation times (reviewed in Drake et al. 1998; Baer et al. 2007; Halligan and Keightley 2009). The nematode *Caenorhabditis elegans* is a particular favorite since it is a selfing hermaphrodite, thus immune to the malign effects of inbreeding, and has a 3–4 day generation cycle (e.g., Keightley and Caballero 1997; Vassilieva and Lynch 1999; Azevedo et al. 2002; Baer et al. 2005; Begin and Schoen 2006; Katju et al. 2015). Most MA studies have focused on gross phenotypes, in particular those related to fitness such as fecundity and longevity (but see Denver et al. 2005). For a variety of reasons, it is interesting to know the mutational properties of fitness-related traits, but, since the networks that regulate such traits are complex and poorly known, we have no way of predicting their mutational structure. To give a concrete instance: the molecular regulation of, and physiological connections between, *C. elegans*' lifespan and reproduction have been much studied (Hansen et al. 2013), yet we still have no causal model that offers anything resembling a prediction as to how spontaneous mutations will affect either trait individually or jointly. This is so even though decades of life-history models have assumed they are bound to each other by an energetic trade-off (Leroi 2001).

If, then, we wish to link mutational properties (VM, **M**, or more ambitiously, **M**) to functional architecture, it seems that we have to start by looking at simpler, lower-level, intermediary phenotypes. Braendle et al. (2010), for example, estimated VM and **M** for cell fates in the exceptionally well-studied nematode vulva; Farhadifar et al. (2015) estimated the full covariance matrix for a set of six traits associated with the first mitotic cell division. In

the same spirit, here we use an MA experiment to estimate the mutational properties of part of the worm's metabolic network. Metabolomic analysis of *C. elegans* has already been widely used to study a range of different biological questions; changes in elements of the metabolome have been associated with organismal development (Swire et al. 2009), aging (Fuchs et al. 2010), responses to physiological stresses (Stupp et al. 2013), and behavior (Srinivasan et al. 2012). Here we use GC-MS (Gas chromatography–mass spectrometry) for metabolic profiling (Swire et al. 2009; Fuchs et al. 2010; Geier et al. 2011; Geier et al. 2013; Stupp et al. 2013) to characterize the evolution of 29 metabolites in an existing set of mutation accumulation lines. Further, since we also have fitness data on the same set of MA lines, we ask whether the mutational structure of fitness can be predicted from features of the metabolome.

MATERIALS AND METHODS

(i) Mutation accumulation lines: assay

The origin and husbandry of the MA lines included in this study are outlined in Baer et al. (2005). By the time the lines were assayed for metabolomic analysis they had accumulated mutations under minimally efficient selection ($N_e \approx 1$) for 250 generations. Of the 100 original MA lines, ~70 remained extant at the time this study was initiated. All remaining MA lines, and their ancestral control (G0), were sent to Imperial College for metabolomic analysis where they were washed and instantly frozen at -80°C . Of those lines, 43 were ultimately included in the assay. Upon thawing, lines were cleaned by bleaching with hypochlorite solution to kill bacterial and fungal contaminants and surviving L1 larvae were chunk-transferred onto NGM plates seeded with *E. coli* OP50 food (Stiernagle 2006). We then grew five replicate plates for each MA line; some replicates were subsequently discarded because they had low biomass and hence poor metabolomic signals, leaving 2–5 replicate samples per line (mean = 3.9). For the ancestral control line we grew 9 replicate plates. All lines were allowed to reproduce until plates had reached a point suitable for synchronization by treatment with hypochlorite (Stiernagle 2006). Following synchronization, worms were incubated at 20°C until young adulthood, defined here as the point at which some eggs were seen on plates but no second generation worms had hatched. At this point, worms were washed from plates with M9 buffer and allowed to settle naturally at the bottom of a 15 ml centrifuge tube for three minutes. The M9 supernatant was then removed and the worm pellet transferred to a microcentrifuge tube, snap-frozen with liquid nitrogen and stored at -80°C .

Fitness assays were carried out in the Baer laboratory at generations 200 (20°C) and 220 (25°C) by counting the number of live progeny produced over a three day period (Baer et al. 2006).

