



Network Architecture and Mutational Sensitivity of the *C. elegans* Metabolome

Lindsay M. Johnson^{1†}, Luke M. Chandler^{2†}, Sarah K. Davies³ and Charles F. Baer^{1,2*}

¹ Department of Biology, University of Florida, Gainesville, FL, United States, ² University of Florida Genetics Institute, Gainesville, FL, United States, ³ Department of Surgery and Cancer, Faculty of Medicine, Imperial College, London, United Kingdom

OPEN ACCESS

Edited by:

Michael Anton Witting,
Helmholtz Zentrum München,
Helmholtz-Gemeinschaft Deutscher
Forschungszentren (HZ), Germany

Reviewed by:

Jonathan Arnold,
University of Georgia, United States
Arthur S. Edison,
University of Georgia, United States

*Correspondence:

Charles F. Baer
cbaer@ufl.edu

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Metabolomics,
a section of the journal
Frontiers in Molecular Biosciences

Received: 23 March 2018

Accepted: 06 July 2018

Published: 31 July 2018

Citation:

Johnson LM, Chandler LM, Davies SK
and Baer CF (2018) Network
Architecture and Mutational Sensitivity
of the *C. elegans* Metabolome.
Front. Mol. Biosci. 5:69.
doi: 10.3389/fmolb.2018.00069

A fundamental issue in evolutionary systems biology is understanding the relationship between the topological architecture of a biological network, such as a metabolic network, and the evolution of the network. The rate at which an element in a metabolic network accumulates genetic variation via new mutations depends on both the size of the mutational target it presents and its robustness to mutational perturbation. Quantifying the relationship between topological properties of network elements and the mutability of those elements will facilitate understanding the variation in and evolution of networks at the level of populations and higher taxa. We report an investigation into the relationship between two topological properties of 29 metabolites in the *C. elegans* metabolic network and the sensitivity of those metabolites to the cumulative effects of spontaneous mutation. The correlations between measures of network centrality and mutability are not statistically significant, but several trends point toward a weak positive association between network centrality and mutational sensitivity. There is a small but significant negative association between the mutational correlation of a pair of metabolites (r_M) and the shortest path length between those metabolites. Positive association between the centrality of a metabolite and its mutational heritability is consistent with centrally-positioned metabolites presenting a larger mutational target than peripheral ones, and is inconsistent with centrality conferring mutational robustness, at least *in toto*. The weakness of the correlation between r_M and the shortest path length between pairs of metabolites suggests that network locality is an important but not overwhelming factor governing mutational pleiotropy. These findings provide necessary background against which the effects of other evolutionary forces, most importantly natural selection, can be interpreted.

Keywords: metabolic network, mutation accumulation, mutational correlation, mutational variance, network centrality

INTRODUCTION

The set of chemical reactions that constitute organismal metabolism is often represented as a network of interacting components, in which individual metabolites are the nodes in the network and the chemical reactions of metabolism are the edges linking the nodes (Jeong et al., 2000). Representation of a complex biological process such as metabolism as a network is conceptually

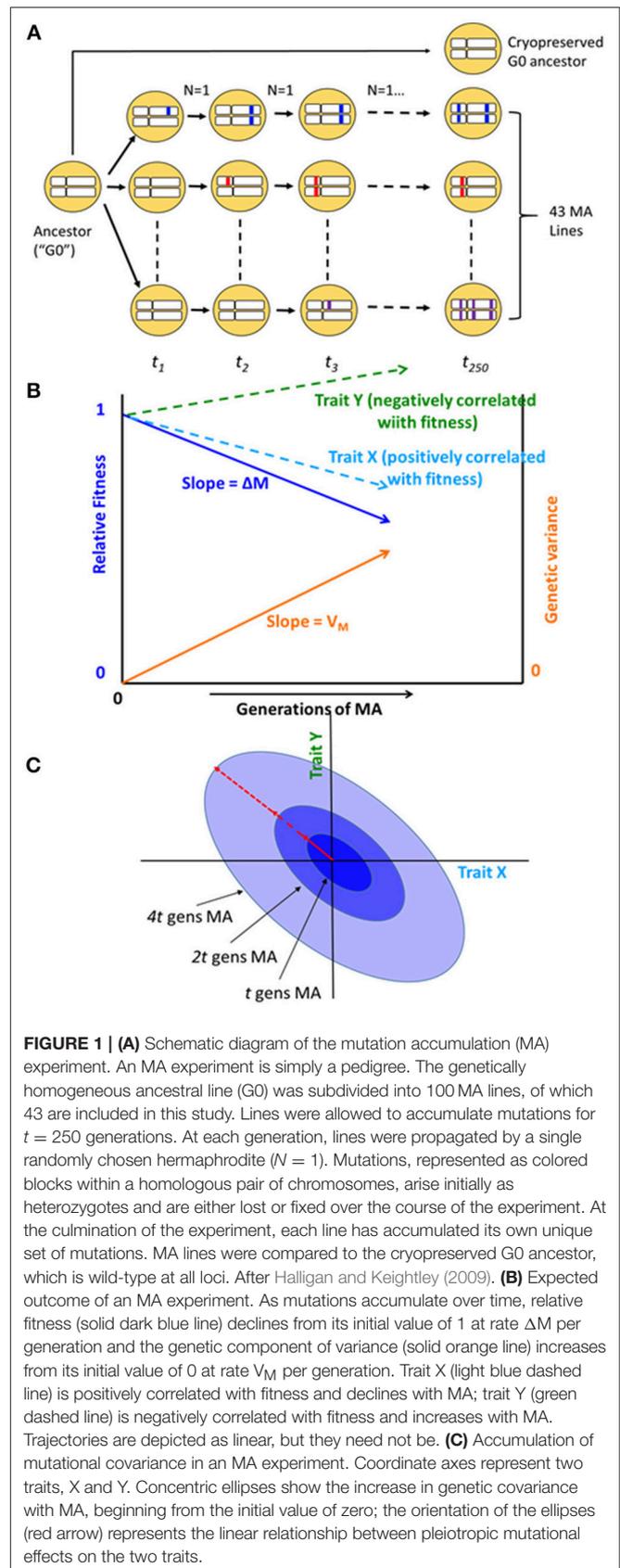
powerful because it offers a convenient and familiar way of visualizing the system, as well as a well-developed mathematical framework for analysis.

If the representation of a biological system as a network is to be useful as more than a metaphor, it must have predictive power (Winterbach et al., 2013). Metabolic networks have been investigated in the context of evolution, toward a variety of ends. Many studies have compared empirical metabolic networks to various random networks, with the goal of inferring adaptive features of network architecture (e.g., Fell and Wagner, 2000; Jeong et al., 2000; Wagner and Fell, 2001; Siegal et al., 2007; Minnhagen and Bernhardsson, 2008; Papp et al., 2009; Bernhardsson and Minnhagen, 2010). Other studies have addressed the relationship between network-level properties of individual elements of the network (e.g., node degree, centrality) and properties such as rates of protein evolution (Vitkup et al., 2006; Greenberg et al., 2008), within-species polymorphism (Hudson and Conant, 2011), and mutational robustness (Levy and Siegal, 2008).

One fundamental evolutionary process that remains essentially unexplored with respect to metabolic networks is mutation. Mutation is the ultimate source of genetic variation, and as such provides the raw material for evolution: the greater the input of genetic variation by mutation, the greater the capacity for evolution. However, in a well-adapted population, most mutations are at least slightly deleterious. At equilibrium, the standing genetic variation in a population represents a balance between the input of new mutations that increase genetic variation and reduce fitness, and natural selection, which removes deleterious variants and thereby increases fitness. Because genetic variation is jointly governed by mutation and selection, understanding the evolution of any biological entity, such as a metabolic network, requires an independent accounting of the effects of mutation and selection.

The cumulative effects of spontaneous mutations can be assessed in the near absence of natural selection by means of a mutation accumulation (MA) experiment (Figure 1). Selection becomes ineffective relative to random genetic drift in small populations, and mutations with effects on fitness smaller than about the reciprocal of the population size (technically, the genetic effective population size, N_e) will be essentially invisible to natural selection (Kimura, 1968). An MA experiment minimizes the efficacy of selection by minimizing N_e , thereby allowing all but the most strongly deleterious mutations to evolve as if they are invisible to selection (Halligan and Keightley, 2009).

