Predicting the Health Impact of Dietary Polyphenols Using a Network Medicine Framework

Italo F. do Valle[1], Harvey G. Roweth[2,3], Michael W. Malloy[2,3], Sofia Moco[4], Denis Barron[4], Elisabeth Battinelli[2,3], Joseph Loscalzo[3,5], Albert-László Barabási[1,6,7]

[1] Network Science Institute and Department of Physics, Northeastern University, Boston, MA, USA
[2] Division of Hematology, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA
[3] Harvard Medical School, Boston, MA, USA
[4] Nestle Institute of Health Sciences, Lausanne, Switzerland
[5] Department of Medicine Brigham and Women's Hospital
[6] Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[7] Department of Network and Data Science, Central European University, Budapest, Hungary

Abstract

Polyphenols, natural products present in most plant-based foods, play a protective role against several complex diseases through their antioxidant activity and by diverse molecular mechanisms. Here we developed a network medicine framework to uncover the mechanistic roles of polyphenols on health by considering the molecular interactions between polyphenol protein targets and proteins associated with diseases. We find that the protein targets of polyphenols cluster in specific neighborhoods of the human interactome, whose network proximity to disease proteins is predictive of the known therapeutic effects of polyphenols. This finding allows us to predict that rosmarinic acid (RA) has a direct impact on platelet function, representing a novel mechanism through which it could affect cardiovascular health, and experimentally confirm that RA inhibits platelet aggregation and alpha granule secretion through inhibition of protein tyrosine phosphorylation. Our framework represents a starting point for mechanistic interpretation of the health effects underlying food-related compounds, allowing us to integrate into a predictive framework knowledge on food metabolism, bioavailability, and drug interaction.

Introduction

Diet plays a defining role in human health. Indeed, while poor diet can significantly increase the risk for coronary heart disease and diabetes mellitus, a healthy diet can play a protective role, even mitigating genetic risk for CHD[1]. Polyphenols are a class of compounds present in plant-based foods, from fruits to vegetables, herbs, spices, tea, and wine, with well documented protective role as antioxidants, which affect several diseases, from cancer to diabetes mellitus, cardiovascular, and neurodegenerative diseases[2,3]. Recent efforts profiling over 500 polyphenols in more than 400 foods[4,5] have documented the high diversity of polyphenols to which humans are exposed through their diet, ranging from flavonoids to phenolic acids, lignans, and stilbenes.

The underlying molecular mechanisms through which specific polyphenols exert their beneficial effects on human health remain largely unexplored. From a mechanistic perspective, dietary polyphenols are not engaged in endogenous metabolic processes of anabolism and catabolism, but rather affect human health through their anti- or pro-oxidant activity[6], by binding to proteins and modulating their activity[7,8], interacting with digestive enzymes[9], and modulating gut microbiota growth[10,11]. Yet, the variety of experimental settings used to explore the molecular effects of polyphenols have, to date, offered a range of often conflicting evidence for interpretation. For example, different clinical trials resulted in conflicting conclusions about the beneficial effects of resveratrol on glycemic control in type 2 diabetes patients[12,13]. We, therefore, need a framework to interpret the evidence present in the literature, and to offer in-depth mechanistic predictions on the molecular pathways responsible for the health implications of polyphenols present in diet. Ultimately, these insights could help us develop novel diagnostic and therapeutic strategies, and may lead to the synthesis of novel drugs.

Here, we address this challenge by developing a network medicine framework to capture the molecular interactions between polyphenols and their cellular binding targets, unveiling their relationship to complex diseases. The developed framework is based on the human interactome, a comprehensive subcellular network consisting of all known physical interactions between human proteins, which has been validated

64 previously as a platform for understanding disease mechanisms[14,15], rational drug target
65 identification, and drug repurposing[16,17].
66 We find that the proteins to which polyphenols bind form identifiable
67 neighborhoods in the human interactome, allowing us to demonstrate that the proximity
68 between polyphenol targets and proteins associated with specific diseases is predictive
69 of the known therapeutic effects of polyphenols. Finally, we unveil the potential
70 therapeutic effects of rosmarinic acid on vascular diseases, predicting that its
71 mechanism of action is related to modulation of platelet function. We confirm this
72 prediction by experiments that indicate that rosmarinic acid modulates platelet function
73 *in vitro* by inhibiting tyrosine protein phosphorylation. Altogether, our results
74 demonstrate that the network-based relationship between disease proteins and
75 polyphenol targets offers a tool to systematically unveil the health effects of
76 polyphenols.
77
78 **Results**
79
80 Polyphenol Targets Cluster in Specific Functional Neighborhoods of the Interactome
81
82 We mapped the targets of 65 polyphenols (see Methods) to the human
83 interactome, consisting of 17,651 proteins and 351,393 interactions (Fig 1a,b). We find
84 that 19 of the 65 polyphenols have only one protein target, while a few polyphenols
85 have an exceptional number of targets (Fig 1c). We computed the Jaccard Index (JI) of
86 the protein targets of each polyphenol pair, finding only a limited similarity of targets
87 among different polyphenols (average JI = 0.0206) (Fig S1a). Even though the average
88 JI is small, it is still significantly higher (Z = 147, Fig S1b) than the JI expected if the
89 polyphenol targets were randomly assigned from the pool of all network proteins with
90 degrees matching the original set. This finding suggests that while each polyphenol
91 targets a specific set of proteins, their targets are confined to a common pool of
92 proteins, likely determined by commonalities in the polyphenol binding domains of the
93 three-dimensional structure of the protein targets[20]. Gene Ontology (GO) Enrichment
94 Analysis of all polyphenol protein targets revealed that they tend to target pathways

related to post-translation protein modifications, regulation, and xenobiotic metabolism (Figs 1d). The enriched GO categories indicate that polyphenols modulate common regulatory processes, but the low similarity in their protein targets, illustrated by the low average JI, indicates that they target different processes within the same pathways.

We next asked whether the polyphenol targets cluster in specific regions of the human interactome. We focused on polyphenols with more than two targets (n=46, Figure 2), and measured the size and significance of the largest connected component (LCC) formed by the targets of each polyphenol. We found that 25 of the 46 polyphenols have a larger LCC than expected by chance (Z-score > 1.95) (Fig 1e, Fig 2). In agreement with experimental evidence documenting the effect of polyphenols on multiple pathways[21], we find that ten polyphenols have their targets organized in multiple connected components of size > 2.

These results indicate that the targets of polyphenols modulate specific well localized neighborhoods of the interactome (Fig 2, S1c). This prompted us to explore if the interactome regions targeted by the polyphenols reside within network neighborhoods associated with specific diseases, seeking a network-based framework to unveil the molecular mechanism through which specific polyphenols modulate health.

Proximity Between Polyphenol Targets and Disease Proteins Reveals their Therapeutic Effects

Polyphenols can be viewed as drugs in that they bind to specific proteins, affecting their ability to perform their normal functions. We, therefore, hypothesized that we can apply the network-based framework used to predict the efficacy of drugs in specific diseases[16,17] to also predict the therapeutic effects of polyphenols. The closer the targets of a polyphenol are to disease proteins, the more likely that the polyphenol will affect the disease phenotype. We, therefore, calculated the network proximity between polyphenol targets and proteins associated with 299 diseases using the closest measure, d_c, representing the average shortest path length between each polyphenol target and the nearest disease protein (see Methods). Consider for example (-)-epigallocatechin 3-O-gallate (EGCG), a polyphenol abundant in green tea.

126 Epidemiological studies have found a positive relationship between green tea
127 consumption and reduced risk of type 2 diabetes mellitus (T2D)[22,23], and physiological
128 and biochemical studies have shown that EGCG presents glucose-lowering effects in
129 both *in vitro* and *in vivo* models[24,25]. We identified 54 experimentally validated EGCG
130 protein targets and mapped them to the interactome, finding that the ECGC targets form
131 an LCC of 17 proteins (Z = 7.61) (Fig 3a). We also computed the network-based
132 distance between EGCG targets and 83 proteins associated with T2D, finding that the
133 two sets are significantly proximal to each other. We ranked all 299 diseases based on
134 the network proximity to the ECGC targets in order to determine whether we can
135 recover the 82 diseases in which ECGC has known therapeutic effects according to the
136 CTD database. By this analysis, we were able to recover 15 previously known
137 therapeutic associations among the top 20 ranked diseases (Table 1), confirming that
138 network-proximity can discriminate between known and unknown disease associations
139 for polyphenols, as previously confirmed among drugs[16,17].
140 We expanded these methods to all polyphenol-disease pairs, to predict diseases
141 for which specific polyphenols might have therapeutic effects. For this analysis, we
142 grouped all 19,435 polyphenol-disease associations between 65 polyphenols and 299
143 diseases into known (1,525) and unknown (17,910) associations. The known
144 polyphenol-disease set was retrieved from CTD, which is limited to manually curated
145 associations for which there is literature-based evidence. For each polyphenol, we
146 tested how well network proximity discriminates between the known and unknown sets
147 by evaluating the area under the Receiving Operating Characteristic (ROC) curve
148 (AUC). For EGCG, network proximity offers good discriminative power (AUC = 0.78, CI:
149 0.70 - 0.86) between diseases with known and unknown therapeutic associations (Table
150 1). We find that network proximity (d_c) offers predictive power with an AUC > 0.7 for 31
151 polyphenols (Fig 3b). In Table 2 we summarize the top 10 polyphenols for which the
152 network medicine framework offers the best predictive power of therapeutic effects,
153 limiting the entries to predictive performance of AUC > 0.6 and performance over top
154 predictions with precision > 0.6.
155 Finally, we performed multiple robustness checks to exclude the role of potential
156 biases in the input data. To test if the predictions are biased by the set of known

