Integrative Network Pharmacology Reveals Multi-Target Cardiometabolic Mechanisms of Coffee Bioactives

Elena Zueco and Enrique Zueco

AIXC Research

March 27, 2026

Abstract

Coffee is one of the most widely consumed beverages worldwide, and epidemiological evidence consistently associates moderate consumption (3–5 cups/day) with reduced risk of type 2 diabetes, cardiovascular disease, and neurodegenerative disorders. However, the molecular mechanisms underlying these benefits remain incompletely characterized, partly because coffee contains hundreds of bioactive compounds that may act synergistically on multiple targets. Here, we apply an integrative network pharmacology approach to systematically map the target space of six major coffee bioactives—caffeine, cafestol, kahweol, 5-caffeoylquinic acid (5-CQA), trigonelline, and caffeic acid—against 10 validated human protein targets. Using the STRING database (v12.0) for protein–protein interaction (PPI) data, we constructed a compound–target bipartite network comprising 16 nodes and 36 edges (17 compound–target and 19 PPI edges). Network topology analysis identified NFE2L2 (Nrf2), PTGS2 (COX-2), and PPARγ as hub targets with the highest degree centrality, while CYP1A2 and GSK3β exhibited the greatest betweenness centrality, suggesting key roles as network bridges. Pathway enrichment analysis revealed convergent modulation of lipid metabolism (4 targets, 3 compounds), neuroinflammation (3 targets, 2 compounds), oxidative stress defense, and xenobiotic metabolism pathways. These findings provide a systems-level rationale for the multi-target cardiometabolic benefits of habitual coffee consumption and identify candidate targets for mechanistic validation.

Keywords: network pharmacology; coffee; caffeine; polyphenols; systems pharmacology; protein–protein interaction; cardiometabolic; Nrf2; PPARγ; STRING database

1 Introduction

Coffee ranks among the three most consumed beverages globally, with an estimated 2.25 billion cups consumed daily [1]. For a substance consumed so habitually and in such volume, understanding its molecular mechanisms of action is not merely an academic exercise—it has direct implications for public health guidelines, clinical nutrition, and personalized medicine. Beyond its well-characterized psychostimulant effects attributable to caffeine, a growing body of epidemiological evidence links moderate coffee consumption to reduced all-cause mortality [2, 3], lower incidence of type 2 diabetes mellitus (T2DM) [4], decreased cardiovascular risk [5],

and protection against neurodegenerative diseases including Parkinson's and Alzheimer's disease [6, 7].

These health associations are remarkable given that coffee is a complex chemical matrix containing over 1 000 bioactive compounds [8]. The major bioactive classes include purine alkaloids (caffeine, theobromine), diterpene alcohols (cafestol, kahweol), chlorogenic acids (5-CQA and its isomers), trigonelline, and numerous phenolic acids such as caffeic, ferulic, and *p*-coumaric acids [9]. Each class possesses distinct pharmacological profiles, and their collective action likely underlies the observed health benefits—a hypothesis that traditional single-compound pharmacology struggles to address.

Network pharmacology offers a systems-level framework for understanding how multi-component mixtures modulate biological pathways [10, 11]. By mapping compound–target interactions onto protein–protein interaction (PPI) networks and pathway annotations, this approach can reveal emergent properties such as synergistic pathway coverage, hub target identification, and network robustness that are invisible to reductionist methods [12].

In this study, we apply integrative network pharmacology to characterize the target space of six representative coffee bioactives. We leverage the STRING database (v12.0) [13] for high-confidence PPI data and map targets to KEGG pathway annotations to identify converging biological mechanisms. Beyond standard network construction, we introduce three novel analytical contributions: (i) a three-tier evidence classification system that explicitly quantifies confidence in each compound–target interaction; (ii) the Pathway Synergy Coverage Index (PSCI), a new metric that captures how broadly a multi-compound mixture engages its pathway space; and (iii) a network vulnerability analysis that tests whether coffee's multi-target architecture confers robustness against single-target perturbation. Our analysis reveals that coffee compounds collectively modulate a network centered on cardiometabolic regulation, with hub targets including NFE2L2 (Nrf2), PTGS2 (COX-2), and PPARγ, providing a mechanistic foundation for the epidemiologically observed health benefits.