Our primary interest is in the relationship between the centrality of a metabolite in the network and the sensitivity of that metabolite to mutation. Roughly speaking, the centrality of a node in a network quantifies some measure of the importance of the node in the network (Koschützki and Schreiber, 2008). A generic property of empirical networks, including metabolic networks, is that they are (approximately) scale-free; scale-free networks are characterized by a topology with a few “hub” nodes (high centrality) and many peripheral nodes (low centrality; Jeong et al., 2000). Scale-free networks are more robust to



random perturbation than are randomly-connected networks (Albert et al., 2000).

Mutation is an important source of perturbation to biological systems, and much effort has gone into theoretical and empirical characterization of the conditions under which mutational robustness will evolve (Wagner et al., 1997; De Visser et al., 2003; Proulx et al., 2007). Mutational robustness can be assessed in two basic ways: top-down, in which a known element of the system is mutated and the downstream effects of the mutation quantified, or bottom-up, in which mutations are introduced at random, either spontaneously or by mutagenesis, and the downstream effects quantified. Top-down experiments are straightforward to interpret: the greater the effects of the mutation (e.g., on a phenotype of interest), the less robust the system. However, the scope of inference is limited to the types of mutations introduced by the investigator (which in practice are almost always gene knockouts), and provide limited insight into natural variation in mutational robustness.

Bottom-up approaches, in which mutations are allowed to accumulate at random, provide insight into the evolution of a system as it actually exists in nature: all else equal, a system, or element of a system (“trait”), that is robust to the effects of mutation will accumulate less genetic variance under MA conditions than one that is not robust (Figure 1B; Stearns et al., 1995). However, the inference is not straightforward, because all else may not be equal: different systems or traits may present different mutational targets (roughly speaking, the number of sites in the genome that potentially affect a trait; Houle, 1998).

Ultimately, disentangling the evolutionary relationship between network architecture, mutational robustness, and mutational target is an empirical enterprise, specific to the system of interest. As a first step, it is necessary to establish the relationship between network architecture (e.g., topology) and the rate of accumulation of genetic variance under MA conditions. If a general relationship emerges, targeted top-down experiments can then be employed to dissect the relationship in more mechanistic detail.

In addition to the relationship between metabolite centrality and mutational variance, we are also interested in the relationship between network topology and the mutational correlation (r_M) between pairs of metabolites (Figure 1C). In principle, mutational correlations reflect pleiotropic relationships between genes underlying pairs of traits (but see below for caveats; Estes et al., 2005). Genetic networks are often modular (Newman, 2006), consisting of groups of genes (modules) within which pleiotropy is strong and between which pleiotropy is weak (Wagner et al., 2007). Genetic modularity implies that mutational correlations will be negatively correlated with the length of the shortest path between network elements. However, it is possible that the network of gene interactions underlying metabolic regulation is not tightly correlated with the metabolic network itself, e.g., if *trans* acting regulation predominates.

Here we report results from a long-term MA experiment in the nematode *Caenorhabditis elegans*, in which replicate MA lines derived from a genetically homogeneous common ancestor (G0) were allowed to evolve under minimally effective selection ($N_e \approx 1$) for approximately 250 generations (Figure 1A).

We previously reported estimates from these MA lines of two key quantitative genetic parameters by which the cumulative effects of mutation on the metabolome can be quantified: the per-generation change in the trait mean (the mutational bias, ΔM) and the per-generation increase in genetic variation (the mutational variance, V_M) for the standing pools of 29 metabolites (Davies et al., 2016); **Supplementary Table 1**. In this report, we interpret those results, and new estimates of mutational correlations (r_M), in the context of the topology of the *C. elegans* metabolic network.

Studies with *C. elegans* have contributed significantly to our understanding of the mutational process, and of short-term evolution more generally (reviewed in Teotónio et al., 2017). However, the key feature of *C. elegans* that makes it such a powerful model system in many contexts—small size, with its many associated benefits—presents a challenge in the context of metabolomics, because current methods require pooling samples from thousands of individual worms. Because metabolomic profiles vary over the course of development, meaningful comparisons between groups require that samples be carefully controlled for timing of development.

MATERIALS AND METHODS

Metabolic Network

The metabolic network of *C. elegans* was constructed following the criteria of Ma and Zeng (2003b), from two reaction databases (i) from Ma and Zeng (2003b); updated at <http://www.ibiodesign.net/kneva/>; we refer to this database as MZ, and (ii) from Yilmaz and Walhout (2016); <http://wormflux.umassmed.edu/>; we refer to this database as YW. Subnetworks that do not contain at least one of the 29 metabolites were excluded from downstream analyses. The method includes several *ad hoc* criteria for retaining or omitting specific metabolites from the analysis (criteria are listed on p. 272 of Ma and Zeng, 2003b). The set of reactions in the MZ and YW databases are approximately 99% congruent; in the few cases in which there is a discrepancy (listed in **Supplementary Table 2**), we chose to use the MZ database because we used the MZ criteria for categorizing currency metabolites (defined below).

To begin, the 29 metabolites of interest were identified and used as starting sites for the network. Next, all forward and reverse reactions stemming from the 29 metabolites were incorporated into the subnetwork until all reactions either looped back to the starting point or reached an endpoint. Currency metabolites were removed following the MZ criteria; a currency metabolite is roughly defined as a molecule such as water, proton, ATP, NADH, etc., that appears in a large fraction of metabolic reactions but is not itself an intermediate in an enzymatic pathway. Metabolic networks in which currency metabolites are included have much shorter paths than networks in which they are excluded. When currency metabolites are included in the network reported here, all shortest paths are reduced to no more than three steps, and most of the shortest paths consist of one or two steps. The biological relevance of path length when currency metabolites are included in the network is unclear (Ma and Zeng, 2003b).

A graphical representation of the network was constructed with the Pajek software package (<http://mrvar.fdv.uni-lj.si/pajek/>) and imported into the networkX Python package (Hagberg et al., 2008). Proper importation from Pajek to networkX was verified by visual inspection.

Network Parameters

Properties of networks can be quantified in many ways, and different measures of network centrality capture different features of network importance (Table 1). We did not have strong prior hypotheses about which specific measure(s) of centrality associated with a given metabolite would prove most informative in terms of a relationship with the mutational properties of that metabolite (i.e., ΔM and/or V_M). Therefore, we assessed the relationship between the mutational properties of a metabolite and several measures of its network centrality: betweenness, closeness, and degree centrality, in- and out-degree, and core number (depicted in Figure 2). These network parameters are all positively correlated. Definitions of the parameters are given in Table 1; correlations between the parameters are included in Table 2. Calculation of network parameters was done using built-in functions in NetworkX.

Mutation Accumulation Lines

A full description of the construction and propagation of the mutation accumulation (MA) lines is given in Baer et al. (2005). Briefly, 100 replicate MA lines were initiated from a nearly-isogenic population of N2-strain *C. elegans* and propagated by single-hermaphrodite descent at 4-day (one generation) intervals for approximately 250 generations. The long-term N_e of the MA lines is very close to one, which means that mutations with a selective effect less than about 25% are effectively neutral (Keightley and Caballero, 1997). The common ancestor of the MA lines (“G0”) was cryopreserved at the outset of the experiment; MA lines were cryopreserved upon completion of the MA phase of the experiment. Based on extensive whole-genome sequencing (Denver et al., 2012; Saxena et al., 2018), we estimate that each MA line carries approximately 70 mutant alleles in the homozygous state.

At the time the metabolomics experiments reported in Davies et al. (2016) were initiated, approximately 70 of the 100 MA lines remained extant, of which 43 ultimately provided sufficient material for Gas Chromatography/Mass Spectrometry (GC-MS). Each MA line was initially replicated five-fold, although not all replicates provided data of sufficient quality to include in subsequent analyses; the mean number of replicates included per MA line is 3.9 (range = 2–5). The G0 ancestor was replicated nine times. However, the G0 ancestor was not subdivided into “pseudolines” (Teotónio et al., 2017), which means that inferences about mutational variances and covariances are necessarily predicated on the assumption that the among-line (co)variance of the ancestor is zero.