157 associations retrieved from CTD, we randomly selected 100 papers from PubMed
158 containing MeSH terms that tag EGCG to diseases. We manually curated the evidence
159 for EGCG's therapeutic effects for the diseases discussed in the published papers,
160 excluding reviews and non-English language publications. The dataset was processed
161 to include implicit associations (see Methods), resulting in a total of 113 diseases
162 associated with EGCG, of which 58 overlap with the associations reported by CTD (Fig
163 3c). We observed that the predictive power of network proximity was unaffected by
164 whether we considered the annotations from CTD, the manually curated list, or the
165 union of both (Fig 3d). To test the role of potential biases in the interactome, we
166 repeated our analysis using only high-quality polyphenol-protein interactions retrieved
167 from ligand-protein 3D resolved structures (Fig S1d) and a subset of the interactome
168 derived from an unbiased high-throughput screening (Fig S1f). We find that the
169 predictive power was largely unchanged, indicating that the literature bias in the
170 interactome does not affect our findings.
171
172 Network Proximity Predicts Gene Expression Perturbation Induced by Polyphenols
173
174 To validate the predicted polyphenol-disease associations, we retrieved expression
175 perturbation signatures from the Connectivity Map database[26] for the treatment of the
176 breast cancer MCF7 cell line with 22 polyphenols (Table S1). We investigated the
177 relationship between the extent to which polyphenols perturb the expression of disease
178 genes, the network proximity between the polyphenol targets and disease proteins, and
179 their known therapeutic effects (Fig 4a). For example, we observe different perturbation
180 profiles for gene pools associated with different diseases: for treatment with genistein (1
181 μM, 6 hours) we observe 10 skin disease (SD) genes with perturbation score > 2, while
182 we observe only one highly perturbed cerebrovascular disorder (CD) gene (Fig 4b).
183 Indeed, network proximity indicates that SD is closer to the genistein targets than CD,
184 suggesting a relationship between network proximity, gene expression perturbation, and
185 the therapeutic effects of the polyphenol (Fig 4a). To test this hypothesis, we computed
186 an enrichment score that measures the overrepresentation of disease genes among the
187 most perturbed genes (see Methods), finding 13 diseases that have their genes

188 significantly enriched among the most deregulated genes by genistein, of which 4 have
189 known therapeutic associations. We find that these four diseases are significantly closer
190 to the genistein targets than the nine diseases with non-therapeutic associations (Fig
191 4c). We observed a similar trend for treatments with other polyphenols, whether we use
192 the same (1μM, Fig 4c) or different (100nM to 10μM, Fig S2) concentrations. This result
193 suggests that changes in gene expression caused by a polyphenol are indicative of its
194 therapeutic effects, but only if the observed expression change is limited to proteins
195 proximal to the polyphenol targets (Fig 4a).
196 Consequently, network proximity should also be predictive of the overall gene
197 expression perturbation caused by a polyphenol on the genes of a given disease. To
198 test this hypothesis, in each experimental combination defined by the polyphenol type
199 and its concentration, we evaluated the maximum perturbation among genes for each
200 disease. We then compared the magnitude of the observed perturbation between
201 diseases that were proximal (d_c < 25th percentile, Z_{d_c}< -0.5) or distal (d_c > 75th
202 percentile, Z_{d_c} > -0.5) to the polyphenol targets. Figures 5a and 5b show the results for
203 the genistein treatment (1μM, 6 hours), indicating that diseases proximal to the
204 polyphenol targets show higher maximum perturbation values than distal diseases. The
205 same trend is observed for other polyphenols (Figs 5b, S3 and S4), confirming that the
206 impact of a polyphenol on cellular signaling pathways is localized in the network space,
207 being greater in the vicinity of the polyphenol targets compared to neighborhoods
208 remote from these targets. These results indicate that network proximity offers a
209 mechanistic interpretation for the gene expression perturbations induced by
210 polyphenols, being also predictive of whether these perturbations result in therapeutic
211 effects.
212
213 Experimental Evidence Confirms that Rosmarinic Acid Modulates Platelet Function
214
215 To demonstrate how the network-based framework can facilitate the mechanistic
216 interpretation of the therapeutic effects of selected polyphenols, we next focus on
217 vascular diseases (VD). Of 65 polyphenols evaluated in this study, we found 27 to have
218 associations to VD, as their targets were within the VD network neighborhood (Table

219 S3). We, therefore, inspected the targets of 15 of the 27 polyphenols with 10 or less
220 targets. The network analysis identified direct links between biological processes related
221 to vascular health and the targets of three polyphenols: gallic acid, rosmarinic acid, and
222 1,4-naphthoquinone (Fig S5, Supplementary Note1). The network neighborhood
223 containing the targets of these polyphenols suggests that gallic acid activity involves
224 thrombus dissolution processes, rosmarinic acid acts on platelet activation and
225 antioxidant pathways through FYN and its neighbors, and 1,4-naphthoquinone acts on
226 signaling pathways of vascular cells through MAP2K1 activity (Fig S5, Supplementary
227 Note 1).
228 To validate the developed framework, we set out to obtain direct experimental
229 evidence of the predicted mechanistic role of rosmarinic acid (RA) in VD. The RA
230 targets are in close proximity to proteins related to platelet function, forming the RA-VD-
231 platelet module: a connected component formed by the RA target FYN and the VD
232 proteins associated with platelet function PDE4D, CD36, and APP (Fig 7a). We,
233 therefore, asked whether RA influenced platelet activation *in vitro*. As platelets can be
234 stimulated through different activation pathways, RA effects can, in principle, occur in
235 any of them. To test these different possibilities, we pretreated platelets with RA and
236 then activated with: 1) glycoprotein VI by collagen or collagen-related peptide
237 (CRP/CRPXL); 2) protease-activated receptors-1,4 by thrombin receptor activator
238 peptide-6 (TRAP-6); 3) prostanoid thromboxane receptor by the thromboxane A_2
239 analogue (U46619); and 4) P2Y1/12 receptor by adenosine diphosphate (ADP)[39]. When
240 we compared the network distance between each stimulant receptor and the RA-VD-
241 platelet module (Fig 7a), we observed that the receptors for CRP/CRPXL, TRAP-6, and
242 U46619 are closer than random expectation, while the receptor for ADP is more distant
243 (Fig 7b). We expected that platelets would be most affected by RA when treated with
244 stimulants whose receptors are most proximal to the RA-VD-platelet module, i.e.,
245 CRP/CRPXL, TRAP-6, and U46619, and as a control, we expect no effect for the
246 distant ADP receptor. The experiments confirm this prediction: RA inhibits collagen-
247 mediated platelet aggregation (Fig 7c) and impairs dense granule secretion induced by
248 CRPXL, TRAP-6, and U46619 (Fig S13). RA-treated platelets also displayed dampened
249 alpha-granule secretion (Fig 7d) and integrin αIIbβ3 activation (Fig S6) in response to

250 U46619. As expected, RA did not affect platelet function when we used an agonist
251 whose receptor is distant from the RA-VD-module, i.e., ADP. These findings suggest
252 that RA impairs basic hallmarks of platelet activation via strong network effects,
253 supporting our hypothesis that the proximity between RA targets and the neighborhood
254 associated with platelet function (Fig 7a) could in part explain RA's impact on VD.
255 We next searched to clarify the molecular mechanisms involved in the impact of
256 RA on platelets. Given that platelet activation is coordinated by several kinases, we
257 hypothesized that RA inhibits platelet function by blocking agonist-induced protein
258 tyrosine phosphorylation. We observed that RA-treated platelets demonstrated a dose-
259 dependent reduction in total tyrosine phosphorylation in response to CRPXL, TRAP-6
260 and U46619 (Fig 7e). Given that RA caused a substantial decrease in phosphorylation
261 of proteins with atomic mass between 50-60 KDa (Fig 7e), we hypothesized that RA
262 may reduce phosphorylation of FYN (59 KDa), or other similarly sized members of the
263 same protein family (*i.e.* src family kinases, SFKs). To test this, we measured the level
264 of phosphorylation within the activation domain (amino acid 416) of SFKs, finding that
265 RA reduced collagen induced phosphorylation of FYN as well as basal tyrosine
266 phosphorylation of SFKs (Fig 7f). This indicates that RA perturbs the phospho-signaling
267 networks that regulate platelet response to extracellular stimuli.
268 Altogether, these findings support our prediction that RA modulates platelet
269 activation and function. It also supports the observation that its mechanism of action
270 involves reduction of phosphorylation at the activation domain of the protein-tyrosine
271 kinase FYN (Fig 7a) and the inhibition of general tyrosine phosphorylation. Finally, while
272 polyphenols are usually associated to their antioxidant function, here we illustrate
273 another mechanistic pathway through which they could benefit health.
274
275 **Discussion**
276
277 Here, we proposed a network-based framework to predict the therapeutic effects of
278 dietary polyphenols in human diseases. We find that polyphenol protein targets cluster
279 in specific functional neighborhoods of the interactome, and we show that the network
280 proximity between polyphenol targets and disease proteins is predictive of the

281 therapeutic effects of polyphenols. We demonstrate that diseases whose proteins are
282 proximal to polyphenol targets tend to have significant changes in gene expression in
283 cell lines treated with the respective polyphenol, while such changes are absent for
284 diseases whose proteins are distal to polyphenol targets. Finally, we find that the
285 network neighborhood around the rosmarinic acid (RA) targets and vascular disease
286 proteins are related to platelet function. We validate this mechanistic prediction by
287 showing that RA modulates platelet function through inhibition of protein tyrosine
288 phosphorylation. These observations suggest a role of RA on prevention of vascular
289 diseases by inhibiting platelet activation and aggregation.

290 The observed results also suggest multiple avenues through which our ability to
291 understand the role of polyphenols could be improved. First, some of the known health
292 benefits of polyphenols might be caused not only by the native molecules, but also by
293 their metabolic byproducts [40,41]. We, however, lack data about colonic degradation, liver
294 metabolism, bioavailability, and interaction with proteins of specific polyphenols or their
295 metabolic byproducts. Future experimental data on protein interactions with polyphenol
296 byproducts and conjugates can be incorporated in the proposed framework, further
297 improving the accuracy of our predictions. The lack of this data does not invalidate the
298 findings presented here, since previous studies report the presence of unmetabolized
299 polyphenols in blood[42–44]; and it has been hypothesized that, in some instances,
300 deconjugation of liver metabolites occurs in specific tissues or cells[45–47]. Second,
301 considering that several experimental studies of polyphenol bioefficacy have been
302 observed in *in vitro* and *in vivo* models, the proposed framework might help us interpret
303 literature evidence, possibly even allowing us to exclude chemical candidates when
304 considering the health benefits provided by a given food in epidemiological association
305 studies.

306 The low bioavailability of some polyphenols in food might still present a major
307 challenge when considering the therapeutic utility of these molecules. However, in the
308 same way that the polyphenol phlorizin led to the discovery of new strategies for
309 disease treatment resulting in the development of new compounds with higher
310 efficacy[48], we believe that the present methodology can helps us identify polyphenol-
311 based candidates for drug development.