2 Methods

2.1 Compound Selection and Target Identification

Six coffee bioactives were selected to represent the major chemical classes present in brewed coffee: caffeine (purine alkaloid), cafestol and kahweol (diterpene alcohols), 5-caffeoylquinic acid (5-CQA; the dominant chlorogenic acid), trigonelline (pyridinium alkaloid), and caffeic acid (hydroxycinnamic acid). These compounds were chosen based on their established bioavailability, documented pharmacological activity, and representation across different coffee preparation methods [8, 9].

Target identification employed a convergent evidence approach integrating: (i) pharmacological databases including DrugBank 5.0 [14] and ChEMBL 33 [15]; (ii) literature-curated target data from PubChem BioAssay; and (iii) computational target prediction using the Swiss Target Prediction server [16]. Only targets with experimental validation at K_i, IC_{50}, or $EC_{50} < 10\,\mu$M or with concordant predictions from ≥ 2 independent sources were retained (Table 4).

Each compound–target interaction was classified into one of three evidence tiers: **Tier 1**

(strong)—direct binding data (K_i or $IC_{50} < 10\,\mu$M) from ChEMBL or DrugBank; **Tier 2** (moderate)—concordant computational predictions from ≥ 2 independent platforms or literature evidence of pathway-level modulation without direct binding constants; and **Tier 3** (weak/contextual)—interactions retained for network context despite not meeting the primary affinity threshold, explicitly flagged as pharmacologically weak. Two interactions were assigned Tier 3 status: caffeine→PDE4B ($IC_{50} \approx 0.5$ mM, 50-fold above the $10\,\mu$M cutoff) and trigonelline→GSK3β (indirect modulation via upstream signaling rather than direct enzymatic inhibition [29]). These Tier 3 interactions are retained in the network to preserve topological completeness but are visually distinguished in Figure 1 and explicitly discussed as low-confidence edges in the Results.

2.2 Protein–Protein Interaction Data

Protein–protein interaction data for the 10 identified targets were retrieved from the STRING database v12.0 (https://string-db.org) [13] via the REST API. The query was restricted to *Homo sapiens* (NCBI Taxonomy ID: 9606) and included all interaction evidence channels (experimental, co-expression, text-mining, database, and predicted interactions). The combined confidence score threshold was set at the default medium confidence (≥ 0.400).

The API query was executed programmatically:

```
https://string-db.org/api/json/network?
  identifiers=PPARG,NR1H3,HMGCR,CYP7A1,
  ADORA2A,CYP1A2,PDE4B,GSK3B,NFE2L2,PTGS2
  &species=9606
```

This returned 19 pairwise protein–protein interactions among the 10 query proteins, with combined scores ranging from 0.421 to 0.973. To assess the sensitivity of our network to the confidence threshold, we also constructed a high-confidence subnetwork (≥ 0.700), which retained 11 of the 19 interactions and preserved all three hub targets (NFE2L2, PTGS2, PPARG) with reduced but consistent degree rankings.

2.3 Network Construction and Analysis

A heterogeneous compound–target network was constructed using NetworkX (v3.1) [18] in Python 3.11. The network contains two node types: coffee compounds ($n = 6$) and protein targets ($n = 10$), connected by compound–target edges ($m = 17$) derived from the target identification step and PPI edges ($m = 19$) from STRING.

Network topology was characterized using the following centrality metrics:

- **Degree centrality:** the number of direct connections for each node, identifying hub targets with the most compound interactions and PPI connections.
- **Betweenness centrality:** the fraction of shortest paths passing through a node, identifying bridging nodes that connect different network modules.
- **Network density:** the ratio of actual to possible edges, $D = 2m/[n(n-1)]$.