Each replicate consisted of stage-synchronized young adult worms taken from a single 10 cm agar plate. Cultures were stage-synchronized by treatment with hypochlorite (“bleaching”) following Stiernagle (2006); details of the synchronization are given in Davies et al. (2016). Following synchronization, worms

were incubated at 20°C until young adulthood, defined 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 and collected for metabolomics. Each sample contained tens of thousands of worms, and although the samples were stage-synchronized, there was almost certainly some variation among samples in both the relative frequency of eggs on the plate and the (small) proportion of worms that had yet to reach adulthood.

Recently, whole-genome sequencing revealed that two MA lines, MA563 and MA564, share approximately 2/3 of their accumulated mutations; the simplest explanation is that the two lines were cross-contaminated around generation 150–175 of the MA protocol. However, averaged over all metabolites, the between-line standard deviation of those two lines is >3X that of either within-line SD, which suggests that the ~1/3 of the mutations in each genome that are unique to each line contribute meaningfully to the differences between those two lines. Accordingly, we chose to include both lines. Further, since only 21 (out of 33) lines that we sequenced are represented in the metabolome dataset, the possibility of further unidentified cross-contamination cannot be ruled out. Comparisons between metabolites will not be biased by shared mutations, although the sampling (co)variance will increase by a factor $k \leq \frac{N}{N-x+1}$, where N is the total number of lines and x is the number of lines that share mutations; $k = \frac{N}{N-x+1}$ if all lines that share mutations share all their mutations.

Metabolomics

Details of the extraction and quantification of metabolites are given in Davies et al. (2016). Briefly, samples were analyzed using an Agilent 5975c quadrupole mass spectrometer with a 7890 gas chromatograph. Metabolites were identified by comparison of GC-MS features to the Fiehn Library (Kind et al., 2009) using the AMDIS deconvolution software (Halket et al., 1999), followed by reintegration of peaks using the GAVIN Matlab script (Behrends et al., 2011). Metabolites were quantified and normalized relative to an external quantitation standard. 34 metabolites were identified, of which 29 were ultimately included in the analyses. Normalized metabolite data are archived in Dryad (<http://dx.doi.org/10.5061/dryad.2dn09>).

Mutational Parameters

In what follows, a “trait” is the (normalized) concentration of a metabolite. There are three mutational parameters of interest: (i) the per-generation proportional change in the trait mean, referred to as the mutational bias, ΔM ; (ii) the per-generation increase in the genetic variance, referred to as the mutational variance, V_M ; and (iii) the genetic correlation between the cumulative effects of mutations affecting pairs of traits, the mutational correlation, r_M . Details of the calculations of ΔM and V_M are reported in Davies et al. (2016); we reprise the basic calculations here.

Mutational Bias (ΔM)

The mutational bias is the change in the trait mean due to the cumulative effects of all mutations accrued over one generation.

TABLE 1 | Definitions of network parameters, following the documentation of NetworkX, v.1.11 (Hagberg et al., 2008) and mutational parameters.

Parameter	Heuristic definition	Formal definition
In Degree (IN ^o), $deg^+(v)$	The number of incoming edges to node v in a directed graph	Self-explanatory
Out Degree (OUT ^o), $deg^-(v)$	The number of outgoing edges from node v in a directed graph	Self-explanatory
Shortest Path Length, $d(v, u)$	Shortest distance from node v to another node u with no repeated walks	Self-explanatory
Betweenness Centrality (BET), $c_B(v)$	Betweenness centrality of node v is the sum of the fraction of all-pairs shortest paths that pass through v . The greater $c_B(v)$, the greater the fraction of shortest paths that pass through node v .	$\frac{c_B(v)}{(n-1)(n-2)}$, where $c_B(v) = \sum_{s,t \in V} \frac{\sigma(s,t v)}{\sigma(s,t)}$, V is the set of nodes, $\sigma(s, t)$ is the number of shortest paths from node s to node t , $\sigma(s, t v)$ is the number of paths from s to t that pass through node v , and n is the number of nodes in the graph. The denominator $(n-1)(n-2)$ is the normalization factor for a directed graph that scales $c_B(v)$ between 0 and 1
Closeness Centrality (CLO), $C(v)$	Closeness centrality of node v is the reciprocal of the sum of the shortest path lengths to all $n-1$ other nodes, normalized by the sum of minimum possible distances $n-1$. The greater $C(v)$, the closer v is to other nodes	$C(v) = \frac{n-1}{\sum_{u=1}^{n-1} d(u,v)}$, where n is the number of nodes and $d(u, v)$ is the shortest path distance between u and v
Degree Centrality (DEG), $C_D(v)$	Degree centrality of node v is the fraction of nodes in the network that node v is connected to	$C_D(v) = \frac{deg^+(v)+deg^-(v)}{n-1}$, where n is the number of nodes in the network
Core Number (CORE)	A k -core is the largest subgraph that contains nodes of at least degree k . The core number of node v is the largest value k of a k -core containing node v	Calculated using the algorithm of Batagelj and Zaversnik (2011).
Mutational Bias (ΔM)	Per-generation rate of change of the trait mean in an MA experiment. Equivalent to the product of the genome-wide mutation rate, μ_G , and the average effect of a mutation on the trait, α	$M_z = \frac{\bar{z}_{MA} - \bar{z}_0}{t}$; \bar{z}_{MA} and \bar{z}_0 represent the MA and ancestral (G0) trait means and t is the number of generations of MA
Mutational Variance (V_M)	Per-generation rate of increase in genetic variance for a trait in an MA experiment. Equivalent to the product of the genome-wide mutation rate, μ_G , and the square of the average effect of a mutation on the trait, α^2	$V_M = V_L = \frac{V_{L,MA} - V_{L,G0}}{2t}$, where $V_{L,MA}$ is the variance among MA lines, $V_{L,G0}$ is the among-line variance in the G0 ancestor, and t is the number of generations of MA
Squared coefficient of variation (I_M , I_E)	I_M is the mutational variance (V_M) scaled by the square of the trait mean, and provides a measure of the evolvability of a trait. I_E is the residual variance (V_E) scaled in the same way	
Mutational heritability (h_M²)	Mutational variance (V_M) scaled as a fraction of the residual variance (V_E). Provides a measure of the short-term response to selection on mutational variance	$h_M^2 = \frac{V_M}{V_E}$
Mutational correlation (r_M)	Genetic correlation between two traits in MA lines. Provides an estimate of pleiotropic effects of new mutations	$r_M = \frac{COV_M(X,Y)}{\sqrt{V_M(X)V_M(Y)}}$, where COV_M is the mutational covariance and V_M is the mutational variance

Abbreviations of the parameters used in **Table 2** follow the parameter name in parentheses in bold type.

$\Delta M_z = \mu_G \alpha_z$, where μ_G is the per-genome mutation rate and α_z is the average effect of a mutation on trait z , and is calculated as $M_z = \frac{\bar{z}_{MA} - \bar{z}_0}{t}$, where \bar{z}_{MA} and \bar{z}_0 represent the MA and ancestral (G0) trait means and t is the number of generations of MA. However, the ΔM was not normally distributed among the 29 metabolites, so for downstream analyses we transformed ΔM as $\Delta M^* = \log_2\left(\frac{MA}{G0}\right)$, where MA and G0 represent the trait values of the MA lines and the G0 ancestor, respectively; $\Delta M = 2^{\Delta M^*} - 1$.