(ii) Metabolomics: extraction and data acquisition

We based our metabolite extraction protocol on a method previously validated by Geier et al. (2011): briefly, we added 1.2 ml methanol to each tube containing an approximately 0.3 ml worm pellet, and used a Precellys Dual bead beater (Bertin Technologies) to homogenize the samples. We repeated this process with an 80% methanol wash; all extraction solvents used

were ice cold. After centrifugation for 5 minutes at 13000 g, we combined supernatants from the two steps, transferred them to new tubes and vacuum-dried them overnight. We then added 20 μ l quantitation standard mixture (1.5mM 2,2,3- $^2\text{H}_3$ -leucine, 1.3mM ^{13}C - [U]-glucose) to the samples before vacuum-drying them in an Agilent 2 ml autosampler vial. We derivatized samples by adding 40 μ l methoxyamine hydrochloride (20 mg ml $^{-1}$ in pyridine) to each tube and incubating at 37 °C for 2 hours with shaking (Lisec et al. 2006). Following incubation, 70 μ l N-Methyl-N-(trimethylsilyl)-trifluoroacetamide (MSTFA) and 10 μ l injection standard (2-fluorobiphenyl) was added to each tube. This was followed by a further incubation period of 30 minutes at 37°C with shaking.

We analyzed our samples (1 μ l injected volumes) on an Agilent 5975c quadrupole mass spectrometer with a 7890 gas chromatograph, equipped with a 30m $\times$ 0.25mm HP-5ms column, and splitless injector. The injector port was heated to 230°C and the source to 200°C; the carrier gas was helium (flow rate 1 ml min $^{-1}$). The temperature program was isothermal at 70°C for 5 minutes, followed by a 5 °C min $^{-1}$ ramp to 230 °C, followed by a 10 °C /min ramp to 330 °C, and then isothermal for 5 minutes. The mass spectrometer was set to scan from 50–600 Da at 5,000 Da/s.

To identify metabolites we used the Fiehn Library (Kind et al. 2009) and AMDIS, a deconvolution program (Halket et al. 1999). We then manually inspected all peaks, set individual integration boundaries, and reintegrated the peaks using the GAVIN Matlab script (Behrends et al. 2011). This gave 34 identified metabolites, of which 29 were included in this study.

Normalized metabolite data have been archived in Dryad (<http://dx.doi.org/10.5061/dryad.2dn09>)

(iii) Estimation of quantitative genetic parameters

To examine the directional mutational bias that each metabolite is subject to we first calculated the mean ratio $\log_2(M_{MAj}/M_0)$ where M_{MAj} is the mean pool size of the j th MA line and M_0 is the mean of the replicate ancestral (G0) samples. The global mean and confidence intervals were then calculated from these $\log_2$ ratios (Figure 1A). The relative per generation change in the mean was calculated as $M = ((M_{MA} - M_0)/t)/M_0 * 100$ where M_{MA} is the global mean of the MA lines (Azevedo et al. 2002).

To examine the rate at which mutational variance accumulated in the MA lines we estimated the between-line variance, V_L , and the environmental variance, V_E , using ANOVA with line as a random variate. Since the residuals were not normally distributed we assessed the significance of V_L by bootstrapping. The among-line variance can be used to estimate the per-generation input of genetic variation from new mutation (the mutational variance, V_M), which is $V_L/2t$ where t is the number of generations of mutation accumulation (Lynch and Walsh 1998, p. 330). Comparisons of mutational variances are complicated by scaling effects (Houle et al. 1996; Fry and Heinsohn 2002) that are not necessarily obviated by log-transformation. We therefore estimated the mutational coefficient of variation $CV_M = (V_M)^{1/2}/M_0 * 100$ as a percentage. However, if M_{MA} changes drastically during the course of the experiment, which it does here in some cases, comparisons of CV_{M_0} will underestimate

the extent of relative variation, so we also scale by the mean of the MA lines, M_{MA} , to give two measures, CV_{M_0} and $CV_{M_{MA}}$. Both M and the CVs were calculated using the median rather than the mean as the measure of central tendency because the median is less sensitive to outliers.

With 43 MA lines, we can jointly estimate the full M -matrix for at most eight variables: the number of observations in the data (i.e., MA lines) must be at least $[n(n+1)/2]$, where n = the dimension of the covariance matrix. To gain insight into the multivariate structure of the metabolome, we employed principal components analysis (PCA). We first calculated PCs based on the individual observations, normalized to zero mean and unit variance, using the PRINCOMP procedure in SAS v. 9.4 with eigenvectors scaled to unit length. We then calculated within and among-line components of covariance with the first eight PCs as response variables, using restricted maximum likelihood (REML) as implemented in the MIXED procedure in SAS v. 9.4, following Fry (2004). We compared the full model with unstructured covariances, using the TYPE=UNR option in the RANDOM and REPEATED statements, respectively, to a model in which the off-diagonal elements of the among-line covariance matrix were constrained equal to zero using the PARMS option in the RANDOM statement. The two models were compared using the corrected Akaike Information Criterion (AICc).