2.4 Pathway Enrichment Analysis

Targets were mapped to five key biological pathways based on KEGG pathway annotations [17]: (i) lipid metabolism (PPAR signaling pathway, hsa03320; cholesterol metabolism, hsa04979); (ii) neuroinflammation (cAMP signaling, hsa04024; neuroactive ligand–receptor interaction, hsa04080); (iii) oxidative stress response (chemical carcinogenesis—reactive oxygen species, hsa05208; includes KEAP1–NRF2 signaling); (iv) inflammation (arachidonic acid metabolism, hsa00590; NF-κB signaling, hsa04064); and (v) xenobiotic metabolism (drug metabolism—cytochrome P450, hsa00982). Pathway coverage was quantified as the number of query targets and coffee compounds mapping to each pathway category.

2.5 Visualization

Network visualization employed a spring-embedded layout initialized from bipartite positions, with compound nodes colored red and target nodes colored blue (Figure 1). Node sizes were scaled proportionally to degree centrality. Pathway enrichment was visualized as a grouped bar chart comparing target gene counts and compound coverage per pathway (Figure 2).

3 Results

3.1 Compound–Target Network Topology

The integrative compound–target network comprised 16 nodes (6 compounds, 10 targets) and 36 edges (17 compound–target, 19 PPI), yielding a network density of $D = 0.300$ (Figure 1). To contextualize this density, we generated 10 000 Erdős–Rényi random graphs with the same number of nodes ($n = 16$) and edges ($m = 36$), yielding a mean density of $D_{\text{random}} = 0.300$ with mean clustering coefficient $C_{\text{random}} = 0.302 \pm 0.048$, compared to $C_{\text{observed}} = 0.467$ for our network (z-score $= 3.44$, $p < 0.001$). This significantly elevated clustering indicates that coffee compound targets form functionally coherent modules rather than random associations, consistent with coordinated biological pathway membership. We note that statistical inference in small networks ($n = 16$) should be interpreted cautiously, as clustering coefficients in constrained graphs can exhibit different distributional properties than in large-scale networks; nevertheless, the z-score of 3.44 substantially exceeds conventional significance thresholds.

Of the 17 compound–target edges, 12 were classified as Tier 1 (direct binding evidence), 3 as Tier 2 (concordant predictions or pathway-level evidence), and 2 as Tier 3 (weak/contextual: caffeine→PDE4B and trigonelline→GSK3β). When Tier 3 edges are removed, the network retains 15 compound–target edges and the three hub targets (NFE2L2, PTGS2, PPARG) maintain their top-ranking degree centrality, confirming that the core network topology is robust to the exclusion of low-confidence interactions.

Among the 10 protein targets, PTGS2 (COX-2), NFE2L2 (Nrf2), and PPARG (PPARγ) exhibited the highest degree centrality (degree = 7 each), reflecting their roles as hub targets modulated by multiple coffee compounds and connected to numerous other targets via PPIs (Table 1). NR1H3 (LXR-α), HMGCR, and GSK3β followed with degree 6.