Mutational Variance (V_M)

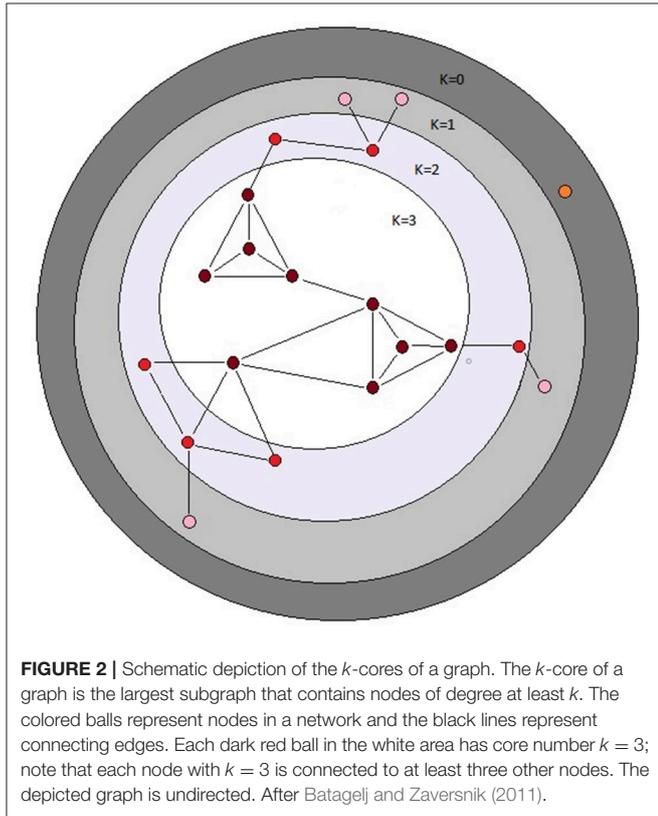
The mutational variance is the increase in the genetic variance due to the cumulative effects of all mutations accrued over one generation. $V_M = \mu_G \alpha_z^2$ and is calculated as $V_M = V_L = \frac{V_{L,MA} - V_{L,G0}}{2t}$, where $V_{L,MA}$ is the variance among MA

lines, $V_{L,G0}$ is the among-line variance in the G0 ancestor, and t is the number of generations of MA (Lynch and Walsh, 1998, p. 330). In this study, we must assume that $V_{L,G0} = 0$.

Comparisons of variation among traits or groups require that the variance be measured on a common scale. V_M is commonly scaled either relative to the trait mean, in which case V_M is the squared coefficient of variation and is often designated I_M , or relative to the residual variance, V_E ; V_M/V_E is the mutational heritability, h_M^2 . I_M and h_M^2 have different statistical properties and evolutionary interpretations (Houle et al., 1996), so we report both. For each metabolite, I_M and I_E are standardized relative to the mean of the MA lines. Both h_M^2 and I_M were natural-log transformed to meet assumptions of normality prior to downstream analyses.

Mutational Correlation, r_M

Pairwise mutational correlations were calculated from the among-line components of (co)variance, which were estimated by REML as implemented in the MIXED procedure of SAS v. 9.4, following Fry (2004). Statistical significance of individual correlations was assessed by Z-test, with a global 5% significance criterion of approximately $P < 0.000167$.



Analysis of the Relationship Between Mutational Parameters and Network Centrality

The six network parameters are all positively correlated, as are the four mutational parameters (Table 2). To assess the overall correlation structure between mutational and network parameters, we employed a hierarchical canonical correlation analysis (CCA), as implemented in the CANCELL procedure of SAS v. 9.4, with the network parameters as the “X” variables and the mutational parameters as the “Y” variables. We initially included all four mutational parameters, resulting in four pairs of canonical variates and four canonical correlations. We then repeated the analysis for each mutational parameter Y_i individually with the full set of six network parameters, resulting in one pair of canonical variates and one canonical correlation for each of the four mutational parameters. Finally, we calculated the pairwise correlation between all mutational parameters and all network parameters. For all analyses except the first, significance was assessed using the False Discovery Rate (FDR) (Benjamini and Hochberg, 1995).

Analysis of the Relationship Between Mutational Correlation (r_M) and Network Architecture

Correlation Between Mutational Correlation (r_M) and Shortest Path Length

Statistical assessment of the correlation between mutational correlation (r_M) and shortest path length presents a problem of non-independence, for two reasons. First, all correlations including the same variable (metabolite) are non-independent; each of the n elements of an $n \times n$ correlation matrix contributes to $n(n-1)/2$ correlations. Second, even though the mutational correlation between metabolites i and j is the same as the mutational correlation between j and i , the shortest path lengths need not be the same, and moreover, the path from i to j

TABLE 2 | Correlations between network parameters (Row/Column 1–5), between mutational parameters (Row/Column 6–9), between network and mutational parameters, and between residual variance (I_E , Row/Column 10) and network and mutational parameters.

	BTW	CLO	DEG	IN°	OUT°	CORE	ΔM	$ \Delta M $	h_M^2	I_M	I_E
BTW		0.43	0.49	0.52	0.39	0.48	−0.16	−0.14	0.03	−0.06	−0.10
CLO			0.52	0.51	0.45	0.52	0.14	0.21	0.21	0.27	0.06
DEG				0.90	0.93	0.79	0.09	0.06	0.25	0.15	0.16
IN°					0.67	0.82	0.22	0.23	0.30	0.21	0.25
OUT°						0.64	−0.04	−0.08	0.17	0.09	0.05
CORE							0.33	0.28	0.53*	0.30	0.28
ΔM								0.84	0.62	0.71	0.81
$ \Delta M $									0.53	0.69	0.84
h_M^2										0.72	0.43
I_M											0.82
I_E											

Abbreviations of network parameters are: BTW, betweenness centrality; CLO, closeness centrality; DEG, degree centrality; IN°, in-degree, OUT°, out-degree; CORE, core number. Abbreviations of mutational parameters are: ΔM , per-generation change in the trait mean; $|\Delta M|$, absolute value of ΔM ; h_M^2 , mutational heritability; I_M , squared mutational CV; I_E , squared residual CV. Network and mutational parameters are defined in Table 1. See text and Supplementary Table 1 for details of mutational parameters.* FDR < 0.1.

may exist whereas the path from j to i may not (depicted in **Supplementary Figure 1**). To account for non-independence of the data, we devised a parametric bootstrap procedure. Three metabolites (L-tryptophan, L-lysine, and Pantothenate) lie outside of the great strong component of the network (Ma and Zeng, 2003a) and are omitted from the analysis. Each off-diagonal element of the 24×24 mutational correlation matrix ($r_{ij} = r_{ji}$) was associated with a random shortest path length sampled with probability equal to its frequency in the empirical distribution of shortest path lengths between all metabolites included in the analysis. Next, we calculated the Spearman's correlation ρ between r_M and the shortest path length. The procedure was repeated 10,000 times to generate an empirical distribution of ρ , to which the observed ρ can be compared. This comparison was done for the raw mutational correlation, r_M , the absolute value, $|r_M|$, and between r_M and the shortest path length in the undirected network (i.e., the shorter of the two paths between metabolites i and j).

RESULTS AND DISCUSSION

Representation of the Metabolic Network

The metabolic network of *C. elegans* was estimated using method of Ma and Zeng (2003b) from two independent but largely congruent databases (Ma and Zeng, 2003b; Yilmaz and Walhout, 2016). Details of the network construction are given in section Metabolic Network of the Materials and Methods; data are presented in **Supplementary Dataset 1**. For the set of metabolites included (see section Materials and Methods), networks constructed from the MZ and YW databases give nearly identical results. In the few cases in which there is a discrepancy ($\sim 1\%$; **Supplementary Table 2**), we use the MZ network, for reasons we explain in the section Materials and Methods. The resulting network is a directed graph including 646 metabolites, with 1203 reactions connecting nearly all metabolites (**Figure 3**).

Network Centrality and Sensitivity to Mutation

Canonical correlation analysis did not identify significant correlation between mutational parameters and network parameters, either collectively (**Figure 4**; **Supplementary Table 3**) or individually. Further, of the 24 pairwise correlations between mutational parameters and network parameters (**Table 2**, **Supplementary Figure 2**), only the correlation between mutational heritability (h_M^2) and core number approaches statistical significance ($r = 0.53$, FDR < 0.1).