312 The methodology introduced here offers a foundation for the mechanistic
313 interpretation of alternative pathways through which polyphenols can affect health, e.g.,
314 the combined effect of different polyphenols and their interaction with drugs[50]. Finally,
315 this methodology can be applied to other food-related chemicals, providing a framework
316 by which to understand their health effects. Future research may help us also account
317 for the way that food-related chemicals affect endogenous metabolic reactions,
318 impacting not only signaling pathways, but also catabolic and anabolic processes.
319 Taken together, the proposed network-based framework has the potential to reveal
320 systematically the mechanism of action underlying the health benefits of polyphenols,
321 offering a logical, rational strategy for mechanism-based drug development of food-
322 based compounds.
323
324 **Methods**
325
326 Building the Interactome
327
328 The human interactome was assembled from 16 databases containing different types of
329 protein-protein interactions (PPIs): 1) binary PPIs tested by high-throughput yeast-two-
330 hybrid (Y2H) experiments[51]; 2) kinase-substrate interactions from literature-derived low-
331 throughput and high-throughput experiments from KinomeNetworkX[52], Human Protein
332 Resource Database (HPRD)[53], and PhosphositePlus[54]; 3) carefully literature-curated
333 PPIs identified by affinity purification followed by mass spectrometry (AP-MS), and from
334 literature-derived low-throughput experiments from InWeb[55], BioGRID[56], PINA[57],
335 HPRD[58], MINT[59], IntAct[59], and InnateDB[60]; 4) high-quality PPIs from three-dimensional
336 (3D) protein structures reported in Instruct[61], Interactome3D[62], and INSIDER[63]; 5)
337 signaling networks from literature-derived low-throughput experiments as annotated in
338 SignaLink2.0[64]; and 6) protein complex from BioPlex2.0[65]. The genes were mapped to
339 their Entrez ID based on the National Center for Biotechnology Information (NCBI)
340 database as well as their official gene symbols. The resulting interactome includes
341 351,444 protein-protein interactions (PPIs) connecting 17,706 unique proteins

342 (Supplementary Data 1). The largest connected component has 351,393 PPIs and
343 17,651 proteins.
344
345 Polyphenols, Polyphenol Targets, and Disease Proteins
346
347 We retrieved 759 polyphenols from the PhenolExplorer database[4]. The database lists
348 polyphenols with food composition data or profiled in biofluids after interventions with
349 polyphenol-rich diets. For our analysis, we only considered polyphenols that: 1) could
350 be mapped in PubChem IDs, 2) were listed in the Comparative Toxicogenomics (CTD)
351 database[18] as having therapeutic effects on human diseases, and 3) had protein-
352 binding information present in the STITCH database[19] with experimental evidence
353 (Figure 1a). After these steps, we considered a final list of 65 polyphenols, for which
354 598 protein targets were retrieved from STITCH (Table S1). We considered 3,173
355 disease proteins corresponding to 299 diseases retrieved from Menche *et al* (2015)[15].
356 Gene ontology enrichment analysis of protein targets was performed using the
357 Bioconductor package clusterProfiler with a significance threshold of $p < 0.05$ and
358 Benjamini-Hochberg multiple testing correction with $q < 0.05$.
359
360 Polyphenol Disease Associations
361
362 We retrieved the polyphenol-disease associations from the Comparative
363 Toxicogenomics Database (CTD). We considered only manually curated associations
364 labeled as therapeutic. By considering the hierarchical structure of diseases along the
365 MeSH tree, we expanded explicit polyphenol-disease associations to include also
366 implicit associations. This procedure was performed by propagating associations in the
367 lower branches of the MeSH tree to consider diseases in the higher levels of the same
368 tree branch. For example, a polyphenol associated with 'heart diseases' would also be
369 associated with the more general category of 'cardiovascular diseases'. By performing
370 this expansion, we obtained a final list of 1,525 known associations between the 65
371 polyphenols and the 299 diseases considered in this study.
372

373 Network Proximity Between Polyphenol Targets and Disease Proteins

374

375 The proximity between a disease and a polyphenol was evaluated using a distance
376 metric that takes into account the shortest path lengths between polyphenol targets and
377 disease proteins[16]. Given S, the set of disease proteins, T, the set of polyphenol targets,
378 and $d(s,t)$, the shortest path length between nodes s and t in the network, we define:

379

380 $$d_c(S,T) = \frac{1}{||T||} \sum_{t \in T} \min_{s \in S} d(s,t) \qquad (1)$$

381 We also calculated a relative distance metric (Z_{dc}) that compares the absolute distance
382 $d_c(S,T)$ between a disease and a polyphenol with a reference distribution describing the
383 random expectation. The reference distribution corresponds to the expected distances
384 between two randomly selected groups of proteins matching the size and degrees of the
385 original disease proteins and polyphenol targets in the network. It was generated by
386 calculating the proximity between these two randomly selected groups across 1,000
387 iterations. The mean $\mu_{d(S,T)}$ and standard deviation $\sigma_{d(S,T)}$ of the reference distribution
388 were used to convert the absolute distance d_c into the relative distance Z_{dc}, defined as:

389

390 $$Z_{d_c} = \frac{d - \mu_{d_c(S,T)}}{\sigma_{d_c(S,T)}} \qquad (2)$$

391

392 Due to the scale-free nature of the human interactome, there are few nodes with high
393 degrees, and to avoid repeatedly choosing the same (high degree) nodes, we
394 performed a degree-preserving random selection. The Supplementary Data 2 reports
395 the proximity scores $d_c$and Z_{d_c} for all pairs of diseases and polyphenols.

396

397 Area Under ROC Curve Analysis

398

399 For each polyphenol, we used AUC to evaluate how well the network proximity
400 distinguishes diseases with known therapeutic associations from all of the others of the
401 set of 299 diseases. The set of known associations (therapeutic) retrieved from CTD
402 were used as positive instances, all unknown associations were defined as negative

instances, and the area under the ROC curve was computed using the implementation in the Scikit-learn Python package. Furthermore, we calculated 95% confidence intervals using the bootstrap technique with 2,000 resamplings with sample sizes of 150 each. Considering that AUC provides an overall performance, we also searched for a metric to evaluate the top ranking predictions. For this analysis, we calculated the precision of the top 10 predictions, considering only the polyphenol-disease associations with relative distance $Z_{dc} < -0.5$[16].

Analysis of Network Proximity and Gene Expression Deregulation

We retrieved perturbation signatures from the Connectivity Map database (https://clue.io/) for the MCF7 cell line after treatment with 22 polyphenols. These signatures reflect the perturbation of the gene expression profile caused by the treatment with that particular polyphenol relative to a reference population, which comprises all other treatments in the same experimental plate[26]. For polyphenols having more than one experimental instance (time of exposure, cell line, dose), we selected the one with highest distil_cc_q75 value (75th quantile of pairwise spearman correlations in landmark genes, https://clue.io/connectopedia/perturbagen_types_and_controls). We performed Gene Set Enrichment Analysis[66] to evaluate the enrichment of disease genes among the top deregulated genes in the perturbation profiles. This analysis offers Enrichment Scores (ES) that have small values when genes are randomly distributed among the ordered list of expression values and high values when they are concentrated at the top or bottom of the list. The ES significance is calculated by creating 1,000 random selection of gene sets with the same size as the original set and calculating an empirical p-value by considering the proportion of random sets resulting in ES smaller than the original case. The p-values were adjusted for multiple testing using the Benjamini-Hochberg method. The network proximity d_c of disease proteins and polyphenol targets for diseases with significant ES were compared according to their therapeutic and non-therapeutic associations using the Student's t-test.

Platelet Isolation

434

435 Human blood collection was performed as previously described in accordance with the
436 Declaration of Helsinki and ethics regulations with Institutional Review Board approval
437 from Brigham and Women's Hospital (P001526). Healthy volunteers did not ingest
438 known platelet inhibitors for at least 10 days prior. Citrated whole blood underwent
439 centrifugation with a slow brake (177 x g, 20 minutes), and the PRP fraction was
440 acquired for subsequent experiments. For washed platelets, PRP was incubated with 1
441 μM prostaglandin E_1 (Sigma, P5515) and immediately underwent centrifugation with a
442 slow brake (1000 x g, 5 minutes). Platelet-poor plasma was aspirated, and pellets
443 resuspended in platelet resuspension buffer (PRB; 10 mM Hepes, 140 mM NaCl, 3 mM
444 KCl, 0.5 mM $MgCl_2$, 5 mM $NaHCO_3$, 10 mM glucose, pH 7.4).
445

446 Platelet Aggregometry

447

448 Platelet aggregation was measured by turbidimetric aggregometry as previously
449 described[67]. Briefly, PRP was pretreated with RA for 1 hour before adding 250 μL to
450 siliconized glass cuvettes containing magnetic stir bars. Samples were placed in
451 Chrono-Log® Model 700 Aggregometers before the addition of various platelet agonists.
452 Platelet aggregation was monitored for 6 minutes at 37°C with a stir speed of 1000 rpm
453 and the maximum extend of aggregation recorded using AGGRO/LINK®8 software. In
454 some cases, dense granule release was simultaneously recorded by supplementing
455 samples with Chrono-Lume® (Chrono-Log®, 395) according to the manufacturer's
456 instructions.
457

458 Platelet Alpha Granule Secretion and Integrin $\alpha_{IIb}\beta_3$ Activation

459

460 Changes in platelet surface expression of P-selectin (CD62P) or binding of Alexa
461 Fluor™ 488-conjugated fibrinogen were used to assess alpha granule secretion and
462 integrin $\alpha_{IIb}\beta_3$ activation, respectively. First, PRP was pre-incubated with RA for 1 hour,
463 followed by stimulation with various platelet agonists under static conditions at 37°C for
464 20 minutes. Samples were then incubated with APC-conjugated anti-human CD62P

