Wiring between close nodes in biological networks evolves more quickly than between distant nodes

Alejandro Gil-Gomez, alejandro.gilgomez@stonybrook.edu
Joshua S. Rest*, joshua.rest@stonybrook.edu, 631-632-1916
Department of Ecology and Evolution
Stony Brook University
650 Life Sciences
Stony Brook, NY 11794-4254

*Corresponding author

Abstract

Protein-protein interaction networks and metabolic networks change across species as a result of a wide range of evolutionary processes; these processes result in the accumulation of network differences between species. The rate at which biological networks evolve is an important question in evolutionary biology. Previous empirical work has focused on interactomes from model organisms to calculate rewiring rates, but this is limited by the relatively small number of species and incomplete nature of network data across species. We present an alternative approach to study network evolution using drug-drug interactions (DDIs) as a proxy for network topology. Here, we propose the rate at which DDIs change across species as an estimate of the rate at which the underlying biological network changes as species diverge. We computed the evolutionary rates of DDIs using previously published data on a high throughput study in gram-negative bacteria. In parallel, we mapped drugs with known targets in biological networks and calculated network distance, connectivity, and adjacency between targets. Our findings indicate that the targets of synergistic DDIs are closer in biological networks than other types of DDIs and that synergistic interactions have a higher evolutionary rate, meaning that nodes that are closer in biological networks evolve at a faster rate. Future studies of network evolution may use DDI data and phylogenetic comparative methods to gain larger-scale perspectives on the details of network evolution within and between species.

Introduction

Biological networks are representations of molecular interactions in the cell. These networks are built by collecting biochemical and genetic interactions from different kinds of biological entities, an approach that has improved drastically in the last decades with the advent of modern high-

throughput methods. However, there are still many limitations to the gathering and analysis of biological networks, specifically for studying network evolution across and within species. This is because many biological networks have poor quality or are incomplete and because the number of organisms for which network data is available is still very limited (Cusick et al. 2005; Jin et al. 2013; Ghadie et al. 2018).

Biological networks evolve as nodes and edges are added or lost in the network. This can be the result of different processes affecting network content, such as gene duplication, gene loss, pseudogenization, horizontal gene transfer, or whole genome duplication (Wagner 2003; Cork and Purugganan 2004), or a result of processes affecting quantitative properties, including non-synonymous substitutions in the nodes affecting their function (Jensen 1976; Ghadie et al. 2018). These mutations may affect the binding of ligands, protein domains, or DNA motifs, which in turn may result in changes in their associated metabolic, signaling, or gene expression networks. The type of network also may affect the rewiring rates of nodes. For example, gene regulatory networks tend to rewire at faster rates than metabolic networks (Shou et al. 2011), this suggests that some network types are less constrained to network evolution as a whole, but the rate at which whole networks change and what factors determine this process remain to be studied in the future.

Moreover, evolutionary events that remove old connections and generate new ones may be random (i.e., neutral or occurring via genetic drift) or may affect fitness. We know from studying static networks that not all motifs are equally abundant in biological networks (Picard et al. 2008), suggesting the action of natural selection. Natural selection may remove deleterious connections and favor advantageous ones. Rewiring rates also vary within networks, such as between different network modules and between nodes depending on their importance for the cell's survival and reproduction, and the essentiality of a node or module may be a constraining factor for network evolution. Previous studies have also shown an inverse relationship between highly connected proteins and their rate of substitution (Fraser et al. 2002), possibly due to strong purifying selection acting on the interfacial sites of interacting proteins (Zotenko et al. 2008).

Part of the key to understanding the forces acting on networks is a characterization of evolutionary rates of network change. Despite advances in studying the connections of individual nodes and network rewiring, the rate and patterns by which quantitative differences between networks accumulate as a function of species divergence remain relatively unexplored. If non-neutral forces act, evolution in networks should be correlated with metrics of inter-node connectivity, including minimum path length (the number of steps along the shortest path between nodes), K-edge connectivity (minimum number of edges that can be removed to disconnect two nodes), and degree (number of edges connected to the node) (Jordan 1869; Erdős and Gallai 1960; Cormen 2001). However, it is unknown whether evolutionary rates of change in inter-node connectivity are associated with these inter-node network measures. This is important for understanding what aspects of network structure that we can quantitate are subject to evolutionary forces acting on variation. This allows us to ask a broader question: which aspects of network structure that we can measure are important to organisms?

In this paper, we use drug-drug interactions (DDIs) as a quantitative proxy measurement of inter-node connectivity, since it has been shown that DDIs are partially dependent on the underlying network topology between targets (Lehár et al. 2007; Yeh et al. 2009). DDIs occur when the effect of two or more drugs is significantly stronger or weaker than their expected combined effect, respectively named synergies and antagonisms (Greco et al. 1995). DDIs are used in the

82 development of novel pharmacological treatments with higher efficiencies at lower doses and to
83 reduce the evolution of drug resistance (Cowen and Steinbach 2008). The reasons why a DDI occurs
84 are varied, but it has been shown that a DDI can occur as a result of factors such as the topology of
85 the underlying network between drug targets, or the essentiality of the metabolites blocked by the
86 perturbation in combination with the inhibition efficiency of the drug on the drug target (Yeh et al.
87 2009). For example, synergies can occur if drugs act on parallel pathways, where the individual
88 effect of each drug is small and an alternative pathway can compensate for the effect of a single
89 drug. Alternatively, antagonistic interactions can occur mainly in two different ways: by causing a
90 partial loss of function in two parallel pathways of an essential product, or as a result of drugs
91 acting sequentially along the same pathway. This relation between DDIs and network topology has
92 been further explored using experiments and metabolic flux simulations in yeast, suggesting that
93 specific combinations of perturbations in the network result in different quantifiable interaction
94 scores (Lehár et al. 2007). More recently, it has been shown that most synergistic interactions are
95 the result of drugs targeting the same cellular process, while antagonistic interactions are the result
96 of drugs targeting different processes (Brochado et al. 2018).

97 Only a few studies have explored interspecific variation of DDIs (Spitzer et al. 2011; Robbins et al.
98 2015; Brochado et al. 2018), but they have demonstrated that DDI measurement could be scalable in
99 the number of species and strains, allowing for the evaluation of network evolution hypotheses
100 within and between species. A high throughput study of DDIs in gram-negative bacteria has shown
101 that synergies are more conserved across species than antagonisms and additive interactions
102 (Brochado et al. 2018).

103 These previous findings suggest several important questions about the evolutionary patterns of
104 biological networks and DDIs across species. Do different parts of biological networks, which in turn
105 are reflected in different types of DDIs, evolve at different rates as species diverge? Do DDIs diverge
106 consistently with neutral processes acting on the underlying network structure? Or is there
107 heterogeneity across the network, for example with differences accumulating slowly in constrained
108 local neighborhoods, and faster between more distant connections? Does the evolutionary rate of
109 DDIs or biological networks depend on evolutionary forces acting on the drug targets and their
110 connectivity?

111 Another way to parse this question is to divide up interactions by type. Do synergistic interactions
112 evolve at a slower rate than other types of drug interactions, and do antagonistic interactions evolve
113 at a faster rate? These differences in rates would be a result of synergistic interactions taking place
114 in local neighborhoods of nodes, while antagonistic interactions act across distant network
115 neighborhoods, which may evolve more slowly due to greater network redundancy between distant
116 nodes. Indeed, if we were to map DDIs to protein targets (when known), are the drug targets of
117 synergistic drug interactions closer in the network than additive and antagonistic interactions? And,
118 more generally, are more closely connected nodes subject to higher rates of evolution than more
119 distantly related nodes? These questions can be addressed by comparing rates of DDI evolution with
120 measures of inter-node connectivity.

121 To address these questions, we used the most complete available dataset of DDIs measured across
122 species and strains (Brochado et al. 2018). We modeled these DDIs under a phylogenetic
123 comparative framework and applied a multivariate Brownian motion model to estimate the
124 evolutionary rate of interaction scores for different clusters of DDIs in six strains (three species) of
125 gram-negative bacteria. We also mapped DDIs to their putative protein targets to evaluate them in

126 known biological networks. We show that DDIs can be used as an effective proxy to evaluate
127 macroevolutionary implications and the role of network evolution in interspecific variation.

128

129 Results

130 DDI divergence between the strains increases as a 131 function of species divergence.

132 We constructed a phylogenetic tree based on highly conserved proteins (**Fig. 1A**), which highlights
133 the close relationship between the three species (including two strains from each species) considered
134 in this study, *E. coli*, *S. enterica*, and *P. aeruginosa*. Other species of Bacteria and Archaea were
135 used to calibrate the nodes for phylogenetic dating and to calibrate the rate of DDI evolution. We
136 estimated the MRCA of the three gammaproteobacteria species in our study at 1049-1128 MYA
137 (95% HPD) and the MRCA of the two *Enterobacteriaceae* (*E. coli* and *S. enterica*) at 57-68.15 MYA
138 (95% HPD).