On the face of it, it appears there is no association between network centrality and any measure of mutational sensitivity. If so, there are various possible explanations. For example, it may be that mutational target and mutational robustness effectively cancel each other out. More worryingly, it may be that the representation of the *C. elegans* metabolic network used here misrepresents the network as it actually exists *in vivo*. For example, the topology of the dynamic metabolic network of the bacterium *E. coli* varies depending on the environmental context (Koschützki et al., 2010), and it seems intuitive that the greater

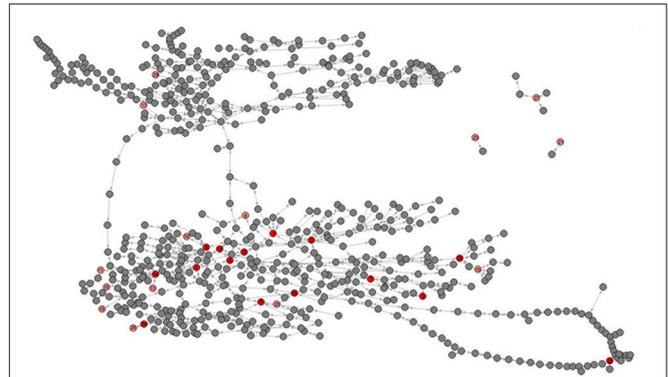


FIGURE 3 | Graphical depiction of the metabolic network including all 29 metabolites. Pink nodes represent included metabolites with core number = 1, red nodes represent included metabolites with core number = 2. Gray nodes represent metabolites with which the included 29 metabolites directly interact. Metabolite identification numbers are: 1, L-Serine; 2, Glycine; 3, Nicotinate; 4, Succinate; 5, Uracil; 6, Fumarate; 7, L-Methionine; 8, L-Alanine; 9, L-Aspartate; 10, L-3-Amino-isobutanoate; 11, trans-4-Hydroxy-L-proline; 12, (S) – Malate; 13, 5-Oxoproline; 14, L-Glutamate; 15, L-Phenylalanine; 16, L-Asparagine; 17, D-Ribose; 18, Putrescine; 19, Citrate; 20, Adenine; 21, L-Lysine; 22, L-Tyrosine; 23, Pantothenate; 24, Xanthine; 25, Hexadecanoic acid; 26, Urate; 27, L-Tryptophan; 28, Adenosine; 29, Alpha, alpha-Trehalose.

spatiotemporal complexity inherent to a multicellular organism would exacerbate that problem. Or, most straightforwardly, it may be that there simply is no functional relationship between the centrality of a metabolite in a network and its sensitivity to mutation.

However, several trends apparent in the results suggest the conservative interpretation may miss meaningful signal emerging from noisy data. First, the point estimates of the canonical correlations are not small (>0.45 in all five cases; e.g., the first canonical correlation in the full analysis is 0.69; **Supplementary Table 3**); it may simply be that the sampling variance associated with the relatively small number of mutations, MA lines and (especially) metabolites overwhelms the signal of a weak but consistently positive association. Second, of the 24 pairwise correlations among mutational and network parameters (**Table 2**), only five are negative, significantly fewer than expected at random if the variables are uncorrelated (cumulative binomial probability = 0.0033). Third, the point estimates of the pairwise correlations are not random with respect to either network or mutational parameters. For all four mutational parameters, the correlation is greatest with core number (exact binomial probability = $(1/6)^4 \approx 0.00077$). Core number is a discrete interval variable, whereas the other measures of network centrality are continuous variables. Quantifying centrality in terms of core number is analogous to categorizing a set of size measurements into “small” and “large”: power is increased, at the cost of losing the ability to discriminate between subtler differences.

Fourth, for five out of six network parameters, the correlation is greatest with h_M^2 (cumulative binomial probability = $(6)(1/4)^5(3/4) + (1/6)^6 \approx 0.0046$). V_M is the numerator of both h_M^2 and I_M ; the difference is the denominator, with h_M^2 scaling

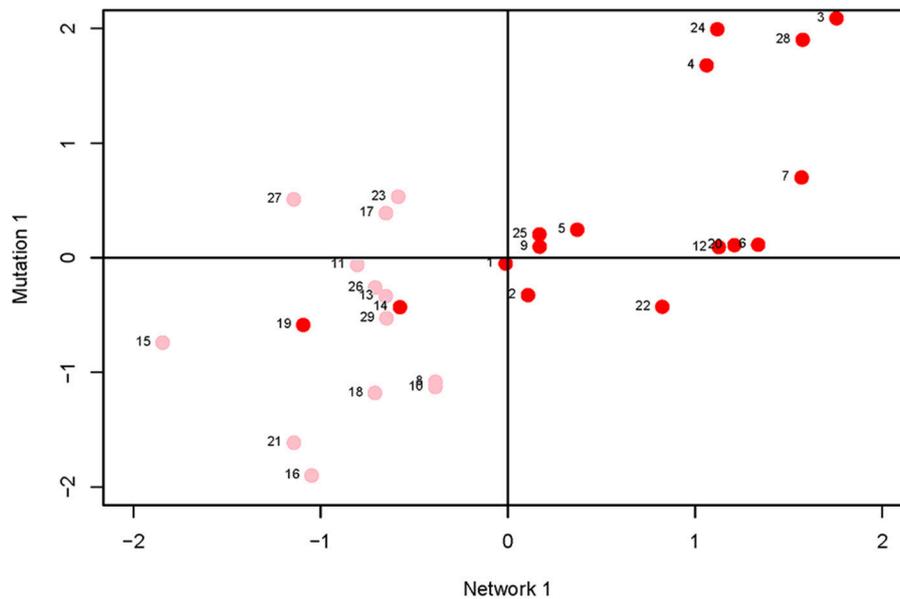


FIGURE 4 | Plot of first canonical variate pair; the network variate is plotted on the X-axis, the mutation variate is plotted on the Y-axis. Each data point represents a metabolite; the numbers are the metabolite identifiers given in the legend to **Figure 3**. Metabolites with core number = 1 are in pink, metabolites with core number = 2 are in red.

V_M by the residual variance, V_E , and I_M scaling V_M by the square of the trait mean. If V_E was more strongly associated with network topology than was V_M , h_M^2 would presumably be more strongly correlated with network parameters than would I_M , analogous to the well-documented V_E -driven negative association between the narrow-sense heritability of a trait and the correlation of the trait with fitness (Houle, 1992). However, I_M and I_E are nearly identically (un)correlated with network parameters (**Table 2**), so that scenario cannot explain the correlation. Coincidence seems as likely an explanation as any.

The Relationship Between Mutational Correlation (r_M) and Shortest Path Length

In an MA experiment, the cumulative effects of mutations on a pair of traits i and j may covary for two, nonexclusive reasons (Estes et al., 2005). More interestingly, individual mutations may have consistently pleiotropic effects, such that mutations that affect trait i also affect trait j in a consistent way. Less interestingly, but unavoidably, individual MA lines will have accumulated different numbers of mutations, and if mutations have consistently directional effects, as would be expected for traits correlated with fitness, lines with more mutations will have more extreme trait values than lines with fewer mutations, even in the absence of consistent pleiotropy. Estes et al. (2005) simulated the sampling process in *C. elegans* MA lines with mutational properties derived from empirical estimates from a variety of traits and concluded that sampling is not likely to lead to large absolute mutational

correlations in the absence of consistent pleiotropy ($|r_M| \leq 0.25$).

Ideally, we would like to estimate the full mutational (co)variance matrix, M , from the joint estimate of the among-line (co)variance matrix. However, with 25 traits, there are $(25 \times 26)/2 = 325$ covariances, and with only 43 MA lines, there is insufficient information to jointly estimate the restricted maximum likelihood of the full M matrix. To proceed, we calculated mutational correlations from pairwise REML estimates of the among-line (co)variances, i.e., $r_M = \frac{COV_L(X,Y)}{\sqrt{VAR_L(X)VAR_L(Y)}}$ (Clark et al., 1995; Mezey and Houle, 2005). Pairwise estimates of r_M are shown in **Supplementary Table 4**. To assess the extent to which the pairwise correlations are sensitive to the underlying covariance structure, we devised a heuristic bootstrap analysis. For a random subset of 12 of the 300 pairs of traits, we randomly sampled six of the remaining 23 traits without replacement and estimated r_M between the two focal traits from the joint REML among-line (co)variance matrix. For each of the 12 pairs of focal traits, we repeated the analysis 100 times.

There is a technical caveat to the preceding bootstrap analysis. Resampling statistics are predicated on the assumption that the variables are exchangeable (Shaw, 1992), which metabolites are not. For that reason, we do not present confidence intervals on the resampled correlations, only the distributions. However, we believe that the analysis provides a meaningful heuristic by which the sensitivity of the pairwise correlations to the underlying covariance structure can be assessed.

Distributions of resampled correlations are shown in **Supplementary Figure 3**. In every case the point estimate of r_M

falls on the mode of the distribution of resampled correlations, and in 11 of the 12 cases, the median of the resampled distribution is very close to the point estimate of r_M . However, in six of the 12 cases, some fraction of the resampled distribution falls outside two standard errors of the point estimate. The most important point that the resampling analysis reveals is this: given that 29 metabolites encompass only a small fraction of the total metabolome of *C. elegans*, even had we been able to estimate the joint likelihood of the full $29 \times 30/2$ M -matrix, the true covariance relationships among those 29 metabolites could conceivably be quite different from those estimated from the data.