465 antibodies (BioLegend®, 304910) and 100 μg/mL Alexa Fluor™ 488-Fibrinogen (Thermo
466 Scientific™, F13191) for 20 minutes before fixation in 2% [v/v] paraformaldehyde
467 (Thermo Scientific™, AAJ19945K2). Fifty thousand platelets were processed per
468 sample using a Cytek™ Aurora spectral flow cytometer. Percent-positive cells were
469 determined by gating on fluorescence intensity compared to unstimulated samples.
470
471 Platelet Cytotoxicity
472
473 Cytotoxicity were tested by measuring lactate dehydrogenase (LDH) release by
474 permeabilized platelets into the supernatant[68]. Briefly, washed platelets were treated
475 with various concentrations of RA for 1 hour, before isolating supernatants via
476 centrifugation (15,000 x g, 5 min). A Pierce LDH Activity Kit (Thermo Scientific™,
477 88953) was then used to assess supernatant levels of LDH.
478
479 Immunoprecipitation and Western blot
480
481 Washed platelets were pre-treated with RA for 1 hour, followed by a 15 minute
482 treatment with Eptifibatide (50 μM). Platelets were then stimulated with various agonists
483 for 5 minutes under stirring conditions (1000 rpm, 37°C). Platelets were lysed on ice
484 with RIPA Lysis Buffer System® (Santa Cruz®, sc-24948) and supernatants clarified via
485 centrifugation (15,000 x g, 10 min, 4°C). For immunoprecipitation of FYN, lysates were
486 first precleared of IgG by incubating with Protein A agarose beads (Cell Signaling
487 Technologies, 9863S) for 30 minutes at 4°C, before isolation of the supernatant via
488 centrifugation (15,000 x g, 10 min, 4°C). Supernatants were incubated with anti-FYN
489 antibodies (Abcam, 2A10) overnight at 4°C before incubation with Protein A beads for 1
490 hour. Beads were then washed 5 times with NP-40 lysis buffer (144 mM Tris, 518 mM
491 NaCl, 6 mM EDTA, 12 mM Na_2VO_3, 33.3% [v/v] NP-40, Halt™ protease inhibitor
492 cocktail (Thermo, 78429)).
493 For Western Blot analysis, total cell lysates or immunoprecipitated FYN were
494 reduced with Laemmli Sample Buffer (Bio-Rad, 1610737) and proteins separated by
495 molecular weight in PROTEAN TGX™ precast gels (Bio-Rad, 4561084). Proteins were

496 transferred to PVDF membranes (Bio-Rad, 1620174) and probed with either 4G10
497 (Milipore, 05-321), a primary antibody clone that recognizes phosphorylated tyrosine
498 residues, or primary antibodies that probe for the site-specific phosphorylation of src
499 family kinases (SFKs, p-Tyr416) within their activation loop. Membranes were incubated
500 with horseradish peroxidase-conjugated secondary antibodies (Cell Signaling
501 Technologies, 7074S) to catalyze an electrochemiluminescent reaction (Thermo
502 Scientific™, PI32109). Membranes were visualized using a Bio-Rad ChemiDoc Imaging
503 System and densitometric analysis of protein lanes conducted using ImageJ (NIH,
504 Version 1.52a).
505
506 Code Availability
507
508 The relevant code for calculating the network proximity, AUCs, and enrichment scores
509 can be found on https://github.com/italodovalle/polyphenols
510
511 **Author Contributions**
512
513 I.F.V and A.L.B designed the study. I.F.V. performed all computational analyses. H.G.R,
514 M.W.M., E.B., and J.L designed and performed experimental validation. J.L. guided
515 I.F.V. on validation case studies. S.M and D.B guided I.F.V for data interpretation and
516 curation of disease associations obtained from literature. I.F.V and A.L.B wrote the
517 paper with input from all authors. All authors read and approved the manuscript.
518
519 **Acknowledgements**
520
521 This study was supported, in part, by NIH grant 1P01HL132825, HG007690, HL108630,
522 and HL119145; American Heart Association grants 151708 and D700382; and ERC
523 grant 810115-DYNASET.
524
525 **Declaration of Interests**
526

527 J.L. and A.L.B are co-scientific founder of Scipher Medicine, Inc., which applies network
528 medicine strategies to biomarker development and personalized drug selection. A.L.B is
529 the founder of Nomix Inc. and Foodome, Inc. that apply data science to health; I.F.V is a
530 scientific consultant for Foodome, Inc.

531 **Table 1 – Top 20 Predicted Therapeutic Associations Between EGCG and Human**
532 **Diseases**

Disease	Distance d_c	Significance Z_{d_c}	Known Therapeutic Effect (References)
nervous system diseases	1.13	-1.72	[69,70]
nutritional and metabolic diseases	1.25	-1.45	[25]
metabolic diseases	1.25	-1.41	[25]
cardiovascular diseases	1.27	-2.67	[71–76]
immune system diseases	1.29	-1.31	[77]
vascular diseases	1.33	-3.47	[71,72,75]
digestive system diseases	1.33	-1.57	[78–82]
neurodegenerative diseases	1.37	-1.71	[83]
central nervous system diseases	1.41	-0.54	[83]
autoimmune diseases	1.41	-1.30	[77]
gastrointestinal diseases	1.43	-1.02	[84]
brain diseases	1.43	-0.89	NA
intestinal diseases	1.49	-1.08	[84]
inflammatory bowel diseases	1.54	-2.10	NA
bone diseases	1.54	-1.18	NA
gastroenteritis	1.54	-1.92	NA
demyelinating diseases	1.54	-1.78	NA
glucose metabolism disorders	1.54	-1.58	[25]
heart diseases	1.56	-1.20	[73,74,76]
diabetes mellitus	1.56	-1.66	[25]

533 Diseases were ordered according to the network distance (d_c) of their proteins to EGCG targets and
534 diseases with relative distance $Z_{dc} > -0.5$ were removed. References reported in CTD for curated
535 'therapeutic associations' are shown.
536

537 **Table 2 – Top Ranked Polyphenols.** Polyphenols for which network proximity to
538 diseases best predicts their therapeutic effects.

Polyphenol	AUC	AUC CI*	Precision**	Concentration in Blood***	N Mapped Targets	LCC Size
Coumarin	0.93	[0.86 - 0.98]	0.6		7	1
Piceatannol	0.86	[0.77 - 0.94]	0.6		39	23
Genistein	0.82	[0.75 - 0.89]	0.7	[0.006 - 0.525 uM]	18	6
Ellagic acid	0.79	[0.63 - 0.92]	0.6		42	19
(-)-epigallocatechin 3-O-gallate	0.78	[0.70 - 0.86]	0.8		51	17
Isoliquiritigenin	0.75	[0.77 - 0.94]	0.6		10	8
Resveratrol	0.75	[0.66 - 0.82]	1		63	25
Pterostilbene	0.73	[0.61 - 0.84]	0.6		5	2
Quercetin	0.73	[0.64 - 0.81]	1	[0.022 - 0.080 uM]	216	140
(-)-epicatechin	0.65	[0.49 - 0.80]	0.8	0.625 uM	11	3

539 Table showing polyphenols with AUC > 0.6 and Precision > 0.6.
540 * Confidence intervals calculated with 2,000 bootstraps with replacement and sample size of 50% of the
541 diseases (150/299)
542 ** Precision was calculated based on the top 10 polyphenols after their ranking based on the distance (d_c)
543 of their targets to the disease proteins and considering only predictions with Z-score < -0.5.
544 *** Concentrations of polyphenols in blood were retrieved from the Human Metabolome Database
545 (HMDB)
546

547

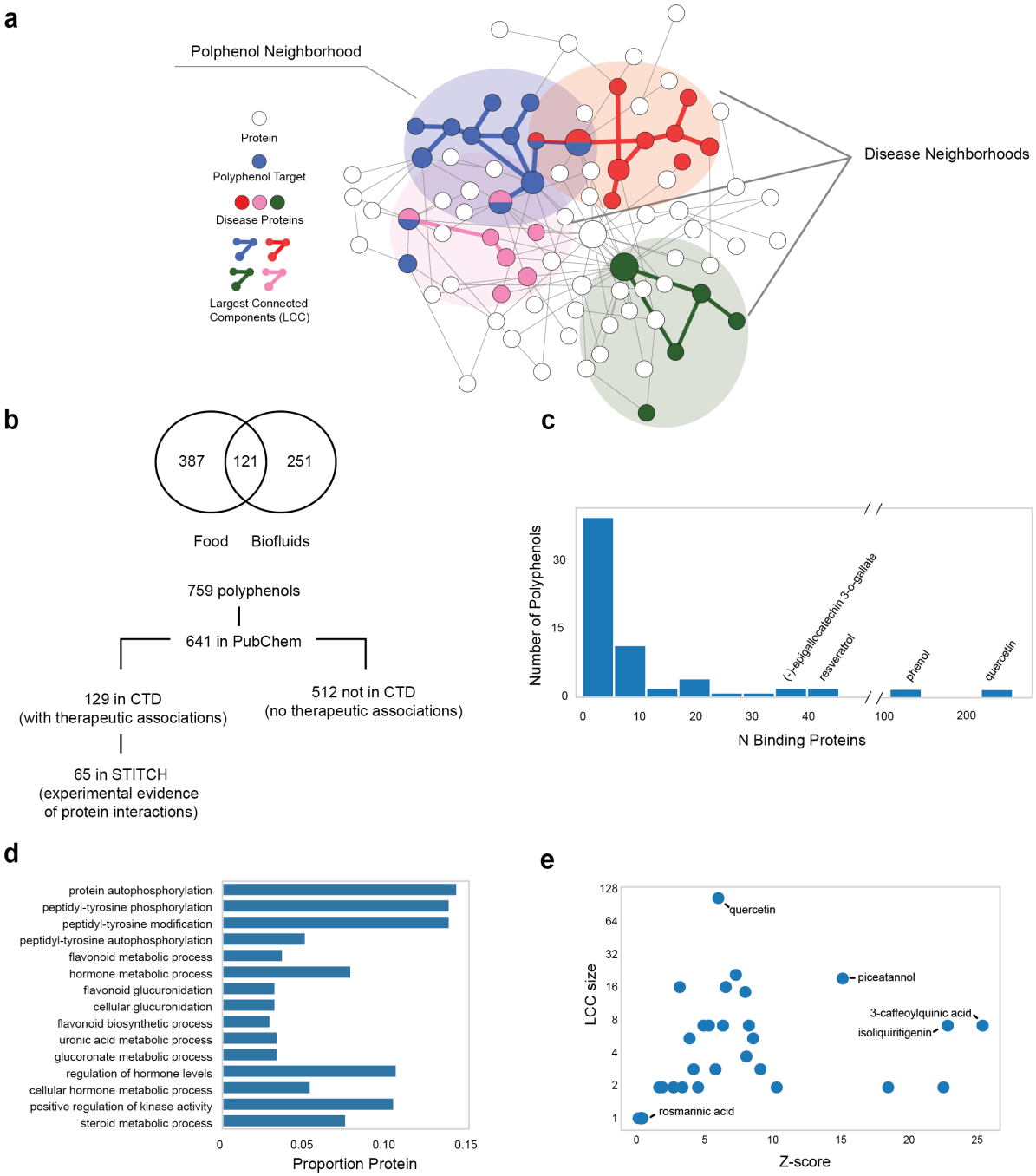
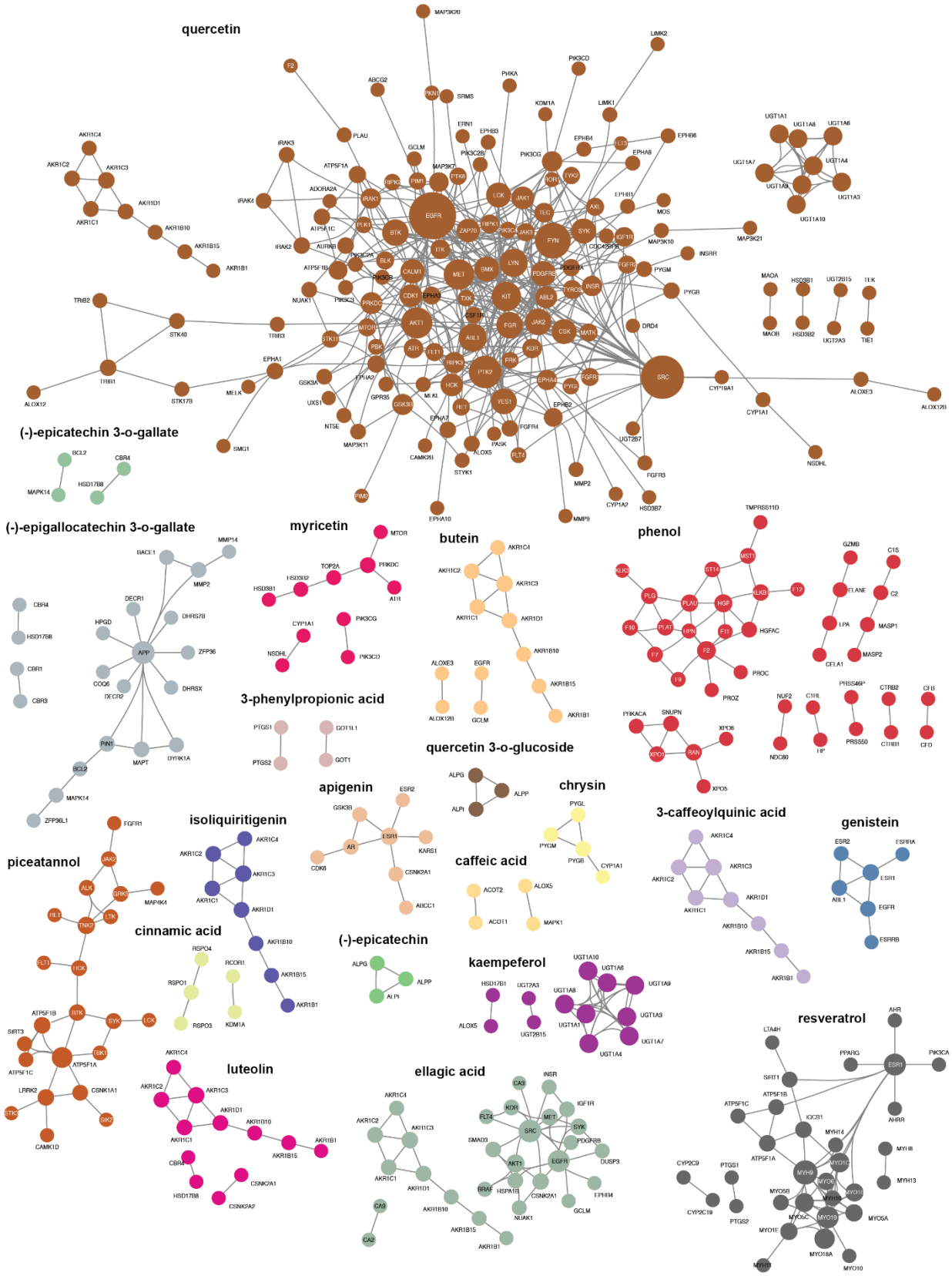