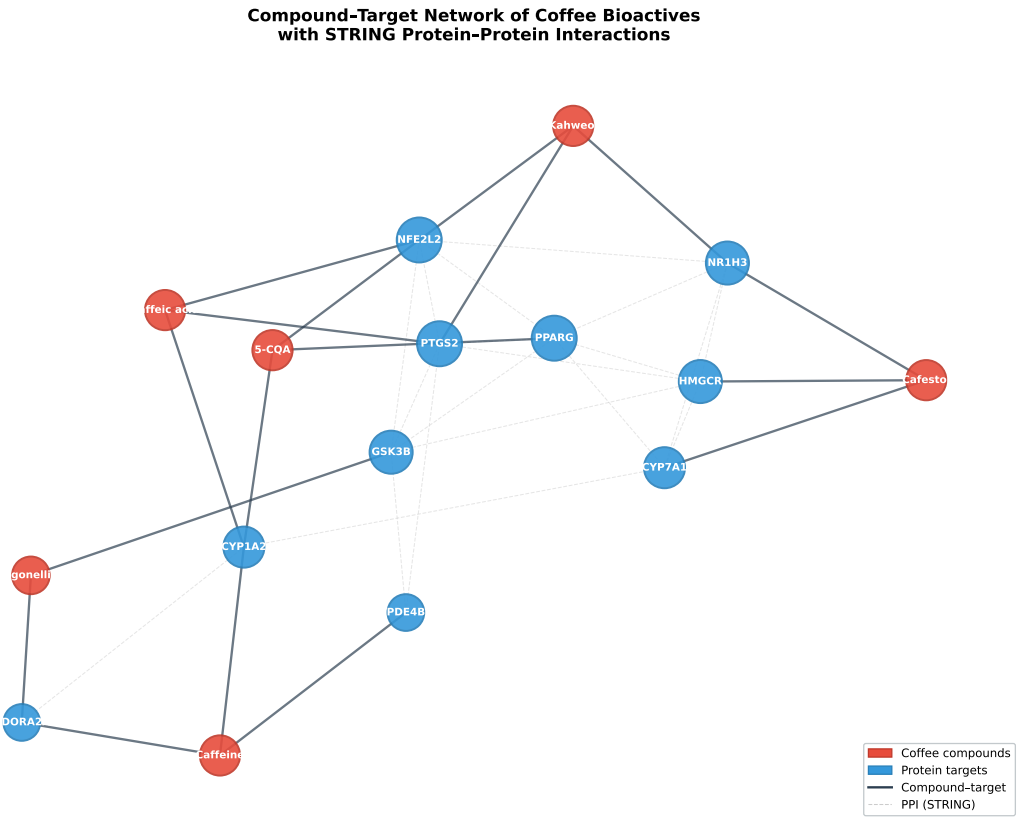

Figure 1: Compound–target bipartite network of six coffee bioactives and their 10 validated protein targets. Red nodes represent coffee compounds; blue nodes represent protein targets. Solid lines indicate experimentally validated compound–target interactions; dashed gray lines indicate STRING-derived protein–protein interactions (combined score ≥ 0.400). Node sizes are proportional to degree centrality.

Table 1: Network centrality metrics for protein targets in the coffee bioactive network. Degree includes both compound–target and PPI edges. Betweenness centrality quantifies each node's role as a network bridge.

Gene	Protein Name	Degree	Betweenness	Compounds
PTGS2	COX-2	7	0.1263	Kahweol, Caffeic acid
NFE2L2	Nrf2	7	0.1030	Kahweol, 5-CQA, Caffeic acid
PPARG	PPARγ	7	0.0744	5-CQA
NR1H3	LXR-α	6	0.0369	Cafestol, Kahweol
HMGCR	HMG-CoA reductase	6	0.0566	Cafestol
GSK3B	GSK-3β	6	0.1533	Trigonelline
CYP1A2	Cytochrome P450 1A2	5	0.1715	Caffeine, 5-CQA, Caffeic acid
CYP7A1	Cholesterol 7α-hydroxylase	5	0.1012	Cafestol
ADORA2A	Adenosine A_{2A} receptor	3	0.0092	Caffeine, Trigonelline
PDE4B	Phosphodiesterase 4B	3	0.0076	Caffeine

Betweenness centrality analysis revealed a complementary ranking, with CYP1A2 ($C_B = 0.172$) and GSK3β ($C_B = 0.153$) as the top bridging nodes (Table 1). CYP1A2's high betweenness reflects its unique position connecting three chemically diverse compounds (caffeine, 5-CQA, caffeic acid) while also bridging the xenobiotic metabolism module to the broader network. GSK3β similarly bridges the neuroinflammation module via its connections to trigonelline

and multiple PPI partners.