The simplest property that describes the relationship between two nodes in a network is the length of the shortest path between them (= number of edges). In a directed network, such as a metabolic network, the shortest path from element i to element j is not necessarily the same as the shortest path from j to i . For each pair of metabolites i and j , we calculated the shortest path length from i to j and from j to i , without repeated walks (Supplementary Table 5). We then calculated Spearman's correlation ρ between the mutational correlation r_M and the shortest path length.

There is a weak, but significant, negative correlation between r_M and the shortest path length between the two metabolites ($\rho = -0.128$, two-tailed $P < 0.03$; Figure 5A), whereas $|r_M|$ is not significantly correlated with shortest path length ($\rho = -0.0058$, two-tailed $P > 0.45$; Figure 5B). The correlation between r_M and the shortest path in the undirected network is similar to the correlation between r_M and the shortest path in the directed network ($\rho = -0.105$, two-tailed $P > 0.10$; Figure 5C).

An intuitive possible cause of the weak negative association between shortest path length and mutational correlation would be if a mutation that perturbs a metabolic pathway toward the beginning of the pathway has effects that propagate downstream in the same pathway, but the effect of the perturbation attenuates. The attenuation could be due either to random noise or to the effects of other inputs into the pathway downstream from the perturbation (or both). The net effect would be a characteristic pathway length past which the mutational effects on two metabolites are uncorrelated, leading to an overall negative correlation between r_M and path length. The finding that the correlations between r_M and the shortest path length in the directed and undirected network are very similar reinforces that conclusion. The negative correlation between r_M and shortest path length is reminiscent of a finding from Arabidopsis, in which sets of metabolites significantly altered by single random gene knockouts are closer in the global metabolic network than expected by chance (Kim et al., 2015).

CONCLUSIONS AND FUTURE DIRECTIONS

The proximate goal of this study was to find out if there are topological properties of the *C. elegans* metabolic network (node centrality, shortest path length) that are correlated with a set of statistical descriptions of the cumulative effects of spontaneous

mutations (ΔM , V_M , r_M). Ultimately, we hope that a deeper understanding of those mathematical relationships will shed light on the mechanistic biology of the organism. Bearing in mind the statistical fragility of the results, we conclude:

Network Centrality May Be Associated With Mutational Sensitivity (V_M); It Is Not Associated With Mutational Robustness ($1/V_M$)

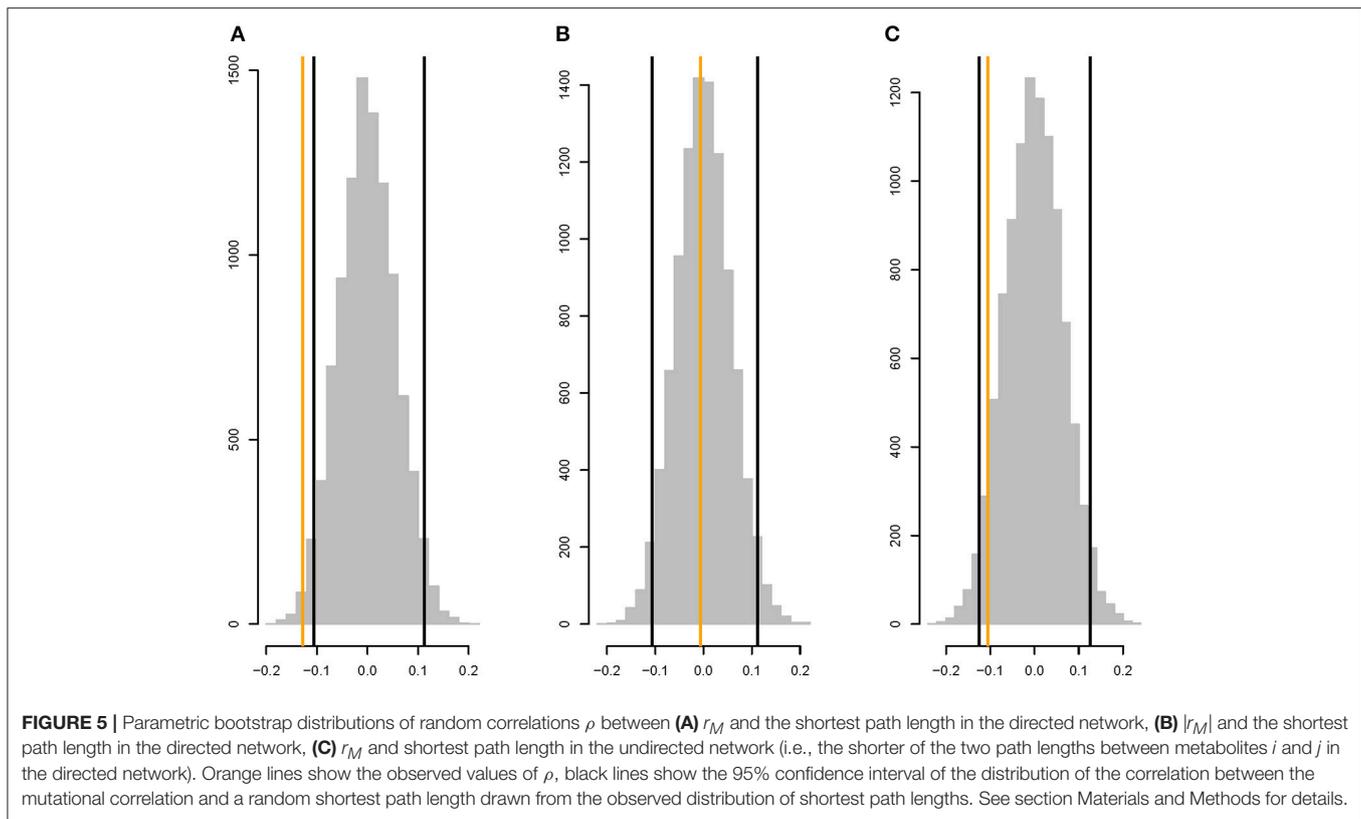
If in fact the apparently non-random features of the data represent a hint of signal emerging from the noise, the most plausible explanation is that metabolites that are central in the network present a larger mutational target than do metabolites that peripherally located. Somewhat analogously, Landry et al. (2007) investigated the mutational properties of transcription in a set of yeast MA lines and found that h_M^2 is positively correlated with both the number of genes with which a given gene interacts ("trans-mutational target") and the number of transcription factor binding sites in a gene's promoter ("cis-mutational target"). Those authors did not formally quantify the network properties of the set of transcripts, although it seems likely that mutational target size as they defined it is positively correlated with centrality in the transcriptional network. It is important to note, however, although $1/V_M$ is a meaningful measure of mutational robustness (Stearns and Kawecki, 1994), it does not necessarily follow that highly-connected metabolites are therefore more robust to the effects of individual mutations (Houle, 1998; Ho and Zhang, 2016).

Pleiotropic Effects of Mutations Affecting the Metabolome Are Predominantly Local

Pleiotropic effects of mutations affecting the metabolome are predominantly local, as evidenced by the significant negative correlation between the mutational correlation, r_M , and the shortest path length between a pair of metabolites. That result is not surprising in hindsight, but the weakness of the correlation suggests that there are other important factors that underlie pleiotropy beyond network proximity.

Future Directions

To advance understanding of the mutability of the *C. elegans* metabolic network, three things are needed. First, it will be important to cover a larger fraction of the metabolic network. Untargeted mass spectrometry of cultures of *C. elegans* reveals many thousands of features (Art Edison, personal communication); 29 metabolites are only the tip of a large iceberg. For example, our intuition leads us to believe that the mutability of a metabolite will depend more on its in-degree (mathematically, the number of edges leading into a node in a directed graph; biochemically, the number of reactions in which the metabolite is a product) than its out-degree. For all four mutational parameters, the point-estimate of the pairwise correlation with in-degree is greater than that with out-degree (Table 2), although that result is not statistically significant (exact binomial probability = 0.0625). We used an *ad hoc* data cloning strategy (Lele et al., 2007) to assess the relationship



between metabolite sample size and statistical power to detect a significant correlation between network features and mutational heritability; the method is explained and results are presented in **Supplementary Figure 4**.