139 We obtained DDI scores (drug-drug interaction Bliss scores) from a previous study (Brochado et al.
140 2018) that assessed almost 3000 combinations of 79 different compounds on six strains, two from
141 each of the tree species. Clustering of the six strains based on their DDI scores revealed a 'chemical
142 distance' tree consistent with the phylogeny ("Strains" in **Fig. 1B**). To test whether DDI scores
143 ("chemical patristic distance") diverge as a function of species divergence, we used linear and
144 logarithmic regressions, which modeled the amount of DDI variation explained by sequence
145 divergence (Manhattan distance on the phylogeny).

146 We sought to estimate the evolutionary rate of DDI change. However, there is low power with only
147 six tips to estimate rate shifts for any individual DDI. To address this limitation, we used tSNE to
148 detect clusters of DDIs that behave similarly across species, allowing for increased power within
149 each cluster. In order to find DDI clusters that are the most biologically relevant from among a
150 large parameter space (after filtering solutions), we first filtered for the top five solutions with the
151 highest degree of phylogenetic modular signal (Adams et al. 2016). From among those five, we
152 chose the solution with the biggest differences in evolutionary rate among clusters, when fitted
153 with a multivariate Brownian motion model. A detailed description of the filters and tests used to
154 select the clusters is in the Methods.

155 The resulting solution, with a perplexity of 245 and 15 clusters, was used for the remainder of the
156 analysis. Cluster 14 had the highest evolutionary rate measured using the Brownian motion
157 standard deviation parameter σ^2, with a value of 0.00784 bliss score^2 per million years, while all
158 the other clusters had rates ranging from 0.00044 to 0.00367. Cluster 15 had the second highest
159 rate with a value of 0.00367 bliss score^2 per million years, and cluster 1 had the third highest rate
160 with a value of 0.00354 bliss score^2 per million years.

161 We examined drug-drug interaction (DDI) variation among these six strains by building a
162 phylomorphospace plot using PCA and found that *Pseudomonas* and non-*Pseudomonas* strains
163 differentiate well along PC1, which explains 47.01% of the variance (**Fig. 1C**). PC2 distinguishes
164 between *E. coli* and *Salmonella* strains and explains 21.54% of the variance. Strains of the same

165 species group closely together in the PCA space, indicating that they share similar global DDI
166 responses. For example, the *Pseudomonas* strains appear to be closer to each other in the
167 phylomorphospace suggesting that they have small DDI variation, alternatively, *E. coli* and *S.*
168 *typhimurium* have greater intraspecific DDI variation.

169 We also correlated the branch length of the dendrogram obtained from clustering DDI data with the
170 branch lengths obtained from the Bayesian phylogeny. We found thatchemical distance increases as
171 a function of phylogenetic distance (linear regression: R^2=0.634, p-value=0.0003; **Fig. 1D**). This
172 relationship is even stronger when phylogenetic distance is log transformed (R^2 =0.918, p-
173 value=$2.4*10^{-8}$; **Fig. 1E**). These results suggest that DDIs, and their underlying network basis,
174 diverge as a function of sequence divergence. DDI divergence plateaus at higher phylogenetic
175 divergence values; this type of saturation has been previously reported in models of network
176 evolution for rewiring rates across species (Shou et al. 2011). Our observation that DDIs diverge as
177 a function of species divergence is consistent with neutral processes acting on the underlying
178 network structure, although selective processes certainly also play a role, for example, clusters with
179 high evolutionary rates may be driven by positive selection.

180

181 Synergistic interactions have the shortest average 182 distance between targets, followed by additive and 183 antagonistic interactions.

184 We observed that highly synergistic DDIs across all species tend to occur when both drugs belong to
185 the same chemical category and target the same cellular process, as can be seen from the
186 accumulation of green bars for category and process across cluster 14 (**Fig. 1B**; in agreement with
187 (Brochado et al. 2018)). This motivated us to formally examine how DDIs map onto biological
188 networks. We therefore leave discussion of DDI rates for now, to first describe our mapping of the
189 DDI targets to known biological networks, and how connectivity on these networks relates to
190 interaction type (antagonism, additivity, synergy); in the next section we will describe how DDI
191 evolutionary rates map onto interaction types network connectivity. We were able to identify
192 protein targets for 39 of the 79 drugs tested by Brochado *et al.* (2018) (see Methods). These drugs
193 had a total of 27 target proteins as identified by their unique protein IDs in *E. coli*. The most
194 common target category was bacterial penicillin-binding protein, a group involved in the
195 biosynthesis of bacterial cell walls. Other target categories mapped include ribosomal RNA, DNA
196 polymerase, topoisomerase, thymidylate synthase, and mitochondrial glycerol-3-phosphate (**Sup.**
197 **Table S1**).

198

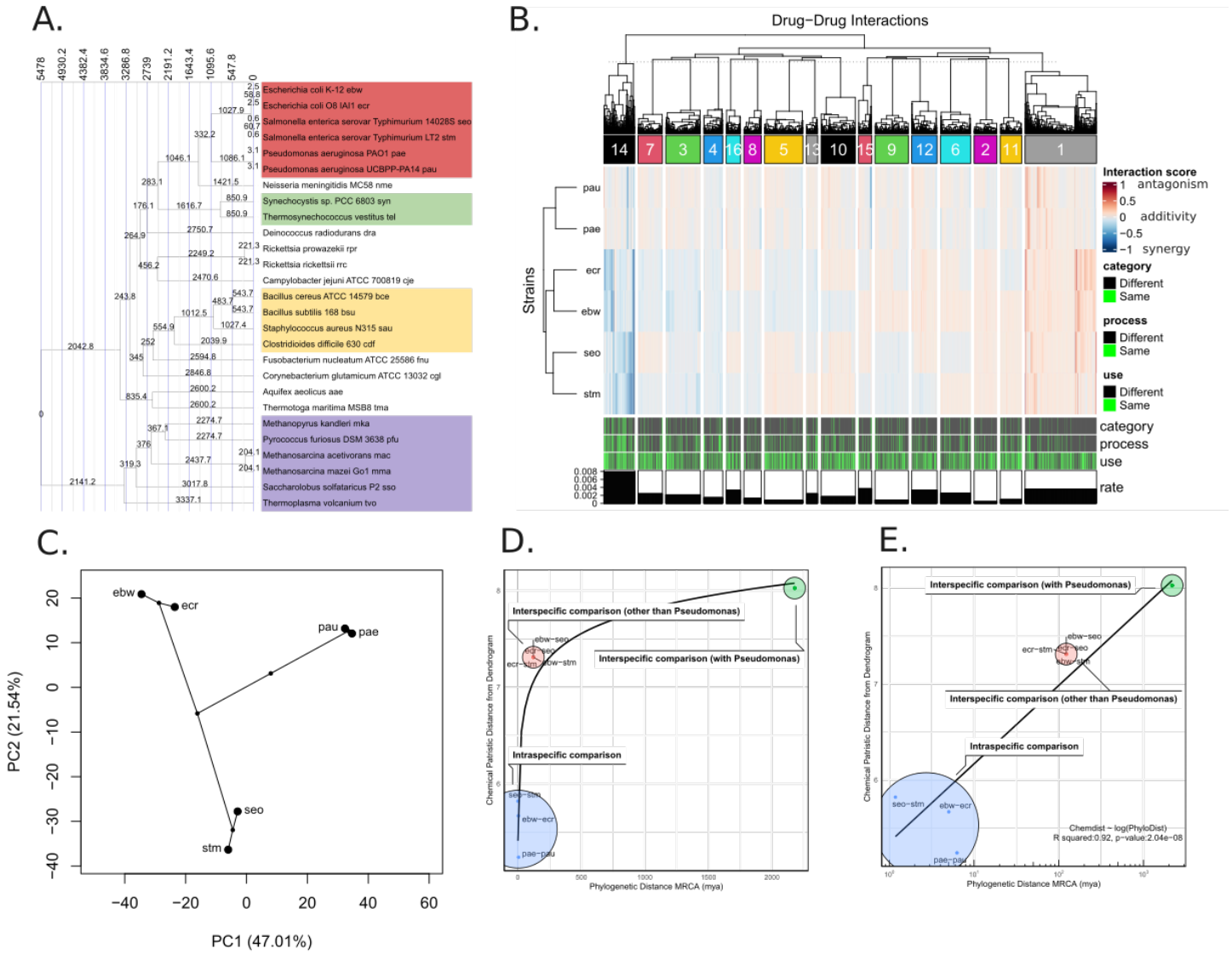

Figure 1: Drug-drug interaction (DDI) chemical distance scales with phylogenetic distance. **A.** Bayesian phylogeny of an alignment of 35 highly conserved proteins in different species of bacteria and Archaea used for phylogenetic dating. Gammaproteobacteria (red), cyanobacteria (green), bacillota (yellow), and Archaea (purple). **B.** Heatmap of DDI scores across strains. Hierarchical clustering based on euclidean distances across columns (strains: *E. coli* K-12 BW2952 (ebw), *Escherichia coli* O8 IAI1 (ecr), *S. enterica subsp. enterica* serovar Typhimurium 14028S (seo), *Salmonella enterica* subsp. enterica serovar Typhimurium LT2 (stm), *Pseudomonas aeruginosa* PAO1 (pae), *P. aeruginosa* UCBPP-PA14 (pau)), and rows (DDIs); the latter was constrained by tSNE clusters (see text). Synergistic interactions are shown in blue and antagonistic interactions are in red, all DDIs are measured in bliss score units. The annotations on the right indicate whether the two drugs involved in the interaction belong to the same drug category, target the same cellular process, or have the same use (in green) or not (black). The bottom row indicates the evolutionary rate calculated in each cluster, measured in bliss score units^2 per million years. **C.** Phylomorphospace plot combining a PCA of the DDI data with the phylogeny; principal components 1 and 2 are shown. **D, E.** DDI-based Euclidean distances between strains scale as a function of phylogenetic distance. This relationship saturates at large phylogenetic distances. Intraspecific comparisons (in blue) have the lowest amount of sequence and chemical divergence, while comparisons with *Pseudomonas* are the most distant in terms of sequence and chemical divergence. **E.** Same comparison is shown when phylogenetic distance is plotted on a log scale. All drug-drug interaction data is from Brochado *et al.* (2018).