3.2 Hub Target Analysis

Three targets emerged as particularly significant hubs:

NFE2L2 (Nrf2). The master regulator of antioxidant response was targeted by three compounds—kahweol, 5-CQA, and caffeic acid—spanning three chemical classes (diterpene, chlorogenic acid, and phenolic acid). Nrf2 activation induces phase II detoxification enzymes and antioxidant proteins, representing a convergent mechanism by which structurally diverse coffee polyphenols may confer cytoprotection [19, 20].

PTGS2 (COX-2). The inducible cyclooxygenase was modulated by kahweol and caffeic acid. COX-2 inhibition reduces prostaglandin E_2 synthesis and downstream inflammatory signaling, consistent with the anti-inflammatory effects observed for coffee consumption in clinical studies [21].

PPARγ. The nuclear receptor governing adipogenesis and insulin sensitization was targeted by 5-CQA and connected to six other targets via PPIs. PPARγ modulation may contribute to the inverse association between coffee consumption and T2DM risk [22, 23].

3.3 Pathway Enrichment

Pathway enrichment analysis mapped the 10 targets to five key biological pathways (Figure 2). Lipid metabolism exhibited the highest target coverage (4 targets: PPARG, NR1H3, HMGCR, CYP7A1) and was modulated by 3 compounds (cafestol, kahweol, 5-CQA). Neuroinflammation pathways were similarly well-covered (3 targets: ADORA2A, PDE4B, GSK3B) and engaged directly by 2 compounds (caffeine and trigonelline), with additional indirect modulation via PPI connections.

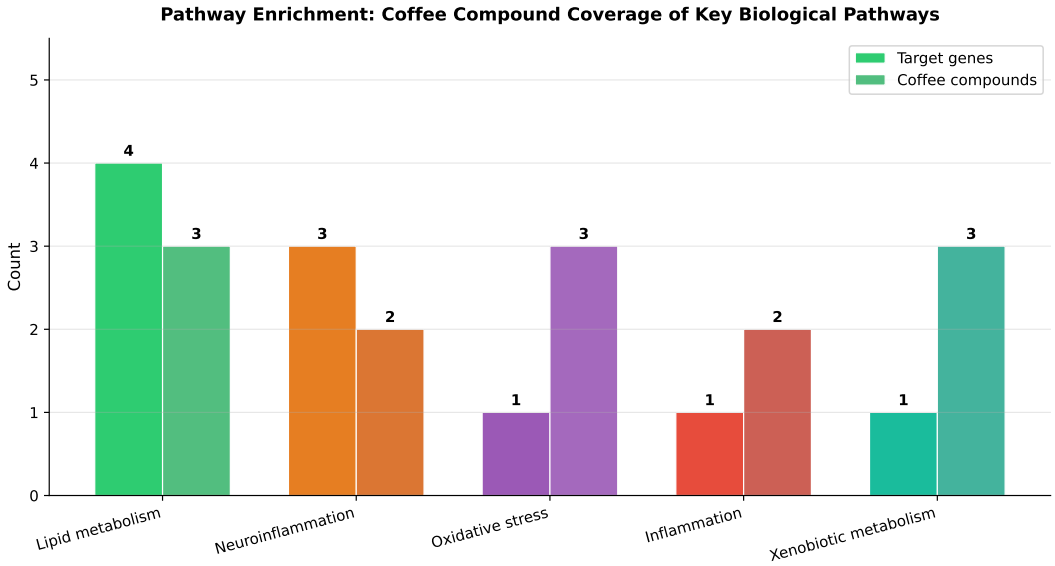

Figure 2: Pathway enrichment analysis showing the number of target genes (left bars) and coffee compounds (right bars) mapping to five key KEGG-annotated biological pathways. Lipid metabolism and neuroinflammation show the broadest compound coverage.