Figure 1 – Properties of Polyphenol Protein Targets. (A) Schematic representation of the human interactome, highlighting regions where polyphenol targets and disease proteins are localized. (B) Diagram showing the selection criteria of the polyphenols evaluated in this study. (C) Distribution of the number of polyphenol protein targets mapped to the human interactome. (D) Top (n=15) enriched GO pathways (Biological Process) among all polyphenol protein targets. The X-axis shows the proportion of targets mapped to each pathway. (E) Size of the Largest Connected Component (LCC) formed by the targets of each polyphenol in the interactome and the corresponding significance (z-score).


quercetin
(-)-epicatechin 3-o-gallate
(-)-epigallocatechin 3-o-gallate
myricetin
butein
phenol
3-phenylpropionic acid
quercetin 3-o-glucoside
apigenin
chrysin
3-caffeoylquinic acid
genistein
isoliquiritigenin
piceatannol
caffeic acid
cinnamic acid
(-)-epicatechin
kaempferol
luteolin
ellagic acid
resveratrol

558

559 **Figure 2 – Protein-protein Interactions of Polyphenol Targets.** The 23 polyphenols whose targets
560 form connected components in the interactome and their respective subgraphs. For example, piceatannol
561 targets form a unique connected component of 23 proteins, while quercetin targets form multiple
562 connected components, the largest with 140 proteins. Polyphenol targets that are not connected to any
563 other target are not shown in the figure. Colors distinguish connected component of different polyphenols.
564

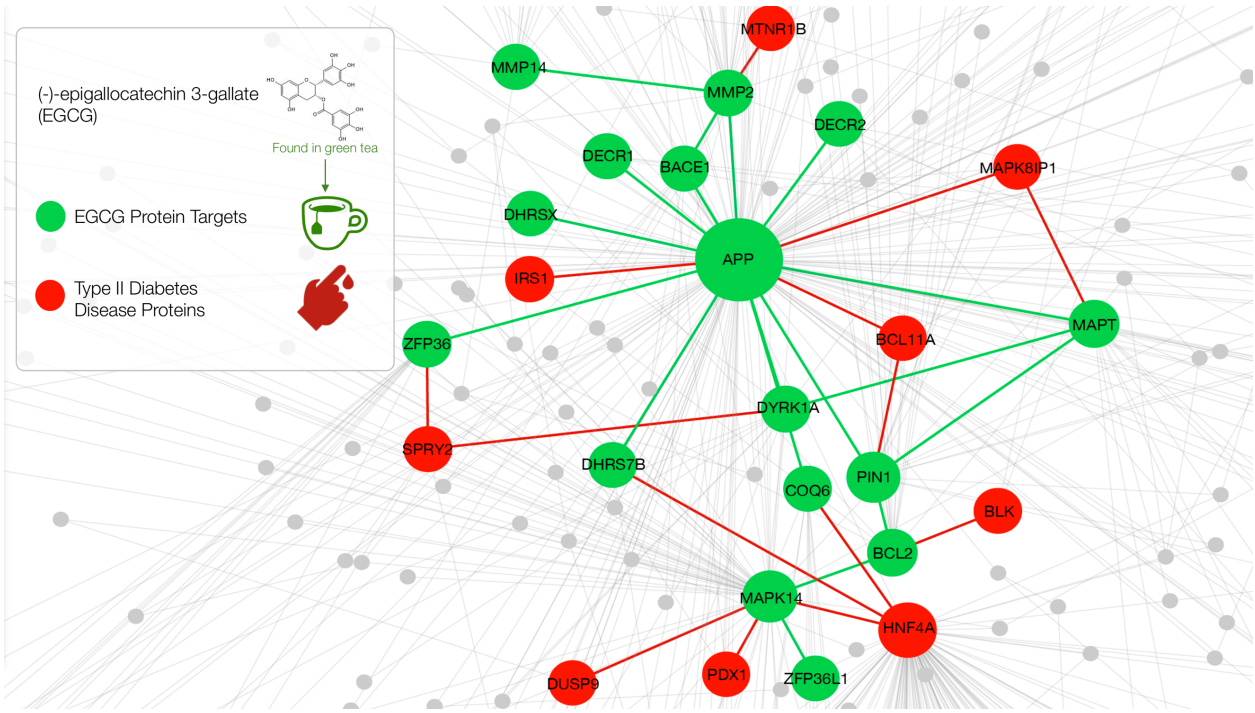


Figure 3 – Proximity Between Polyphenol Targets and Disease Proteins is Predictive of Therapeutic Effects of the Polyphenol. (a) Interactome neighborhood showing the EGCG protein targets and their interactions with type 2 diabetes (T2D)-associated proteins. (b) Distribution of AUC values considering the predictions of therapeutic effects for 65 polyphenols. (c) Comparison of the ECGC-disease associations considering the CTD database and the in-house database derived from the manual curation of the literature. (d) Comparison of the prediction performance when considering known EGCG-disease associations from the CTD, in-house manually curated database, or combined datasets.