Second, to more precisely partition mutational (co)variance into within- and among-line components, more MA lines are needed. We estimate that each MA line carries about 70 unique mutations (see section Materials and Methods), thus the mutational (co)variance is the result of about 3000 total mutations, distributed among 43 MA lines. The MA lines were a preexisting resource, and the sample size was predetermined. Mutational heritability of these metabolite traits is typical of many types of traits ($\bar{h}_M^2 \approx 0.001/\text{generation}$), and it is encouraging that we were able to detect significant mutational variance for 25/29 metabolites (**Supplementary Table 1**). However, only 14% (42/300) of pairwise mutational correlations are significantly different from zero at the experiment-wide 5% significance level, roughly corresponding to $|r_M| > 0.5$ (**Supplementary Table 4**); 18 of the 42 significant mutational correlations are not significantly different from $|r_M| = 1$. The issue of statistical power to detect genetic correlations is of course not unique to metabolomic traits (Phillips, 1998). Moreover, it remains uncertain how sensitive estimates of mutational correlations are to the underlying covariance structure of the metabolome. It also remains to be seen if the mutability of specific features of metabolic networks are genotype or species-specific, and the extent to which mutability depends on environmental context.

Third, it will be important to quantify metabolites (static concentrations and fluxes) with more precision. The metabolite data analyzed in this study were collected from large cultures ($n > 10,000$ individuals) of approximately stage-synchronized worms, and were normalized relative to an external quantitation standard (Davies et al., 2016). Metabolic phenotypes are inherently properties of individual organisms; an ideal experiment would characterize the metabolomes of single individuals, assayed at the identical stage of development. Single-worm metabolomics is on the near horizon (M. Witting, personal communication). Minimizing the number of individuals in a sample is important for two reasons; (1) the smaller the sample, the easier it is to be certain the individuals are at the same developmental stage, and (2) knowing the exact number of individuals in a sample makes normalization relative to an external standard more interpretable. Ideally, data would be normalized relative to both an external standard and an internal standard (e.g., total protein; Clark et al., 1995).

This study provides an initial assessment of the relationship between mutation and metabolic network architecture. To begin to uncover the relationship between metabolic architecture and natural selection, the next step is to repeat these analyses with respect to the standing genetic variation (V_G). There is some reason to think that more centrally-positioned metabolites will be more evolutionarily constrained (i.e., under stronger purifying selection) than peripheral metabolites (Vitkup et al., 2006), in which case the ratio of the mutational variance to the

standing genetic variance (V_M/V_G) will increase with increasing centrality.

Finally, although most mutations are deleterious, not all are, and the network properties of beneficial mutations remain unexplored. One potentially fruitful avenue of research would be to do a study analogous to this one in a set of lines initiated from a mutationally-degraded progenitor that have been allowed to recover fitness at large population size (“recovery lines”; Estes and Lynch, 2003; Estes et al., 2011).

AUTHOR CONTRIBUTIONS

CB conceived the research, analyzed data and wrote the paper. LJ and LC analyzed data. SD performed GC-MS, analyzed data.

ACKNOWLEDGMENTS

This work was initially conceived by Armand Leroi and Jake Bundy. We thank Art Edison, Dan Hahn, Tom Hladish, Marta Wayne, Michael Witting, and several reviewers for their generosity and helpful advice. We especially thank Hongwu Ma for leading us to and through his metabolite database

REFERENCES

- Albert, R., Jeong, H., and Barabasi, A. L. (2000). Error and attack tolerance of complex networks. *Nature* 406, 378–382. doi: 10.1038/35019019
- Baer, C. F., Shaw, F., Steding, C., Baumgartner, M., Hawkins, A., Houppert, A., et al. (2005). Comparative evolutionary genetics of spontaneous mutations affecting fitness in rhabditid nematodes. *Proc. Natl. Acad. Sci. U.S.A.* 102, 5785–5790. doi: 10.1073/pnas.0406056102
- Batagelj, V., and Zaversnik, M. (2011). Fast algorithms for determining (generalized) core groups in social networks. *Adv. Data Anal. Classif.* 5, 129–145. doi: 10.1007/s11634-010-0079-y
- Behrends, V., Tredwell, G. D., and Bundy, J. G. (2011). Gavin, a add-on to AMDIS, new GUI-driven version. *Anal. Biochem.* 415, 206–208. doi: 10.1016/j.ab.2011.04.009
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* 57, 289–300.
- Bernhardsson, S., and Minnhagen, P. (2010). Selective pressure on metabolic network structures as measured from the random blind-watchmaker network. *New J. Phys.* 12:103047. doi: 10.1088/1367-2630/12/10/103047
- Clark, A. G., Wang, L., and Hulleberg, T. (1995). Spontaneous mutation rate of modifiers of metabolism in *Drosophila*. *Genetics* 139, 767–779.
- Davies, S. K., Leroi, A. M., Burt, A., Bundy, J., and Baer, C. F. (2016). The mutational structure of metabolism in *Caenorhabditis elegans*. *Evolution* 70, 2239–2246. doi: 10.1111/evo.13020
- Denver, D. R., Wilhelm, L. J., Howe, D. K., Gafner, K., Dolan, P. C., and Baer, C. F. (2012). Variation in base-substitution mutation in experimental and natural lineages of *Caenorhabditis* nematodes. *Genome Biol. Evol.* 4, 513–522. doi: 10.1093/gbe/evs028
- de Visser, J., Hermisson, J., Wagner, G. P., Meyers, L. A., Bagheri, H. C., Blanchard, J. L., et al. (2003). Perspective: evolution and detection of genetic robustness. *Evolution* 57, 1959–1972. doi: 10.1111/j.0014-3820.2003.tb00377.x
- Estes, S., Ajie, B. C., Lynch, M., and Phillips, P. C. (2005). Spontaneous mutational correlations for life-history, morphological and behavioral characters in *Caenorhabditis elegans*. *Genetics* 170, 645–653. doi: 10.1534/genetics.104.040022

and Reviewer #3 for his/her many insightful comments and suggestions. Support was provided by NIH grant R01GM107227 to CFB and E. C. Andersen.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2018.00069/full#supplementary-material>

Supplementary Figure 1 | Depiction of shortest path length in a directed network.

Supplementary Figure 2 | Plots of mutational parameters vs. network statistics.

Supplementary Figure 3 | Bootstrap distributions of r_M with six randomly chosen covariates.

Supplementary Figure 4 | Power curve for metabolite sample size.

Supplementary Table 1 | Network and mutational parameters of metabolites.

Supplementary Table 2 | Table of discrepancies between MZ and YW methods.

Supplementary Table 3 | Canonical correlation analysis.

Supplementary Table 4 | Mutational and environmental correlations.

Supplementary Table 5 | Shortest network path lengths.

Supplementary Data Set 1 | Metabolic network data.

- Estes, S., and Lynch, M. (2003). Rapid fitness recovery in mutationally degraded lines of *Caenorhabditis elegans*. *Evolution* 57, 1022–1030. doi: 10.1111/j.0014-3820.2003.tb00313.x
- Estes, S., Phillips, P. C., and Denver, D. R. (2011). Fitness recovery and compensatory evolution in natural mutant lines of *C. elegans*. *Evolution* 65, 2335–2344. doi: 10.1111/j.1558-5646.2011.01276.x
- Fell, D. A., and Wagner, A. (2000). The small world of metabolism. *Nat. Biotechnol.* 18, 1121–1122. doi: 10.1038/81025
- Fry, J. D. (2004). “Estimation of genetic variances and covariances by restricted maximum likelihood using PROC MIXED,” in *Genetic Analysis of Complex Traits Using*, S. A. S. ed A. M. Saxton (Cary, NC: SAS Institute, Inc.), 11–34.
- Greenberg, A. J., Stockwell, S. R., and Clark, A. G. (2008). Evolutionary constraint and adaptation in the metabolic network of *Drosophila*. *Mol. Biol. Evol.* 25, 2537–2546. doi: 10.1093/molbev/msn205
- Hagberg, A. A., Schult, D. A., and Swart, P. J. (2008). “Exploring network structure, dynamics, and function using NetworkX,” in *7th Python in Science Conference (SciPy 2008)*, eds G. Varoquaux, T. Vaught, and J. Millman (Los Alamos, NM: United States Dept. of Energy), 11–15.
- Halket, J. M., Przyborowska, A., Stein, S. E., Mallard, W. G., Down, S., and Chalmers, R. A. (1999). Deconvolution gas chromatography/mass spectrometry of urinary organic acids – potential for pattern recognition and automated identification of metabolic disorders. *Rapid Commun. Mass Spectrom.* 13, 279–284. doi: 10.1002/(SICI)1097-0231(19990228)13:4<279::AID-RCM478>3.0.CO;2-I
- Halligan, D. L., and Keightley, P. D. (2009). Spontaneous mutation accumulation studies in evolutionary genetics. *Ann. Rev. Ecol. Evol. Syst.* 40, 151–172. doi: 10.1146/annurev.ecolsys.39.110707.173437
- Ho, W. C., and Zhang, J. Z. (2016). Adaptive genetic robustness of *Escherichia coli* metabolic fluxes. *Mol. Biol. Evol.* 33, 1164–1176. doi: 10.1093/molbev/msw002
- Houle, D. (1992). Comparing evolvability and variability of quantitative traits. *Genetics* 130, 195–204.
- Houle, D. (1998). How should we explain variation in the genetic variance of traits? *Genetica* 103, 241–253. doi: 10.1023/A:1017034925212
- Houle, D., Morikawa, B., and Lynch, M. (1996). Comparing mutational variabilities. *Genetics* 143, 1467–1483.