To assess the relationship between the metabolome and fitness while taking into account the correlation structure of the metabolome, we performed PCA on the line means as outlined in the previous paragraph. Lifetime reproduction weighted by survival ("Total fitness", W) of some of these MA lines has been previously assayed at generation 200 (31/43 lines) at 20°C and 220 (35/43 lines) at 25°C (Baer et al. 2006). Scaling W by the mean of the ancestral control provides a measure of relative fitness, w . Averaging the two assays, mean W of the 36 MA lines for which we have measurements of both metabolites and fitness was 71% that of the ancestral control (range 0.3%–123%). With 36 MA lines, we can jointly estimate the among-line covariances of at most eight variables, so we calculated the correlation between line mean relative fitness and PC1-PC7, using the same linear model strategy outlined in the previous paragraph but with only a single hierarchical level (= line means). Because fitness was measured after 200 or 220 generations of MA and metabolite traits were measured after 250 generations, the among-line correlation calculated from the data will underestimate the mutational correlation, because some fraction (50/250 or 30/250 respectively) of the among-line variance in the metabolite traits will be due to mutations that occurred subsequent to measurement of fitness and thus cannot contribute to the correlation. The correlation between metabolic traits and w was calculated from the equation

$$r_{xW} = \frac{Cov(x, w)}{\sqrt{\left(\frac{t_w}{t_x}\right) Var(x) Var(w)}}$$

where the index x designates the metabolite, t_x is 250 generations of MA and t_w is 210 generations, the unweighted average of the two fitness assays (Braendle et al. 2010). Statistical significance of correlations of relative fitness with PC1-PC7 was assessed pairwise by Likelihood Ratio Test of a model with unstructured covariance compared to a model with the off-diagonal elements constrained to zero. The models are nested and differ by a single parameter, so twice the difference in the log-

likelihoods is asymptotically chi-square distributed with a single degree of freedom. We accounted for multiple tests by applying the sequential Bonferroni correction.

RESULTS

Some metabolites are vulnerable to mutations of very large effect

To initially examine the directional effects of mutation we calculated the mean $\log_2(\text{MA/G0})$ for each metabolite. We found that most metabolites showed a significant directional bias (Fig. 1A, Supplementary Table S1): 7 increased, 18 decreased; we failed to detect a directional change in 4 (one-sample t-tests on $\log_2$ ratios, $\mu=0$, $df = 42$). Succinic acid showed the greatest increase (12-fold); asparagine showed the greatest decrease (66%) relative to the ancestral strain. We calculated the per-generation relative change, M , for each metabolite and found that, on average, the metabolites changed by 0.73% (± 0.59 , 95%CI) per generation, independent of the sign of the change (Supplementary Table S1, $\text{abs}(M)$ (%)). By way of comparison, lifetime reproduction (“Total Fitness”) in the same set of lines decreases by about 0.1%/generation (Baer et al. 2006). Some metabolites, such as succinic acid, xanthine, nicotinic acid and adenosine, increased by more than 2% per generation, but others showed far smaller changes, or no detectable change at all. Thus it appears that metabolite pool sizes vary greatly in their susceptibility to mutation.

The per-generation increase in the among-line component of variance reflects the mutational variance, V_M . Of the 29 metabolites, seven (glycine, glutamate, putrescine, 3-aminoisobutyric acid, lysine, alanine, asparagine) showed no significant V_M insofar that their 95% confidence intervals overlapped zero (Supplementary Table S1).

Comparisons of variances are only meaningful when the different groups are measured on comparable scales. Metabolite pools differ greatly in size so in order to compare the rates at which they accumulate mutational variance we need to scale V_M in some fashion (see Materials and Methods). V_M is typically scaled either relative to the trait mean (i.e., the mutational coefficient of variation, CV_M) or relative to the environmental component of variance (the mutational heritability, h_M^2). CV_M can be scaled relative either to the mean of the ancestral control (CV_{M_C}) or to the mean of the MA lines (CV_{M_MA} ; Baer et al. 2006). If it is assumed that the effects of mutations are multiplicative (Barton 1990), it is appropriate to scale by the mean of the MA lines; scaling by the MA mean is approximately equivalent to calculating variances of log-transformed data (Fry and Heinsohn 2002). Since the mean pool sizes of some metabolites changed so much in the course of the experiment, in some cases the two methods give very different estimates of CV_M , and so we report both (Supplementary Table S1). Scaling by the ancestral control line means gives a mean CV_{M_C} across metabolites of 7.00% ($\pm 95\%CI$: 5.90; range: 0.25 – 72.38%); scaling by the MA line mean gives a mean CV_{M_MA} across metabolites of 1.60% ($\pm 95\%CI$: ± 0.33 ; range: 0.42 – 4.11%). The correlation between these two ways of estimating CV_M is moderate but significant: $r^2 = 0.64$ ($P < 0.01$). Thus, depending on scaling, by this measure, metabolites vary between one and two orders of magnitude in their vulnerability to mutation.