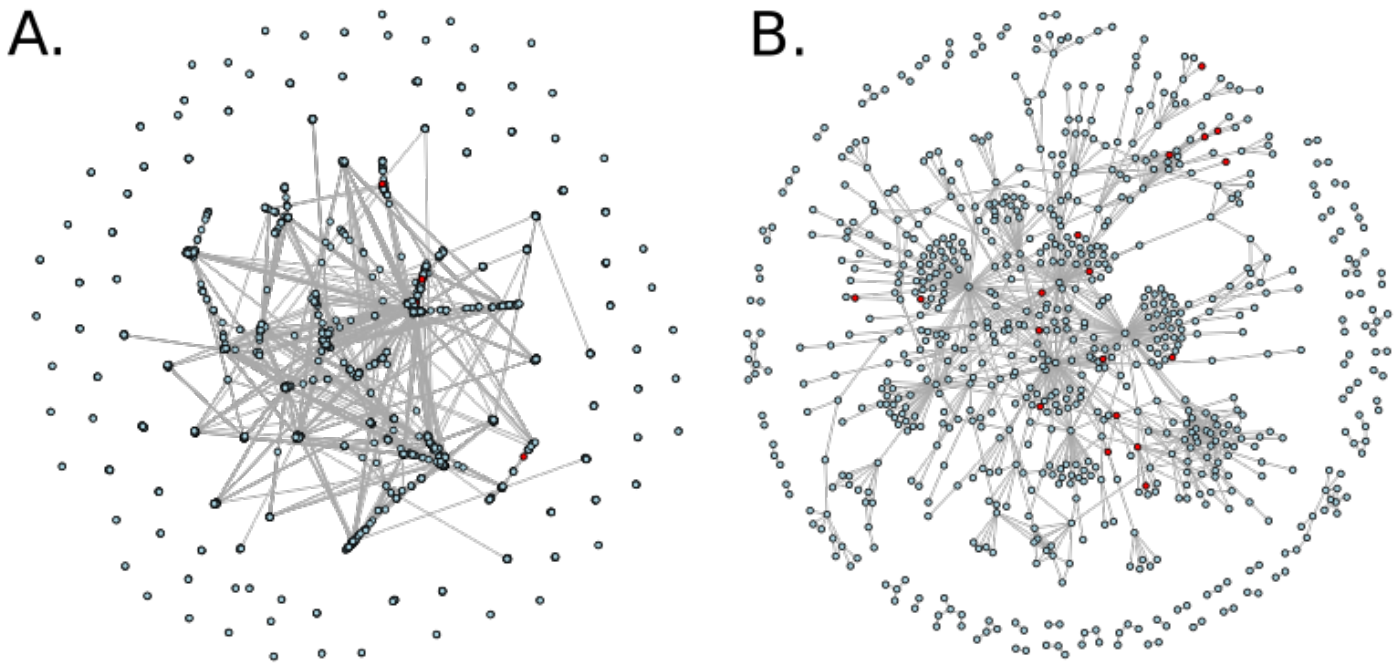

Figure 2: **A.** Graphical representations of metabolic interactions via the co-functional gene pair network of *E. coli* (EcoliNet: EcoCyc/GO-BP). This network contains 1835 nodes with an average path length of 4.8 and contains 20 proteins targeted by 36 drugs in our analysis. **B.** Graphical representations of the PPI network of *E. coli*, as determined by small and medium-scale experiments (EcoliNet: LC. Small/medium-scale PPI). This network contains 764 nodes with an average path length of 4.9, with 18 proteins targeted by 27 drugs in our analysis. In **A** and **B**, each node represents a unique protein in *E. coli* (KEGG ID); the red nodes are target proteins identified as participating in DDIs in our analysis.

We then examined the *E. coli* DDI scores as a function of the connectivity of their targets. We used a co-functional gene pair network of *E. coli* derived from EcoCyc and GO-BP data, which contains genes, proteins, pathways, and molecular interactions in *E.coli* and Gene Ontology Biological Processes, respectively (Ashburner et al. 2000; Karp et al. 2018) (**Fig. 2A**). We also used a small/medium scale protein-protein interaction (PPI) network, also from *E. coli* (Kim et al. 2015) (**Fig. 2B**). We measured the length of minimum distance paths between all DDI targets in the networks and found that synergistic interactions have the shortest average distance, followed by additive interactions, and antagonisms. Synergistic interactions have a significantly shorter average path length between protein targets than antagonisms and additive interactions in both networks (**Fig. 3A, B**). This result is consistent with our expectations, given that synergistic interactions tend to be more common between drugs that target the same cellular process and thus should be closer to each other in biological networks (Brochado et al. 2018). We next examined K-edge connectivity, where two nodes are K-edge-connected if after removing k edges or less the nodes remain connected. For the co-functional network, we found that K-edge connectivity was higher in synergies than additivities (Wilcoxon p-value=0.013), and higher in additivities than antagonisms (Wilcoxon p-value=0.045) (**Fig. 3C**)**.** This result was not significant for the PPI network (**Fig. 3D**). We found that node degree (the number of connected edges) was significantly higher in synergies than additivities (Wilcoxon p-value=0.0025) or antagonistic interactions (Wilcoxon p-value=0.00018) in the co-functional network (**Fig. 3E**), but not in the PPI network (**Fig. 3F**). We also found that average betweenness centrality was higher for synergies than for additivities (Wilcoxon p-value=0.0015) and antagonisms (Wilcoxon p-value=0.013) in the co-functional network, but not in the PPI network.

Average eigenvector centrality was also higher in synergies than additivities (Wilcoxon p-value=0.0002) and antagonisms (Wilcoxon p-value=0.0005) in the co-functional network, but these differences were not significant in the PPI network.

Synergistic interactions evolve faster than additive and antagonistic interactions.

Cluster 14, which had the highest evolutionary rate (σ^2=0.00784), appears to be rich in synergies in *E. coli* and *Salmonella,* while *Pseudomonas* has more additive DDIs in this cluster (**Fig. 1B).** In contrast, cluster 15, which was the second highest rate cluster (σ^2=0.00367), appears to be rich in synergies in *Pseudomonas* while the DDIs of *E. coli* and *Salmonella* are primarily additive. The third cluster with the highest rate, cluster 1 (σ^2=0.00354), contains highly antagonistic DDIs in all the species. These observations suggest that the evolutionary rate of DDIs may vary as a function of the biological relatedness (i.e. interaction type and network connectivity) of their targets. To more formally investigate this, we used the evolutionary rates determined for DDIs on a per-cluster basis (via tSNE, described above) and examined how the rates vary based on the targets' connectivity (in *E. coli*) and type of interaction. None of the DDIs were exclusively antagonistic across all strains and species, although we found all other combinations (i.e., additivity, synergy, additivity↔antagonism, additivity↔synergy, antagonism↔synergy and additivity↔antagonism↔synergy, where the arrow indicates that some strains or species have one interaction type and other strains or species have another interaction type).

Overall, synergistic and additive↔synergistic interactions have faster evolutionary rates than any other interaction combination (**Fig. 4A**; all possible group comparisons were significantly different from each other with p-values<0.001 except additivity↔antagonism↔synergy vs. additivity↔antagonism, and excluding antagonism↔synergy, with only a single observation; **Sup. Table S2**). We also scored each target pair's interaction type as a simple sum of its interaction types across strains and species (+1 for synergies, -1 for antagonisms), and found again that more synergistic interactions (sum ≤ -3) have faster evolutionary rates than additivities (3 > sum > -3) (p-value<0.001) and antagonisms (sum ≥ 3) (p-value<0.001) (**Fig. 4B**). This result of higher rates in synergistic interactions is consistent with the result for interaction types from a single strain, *E. coli* K-12 EBW2952 (**Fig. 4C**; all categories are different at p<0.001, except for additivities vs antagonisms in the PPI network**)**.