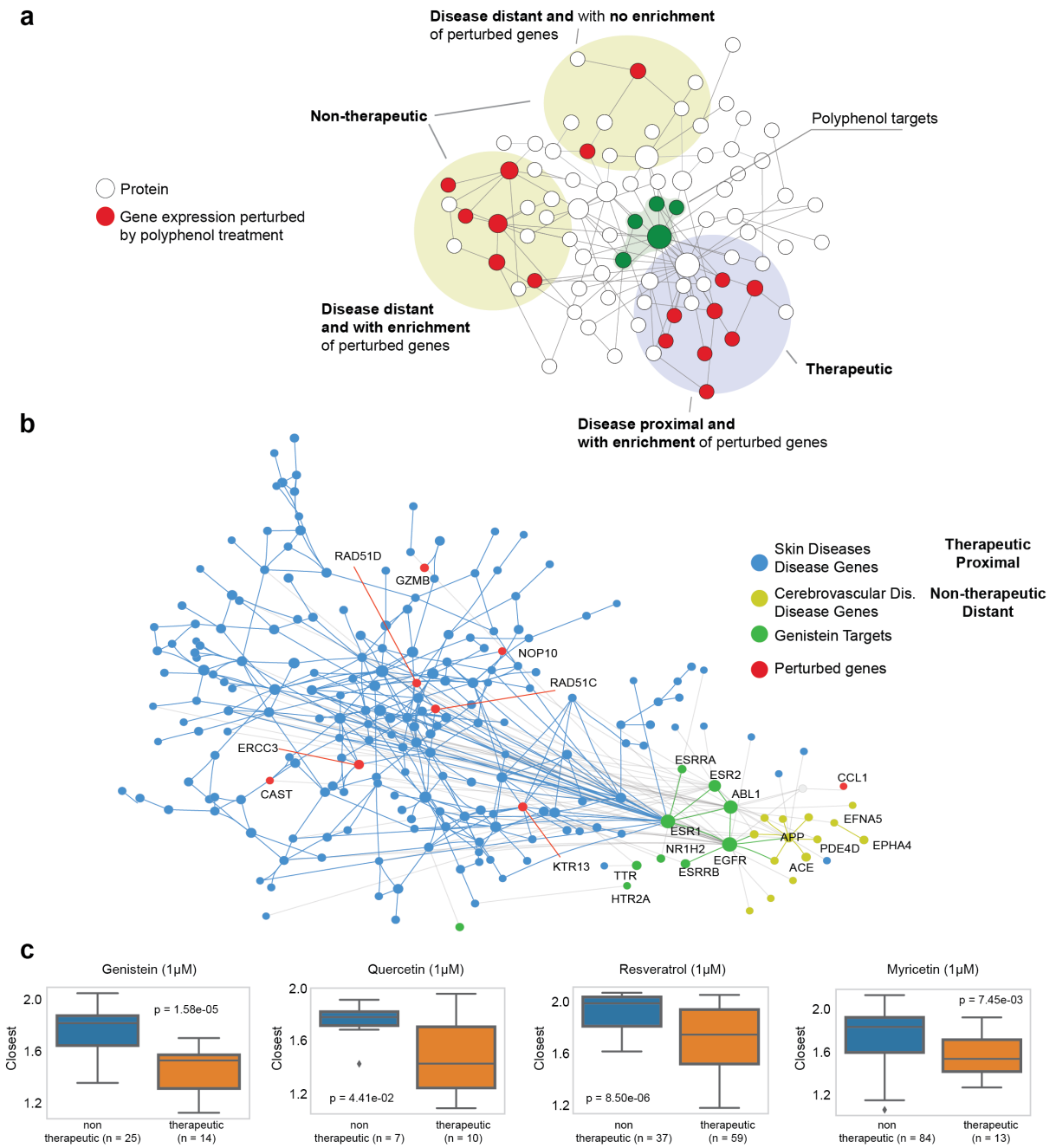


Figure 4 – Relationships among Gene Expression Perturbation, Network Proximity, and Therapeutic Effects of Polyphenols on Diseases. (a) Schematic representation of the relationship between the extent to which a polyphenol perturbs disease genes expression, its proximity to the disease genes, and its therapeutic effects. (b) Interactome neighborhood showing the modules of skin diseases (SD, genistein, and cerebrovascular disorders (CD). The SD module has 10 proteins with high perturbation scores (>2) in the treatment of the MCF7 cell line with 1 μM of genistein. Genes associated to SD are significantly enriched among the most differentially expressed genes, and the maximum perturbation score among disease genes is higher in SD than CD. (c) Among the diseases in which genes are enriched with highly perturbed genes, those with therapeutic associations show smaller network distances to the polyphenol targets than those without. The same trend is observed in treatments of the polyphenols quercetin, resveratrol, and myricetin.

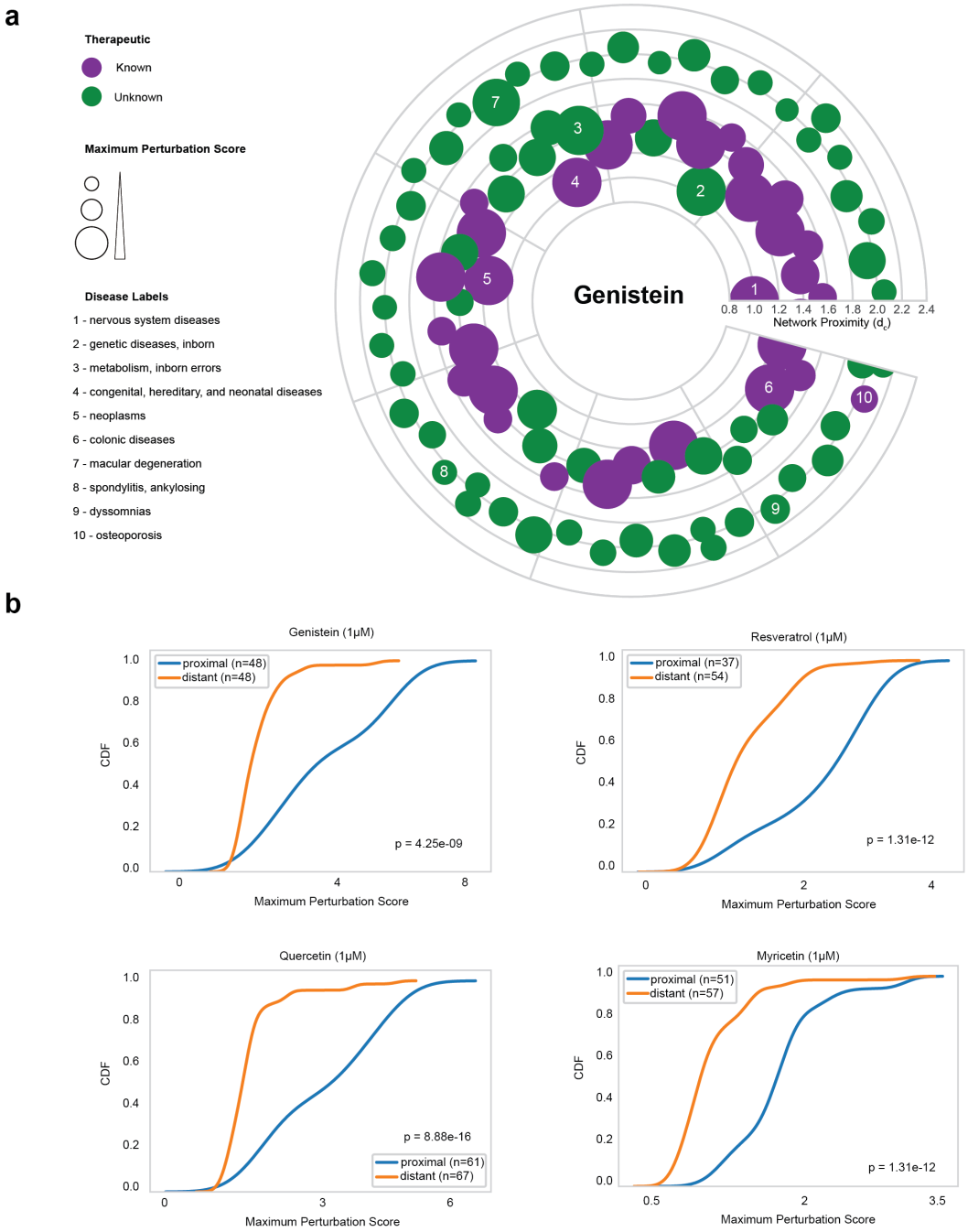


Figure 5 – Diseases Proximal to Polyphenol Targets Have Higher Gene Expression Perturbation Profiles. (a) Schematic representation of proximal and distal diseases in relation to genistein targets. Each node represents a disease and the node size proportional to the perturbation score after treatment with genistein (1 μM, 6 hours). Distance from the origin represents the network proximity (d_c) to genistein targets. Purple nodes represent diseases in which the therapeutic association was previously known. (c) Cumulative distribution of the maximum perturbation scores of genes from diseases that are distal or proximal to polyphenol targets considering different polyphenols (1 μM, 6 hours): genistein, quercetin, resveratrol, and myricetin. Statistical significance was evaluated with the Kolmogorov-Smirnov test.

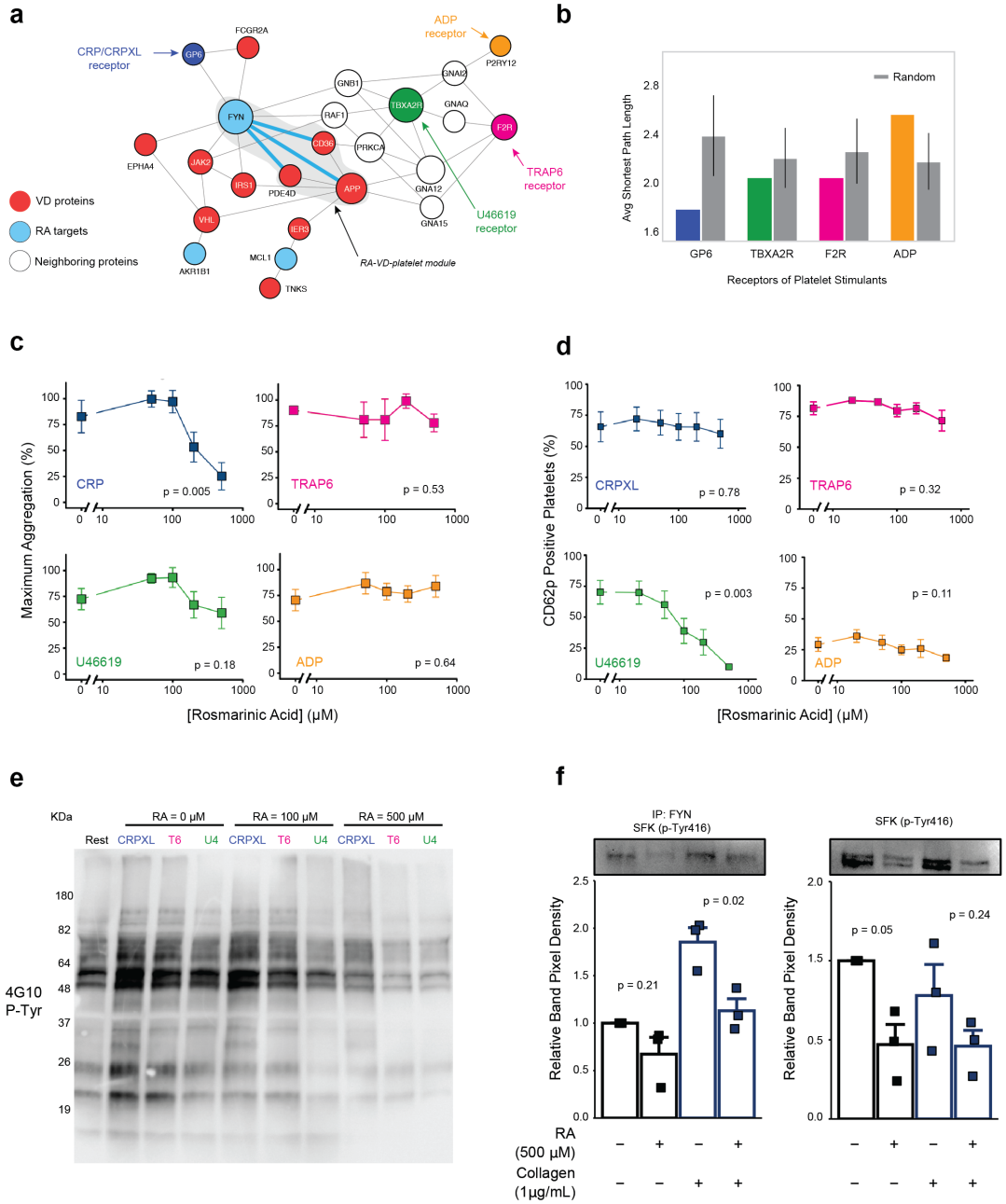


Figure 6 – Rosmarinic Acid Modulates Platelet Function. (a) Interactome neighborhood showing rosmarinic acid (RA) targets and the RA-VD-platelet module - the connected component formed by the RA target FYN and the VD proteins associated with platelet function PDE4D, CD36, and APP – and the receptor for platelet agonists used in our experiments (collagen/CRPXL, TRAP6, U46619, and ADP). (b) Average shortest path length from each platelet agonist receptor and the RA-VD-platelet module formed by the proteins FYN, PDE4D, CD36, APP. Bars represent standard deviation of that same measure over 1000 iterations of random selection of nodes in a degree preserving fashion. c-e) Platelet-rich plasma (PRP) or washed platelets were pre-treated with RA for 1 hour before stimulation with either collagen (1