As noted, there are estimates of mutational parameters in *C. elegans* for a variety of phenotypes such as fitness components (productivity, longevity, and survival to maturity),

body size and various behavioral traits and traits related to mitotic cell division (Supplementary Table S2). For gross phenotypes M is typically a fraction of a percent per generation; the most rapid relative decline recorded is ~0.2% per generation for productivity (Vassilieva et al. 2000; Joyner-Matos et al. 2011) and lifespan (Joyner-Matos et al. 2009). Although the overall rate of directional change in metabolite pool levels is somewhat larger than that of gross phenotypes (0.73%), some metabolites show rates of directional change on the order 2–6%. These extraordinary changes in pool size are not due to a few extreme lines (Figure 1B). On average, coefficients of mutational variation (CV_{M_MA}) for gross phenotypes and metabolome pool sizes are similar to those previously observed for life-history traits (around 1–2% generation), but a few metabolites show increases of up to 4%, among the largest reported for any trait (Houle et al. 1996). Mutational heritabilities average about 10^{-3} , a typical value for gross phenotypes (Houle et al. 1996) but approximately an order of magnitude greater than the typical h_M^2 estimated for transcript abundance (Rifkin et al. 2005; Landry et al. 2007; McGuigan et al. 2014).

These results imply that the metabolome is much more subject to large effects of mutation than are gross phenotypes such as fitness, growth or behavior. Another striking difference can be seen in the direction of the mutational effects. Whereas M of gross phenotypes is almost always negative, that is, mutations tend to cause the trait to decline in magnitude, for these extremely sensitive metabolites, M is positive, that is, pool size increases.

Correlation and modularity are both hallmarks of the metabolic mutational structure

To assess the multivariate architecture of the metabolome, we first carried out PCA on the normalized metabolite concentrations (Supplementary Figure S1; Supplementary Table S3) and then addressed whether the principal components vary significantly among lines using the general linear model outlined in the Methods. Of the first eight principal components, there is significant among-line variance (i.e., V_M) for PC2-PC7 (Z-test, $P < 0.003$ in all cases); only PC1 does not vary significantly among MA lines. The non-heritable PC1 explains 25.8% of the total phenotypic variance, and presumably captures the variation due to microenvironmental fluctuations and/or experimental error. PC2 explains 18% of the phenotypic variance, and the first eight PCs collectively explain about 78% of the variance. Mutational heritabilities of PC2-PC7 are all on the order of 10^{-3} /generation (Supplementary Table S4).

The substantial departure from uniformity of the eigenvalues of the PCs indicates the presence of underlying phenotypic correlations between metabolites (Supplementary Figure S2). To investigate the extent to which the PCs are themselves *genetically* correlated (i.e., coefficients of different PCs covary among MA lines), we compared a model in which the first eight PCs were allowed to covary among lines (unstructured covariance) to a model in which the off-diagonal elements of the among-line covariance matrix were constrained to zero, using AICc as the criterion for goodness-of-fit. The simpler model with the off-diagonals constrained to zero provides a more parsimonious fit (AICc = 4515.1 vs. AICc = 4530.6, relative likelihood >2300). This finding suggests that there are many potential avenues available for the evolution of the metabolome that are relatively unentangled by genetic constraints - or in other words, the metabolome has a modular structure. However,

caution is in order; some of the point estimates of the among-line correlations between PC eigenvectors are large (Supplementary Table S5), and the significantly better fit of the constrained model may simply reflect the limitations of power.

The association of metabolite pool size with fitness

Given that the production of progeny requires the synthesis and utilization of many metabolites, we might expect that metabolite pool sizes are correlated with fitness. Because the metabolites are themselves correlated, we initially carried out a PCA based on MA line means (Supplementary Table S6) and then calculated the correlations of relative fitness with the first seven PCs (Supplementary Table S7). PC3 is marginally significantly correlated with relative fitness ($r_{PC3,w} = 0.495$, $df=34$, $P<0.0069$, sequential Bonferroni-adjusted $P_{\alpha=0.05} = 0.0071$).