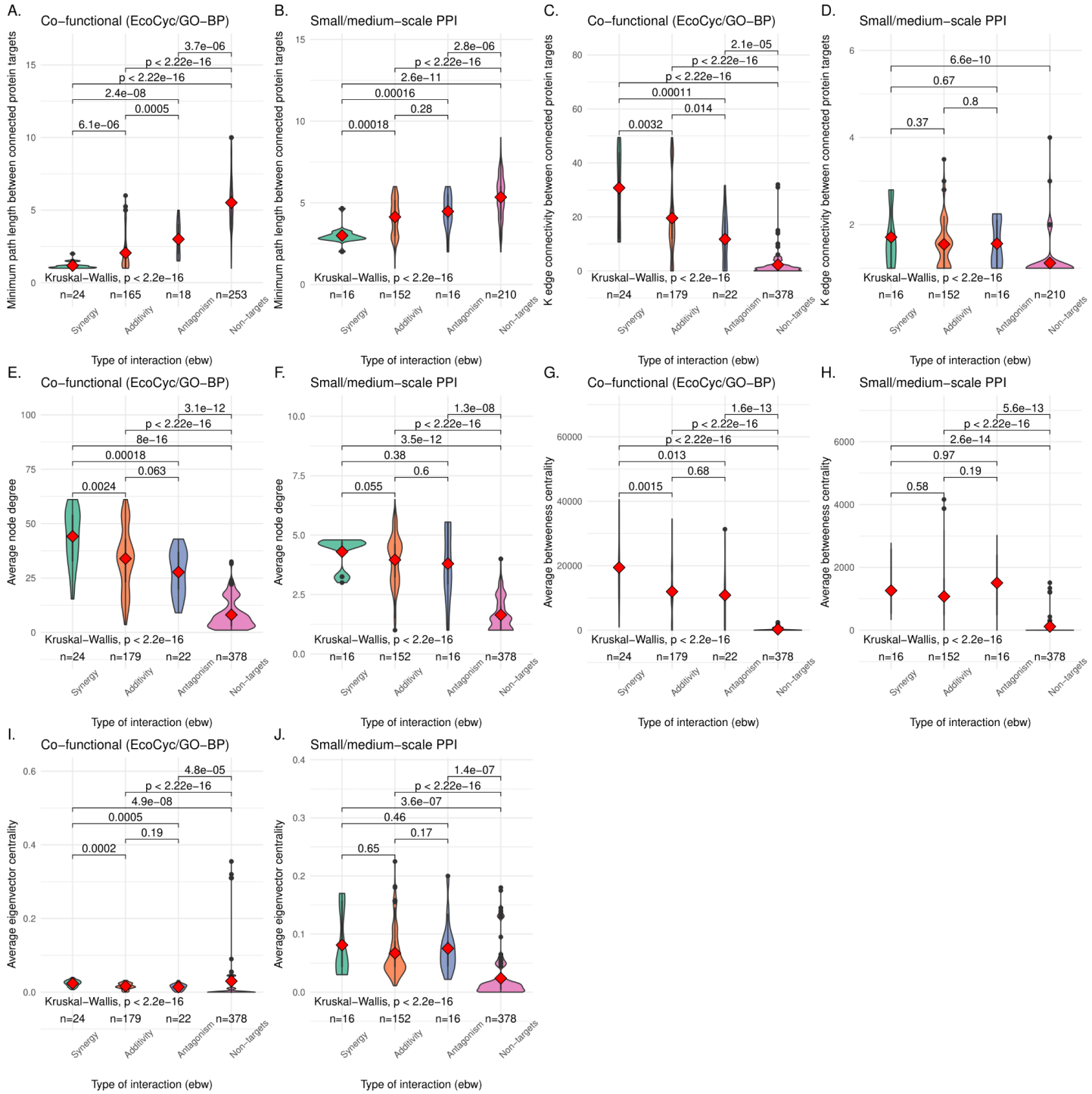

Figure 3: Differences in minimum path length, k edge connectivity, mean node degree, mean betweenness centrality, and mean eigenvector centrality between targets for different types of drug interactions in *E. coli* for co-functional and PPI networks. **A-B**. Synergistic interactions have a lower average path length than additive interactions and antagonistic interactions. **C-D.** K-edge connectivity is higher in synergies than in additive and antagonistic interactions in the co-functional network, but not in the PPI network. **E-F.** The mean node degree is higher in synergies than in additive or antagonistic interactions in the co-functional network, but not in the PPI network. **G-H**. Mean betweenness centrality is higher in synergies than in additive or antagonistic interactions in the co-functional network, but not in the PPI network. **I-J**. Mean eigenvector centrality is higher in synergies than in additive or antagonistic interactions in the co-functional network, but not in the PPI network.

In concert with this, DDIs whose targets are nearby in the network have the highest evolutionary rates (this is consistent with the fact that cluster 14 has the highest rate, and contains a majority of synergies; **Fig. 4D**). This result is significant for both the PPI network (p-value=0.008) and the co-functional (EcoCyc/GO-BP) network (p-value=0.022) (**Sup. Table S2**). We interpret this to mean that wiring between close nodes in biological networks evolves faster than between distant nodes. The relationship between evolutionary rates and other network connectivity measures such as K-edge connectivity appears to be less obvious: evolutionary rates increase sharply as a function of k edge connectivity in the co-functional (EcoCyc/GO-BP) network when connectivity is greater than 20, however, this pattern doesn't hold for the PPI network, where all the protein targets have connectivities lower than 5 (**Sup. Fig. S1K**). Evolutionary rates drop sharply and increase for node degrees greater than 20 in the co-functional (EcoCyc/GO-BP) network, but this result is not seen in the PPI network where the maximum node degree is 10 (**Sup. Fig. S1L**). Rates do not increase or decrease as a function of betweenness centrality (**Sup. Fig. S1M**), however, we see a sharp decrease in rates for interactions with higher average eigenvector centralities in the PPI network.

We also evaluated adjacency (whether two nodes are directly linked) and connectedness (whether two nodes are connected through a path). We found that most target interactions are not adjacent, however, most adjacent nodes have higher evolutionary rates than non-adjacent nodes, both in the co-functional (EcoCyc) (Wilcoxon p-value=0.00025; **Sup. Fig. S1A**) and PPI networks (Wilcoxon p-value=0.0047; **Sup. Fig. S1B**; note that this result was not significant in the EcoCyc/GO-BP network; Wilcoxon p-value=0.2). We also examined whether interaction type changes across species are constrained by different network parameters. We found that some interaction transitions, such as additive interaction to synergistic interaction, tend to have greater connectivity, degree, and betweenness centrality values (**Sup. Fig. S1DEFG**). In addition, connected targets had higher evolutionary rates than disconnected targets in the co-functional (Ecocyc) network (Wilcoxon p-value=0.00049; **Sup. Fig. S1C**), but not in the (EcoCyc/GO-BP) network (Wilcoxon p-value=0.061). At the same time, the co-functional (EcoCyc/GO-BP) network had more connected targets in antagonisms (18 connected, 4 disconnected) and additive interactions (165 connected, 14 disconnected) than disconnected targets, while all 24 synergistic interactions were connected (**Sup. Fig. S1I**). In contrast all evaluated targets were connected in the PPI network (**Sup. Fig. S1J**).

Discussion

We studied the evolution of drug-drug interactions (DDIs) as a proxy for studying the evolution of network topology. We found that the evolutionary rates of DDIs among gram-negative bacteria are correlated with phylogenetic distance, but that they plateau at the largest distances, suggestive of neutral processes. We then mapped drugs with known targets to different biological networks in Escherichia coli K-12 and found that the targets of synergistic DDIs are closer in biological networks than other types of DDIs, and that synergistic interactions have a higher evolutionary rate, meaning that nodes that are closer in biological networks evolve at a faster rate.