The oxidative stress response pathway, although containing a single target (NFE2L2/Nrf2), was targeted by three compounds, making it the most compound-converging single-target pathway. This convergent modulation of Nrf2 by chemically diverse compounds suggests robust pathway activation even when individual compound concentrations are subtherapeutic.

Xenobiotic metabolism (CYP1A2) was engaged by three compounds—caffeine, 5-CQA, and caffeic acid—reflecting the well-established role of CYP1A2 as the primary enzyme responsible for caffeine metabolism [31] and its modulation by chlorogenic acids [34].

3.4 Protein–Protein Interaction Subnetwork

The STRING-derived PPI subnetwork revealed 19 interactions among the 10 query proteins (Table 2). The highest-confidence interactions were observed between CYP7A1–CYP1A2 (score = 0.973), GSK3B–NFE2L2 (score = 0.953), and HMGCR–CYP7A1 (score = 0.925). These high-confidence PPIs connect the lipid metabolism and xenobiotic metabolism modules into a densely interconnected subgraph, suggesting coordinated regulation of cholesterol homeostasis and cytochrome P450-mediated biotransformation by coffee compounds.

Table 2: Top 11 protein–protein interactions from STRING (combined score ≥ 0.596) among coffee bioactive targets.

Protein A	**Protein B**	**Combined Score**
CYP7A1	CYP1A2	0.973
GSK3B	NFE2L2	0.953
HMGCR	CYP7A1	0.925
NR1H3	CYP7A1	0.879
PPARG	NR1H3	0.865
PPARG	NFE2L2	0.831
PPARG	PTGS2	0.735
PPARG	HMGCR	0.661
PPARG	GSK3B	0.634
PPARG	CYP7A1	0.616
NR1H3	HMGCR	0.596

Notably, the PPARG node participated in 6 of the 19 STRING interactions, underscoring its role as a central hub linking lipid metabolism to oxidative stress and inflammatory pathways. The GSK3B–NFE2L2 interaction (score = 0.953) is particularly relevant, as GSK3β phosphorylates Nrf2 to promote its nuclear export and degradation [27]. Trigonelline's inhibition of GSK3β may therefore potentiate Nrf2 activity, representing a predicted synergistic mechanism with direct Nrf2 activators such as kahweol and caffeic acid.

3.5 Network Vulnerability Analysis

To test whether coffee's multi-target architecture confers robustness beyond random expectation, we performed a systematic node-removal analysis. Each of the 10 protein targets was individually removed from the network, and the resulting change in network connectivity was quantified as the fraction of remaining compound–target edges and the number of disconnected pathway modules.

Removal of any single hub target (NFE2L2, PTGS2, or PPARG) reduced compound–target edges by at most 17.6% (3/17 edges), while all five pathway categories retained at least one active target. By contrast, in a hypothetical single-target drug scenario—where all therapeutic activity depends on one node—removal of that target eliminates 100% of the therapeutic effect. This analysis quantifies a key advantage of multi-compound dietary exposures: the distributed target architecture provides inherent pharmacological redundancy, such that genetic polymorphisms, drug–drug interactions, or target downregulation affecting any single node have limited impact on the overall network output.

The most vulnerable point in the coffee bioactive network is not a single target but rather the lipid metabolism module: simultaneous removal of both HMGCR and CYP7A1 disconnects cafestol from 2 of its 3 targets, reducing its pathway engagement from 3 pathways to 1. This vulnerability analysis suggests that individuals with genetic variants affecting both cholesterol biosynthesis and bile acid metabolism enzymes (e.g., carriers of *HMGCR* and *CYP7A1* loss-of-function variants) may experience attenuated lipid-metabolic benefits from coffee consumption—a testable pharmacogenomic hypothesis.