These results suggest that it might be possible to predict the relative fitness of a given MA line from knowing its metabolome. We investigated the relationship between metabolites and relative fitness by means of stepwise multiple regression, using the scores of the first seven principal components calculated from line means as explanatory variables. One line (MA577) had a relative fitness of only about 1%, which is probably an anomalous measurement, so we omitted that line from the analysis. Beginning with a model with all variables (PC1-PC7), we removed variables stepwise by their F-value to find the best-fitting model, using AICc as the criterion for goodness-of-fit. The model with only PC3 has the smallest (best) AICc (6.1), but a model with only a random intercept fits almost as well (AICc=8.6, relative probability $\approx 3.5:1$).

DISCUSSION

Here we estimate, for the first time, the sensitivity of a metabolic network, or part of one, to spontaneous mutations (for a different but related perspective see Clark et al. 1995), who reported the effects of spontaneous mutations on the activities of 12 metabolic enzymes in *Drosophila melanogaster*). Obviously, there are many more than 29 metabolites present in a worm, so our view of the worm's metabolome is restricted. Nevertheless, this study lays the groundwork for future studies with greater coverage of the metabolome. More generally, it is one of the first phenomic studies of **M** beyond the transcriptome in any species.

By two criteria – the per-generational relative rate of directional change, M , and the standardized per-generation rate of increase in genetic variation, CV_M – we find that the pool sizes of metabolites vary greatly in their vulnerability to mutation. We also find that some metabolites are much more vulnerable to the directional effects of mutation than any fitness-related, morphological or behavioral trait studied until now. In other words, some metabolites show relatively little genetic canalization (*cf* Fraser and Schadt 2010).

Several studies have reported mutational heritabilities for large numbers of gene transcripts (Rifkin et al. 2005; Landry et al. 2007; McGuigan et al. 2014), and the average h_M^2 for transcript abundance is typically on the order of 10^{-4} or less. The contrast in mutational heritability between the transcriptome and the metabolome (or at least a small fraction of the metabolome) is noteworthy, inasmuch as the relative influence of genes and environment on

the metabolome appear more similar to those of gross phenotypic traits than of gene transcripts.

Although the mutational vulnerability of the metabolome *per se* has not been previously studied, the relative sensitivity of the metabolome to mutational perturbation is not unexpected (Clark et al. 1995). A variety of circumstantial evidence suggests that the metabolome is more sensitive to perturbation than are gross phenotypes. Indeed, metabolomics has been advocated as a way to study the functions of otherwise phenotypically “silent” mutations (Raamsdonk et al. 2001; Blaise et al. 2007). Metabolic networks function to maintain the flux of metabolites critical to fitness in the face of environmental and mutational perturbation, and they do so by redirecting flux through different parts of the network, which results in altered metabolite pool sizes.

The discovery of strong directional mutational biases in some metabolites prompts the question of how the integrity of the metabolome is preserved in nature. We found that several metabolites showed modest correlations with relative fitness across MA lines, collectively manifested in PC3, which implies that the metabolome is under selection. If metabolite pool sizes are at evolutionary equilibrium, however, then selection must act in a particular way: it must oppose the mutational bias for any given metabolite. For example, if mutational bias tends to increase pool size, then at equilibrium selection must oppose this increase, that is, there must be a negative genetic correlation between pool size and relative fitness. Is this in fact so? Of the 29 metabolites, 20 showed a significant directional bias and a significant mutational variance. For each of these 20, we asked whether M and $r_{w,metabolite}$ were in the same or opposite directions. We found that in 10 cases, they were the same direction and 10 they were not; so there is no tendency for natural selection to oppose the mutational bias. Looking at it another way, of the four metabolites – succinic acid, xanthine, nicotinic acid and adenosine – that were most vulnerable to directional mutation, each having $M > 2\%$, none showed significant pairwise correlations with fitness, much less in a way implying selection against the mutations that altered their pool sizes so spectacularly. This suggests one of two possibilities: firstly, that some metabolite pool sizes are evolving rapidly under mutation pressure; or, secondly, and more plausibly, that our laboratory studies have not captured all of the fitness consequences of these mutations, and hence all of the selective forces that maintain the integrity of the metabolome. It should of course also be borne in mind that it is not generally meaningful to assign a single biological interpretation to the pool size of any one metabolite. Furthermore, our current analysis contains no information on cellular or tissue localization specific changes in metabolism, and while this a natural consequence of the methods we used, it does mean that future research could still pick up effects based on changes in metabolite concentrations within an individual that are not detected when looking at whole organism extracts.

Since evolution ultimately depends on new mutations, the presence of strong directional mutational biases and large mutational variances suggests that the evolution of the metabolome may be far less constrained than gross phenotypes. The presence of significant mutational heritability of seven of the first eight phenotypic principal components reinforces the conclusion that there is substantial scope for evolution in metabolic space, as does the lack of significant among-line covariance among those PCs. On the other hand, the skewed

distribution of the eigenvalues indicates the presence of significant phenotypic correlations, which may reflect underlying genetic constraint.