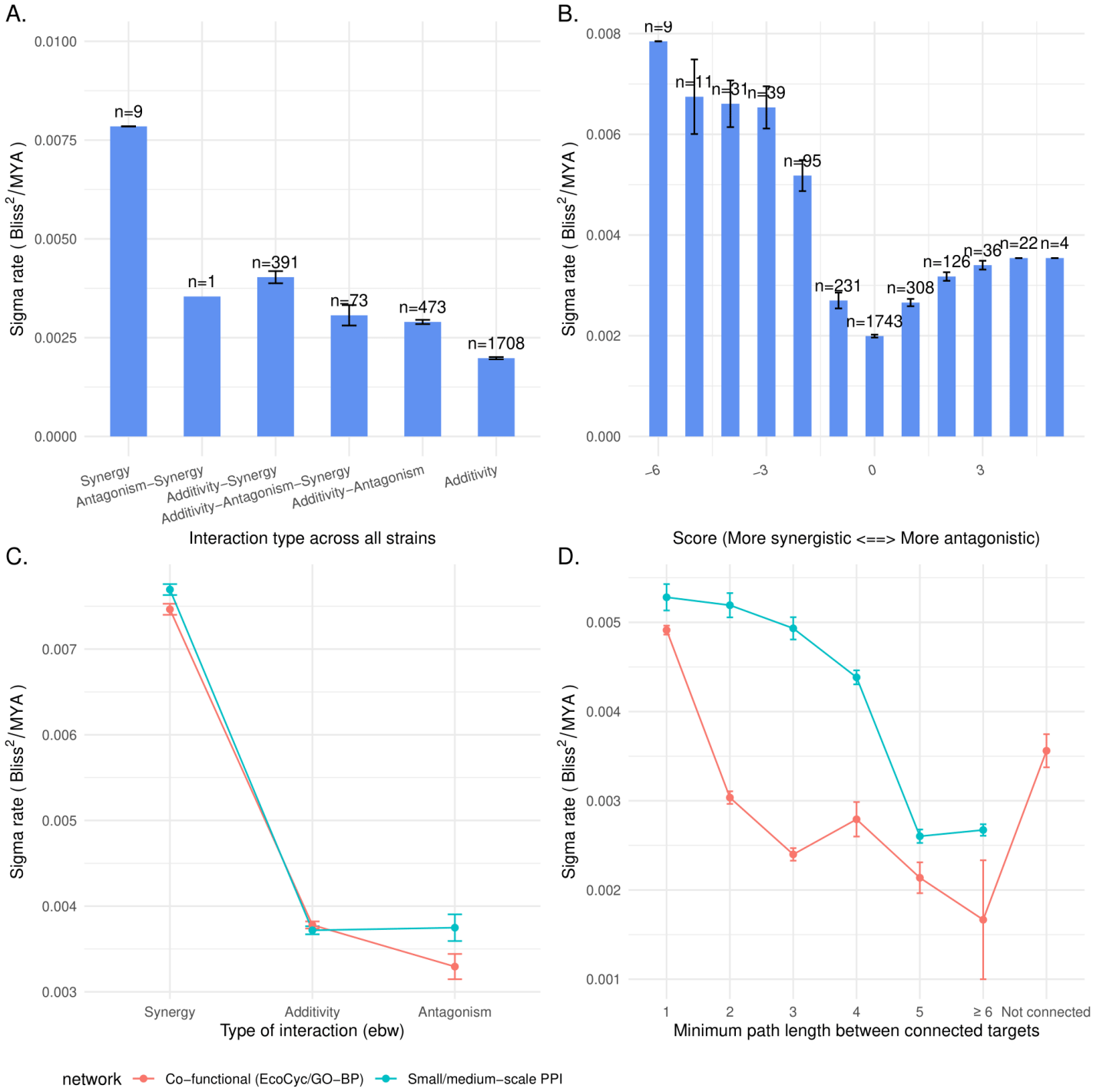

Figure 4: Evolutionary rates of drug-drug interactions vary by interaction type and path length. **A.** DDIs containing only synergies or synergy-antagonism combinations across all strains have faster evolutionary rates. Error bars represent the standard error of the mean. **B.** DDIs that are more synergistic across all strains have higher evolutionary rates. The score is the aggregated sum across strains, where DDIs were scored as +1 for synergies, 0 for additive, and -1 for antagonistic interactions. Error bars represent the standard error of the mean. **C.** Evolutionary rate as a function of interaction type for the co-functional (pink) and small/medium scale PPI network (blue). Synergistic interactions have higher evolutionary rates than additive and antagonistic interactions in *E. coli*. **D.** Evolutionary rate as a function of the minimum distance between targets. Wiring between close nodes in biological networks evolves more quickly than between distant nodes. See **Sup. Table S2** for statistical tests for differences among groups shown here.

339 Previous studies on interspecific comparisons of DDIs focused on the identification of novel
340 combinations of drugs that can be used to treat different bacterial (Brochado et al. 2018) and fungal
341 infections (Spitzer et al. 2011; Robbins et al. 2015). These pharmacological studies provide a static
342 view of DDIs in species of interest, not assessing the factors that cause interspecific divergence, and
343 how these factors are related to differences in interactions among drug targets across species (e.g.
344 via divergence in metabolic or signaling pathways). These studies suggested that chemogenomic
345 variation might rapidly accumulate; for example, in fungi very few combinations of antifungals act
346 across species (Brown et al. 2014). Similarly, the species of gram-negative bacteria in Brochado *et*
347 *al.* (2018) showed that 70% of DDIs detected across six strains were species-specific, and 20% were
348 strain specific.
349

350 Here we use this chemogenomic data to evaluate hypotheses of network evolution. Examining
351 pairwise distances of DDIs between species, we found that DDIs diverge more rapidly than protein
352 sequence between species (**Fig. 1D).** We observed a rapid accumulation of differences and overall
353 dissimilarity in DDI responses among species. Indeed the chemical distances curve appears
354 saturated among distantly related species; just as network rewiring rates slow down at this scale,
355 according to Shou *et al.* (2011).

356 We also clustered DDIs by their similarity across strains and species and then evaluated the rate of
357 evolution of these clusters under a Brownian motion model. This approach is analogous to methods
358 used in geometric morphometrics, in which specimen landmarks are grouped to evaluate different
359 features of an organ such as the human brain, and modularity tests are used to assess whether there
360 are differences in the evolution rate of different brain regions (Smaers and Vanier 2019). Modularity
361 among traits mirrors shared functional, genetic, and developmental pathways (Cheverud 1996);
362 observing high modularity among clusters indicates differences among groups of DDIs due to the
363 variance between clusters being greater than the variance within clusters. We observed that the
364 distribution of synergies, additivities and antagonisms among clusters partially explains different
365 rate estimates. Clusters with low rates are composed of mostly additive DDIs; however some of these
366 low-rate estimates are caused by resistance to one or both antibiotics, and a low phylogenetic signal
367 is expected when multiple strains are resistant (Michel et al. 2008). Other clusters with high rates
368 (e.g., clusters 14, 15, and 1) have some highly synergistic or antagonistic interactions in some strains
369 and additivities in other strains. With these rate estimates and interaction patterns, we asked
370 whether network-based metrics are correlated with DDI divergence, by linking a subset of individual
371 DDIs with their respective protein targets in *E. coli* co-functional (EcoCyc/GO-BP) and PPI
372 networks.

373 We first made several observations about the *E. coli* DDI types and network: synergies occur when
374 the targets are closer in the network, while additive interactions and antagonisms target on average
375 more distant proteins (**Fig. 3A, B**). This result is not surprising given the fact that synergies are
376 expected to be more prevalent in drug combinations that target the same pathway or that belong to
377 the same chemical class (Cowen and Steinbach 2008; Yeh et al. 2009), however, we think this is the
378 first time this pattern has been shown using network data. We also found that synergistic
379 interactions are more prevalent among targets that are part of the same cellular process or function,
380 while antagonisms can take place between targets that are in different parts of the cell (Yeh et al.
381 2006; Wang et al. 2012; Brochado et al. 2018). We also found targets involved in synergies have a
382 higher node degree (for metabolic networks), and K-edge connectivity than targets involved in

383 additive or antagonistic interactions (**Fig. 3C-F**). This result agrees with a previous study that
384 reports on the positive contribution of node degree in synergy prediction (Liu et al. 2022). Average
385 node betweenness and eigenvector centralities were also found to be more significant among
386 synergistic DDIs compared to additivities and antagonisms in the metabolic network (**Fig. 3G-J**).
387 This result suggests that targets affected by synergistic interactions tend to be "hubs" in the
388 network (Li et al. 2011; Tilli et al. 2016), or that they are more central to different pathways than
389 antagonistic or additive interactions.
390
391 We examined the rate of evolution of DDIs whose targets are closer in the PPI and metabolic
392 networks and found that their rate of evolution is higher. This is correlated with our observation
393 that synergies tend to occur closer in the PPI and metabolic networks and have a high evolutionary
394 rate (accounting for the two DDI clusters with the highest rate). Antagonistic interactions are
395 overrepresented in the cluster with the third highest rate of evolution. Correspondingly, as the path
396 length between targets decreases, the evolutionary rate of the DDI increases (**Fig. 4D**) (although
397 completely disconnected proteins do not follow this trend). This is consistent with the positive
398 correlation observed between path length distances between protein interaction and gene
399 interaction/congruence networks in yeast (Ye et al. 2005; VanderSluis et al. 2018), because the gene
400 interaction and DDI networks should have a similar structure (Wang et al. 2012).
401
402 What do these observations mean about the overall pattern of network evolution? A possible result
403 of why synergies have a higher rate of evolution is that the underlying network is also evolving at a
404 faster rate resulting in big differences between species, meaning that a network effect can result in a
405 higher rate of DDI evolution. This may mean that functional modules of the network, such as
406 biological pathways for which synergistic DDI act, undergo a faster rate of rewiring than parts of the
407 network that are more distantly connected (additive and antagonistic interactions). Different
408 processes may be driving this pattern, for instance, positive selection and gain of function mutations
409 may result in novel network combinations within an already existing pathway, that wouldn't be
410 possible across proteins involved in different processes, expressed at different times or found in
411 different cellular components. At the same time, purifying selection may constrain the rewiring of
412 the network in co-functional modules, since connection losses would result in a deleterious effect on
413 the metabolic pathway, this same situation may not be as deleterious if the targets are far apart in
414 the network and they have very few interactions among each other. However, next, we highlight a
415 different type of bias in DDI data that may contribute to our observed signal (high evolutionary rate
416 of synergistic, well-connected targets in *E. coli*) and cause difficulty in the interpretation of DDI
417 data.
418
419 ***Structural vs. network evolution:*** Another important explanation for our observed pattern (high
420 evolutionary rate of synergistic, well-connected targets) lies in the fact that some strains are
421 resistant to particular antibiotics. We suspect that some changes in DDI that we correlated with
422 changes in network structure are actually caused by changes in resistance, rather than changes in
423 the DDI. Changes in resistance are likely based on structural changes in target proteins,
424 transporters, or proteins, and don't represent the types of biological network changes that we hoped
425 to measure with the DDI. Note that Brochado *et al.* (2018) reported any strain with fitness values
426 (e.g., ratio of drug/non-drug treated growth) of >0.7 as being resistant to that drug. However, some
427 drugs that are known to be clinically bacteriostatic (inhibits growth) fall in this range, and strains
428 are reported to be resistant; e.g. see the case of SMM resistance in all species and strains, below.