607 μg/mL), collagen-related peptide (CRP-XL, 1μg/mL), thrombin receptor activator peptide-6 (TRAP-6, 20
608 μM), U46619 (1 μM), or ADP (10 μM). Platelets were assessed for either (c) aggregation, (d) alpha
609 granule secretion. Platelet lysates were also probed for either (e) non-specific tyrosine phosphorylation
610 (p-Tyr) of the whole cell lysate, or (d) site-specific phosphorylation of src family kinases (SFKs) and FYN
611 at residue 416. n = 3-6 separate blood donations, mean +/- SEM. p-values in (c) and (d) were determined
612 by Kruskal-Wallis test and by unpaired t.tests in (f).
613

614 References

615

616 1. Khera, A. V. *et al.* Genetic Risk, Adherence to a Healthy Lifestyle, and Coronary
617 Disease. *N. Engl. J. Med.* **375**, 2349–2358 (2016).
618 2. Arts, I. C. W. & Hollman, P. C. H. Polyphenols and disease risk in epidemiologic
619 studies. *Am. J. Clin. Nutr.* **81**, 317S-325S (2005).
620 3. Wang, X., Ouyang, Y. Y., Liu, J. & Zhao, G. Flavonoid intake and risk of CVD: A
621 systematic review and meta-analysis of prospective cohort studies. *Br. J. Nutr.*
622 **111**, 1–11 (2014).
623 4. Neveu, V. *et al.* Phenol-Explorer: an online comprehensive database on
624 polyphenol contents in foods. *Database* **2010**, bap024–bap024 (2010).
625 5. Pérez-Jiménez, J., Neveu, V., Vos, F. & Scalbert, A. Systematic analysis of the
626 content of 502 Polyphenols in 452 foods and beverages: An application of the
627 phenol-explorer database. *J. Agric. Food Chem.* **58**, 4959–4969 (2010).
628 6. Zhang, H. & Tsao, R. Dietary polyphenols, oxidative stress and antioxidant and
629 anti-inflammatory effects. *Curr. Opin. Food Sci.* (2016)
630 doi:10.1016/j.cofs.2016.02.002.
631 7. Boly, R. *et al.* Quercetin inhibits a large panel of kinases implicated in cancer cell
632 biology. *Int. J. Oncol.* **38**, 833–842 (2011).
633 8. Lacroix, S. *et al.* A computationally driven analysis of the polyphenol-protein
634 interactome. *Sci. Rep.* **8**, 2232 (2018).
635 9. Hanhineva, K. *et al.* Impact of dietary polyphenols on carbohydrate metabolism.
636 *Int. J. Mol. Sci.* **11**, 1365–402 (2010).
637 10. Hervert-Hernández, D. & Goñi, I. Dietary polyphenols and human gut microbiota:
638 A review. *Food Rev. Int.* **27**, 154–169 (2011).
639 11. Zhang, S. *et al.* Dietary pomegranate extract and inulin affect gut microbiome
640 differentially in mice fed an obesogenic diet. *Anaerobe* **48**, 184–193 (2017).
641 12. Thazhath, S. S. *et al.* Administration of resveratrol for 5 wk has no effect on
642 glucagon-like peptide 1 secretion, gastric emptying, or glycemic control in type 2
643 diabetes: a randomized controlled trial. *Am. J. Clin. Nutr.* **103**, 66–70 (2016).
644 13. Bhatt, J. K., Thomas, S. & Nanjan, M. J. Resveratrol supplementation improves

645 glycemic control in type 2 diabetes mellitus. *Nutr. Res.* **32**, 537–41 (2012).
646 14. Sharma, A. *et al.* A disease module in the interactome explains disease
647 heterogeneity, drug response and captures novel pathways and genes in asthma.
648 *Hum. Mol. Genet.* **24**, 3005–3020 (2014).
649 15. Menche, J. *et al.* Disease networks. Uncovering disease-disease relationships
650 through the incomplete interactome. *Science* **347**, 1257601 (2015).
651 16. Guney, E., Menche, J., Vidal, M. & Barabási, A.-L. Network-based in silico drug
652 efficacy screening. *Nat. Commun.* **7**, 10331 (2016).
653 17. Cheng, F. *et al.* Network-based approach to prediction and population-based
654 validation of in silico drug repurposing. *Nat. Commun.* **9**, 1–12 (2018).
655 18. Davis, A. P. *et al.* The Comparative Toxicogenomics Database: update 2019.
656 *Nucleic Acids Res.* **47**, D948–D954 (2019).
657 19. Szklarczyk, D. *et al.* STRING v10: protein-protein interaction networks, integrated
658 over the tree of life. *Nucleic Acids Res.* **43**, D447-52 (2015).
659 20. Kovács, I. A. *et al.* Network-based prediction of protein interactions. *Nat.*
660 *Commun.* **10**, 1240 (2019).
661 21. Sarkar, F. H., Li, Y., Wang, Z. & Kong, D. Cellular signaling perturbation by
662 natural products. *Cell. Signal.* **21**, 1541–7 (2009).
663 22. Iso, H. *et al.* The relationship between green tea and total caffeine intake and risk
664 for self-reported type 2 diabetes among Japanese adults. *Ann. Intern. Med.* **144**,
665 554–62 (2006).
666 23. Song, Y., Manson, J. E., Buring, J. E., Sesso, H. D. & Liu, S. Associations of
667 dietary flavonoids with risk of type 2 diabetes, and markers of insulin resistance
668 and systemic inflammation in women: a prospective study and cross-sectional
669 analysis. *J. Am. Coll. Nutr.* **24**, 376–84 (2005).
670 24. Keske, M. A. *et al.* Vascular and metabolic actions of the green tea polyphenol
671 epigallocatechin gallate. *Curr. Med. Chem.* **22**, 59–69 (2015).
672 25. Wolfram, S. *et al.* Epigallocatechin gallate supplementation alleviates diabetes in
673 rodents. *J. Nutr.* **136**, 2512–8 (2006).
674 26. Subramanian, A. *et al.* A Next Generation Connectivity Map: L1000 Platform and
675 the First 1,000,000 Profiles. *Cell* **171**, 1437-1452.e17 (2017).

676 27. Gil-Longo, J. & González-Vázquez, C. Vascular pro-oxidant effects secondary to
677 the autoxidation of gallic acid in rat aorta. *J. Nutr. Biochem.* **21**, 304–9 (2010).
678 28. Badavi, M., Bazaz, A., Dianat, M. & Sarkaki, A. Gallic acid improves endothelium-
679 dependent vasodilatory response to histamine in the mesenteric vascular bed of
680 diabetic rats. *J. Diabetes* **9**, 1003–1011 (2017).
681 29. Jin, L. *et al.* Gallic Acid Reduces Blood Pressure and Attenuates Oxidative Stress
682 and Cardiac Hypertrophy in Spontaneously Hypertensive Rats. *Sci. Rep.* **7**, 1–14
683 (2017).
684 30. de Oliveira, L. M. *et al.* The vasorelaxant effect of gallic acid involves
685 endothelium-dependent and -independent mechanisms. *Vascul. Pharmacol.* **81**,
686 69–74 (2016).
687 31. Badhani, B., Sharma, N. & Kakkar, R. Gallic acid: a versatile antioxidant with
688 promising therapeutic and industrial applications. *RSC Adv.* **5**, 27540–27557
689 (2015).
690 32. Couto, J. A. *et al.* Somatic MAP2K1 Mutations Are Associated with Extracranial
691 Arteriovenous Malformation. *Am. J. Hum. Genet.* **100**, 546–554 (2017).
692 33. Séverin, S. *et al.* Distinct and overlapping functional roles of Src family kinases in
693 mouse platelets. *J. Thromb. Haemost.* **10**, 1631–45 (2012).
694 34. Reddy, K. B., Smith, D. M. & Plow, E. F. Analysis of Fyn function in hemostasis
695 and alphaIIbbeta3-integrin signaling. *J. Cell Sci.* **121**, 1641–8 (2008).
696 35. Quek, L. S. *et al.* Fyn and Lyn phosphorylate the Fc receptor gamma chain
697 downstream of glycoprotein VI in murine platelets, and Lyn regulates a novel
698 feedback pathway. *Blood* **96**, 4246–53 (2000).
699 36. Ji, R. *et al.* Rosmarinic acid exerts an antagonistic effect on vascular calcification
700 by regulating the Nrf2 signalling pathway. *Free Radic. Res.* 1–11 (2019)
701 doi:10.1080/10715762.2018.1558447.
702 37. Sotnikova, R. *et al.* Rosmarinic acid administration attenuates diabetes-induced
703 vascular dysfunction of the rat aorta. *J. Pharm. Pharmacol.* **65**, 713–23 (2013).
704 38. Han, J. *et al.* Rosmarinic Acid Protects against Inflammation and Cardiomyocyte
705 Apoptosis during Myocardial Ischemia/Reperfusion Injury by Activating
706 Peroxisome Proliferator-Activated Receptor Gamma. *Front. Pharmacol.* **8**, 456