Recently, several studies have compared the vulnerability to mutation of various traits such as body size, fitness and vulval morphology among different strains and species of nematodes (Baer et al. 2005; Ostrow et al. 2007; Braendle et al. 2010). They have shown that species potentially vary substantially in the quality and quantity of new mutations affecting these traits, and that the cumulative mutational effects may be context-dependent (Matsuba et al. 2013). Given the size of the mutational effects detected here, it would be of great interest to know whether this is also true for the metabolome.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Art Edison and Joe Travis for helpful discussions, and Thomas Bataillon, Annalise Paaby, Ruth Shaw, and an anonymous reviewer for their many helpful comments. SKD was funded by the Natural Environment Research Council (NERC), UK, for a PhD scholarship. CFB was funded by NIH grants R01GM072639 and R01GM107227.

LITERATURE CITED

- Ajje BC, Estes S, Lynch M, Phillips PC. Behavioral degradation under mutation accumulation in *Caenorhabditis elegans*. *Genetics*. 2005; 170:655–660. [PubMed: 15834141]
- Arnold SJ. Constraints on phenotypic evolution. *The American Naturalist*. 1992; 140:S85–S107.
- Azevedo RBR, Keightley PD, Lauren-Maatta C, Vassilieva LL, Lynch M, Leroi AM. Spontaneous mutational variation for body size in *Caenorhabditis elegans*. *Genetics*. 2002; 162:755–765. [PubMed: 12399386]
- Baer CF, Miyamoto MM, Denver DR. Mutation rate variation in multicellular eukaryotes: causes and consequences. *Nature Reviews Genetics*. 2007; 8:619–631.
- Baer CF, Phillips N, Ostrow D, Avalos A, Blanton D, Boggs A, Keller T, Levy L, Mezerhane E. Cumulative effects of spontaneous mutations for fitness in *Caenorhabditis*: Role of genotype, environment and stress. *Genetics*. 2006; 174:1387–1395. [PubMed: 16888328]
- Baer CF, Shaw F, Steding C, Baurgartner M, Hawkins A, Houppert A, Mason N, Reed M, Sinnonelic K, Woodard W, Lynch M. Comparative evolutionary genetics of spontaneous mutations affecting fitness in rhabditid nematodes. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102:5785–5790. [PubMed: 15809433]
- Begin M, Schoen DJ. Low impact of germline transposition on the rate of mildly deleterious mutation in *Caenorhabditis elegans*. *Genetics*. 2006; 174:2129–2136. [PubMed: 17057249]
- Behrends V, Tredwell GD, Bundy JG. Gavin, a add-on to AMDIS, new GUI-driven version. *Analytical Biochemistry*. 2011; 415:206–208. [PubMed: 21575589]
- Blaise BJ, Giacomotto J, Elena B, Dumas ME, Toulhoat P, Segalat L, Emsley L. Metabotyping of *Caenorhabditis elegans* reveals latent phenotypes. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104:19808–19812. [PubMed: 18077412]
- Braendle C, Baer CF, Felix M-A. Bias and evolution of the mutationally accessible phenotypic space in a developmental system. *PLoS Genetics*. 2010; 6
- Brakefield PM. Evo-devo and constraints on selection. *Trends in Ecology & Evolution*. 2006; 21:362–368. [PubMed: 16713653]
- Clark AG, Wang L, Hulleberg T. Spontaneous mutation rate of modifiers of metabolism in *Drosophila*. *Genetics*. 1995; 139:767–779. [PubMed: 7713431]