429 Indeed, Bushby (1973) reported "resistance to one of the drugs as measured by conventional tests
430 may not abolish synergy." To consider the effect of structural differences on our results, we,
431 therefore, repeated the entire analysis filtering for DDIs that are "susceptible" (fitness < 0.7) in at
432 least one of the species for both drugs, assuming that this decreases the proportion of true resistance
433 in the dataset. This set included 1127/2655 DDIs (42.2% of the total DDIs). Our results for this
434 smaller set are consistent with the full dataset (**Sup. Figs. S2-S4**). A minor difference in the subset
435 is that antagonisms have a slightly lower (instead of equivalent) rate than additivities. Although the
436 consistency between the full and reduced dataset is desirable, it doesn't remove the possibility that
437 differences in resistance among strains are a substantial part of the evolutionary signal.
438
439 **Network and Evolutionary Insights into Drug-drug interactions**
440 We used DDIs to learn about the evolution of networks, so discussion of some of the individual DDIs
441 may illustrate limitations of the method. The evolutionary and network parameters also provide
442 some insight into individual DDIs with clinical relevance.
443
444 ***A22-NOV***
445 For example, A22 and novobiocin (A22-NOV) (**Sup. Table S3**): A22 binds the ATP-binding domain
446 of MreB (eco:b3251) (Bean et al. 2009), the actin homolog in prokaryotes, while novobiocin binds
447 gyrase subunit B (gyrB, eco:b3699) and topoisomerase IV (parC, eco:b3019), which are two type II
448 topoisomerases involved in DNA unwinding and DNA duplication (Gellert et al. 1976; Hardy and
449 Cozzarelli 2003). Both targets, MreB and the ParC subunit of topoisomerase IV are known to
450 physically interact (Madabhushi and Marians 2009) and are required for chromosome segregation
451 (Huang et al. 1998). A22 and novobiocin have previously been reported to be synergistic in *E. coli*
452 and *P. aeruginosa (Taylor 2011; Taylor et al. 2012)*; however, in the dataset from Brochado *et al.*
453 *(2018)*, (**Sup. Table S3**), the interaction type is variable within species. Both *E. coli* strains are
454 susceptible to both drugs, but they have different interaction terms, one is additive, the other
455 synergistic. *Salmonella* also shows intraspecific differences, where one of the strains is synergistic
456 and the other additive, although in this case, the synergistic strain is resistant to novobiocin, with a
457 fitness of 0.85. Both *Pseudomonas* strains are resistant to novobiocin, with fitnesses greater than
458 0.86, and they have additive interactions. This high variability of interaction type within species
459 resulted in an evolutionary rate of 0.008 bliss^2/MYA for this DDI, within the cluster with the highest
460 evolutionary rate. The distance between these targets in the co-functional (EcoCyc/GO-BP) network
461 was 2, and in the DDI network the distance between targets was 4. So why are the rates so high
462 among strains, for this DDI? While not directly explanatory, strain-specific network predictions at
463 string-DB (not used in this study) also suggest strain-specific differences in local network content.
464 Of the genes all connected to each other in the local network (mreB, gyrA, parC, gyrB), parC is
465 missing in three strains, while a fourth strain has reduced connectivity for MreB (only connected to
466 gyrB). This type of gene content variation suggests the type of intraspecific variation that may lead
467 to high rates of network evolution.
468
469 ***CIP-MER***
470 A DDI with targets in the same local network is between ciprofloxacin (CIP) and meropenem (MER).
471 CIP targets DNA gyrase subunit A (gyrA, eco:b2231), and MER targets D-alanyl-D-alanine
472 carboxypeptidase DacC (dacC, eco:b0839). The mechanism of action of this interaction is very
473 similar to A22-NOV. Both meropenem and A22 affect the bacterial cell wall, and ciprofloxacin and
474 novobiocin inhibit topoisomerases and gyrases, however, the interaction between meropenem and

ciprofloxacin is additive in all species studied in the dataset, and all strains were deemed susceptible to the drug in this dataset (however, synergistic effects have been reported in *Pseudomonas (Ermertcan et al. 2001)*. For this interaction, we calculated an evolutionary rate of 0.002 bliss^2/MYA, part of the cluster that is in the middle of the other clusters when sorted by rate (9/16). This rate is 4 times lower than that of SMM-TMP (below) and A22-NOV, which makes sense given that the interaction is additive in all the species of the set, thus having a low trait value across species. The distance between targets in the co-functional (EcoCyc/GO-BP) network was 1.5, while in the PPI network, the distance between targets was 5.5.

SMM-TMP

Another illustrative example of a known DDI used in the clinic found in our data is the combination of sulfamonomethoxine and trimethoprim (SMM-TMP) (**Sup. Table S1**) (Bushby 1973). These drugs inhibit successive steps in the synthesis of tetrahydrofolic acid synthesis, necessary for the biosynthesis of amino acids, purines, and thymidine. Sulfamonomethoxine competes against the enzymatic substrate of dihydropteroate synthase (folP, eco:b3177); this blocks the production of dihydrofolic acid, which in turn is a substrate for dihydrofolate reductase. Trimethoprim also inhibits dihydrofolate reductase (folA, eco:b0048) by competing against the substrate for the binding site of the enzyme. Even on its own, trimethoprim can contribute to "thymineless death" (Then and Angehrn 1973); but while the effect of each drug on its own is bacteriostatic (prevents growth; Ocampo et al. 2014), when used in combination, the combined effect is bactericidal (kills bacteria), thus suggesting a strong synergistic effect between the two drugs (Bushby 1973; Ocampo et al. 2014).

We measured an evolutionary rate of 0.008 bliss^2/MYA for this DDI, as it belongs to the cluster with the highest evolutionary rate. This rate can be explained by the difference in interaction type between *Pseudomonas* and non-*Pseudomonas* strains. In the data used in our study, the interaction was synergistic for *E. coli* and *S. typhimurium*, but the interaction was additive in *Pseudomonas*. The DDI may be additive in *Pseudomonas* because it is resistant to both drugs due in part to both differences in its permeability to the drugs and efflux pumps that remove the drugs (Eliopoulos and Huovinen 2001). *E. coli* and *S. typhimurium* are only reported by Brochado *et al. (2018)* as resistant to sulfamonomethoxine, having a fitness between 77-88% of non drug-treated.

Our goal was to capture the rate of evolution associated with differences in the folate biosynthesis pathway across species. The fact that the two targets act as successive steps in all studied species suggests several possible causes for this observed rate. There are some differences in the presence of pathways components that are peripheral to these steps (Pribat et al. 2010), and it is possible that such differences in the network beyond the direct connection contribute to DDI differences. Indeed, although directly linked in the co-functional (EcoCyc/GO-BP) network, the targets are highly connected to each other, with a K-edge connectivity of 21 (68% quantile), meaning that 21 edges have to be removed to disconnect the two targets. It is also possible that there are differences in the direct target enzymes and their connectivity (e.g., concentration, kmax). An alternative explanation is resistance: resistance to sulfamonomethoxine and trimethoprim was detected soon after they were introduced in clinical treatments, and several point mutations have been described in the target genes across different bacteria, and the most common mechanism of resistance is due to the bacteria having an extra copy of the target gene (Estrada et al. 2016). All of this suggests that, for SMM-TMP, evolution of the DDI is the result of many types of biological differences, not just specific

521 differences in network interactions between the drug targets. This is an important limitation to the
522 DDI approach to studying network evolution.
523

524 **Conclusions**

525 Previous research on network evolution has focused primarily on the study of individual nodes,
526 either by comparing rewiring rates across networks or by comparing the role of sequence
527 conservation to the number of connections per node. Here, we introduce a novel framework to
528 compare evolutionary rates across nodes using network measurements such as minimum distance,
529 K-edge connectivity, and node degree between proteins. To do so, we examine changes in DDIs terms
530 among strains and species, and compare these rates of change with the interaction types and
531 network topology of the drug's targets (Lehár et al. 2007). This approach has some advantages over
532 direct network analyses, such as increasing the number of species under study, without the need of
533 obtaining the underlying network in each species. DDIs can be quantified using high-throughput
534 experiments across different species and strains (Brochado et al. 2018), and are phenotypic
535 quantitative traits that can be modeled in a phylogenetic comparative framework, allowing for an
536 independent measurement of evolutionary rates between nodes in the network. An important
537 limitation of our approach is that evolution of resistance to a drug may obscure underlying network-
538 based effects. Still, our approach suggests a general picture of network evolution, where close nodes
539 in the network (which tend to respond synergistically in response to targeting both at once) evolve at
540 faster rates than more distant nodes.