- Hudson, C. M., and Conant, G. C. (2011). Expression level, cellular compartment and metabolic network position all influence the average selective constraint on mammalian enzymes. *BMC Evol. Biol.* 11:89. doi: 10.1186/1471-2148-11-89
- Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., and Barabasi, A. L. (2000). The large-scale organization of metabolic networks. *Nature* 407, 651–654. doi: 10.1038/35036627
- Keightley, P. D., and Caballero, A. (1997). Genomic mutation rates for lifetime reproductive output and lifespan in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U.S.A.* 94, 3823–3827. doi: 10.1073/pnas.94.8.3823
- Kim, T., Dreher, K., Nilo-Poyanco, R., Lee, I., Fiehn, O., Lange, B. M., et al. (2015). Patterns of metabolite changes identified from large-scale gene perturbations in arabidopsis using a genome-scale metabolic network. *Plant Physiol.* 167, 1685–1698. doi: 10.1104/pp.114.252361
- Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature* 217, 624–626. doi: 10.1038/217624a0
- Kind, T., Wohlgemuth, G., Lee Do, Y., Lu, Y., Palazoglu, M., Shahbaz, S., et al. (2009). Mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Anal. Chem.* 81, 10038–10048. doi: 10.1021/ac9019522
- Koschützki, D., Junker, B. H., Schwender, J., and Schreiber, F. (2010). Structural analysis of metabolic networks based on flux centrality. *J. Theor. Biol.* 265, 261–269. doi: 10.1016/j.jtbi.2010.05.009
- Koschützki, D., and Schreiber, F. (2008). Centrality analysis methods for biological networks and their application to gene regulatory networks. *Gene Regul. Syst. Bio.* 2008, 193–201. doi: 10.4137/GRSB.S702
- Landry, C. R., Lemos, B., Rifkin, S. A., Dickinson, W. J., and Hartl, D. L. (2007). Genetic properties influencing the evolvability of gene expression. *Science* 317, 118–121. doi: 10.1126/science.1140247
- Lele, S. R., Dennis, B., and Lutscher, F. (2007). Data cloning: easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods. *Ecol. Lett.* 10, 551–563. doi: 10.1111/j.1461-0248.2007.01047.x
- Levy, S. F., and Siegal, M. L. (2008). Network hubs buffer environmental variation in *Saccharomyces cerevisiae*. *PLoS Biol.* 6:e264. doi: 10.1371/journal.pbio.0060264
- Lynch, M., and Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sunderland, MA: Sinauer.
- Ma, H. W., and Zeng, A. P. (2003a). The connectivity structure, giant strong component and centrality of metabolic networks. *Bioinformatics* 19, 1423–1430. doi: 10.1093/bioinformatics/btg177
- Ma, H. W., and Zeng, A. P. (2003b). Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms. *Bioinformatics* 19, 270–277. doi: 10.1093/bioinformatics/19.2.270
- Mezey, J. G., and Houle, D. (2005). The dimensionality of genetic variation for wing shape in *Drosophila melanogaster*. *Evolution* 59, 1027–1038. doi: 10.1111/j.0014-3820.2005.tb01041.x
- Minnhagen, P., and Bernhardsson, S. (2008). The blind watchmaker network: scale-freeness and evolution. *PLoS ONE* 3:e1690. doi: 10.1371/journal.pone.0001690
- Newman, M. E. J. (2006). Modularity and community structure in networks. *Proc. Natl. Acad. Sci. U.S.A.* 103, 8577–8582. doi: 10.1073/pnas.0601602103
- Papp, B., Teusink, B., and Notebaart, R. A. (2009). A critical view of metabolic network adaptations. *HFSP J.* 3, 24–35. doi: 10.2976/1.3020599
- Phillips, P. C. (1998). Designing experiments to maximize the power of detecting correlations. *Evolution* 52, 251–255. doi: 10.1111/j.1558-5646.1998.tb05158.x
- Proulx, S. R., Nuzhdin, S. V., and Promislow, D. E. L. (2007). Direct selection on genetic robustness revealed in the yeast transcriptome. *PLoS ONE* 2:e911. doi: 10.1371/journal.pone.0000911
- Saxena, A. S., Salomon, M. P., Matsuba, C., and Yeh, S.-D. (2018). Tempo, Mode, and Fitness Effects of Mutation in *C. elegans*. *BioRxiv*. doi: 10.1101/280826
- Shaw, R. G. (1992). Comparison of quantitative genetic parameters - reply. *Evolution* 46, 1967–1969. doi: 10.1111/j.1558-5646.1992.tb01185.x
- Siegal, M. L., Promislow, D. E. L., and Bergman, A. (2007). Functional and evolutionary inference in gene networks: does topology matter? *Genetica* 129, 83–103. doi: 10.1007/s10709-006-0035-0
- Stearns, S. C., Kaiser, M., and Kawecki, T. J. (1995). The differential genetic and environmental canalization of fitness components in *Drosophila melanogaster*. *J. Evol. Biol.* 8, 539–557. doi: 10.1046/j.1420-9101.1995.8050539.x
- Stearns, S. C., and Kawecki, T. J. (1994). Fitness sensitivity and the canalization of life-history traits. *Evolution* 48, 1438–1450. doi: 10.1111/j.1558-5646.1994.tb02186.x
- Stiernagle, T. (2006). Maintenance of *C. elegans*. *WormBook* 11, 1–11. doi: 10.1895/wormbook.1.101.1
- Teotônio, H., Estes, S., Phillips, P. C., and Baer, C. F. (2017). Experimental evolution with *Caenorhabditis nematodes*. *Genetics* 206, 691–716. doi: 10.1534/genetics.115.186288
- Vitkup, D., Kharchenko, P., and Wagner, A. (2006). Influence of metabolic network structure and function on enzyme evolution. *Genome Biol.* 7:R39. doi: 10.1186/gb-2006-7-5-r39
- Wagner, A., and Fell, D. A. (2001). The small world inside large metabolic networks. *Proc. R. Soc. B Biol. Sci.* 268, 1803–1810. doi: 10.1098/rspb.2001.1711
- Wagner, G. P., Booth, G., and Bagheri, H. C. (1997). A population genetic theory of canalization. *Evolution* 51, 329–347. doi: 10.1111/j.1558-5646.1997.tb02420.x
- Wagner, G. P., Pavlicev, M., and Cheverud, J. M. (2007). The road to modularity. *Nat. Rev. Genet.* 8, 921–931. doi: 10.1038/nrg2267
- Winterbach, W., Van Mieghem, P., Reinders, M., Wang, H. J., and De Ridder, D. (2013). Topology of molecular interaction networks. *BMC Syst. Biol.* 7:90. doi: 10.1186/1752-0509-7-90
- Yilmaz, L. S., and Walhout, A. J. (2016). A *Caenorhabditis elegans* genome-scale metabolic network model. *Cell Syst.* 2, 297–311. doi: 10.1016/j.cels.2016.04.012

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling Editor and reviewer AE declared their involvement as co-editors in the Research Topic, and confirm the absence of any other collaboration.

Copyright © 2018 Johnson, Chandler, Davies and Baer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.