707 (2017).
708 39. Grover, S. P., Bergmeier, W. & Mackman, N. Platelet Signaling Pathways and
709 New Inhibitors. *Arterioscler. Thromb. Vasc. Biol.* **38**, e28–e35 (2018).
710 40. Moco, S., Martin, F. P. J. & Rezzi, S. Metabolomics view on gut microbiome
711 modulation by polyphenol-rich foods. *J. Proteome Res.* **11**, 4781–4790 (2012).
712 41. van Duynhoven, J. *et al.* Metabolic fate of polyphenols in the human
713 superorganism. *Proc. Natl. Acad. Sci.* **108**, 4531–4538 (2011).
714 42. Ottaviani, J. I., Heiss, C., Spencer, J. P. E., Kelm, M. & Schroeter, H.
715 Recommending flavanols and procyanidins for cardiovascular health: Revisited.
716 *Mol. Aspects Med.* **61**, 63–75 (2018).
717 43. Stalmach, A., Troufflard, S., Serafini, M. & Crozier, A. Absorption, metabolism and
718 excretion of Choladi green tea flavan-3-ols by humans. *Mol. Nutr. Food Res.*
719 (2009) doi:10.1002/mnfr.200800169.
720 44. Meng, X. *et al.* Identification and characterization of methylated and ring-fission
721 metabolites of tea catechins formed in humans, mice, and rats. *Chem. Res.*
722 *Toxicol.* (2002) doi:10.1021/tx010184a.
723 45. Perez-Vizcaino, F., Duarte, J. & Santos-Buelga, C. The flavonoid paradox:
724 Conjugation and deconjugation as key steps for the biological activity of
725 flavonoids. *J. Sci. Food Agric.* **92**, 1822–1825 (2012).
726 46. Shimoi, K. & Nakayama, T. Glucuronidase deconjugation in inflammation.
727 *Methods Enzymol.* **400**, 263–272 (2005).
728 47. Kaneko, A. *et al.* Glucuronides of phytoestrogen flavonoid enhance macrophage
729 function via conversion to aglycones by β-glucuronidase in macrophages. *Immun.*
730 *Inflamm. Dis.* **5**, 265–279 (2017).
731 48. Chao, E. C. & Henry, R. R. SGLT2 inhibition-A novel strategy for diabetes
732 treatment. *Nat. Rev. Drug Discov.* **9**, 551–559 (2010).
733 49. Mereles, D. & Hunstein, W. Epigallocatechin-3-gallate (EGCG) for clinical trials:
734 More Pitfalls than Promises? *Int. J. Mol. Sci.* **12**, 5592–5603 (2011).
735 50. Jensen, K., Ni, Y., Panagiotou, G. & Kouskoumvekaki, I. Developing a Molecular
736 Roadmap of Drug-Food Interactions. *PLoS Comput. Biol.* **11**, 1–15 (2015).
737 51. Rolland, T. *et al.* A proteome-scale map of the human interactome network. *Cell*

738 **159**, 1212–1226 (2014).
739 52. Cheng, F., Jia, P., Wang, Q. & Zhao, Z. Quantitative network mapping of the
740 human kinome interactome reveals new clues for rational kinase inhibitor
741 discovery and individualized cancer therapy. *Oncotarget* **5**, 3697–710 (2014).
742 53. Calçada, D. *et al.* The role of low-grade inflammation and metabolic flexibility in
743 aging and nutritional modulation thereof: A systems biology approach. *Mech.*
744 *Ageing Dev.* (2014) doi:10.1016/j.mad.2014.01.004.
745 54. Hornbeck, P. V *et al.* PhosphoSitePlus, 2014: mutations, PTMs and
746 recalibrations. *Nucleic Acids Res.* **43**, D512-20 (2015).
747 55. Li, T. *et al.* A scored human protein-protein interaction network to catalyze
748 genomic interpretation. *Nat. Methods* **14**, 61–64 (2016).
749 56. Chatr-Aryamontri, A. *et al.* The BioGRID interaction database: 2017 update.
750 *Nucleic Acids Res.* **45**, D369–D379 (2017).
751 57. Cowley, M. J. *et al.* PINA v2.0: mining interactome modules. *Nucleic Acids Res.*
752 **40**, D862-5 (2012).
753 58. Peri, S. *et al.* Human protein reference database as a discovery resource for
754 proteomics. *Nucleic Acids Res.* **32**, D497-501 (2004).
755 59. Orchard, S. *et al.* The MIntAct project--IntAct as a common curation platform for
756 11 molecular interaction databases. *Nucleic Acids Res.* **42**, D358-63 (2014).
757 60. Breuer, K. *et al.* InnateDB: systems biology of innate immunity and beyond--
758 recent updates and continuing curation. *Nucleic Acids Res.* **41**, D1228-33 (2013).
759 61. Meyer, M. J., Das, J., Wang, X. & Yu, H. INstruct: a database of high-quality 3D
760 structurally resolved protein interactome networks. *Bioinformatics* **29**, 1577–9
761 (2013).
762 62. Mosca, R., Céol, A. & Aloy, P. Interactome3D: adding structural details to protein
763 networks. *Nat. Methods* **10**, 47–53 (2013).
764 63. Meyer, M. J. *et al.* Interactome INSIDER: a structural interactome browser for
765 genomic studies. *Nat. Methods* **15**, 107–114 (2018).
766 64. Fazekas, D. *et al.* SignaLink 2 - a signaling pathway resource with multi-layered
767 regulatory networks. *BMC Syst. Biol.* **7**, 7 (2013).
768 65. Huttlin, E. L. *et al.* Architecture of the human interactome defines protein

769 communities and disease networks. *Nature* **545**, 505–509 (2017).
770 66. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based
771 approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.*
772 *U. S. A.* **102**, 15545–50 (2005).
773 67. Roweth, H. G. *et al.* Two novel, putative mechanisms of action for citalopram-
774 induced platelet inhibition. *Sci. Rep.* **8**, 1–14 (2018).
775 68. Roweth, H. G. *et al.* Citalopram inhibits platelet function independently of SERT-
776 mediated 5-HT transport. *Sci. Rep.* **8**, 1–14 (2018).
777 69. Nath, S., Bachani, M., Harshavardhana, D. & Steiner, J. P. Catechins protect
778 neurons against mitochondrial toxins and HIV proteins via activation of the BDNF
779 pathway. *J. Neurovirol.* **18**, 445–455 (2012).
780 70. Park, K.-S. *et al.* (-)-Epigallocatethin-3-O-gallate counteracts caffeine-induced
781 hyperactivity: evidence of dopaminergic blockade. *Behav. Pharmacol.* **21**, 572–
782 575 (2010).
783 71. Ramesh, E., Geraldine, P. & Thomas, P. A. Regulatory effect of epigallocatechin
784 gallate on the expression of C-reactive protein and other inflammatory markers in
785 an experimental model of atherosclerosis. *Chem. Biol. Interact.* **183**, 125–32
786 (2010).
787 72. Han, S. G., Han, S.-S., Toborek, M. & Hennig, B. EGCG protects endothelial cells
788 against PCB 126-induced inflammation through inhibition of AhR and induction of
789 Nrf2-regulated genes. *Toxicol. Appl. Pharmacol.* **261**, 181–8 (2012).
790 73. Sheng, R., Gu, Z.-L. & Xie, M.-L. Epigallocatechin gallate, the major component of
791 polyphenols in green tea, inhibits telomere attrition mediated cardiomyocyte
792 apoptosis in cardiac hypertrophy. *Int. J. Cardiol.* **162**, 199–209 (2013).
793 74. Devika, P. T. & Stanely Mainzen Prince, P. (-)Epigallocatechingallate protects the
794 mitochondria against the deleterious effects of lipids, calcium and adenosine
795 triphosphate in isoproterenol induced myocardial infarcted male Wistar rats. *J.*
796 *Appl. Toxicol.* **28**, 938–44 (2008).
797 75. Yi, Q.-Y. *et al.* Chronic infusion of epigallocatechin-3-O-gallate into the
798 hypothalamic paraventricular nucleus attenuates hypertension and
799 sympathoexcitation by restoring neurotransmitters and cytokines. *Toxicol. Lett.*

800 **262**, 105–113 (2016).
801 76. Devika, P. T. & Prince, P. S. M. Preventive effect of (-)epigallocatechin-gallate
802 (EGCG) on lysosomal enzymes in heart and subcellular fractions in isoproterenol-
803 induced myocardial infarcted Wistar rats. *Chem. Biol. Interact.* **172**, 245–52
804 (2008).
805 77. Hushmendy, S. *et al.* Select phytochemicals suppress human T-lymphocytes and
806 mouse splenocytes suggesting their use in autoimmunity and transplantation.
807 *Nutr. Res.* **29**, 568–78 (2009).
808 78. Shen, K. *et al.* Epigallocatechin 3-Gallate Ameliorates Bile Duct Ligation Induced
809 Liver Injury in Mice by Modulation of Mitochondrial Oxidative Stress and
810 Inflammation. *PLoS One* **10**, e0126278 (2015).
811 79. ZHEN, M. *et al.* Green tea polyphenol epigallocatechin-3-gallate inhibits oxidative
812 damage and preventive effects on carbon tetrachloride–induced hepatic fibrosis.
813 *J. Nutr. Biochem.* **18**, 795–805 (2007).
814 80. Yasuda, Y. *et al.* (−)-Epigallocatechin gallate prevents carbon tetrachloride-
815 induced rat hepatic fibrosis by inhibiting the expression of the PDGFRβ and IGF-
816 1R. *Chem. Biol. Interact.* **182**, 159–164 (2009).
817 81. Cao, W. *et al.* iTRAQ-based proteomic analysis of combination therapy with
818 taurine, epigallocatechin gallate, and genistein on carbon tetrachloride-induced
819 liver fibrosis in rats. *Toxicol. Lett.* **232**, 233–245 (2015).
820 82. Kitamura, M. *et al.* Epigallocatechin gallate suppresses peritoneal fibrosis in mice.
821 *Chem. Biol. Interact.* **195**, 95–104 (2012).
822 83. Sakla, M. S. & Lorson, C. L. Induction of full-length survival motor neuron by
823 polyphenol botanical compounds. *Hum. Genet.* **122**, 635–643 (2008).
824 84. Shimizu, M. *et al.* (−)-Epigallocatechin gallate inhibits growth and activation of the
825 VEGF/VEGFR axis in human colorectal cancer cells. *Chem. Biol. Interact.* **185**,
826 247–252 (2010).
827