541

542 Materials and Methods

543 We obtained DDI scores (Bliss scores) from a previous study (Brochado et al. 2018) that assessed
544 almost 3000 combinations of 79 different compounds on six strains of three species of gram-negative
545 bacteria: *E. coli* K-12 BW2952 (ebw), *Escherichia coli* O8 IAI1 (ecr), *S. enterica subsp. enterica* serovar
546 Typhimurium 14028S (seo), *Salmonella enterica* subsp. enterica serovar Typhimurium LT2 (stm),
547 *Pseudomonas aeruginosa* PAO1 (pae), *P. aeruginosa* UCBPP-PA14 (pau). We used the following
548 datasets from Brochado (2018): interaction scores from table ED09C as input for our phylogenetic
549 comparative analysis, strain susceptibility to antibiotics from Sup. Table 1 (used for our **Sup. Figs.**
550 **S2-S4**), and the categories that DDI interactions belong to from Sup. Table 2. We built a maximum
551 likelihood phylogenetic tree of the six strains using PhySpeTree (Fang et al. 2019) to obtain
552 concatenated alignments of 35 highly conserved protein sequences (HCP). We use this concatenated
553 alignment of HCP sequences to construct a bayesian phylogenetic tree in BEAST2 (Bouckaert et al.
554 2014) using a birth-death tree with a strict molecular clock, and calibration nodes from the literature
555 to estimate divergence times between species (Battistuzzi et al. 2004).
556

557 Clustering DDI using tSNE.

558 To reduce data dimensionality of the trait space (i.e., DDI data), we classified DDIs into clusters
559 using tSNE in the R packages bigMap (Garriga and Bartumeus 2018) and bigmemory (Kane et al.
560 2013), with parameters: 80 threads, 80 layers, and 9 rounds. Dimensionality reduction was
561 performed to increase the robustness of rate estimators given the small number of species used.
562 DDIs were clustered by similarity across strains, treating each strain as a dimension, thus DDIs

563 that are more similar to each other across strains would cluster together under tSNE in the same
564 cluster. Our tSNE analyses were performed for a range of 251 perplexity values between 5 and 2505
565 (**Sup. Doc. S1**). The clustering output was visualized and evaluated based on the stability and
566 plateauing of cost and effect size curves, and the variance between threads. We obtained stable
567 solutions between 175 and 1455 perplexity values. The pakde algorithm was applied with a
568 perplexity of 1/3 the respective tSNE perplexity. To find which of the clustering solutions was the
569 most modular, we tested for modularity in the data using the function phylo.modularity within the
570 R package geomorph (Adams et al. 2016), the five most modular clustering patterns were selected
571 due to their most negative ZCR coefficient, which indicates a higher degree of modularity. These
572 clustering patterns had the following perplexity values: 355, 195, 255, 245, and 335; and the
573 following number of clusters: 9, 15, 16, 10, and 14. All five clustering patterns had a strong modular
574 signal, with multivariate effect sizes under -26.4, p-value=0.001 and covariance ratios below 0.92.
575 Lastly, we fitted a multivariate Brownian motion model to each of these 5 clustering patterns and
576 calculated the evolutionary rates per cluster using compare.multi.evol.rates, also using the R
577 package geomorph (Adams et al. 2016). Out of the 5 most modular clustering models, the model with
578 a perplexity of 245 was the one with a higher Z effect in the test, showing the biggest differences
579 between groups. Thus, this clustering pattern was used in the following steps. In addition, the
580 sigma rates calculated for each of the clusters were used as approximations for DDI evolutionary
581 rates that are part of that cluster.
582

583 Identification of drug targets and their biological networks.

584 Each drug was identified with unique Pubchem and CHEMBL IDs using webchem (Szöcs et al.
585 2020), which were then used to retrieve their mode of action from IUPHAR (Armstrong et al. 2019).
586 We also compared our targets to a previously published dataset on drugs and drug targets (Santos et
587 al. 2017), identified unique Uniprot IDs and KO IDs for each target protein, and converted these IDs
588 into *E. coli* Uniprot IDs using the KEGG Orthology (Kanehisa and Goto 2000). We didn't include in
589 the analysis drugs whose mode of action was unknown, or that had non-protein molecules as
590 targets, such as small molecules, RNA, or DNA.

591 Two biological networks of *E. coli* were downloaded from EcoliNet (Kim et al. 2015), small/medium-
592 scale protein-protein interactions (LC) and the co-functional gene pair network of E. coli derived
593 from EcoCyc and GO-BP.
594

595 Calculating inter-node network metrics.

596 For each of the networks, we calculated the average path length and node degree distributions. In
597 addition, the minimum distance between each of the nodes in the network was calculated, as well
598 as the node degree (number of connections per node), the K-edge connectivity between each pair of
599 nodes (i.e. the minimum number of edges that can be removed to disconnect the nodes),
600 betweenness centrality (i.e. a measure of centrality in the network based on shortest paths) and
601 eigenvector centrality (i.e. a measure of the influence of the node in the network). We used the R
602 package igraph (Csardi et al. 2006) to calculate these values in each one of the biological networks
603 and for each node or pair of nodes. We also generated an adjacency matrix, which contains
604 information on whether two nodes are connected directly by an edge or not. In addition, we used K-
605 edge connectivity as a proxy for connectedness between proteins (i.e., a pair with K-edge

connectivity equal to zero is disconnected, and proteins with K-edge connectivity different than zero are connected). Connectedness is used due to the limitations of adjacency in describing the overall connection between nodes.

Data and code availability. The code used for data analysis is available from: https://github.com/phylogenomic/ch3-netbio

References

Adams DC, Collyer M, Kaliontzopoulou A, Sherratt E. 2016. geomorph: Software for geometric morphometric analyses. Available from: https://rune.une.edu.au/web/handle/1959.11/21330

Armstrong JF, Faccenda E, Harding SD, Pawson AJ, Southan C, Sharman JL, Campo B, Cavanagh DR, Alexander SPH, Davenport AP, et al. 2019. The IUPHAR/BPS Guide to PHARMACOLOGY in 2020: extending immunopharmacology content and introducing the IUPHAR/MMV Guide to MALARIA PHARMACOLOGY. *Nucleic Acids Res.*:gkz951.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. 2000. Gene Ontology: tool for the unification of biology. *Nat. Genet.* 25:25–29.

Battistuzzi FU, Feijao A, Hedges SB. 2004. A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol. Biol.* 4:44.

Bean GJ, Flickinger ST, Westler WM, McCully ME, Sept D, Weibel DB, Amann KJ. 2009. A22 Disrupts the Bacterial Actin Cytoskeleton by Directly Binding and Inducing a Low-Affinity State in MreB. *Biochemistry* 48:4852–4857.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 10:e1003537.

Brochado AR, Telzerow A, Bobonis J, Banzhaf M, Mateus A, Selkrig J, Huth E, Bassler S, Zamarreño Beas J, Zietek M, et al. 2018. Species-specific activity of antibacterial drug combinations. *Nature* 559:259–263.

Brown JCS, Nelson J, VanderSluis B, Deshpande R, Butts A, Kagan S, Polacheck I, Krysan DJ, Myers CL, Madhani HD. 2014. Unraveling the Biology of a Fungal Meningitis Pathogen Using Chemical Genetics. *Cell* 159:1168–1187.

Bushby SR. 1973. Trimethoprim-sulfamethoxazole: in vitro microbiological aspects. *J. Infect. Dis.* 128:Suppl:442–462 p.

Cheverud JM. 1996. Developmental integration and the evolution of pleiotropy. *Am. Zool.* 36:44–50.

Cork JM, Purugganan MD. 2004. The evolution of molecular genetic pathways and networks. *Bioessays* 26:479–484.

Cormen TH. 2001. Single-source shortest paths. *Introduction to Algorithms* [Internet]. Available from: https://cir.nii.ac.jp/crid/1572543025240409984

Cowen LE, Steinbach WJ. 2008. Stress, drugs, and evolution: the role of cellular signaling in fungal drug resistance. *Eukaryot. Cell* 7:747–764.

Csardi G, Nepusz T, Others. 2006. The igraph software package for complex network research. *InterJournal, complex systems* 1695:1–9.

Cusick ME, Klitgord N, Vidal M, Hill DE. 2005. Interactome: gateway into systems biology. *Hum. Mol. Genet.* 14 Spec No. 2:R171–R181.

Eliopoulos GM, Huovinen P. 2001. Resistance to Trimethoprim-Sulfamethoxazole. *Clin. Infect. Dis.* 32:1608–1614.

649 Erdős P, Gallai T. 1960. Gráfok előírt fokszámú pontokkal. *Mat. Lapok* .

650 Ermertcan S, Hoşgör M, Tünger O, Coşar G. 2001. Investigation of synergism of meropenem and
651 ciprofloxacin against Pseudomonas aeruginosa and Acinetobacter strains isolated from intensive
652 care unit infections. *Scand. J. Infect. Dis.* 33:818–821.