- Dean AM, Thornton JW. Mechanistic approaches to the study of evolution: the functional synthesis. *Nature Reviews Genetics*. 2007; 8:675–688.
- Denver DR, Morris K, Streelman JT, Kim SK, Lynch M, Thomas WK. The transcriptional consequences of mutation and natural selection in *Caenorhabditis elegans*. *Nature Genetics*. 2005; 37:544–548. [PubMed: 15852004]
- Drake JW, Charlesworth B, Charlesworth D, Crow JF. Rates of spontaneous mutation. *Genetics*. 1998; 148:1667–1686. [PubMed: 9560386]
- Estes S, Ajie BC, Lynch M, Phillips PC. Spontaneous mutational correlations for life-history, morphological and behavioral characters in *Caenorhabditis elegans*. *Genetics*. 2005; 170:645–653. [PubMed: 15834140]
- Etienne V, Andersen EC, Ponciano JM, Blanton D, Cadavid A, Joyner-Matos J, Matsuba C, Tabman B, Baer CF. The Red Death meets the abdominal bristle: Polygenic mutation for susceptibility to a bacterial pathogen in *Caenorhabditis elegans*. *Evolution*. 2015; 69:508–519. [PubMed: 25495240]
- Farhadifar R, Baer CF, Valfort A-C, Andersen EC, Mueller-Reichert T, Delattre M, Needleman DJ. Scaling, selection, and evolutionary dynamics of the mitotic spindle. *Current Biology*. 2015; 25:732–740. [PubMed: 25683802]
- Fraser HB, Schadt EE. The quantitative genetics of phenotypic robustness. *PLoS One*. 2010; 5:e8635. [PubMed: 20072615]
- Fry, JD. Estimation of genetic variances and covariances by restricted maximum likelihood using PROC MIXED. In: Saxton, AM., editor. *Genetic Analysis of Complex Traits Using SAS*. SAS Institute, Inc; Cary, NC: 2004.
- Fry JD, Heinsohn SL. Environment dependence of mutational parameters for viability in *Drosophila melanogaster*. *Genetics*. 2002; 161:1155–1167. [PubMed: 12136018]
- Fuchs S, Bundy JG, Davies SK, Viney JM, Swire JS, Leroi AM. A metabolic signature of long life in *Caenorhabditis elegans*. *BMC Biol*. 2010; 8
- Geier FM, Fuchs S, Valbuena G, Leroi AM, Bundy JG. Profiling the metabolic signature of senescence. *Methods in molecular biology* (Clifton, N.J.). 2013; 965:355–371.
- Geier FM, Want EJ, Leroi AM, Bundy JG. Cross-platform comparison of *Caenorhabditis elegans* tissue extraction strategies for comprehensive metabolome coverage. *Analytical Chemistry*. 2011; 83:3730–3736. [PubMed: 21480661]
- Gould, SJ. *The Structure of Evolutionary Theory*. Harvard; Cambridge, Mass: 2002.
- Halket JM, Przyborowska A, Stein SE, Mallard WG, Down S, Chalmers RA. Deconvolution gas chromatography/mass spectrometry of urinary organic acids -- potential for pattern recognition and automated identification of metabolic disorders. *Rapid Communications in Mass Spectrometry*. 1999; 13:279–284. [PubMed: 10097403]
- Halligan DL, Keightley PD. Spontaneous mutation accumulation studies in evolutionary genetics. *Annual Review of Ecology Evolution and Systematics*. 2009; 40:151–172.
- Hansen M, Flatt T, Aguilaniu H. Reproduction, fat metabolism, and life span: what is the connection? *Cell Metabolism*. 2013; 17:10–19. [PubMed: 23312280]
- Hendrikse JL, Parsons TE, Hallgrímsson B. Evolvability as the proper focus of evolutionary developmental biology. *Evolution & Development*. 2007; 9:393–401. [PubMed: 17651363]
- Houle D. Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters. *Evolution*. 1991:630–648.
- Houle D, Morikawa B, Lynch M. Comparing mutational variabilities. *Genetics*. 1996; 143:1467–1483. [PubMed: 8807316]
- Ibarra RU, Edwards JS, Palsson BO. *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature*. 2002; 420:186–189. [PubMed: 12432395]
- Joyner-Matos J, Bean LC, Richardson HL, Sammeli T, Baer CF. No evidence of elevated germline mutation accumulation under oxidative stress in *Caenorhabditis elegans*. *Genetics*. 2011; 189:1439–1447. [PubMed: 21979932]
- Joyner-Matos J, Upadhyay A, Salomon MP, Grigaltchik V, Baer CF. Genetic (co)variation for life span in rhabditid nematodes: role of mutation, selection, and history. *Journals of Gerontology Series a-Biological Sciences and Medical Sciences*. 2009; 64:1134–1145.