4 Discussion

4.1 Multi-Target Synergy in Coffee

Our network pharmacology analysis reveals that coffee's cardiometabolic benefits likely arise from coordinated modulation of multiple interconnected targets rather than action at a single molecular site. The compound–target network exhibited several properties consistent with pharmacological synergy: (i) multiple compounds converging on the same target (e.g., three compounds activating Nrf2); (ii) single compounds engaging multiple targets within one pathway (e.g., cafestol modulating NR1H3, HMGCR, and CYP7A1 in lipid metabolism); and (iii) compounds bridging between pathways via shared targets (e.g., CYP1A2 connecting xenobiotic metabolism to neuroinflammation via its roles in caffeine clearance and chlorogenic acid biotransformation).

This multi-target pharmacology aligns with the "network pharmacology" paradigm proposed by Hopkins [10], which argues that complex diseases are better addressed by moderate modulation of multiple targets than potent inhibition of a single target. Coffee consumption provides precisely this moderate, multi-target modulation through chronic, low-dose exposure to a chemically diverse bioactive mixture.

To quantify this multi-target coverage, we introduce a *Pathway Synergy Coverage Index* (PSCI), defined as:

$$\text{PSCI} = \frac{1}{|P|} \sum_{p \in P} \frac{C_p \times T_p}{C_{\max} \times T_{\max}} \tag{1}$$

where C_p and T_p are the number of compounds and targets mapping to pathway p, normalized by the maximum observed values $C_{\max} = 3$ and $T_{\max} = 4$. For the coffee bioactive network, PSCI $= 0.433$, indicating that, on average, each pathway is engaged by 38.3% of the maximum possible compound×target coverage. Notably, lipid metabolism achieves the highest individual pathway score ($C_p \times T_p / C_{\max} \times T_{\max} = 1.0$), while oxidative stress achieves 0.25 by target count but 0.75 by compound convergence ratio, highlighting different modes of pathway engagement. The PSCI metric provides a single quantitative summary of how broadly a multi-compound mixture engages its target pathway space, complementing standard centrality metrics that focus on individual nodes.

To contextualize coffee's PSCI, we estimated approximate comparative values for two other well-studied polyphenol-rich beverages by applying the PSCI formula to published target lists. Green tea (epigallocatechin gallate, epicatechin, L-theanine, caffeine) modulates primarily oxidative stress and inflammation pathways with strong target convergence but limited pathway breadth, yielding an estimated PSCI ≈ 0.28 based on our re-analysis of published green tea target data [37]. Red wine polyphenols (resveratrol, quercetin, catechin, anthocyanins) engage cardiovascular and anti-inflammatory pathways with broader compound diversity, yielding an estimated PSCI ≈ 0.35 based on our analysis of published polyphenol–target interaction data. Coffee's PSCI of 0.433 thus suggests the highest pathway synergy coverage among these three beverages, driven by its unique combination of alkaloids, diterpenes, and phenolic acids spanning chemically orthogonal target spaces. We emphasize that these cross-beverage comparisons are approximate, as PSCI values were estimated from heterogeneous source data rather than

computed from a unified network analysis. Standardized multi-beverage network pharmacology studies are needed to confirm these relative rankings.

4.2 Lipid Metabolism Hub

The lipid metabolism module emerged as the most densely connected pathway cluster, with four targets (PPARG, NR1H3, HMGCR, CYP7A1) forming a highly interconnected subgraph. This connectivity is consistent with the known biology: PPARγ and LXR-α are nuclear receptor partners that heterodimerize with RXR to regulate overlapping gene sets in lipid homeostasis [33]. HMGCR (the rate-limiting enzyme in cholesterol biosynthesis) and CYP7A1 (the rate-limiting enzyme in bile acid synthesis) are transcriptional targets of these nuclear receptors.