653 Estrada A, Wright DL, Anderson AC. 2016. Antibacterial Antifolates: From Development through
654 Resistance to the Next Generation. *Cold Spring Harb. Perspect. Med.* [Internet] 6. Available
655 from: http://dx.doi.org/10.1101/cshperspect.a028324

656 Fang Y, Liu C, Lin J, Li X, Alavian KN, Yang Y, Niu Y. 2019. PhySpeTree: an automated pipeline for
657 reconstructing phylogenetic species trees. *BMC Evol. Biol.* 19:219.

658 Fraser HB, Hirsh AE, Steinmetz LM, Scharfe C, Feldman MW. 2002. Evolutionary rate in the protein
659 interaction network. *Science* 296:750–752.

660 Garriga J, Bartumeus F. 2018. bigMap: Big Data Mapping with Parallelized t-SNE. *arXiv [cs.LG]*
661 [Internet]. Available from: http://arxiv.org/abs/1812.09869

662 Gellert M, O'Dea MH, Itoh T, Tomizawa J. 1976. Novobiocin and coumermycin inhibit DNA
663 supercoiling catalyzed by DNA gyrase. *Proc. Natl. Acad. Sci. U. S. A.* 73:4474–4478.

664 Ghadie MA, Coulombe-Huntington J, Xia Y. 2018. Interactome evolution: insights from genome-wide
665 analyses of protein–protein interactions. *Curr. Opin. Struct. Biol.* 50:42–48.

666 Greco WR, Bravo G, Parsons JC. 1995. The search for synergy: a critical review from a response
667 surface perspective. *Pharmacol. Rev.* 47:331–385.

668 Hardy CD, Cozzarelli NR. 2003. Alteration of Escherichia coli topoisomerase IV to novobiocin
669 resistance. *Antimicrob. Agents Chemother.* 47:941–947.

670 Huang WM, Libbey JL, van der Hoeven P, Yu SX. 1998. Bipolar localization of Bacillus subtilis
671 topoisomerase IV, an enzyme required for chromosome segregation. *Proc. Natl. Acad. Sci. U. S.*
672 *A.* 95:4652–4657.

673 Jensen RA. 1976. Enzyme Recruitment in Evolution of New Function. *Annu. Rev. Microbiol.* 30:409–
674 425.

675 Jin Y, Turaev D, Weinmaier T, Rattei T, Makse HA. 2013. The evolutionary dynamics of protein-
676 protein interaction networks inferred from the reconstruction of ancient networks. *PLoS One*
677 8:e58134.

678 Jordan C. 1869. Sur les assemblages de lignes. 1869:185–190.

679 Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.*
680 28:27–30.

681 Kane M, Emerson JW, Weston S. 2013. Scalable Strategies for Computing with Massive Data. *J.*
682 *Stat. Softw.* 55:1–19.

683 Karp PD, Ong WK, Paley S, Billington R, Caspi R, Fulcher C, Kothari A, Krummenacker M,
684 Latendresse M, Midford PE, et al. 2018. The EcoCyc Database. *EcoSal Plus* [Internet] 8.
685 Available from: http://dx.doi.org/10.1128/ecosalplus.ESP-0006-2018

686 Kim H, Shim JE, Shin J, Lee I. 2015. EcoliNet: a database of cofunctional gene network for
687 Escherichia coli. *Database* [Internet] 2015. Available from:

688 http://dx.doi.org/10.1093/database/bav001

689 Lehár J, Zimmermann GR, Krueger AS, Molnar RA, Ledell JT, Heilbut AM, Short GF 3rd, Giusti LC,
690 Nolan GP, Magid OA, et al. 2007. Chemical combination effects predict connectivity in biological
691 systems. *Mol. Syst. Biol.* 3:80.

692 Li S, Zhang B, Zhang N. 2011. Network target for screening synergistic drug combinations with
693 application to traditional Chinese medicine. *BMC Syst. Biol.* 5 Suppl 1:S10.

694 Liu X, Song C, Liu S, Li M, Zhou X, Zhang W. 2022. Multi-way relation-enhanced hypergraph
695 representation learning for anti-cancer drug synergy prediction. *Bioinformatics* 38:4782–4789.

696 Madabhushi R, Marians KJ. 2009. Actin homolog MreB affects chromosome segregation by regulating
697 topoisomerase IV in Escherichia coli. *Mol. Cell* 33:171–180.

698 Michel J-B, Yeh PJ, Chait R, Moellering RC Jr, Kishony R. 2008. Drug interactions modulate the
699 potential for evolution of resistance. *Proc. Natl. Acad. Sci. U. S. A.* 105:14918–14923.

700 Ocampo PS, Lázár V, Papp B, Arnoldini M, Abel zur Wiesch P, Busa-Fekete R, Fekete G, Pál C,
701 Ackermann M, Bonhoeffer S. 2014. Antagonism between bacteriostatic and bactericidal
702 antibiotics is prevalent. *Antimicrob. Agents Chemother.* 58:4573–4582.

703 Picard F, Daudin J-J, Koskas M, Schbath S, Robin S. 2008. Assessing the exceptionality of network
704 motifs. *J. Comput. Biol.* 15:1–20.

705 Pribat A, Blaby IK, Lara-Núñez A, Gregory JF 3rd, de Crécy-Lagard V, Hanson AD. 2010. FolX and
706 FolM are essential for tetrahydromonapterin synthesis in Escherichia coli and Pseudomonas
707 aeruginosa. *J. Bacteriol.* 192:475–482.

708 Robbins N, Spitzer M, Yu T, Cerone RP, Averette AK, Bahn Y-S, Heitman J, Sheppard DC, Tyers M,
709 Wright GD. 2015. An Antifungal Combination Matrix Identifies a Rich Pool of Adjuvant
710 Molecules that Enhance Drug Activity against Diverse Fungal Pathogens. *Cell Rep.* 13:1481–
711 1492.

712 Shou C, Bhardwaj N, Lam HYK, Yan K-K, Kim PM, Snyder M, Gerstein MB. 2011. Measuring the
713 evolutionary rewiring of biological networks. *PLoS Comput. Biol.* 7:e1001050.

714 Smaers JB, Vanier DR. 2019. Brain size expansion in primates and humans is explained by a
715 selective modular expansion of the cortico-cerebellar system. *Cortex* 118:292–305.

716 Spitzer M, Griffiths E, Blakely KM, Wildenhain J, Ejim L, Rossi L, De Pascale G, Curak J, Brown E,
717 Tyers M, et al. 2011. Cross-species discovery of syncretic drug combinations that potentiate the
718 antifungal fluconazole. *Mol. Syst. Biol.* 7:499–499.

719 Szöcs E, Stirling T, Scott ER, Scharmüller A, Schäfer RB. 2020. webchem: An R Package to Retrieve
720 Chemical Information from the Web. *J. Stat. Softw.* 93:1–17.

721 Taylor P. 2011. Combating intrinsic antibiotic resistance in Gram-negative bacteria. Available from:
722 https://macsphere.mcmaster.ca/handle/11375/11329

723 Taylor PL, Rossi L, De Pascale G, Wright GD. 2012. A forward chemical screen identifies antibiotic
724 adjuvants in Escherichia coli. *ACS Chem. Biol.* 7:1547–1555.

725 Then R, Angehrn P. 1973. Nature of the Bactericidal Action of Sulfonamides and Trimethoprim,
726 Alone and in Combination. *J. Infect. Dis.* 128:S498–S501.

727 Tilli TM, Carels N, Tuszynski JA, Pasdar M. 2016. Validation of a network-based strategy for the
728 optimization of combinatorial target selection in breast cancer therapy: siRNA knockdown of
729 network targets in MDA-MB-231 cells as an in vitro model for inhibition of tumor development.
730 *Oncotarget* 7:63189–63203.

731 VanderSluis B, Costanzo M, Billmann M, Ward HN, Myers CL, Andrews BJ, Boone C. 2018.
732 Integrating genetic and protein–protein interaction networks maps a functional wiring diagram
733 of a cell. *Curr. Opin. Microbiol.* 45:170–179.

734 Wagner A. 2003. How the global structure of protein interaction networks evolves. *Proceedings of the*
735 *Royal Society of London. Series B: Biological Sciences* 270:457–466.

736 Wang Y-Y, Xu K-J, Song J, Zhao X-M. 2012. Exploring drug combinations in genetic interaction
737 network. *BMC Bioinformatics* 13 Suppl 7:S7.

738 Yeh PJ, Hegreness MJ, Aiden AP, Kishony R. 2009. Drug interactions and the evolution of antibiotic
739 resistance. *Nat. Rev. Microbiol.* 7:460–466.

740 Yeh P, Tschumi AI, Kishony R. 2006. Functional classification of drugs by properties of their pairwise
741 interactions. *Nat. Genet.* 38:489–494.

742 Ye P, Peyser BD, Spencer FA, Bader JS. 2005. Commensurate distances and similar motifs in genetic
743 congruence and protein interaction networks in yeast. *BMC Bioinformatics* 6:270.

744 Zotenko E, Mestre J, O'Leary DP, Przytycka TM. 2008. Why do hubs in the yeast protein interaction
745 network tend to be essential: reexamining the connection between the network topology and
746 essentiality. *PLoS Comput. Biol.* 4:e1000140.

747