- Katju V, Packard LB, Bu L, Keightley PD, Bergthorsson U. Fitness decline in spontaneous mutation accumulation lines of *Caenorhabditis elegans* with varying effective population sizes. *Evolution*. 2015; 69:104–116. [PubMed: 25338758]
- Keightley PD, Caballero A. Genomic mutation rates for lifetime reproductive output and lifespan in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America*. 1997; 94:3823–3827. [PubMed: 9108062]
- Kind T, Wohlgemuth G, Lee do Y, Lu Y, Palazoglu M, S S, Fiehn O. Mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Analytical Chemistry*. 2009; 81::10038–10048. [PubMed: 19928838]
- Landry CR, Lemos B, Rifkin SA, Dickinson WJ, Hartl DL. Genetic properties influencing the evolvability of gene expression. *Science*. 2007; 317:118–121. [PubMed: 17525304]
- Leroi AM. Molecular signals versus the Loi de Balancement. *Trends in Ecology & Evolution*. 2001; 16:24–29. [PubMed: 11146141]
- Lewis NE, Hixson KK, Conrad TM, Lerman JA, Charusanti P, Polpitiya AD, Adkins JN, Schramm G, Purvine SO, Lopez-Ferrer D, Weitz KK, Eils R, Konig R, Smith RD, Palsson BO. Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. *Molecular Systems Biology*. 2010; 6
- Lisec J, Schauer N, Kopka J, Willmitzer L, Fernie AR. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nature Protocols*. 2006; 1:387–396. [PubMed: 17406261]
- Lynch, M.; Walsh, B. *Genetics and Analysis of Quantitative Traits*. Sinauer; Sunderland, MA: 1998.
- Matsuba C, Ostrow DG, Salomon MP, Tolani A, Baer CF. Temperature, stress and spontaneous mutation in *Caenorhabditis briggsae* and *Caenorhabditis elegans*. *Biology Letters*. 2013; 9:20120334. [PubMed: 22875817]
- Maynard Smith J, Burian R, Kauffman S, Alberch P, Campbell J, Goodwin B, Lande R, Raup D, Wolpert L. Developmental constraints and evolution: a perspective from the Mountain Lake conference on development and evolution. *Quarterly Review of Biology*. 1985; 60:265–287.
- McGuigan K, Collet JM, McGraw EA, Ye YH, Allen SL, Chenoweth SF, Blows MW. The nature and extent of mutational pleiotropy in gene expression of male *Drosophila serrata*. *Genetics*. 2014; 196:911–921. [PubMed: 24402375]
- Ostrow D, Phillips N, Avalos A, Blanton D, Boggs A, Keller T, Levy L, Rosenbloom J, Baer CF. Mutational bias for body size in rhabditid nematodes. *Genetics*. 2007;1653–1661. [PubMed: 17483403]
- Papp B, Notebaart RA, Pal C. Systems-biology approaches for predicting genomic evolution. *Nature Reviews Genetics*. 2011; 12:591–602.
- Psujek S, Beer RD. Developmental bias in evolution: evolutionary accessibility of phenotypes in a model evo-devo system. *Evolution and Development*. 2008; 10:375–390. [PubMed: 18460098]
- Raamsdonk LM, Teusink B, Broadhurst D, Zhang NS, Hayes A, Walsh MC, Berden JA, Brindle KM, Kell DB, Rowland JJ, Westerhoff HV, van Dam K, Oliver SG. A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nature Biotechnology*. 2001; 19:45–50.
- Rifkin SA, Houle D, Kim J, White KP. A mutation accumulation assay reveals a broad capacity for rapid evolution of gene expression. *Nature*. 2005; 438:220–223. [PubMed: 16281035]
- Srinivasan J, von Reuss SH, Bose N, Zaslaver A, Mahanti P, Ho MC, O'Doherty OG, Edison AS, Sternberg PW, Schroeder FC. A modular library of small molecule signals regulates social behaviors in *Caenorhabditis elegans*. *PLoS Biology*. 2012; 10
- Stiernagle, T. Maintenance of *C. elegans*. In: T. C. e. R. Community. , editor. *WormBook*. 2006.
- Stoltzfus A. Mutationism and the dual causation of evolutionary change. *Evolution and Development*. 2006; 8:304–317. [PubMed: 16686641]
- Stupp GS, von Reuss SH, Izrayelit Y, Ajredini R, Schroeder FC, Edison AS. Chemical detoxification of small molecules by *Caenorhabditis elegans*. *ACS Chemical Biology*. 2013; 8:309–313. [PubMed: 23163740]
- Swire J, Fuchs S, Bundy JG, Leroi AM. The cellular geometry of growth drives the amino acid economy of *Caenorhabditis elegans*. *Proceedings of the Royal Society B-Biological Sciences*. 2009; 276:2747–2754.

- Vassilieva LL, Hook AM, Lynch M. The fitness effects of spontaneous mutations in *Caenorhabditis elegans*. *Evolution*. 2000; 54:1234–1246. [PubMed: 11005291]
- Vassilieva LL, Lynch M. The rate of spontaneous mutation for life-history traits in *Caenorhabditis elegans*. *Genetics*. 1999; 151::119–129. [PubMed: 9872953]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

M. B. Frequency distribution of pool sizes for individual lines for the four most rapidly changing metabolites: orange arrowheads indicate means of ancestral control line; green arrowheads indicate means of MA lines.