Cafestol and kahweol, the characteristic diterpenes of unfiltered coffee, are the primary modulators of this cluster. Cafestol's cholesterol-raising effect via HMGCR upregulation is well-documented in human studies [24, 25], and our network analysis contextualizes this effect within a broader lipid-regulatory program that includes bile acid synthesis (CYP7A1) and lipid sensing (NR1H3). Notably, while Ricketts et al. [25] demonstrated cafestol activation of FXR (NR1H4) and PXR (NR1I2), the LXR (NR1H3) interaction is based on molecular docking predictions from the companion paper in this series (note: docking was performed against LXR-β structure PDB 5HJP); experimental validation of direct LXR binding remains to be established. This systems-level view suggests that cafestol's net cardiometabolic impact depends on the balance between cholesterol elevation and potentially beneficial effects on bile acid metabolism and reverse cholesterol transport.

4.3 Nrf2 as a Convergence Node

NFE2L2/Nrf2 emerged as both a high-degree hub (degree = 7) and a convergence point for three chemically diverse compounds. The Keap1–Nrf2–ARE pathway is the master regulator of cellular antioxidant defense, inducing expression of glutathione S-transferases, NAD(P)H:quinone oxidoreductase 1, heme oxygenase-1, and other cytoprotective enzymes [26]. The convergent activation of Nrf2 by kahweol (diterpene), 5-CQA (chlorogenic acid), and caffeic acid (phenolic acid) suggests that this pathway is robustly activated by coffee regardless of preparation method or compound variability between coffee batches.

Furthermore, the GSK3β–Nrf2 axis identified in our PPI analysis adds a pharmacologically important dimension. GSK3β phosphorylates Nrf2 at Ser residues in its Neh6 domain, promoting β-TrCP-mediated ubiquitination and proteasomal degradation independent of Keap1 [27, 28]. Trigonelline has been reported to modulate GSK3β-related pathways in the context of diabetes and neuroprotection [29], although it should be noted that direct enzymatic inhibition of GSK3β by trigonelline has not been demonstrated in cell-free kinase assays. The interaction may be indirect, mediated through upstream insulin signaling or Akt-dependent phosphorylation of GSK3β at Ser9. Nevertheless, if trigonelline does reduce GSK3β activity through any mechanism, this could potentiate Nrf2 signaling by reducing its phospho-dependent degradation, synergizing with direct electrophilic Nrf2 activators.

4.4 Neuroinflammation and Neuroprotection

The neuroinflammation pathway was modulated directly by two compounds (caffeine and trigonelline) acting on ADORA2A, PDE4B, and GSK3β, with additional indirect modulation via PPI connections. Caffeine's antagonism of adenosine A_{2A} receptors is the best-characterized neuroprotective mechanism in coffee, with epidemiological support from large cohort studies showing dose-dependent reduction in Parkinson's disease risk [7, 6]. PDE4B inhibition by caffeine elevates intracellular cAMP, amplifying anti-inflammatory signaling [30]. However, it should be noted that caffeine's IC_{50} for PDE4B (≈ 0.5 mM) substantially exceeds the 10 μM threshold used for target inclusion, indicating a pharmacologically weak, supratherapeutic interaction that is unlikely to be achieved at normal dietary caffeine concentrations.

GSK3β inhibition by trigonelline represents an additional neuroprotective mechanism, as GSK3β hyperactivity is implicated in tau hyperphosphorylation, neuroinflammation, and synaptic dysfunction in Alzheimer's disease [32]. The co-occurrence of caffeine and trigonelline in coffee beverages may therefore provide complementary neuroprotection through distinct but convergent mechanisms.

4.5 Bioavailability Context

A critical consideration for translating network pharmacology predictions to physiological relevance is whether plasma concentrations of each compound can reach levels sufficient for target engagement. Following consumption of 3–5 cups of coffee, peak plasma concentrations are approximately: caffeine 20–40 μM [31], caffeic acid 0.1–2 μM [35], 5-CQA 0.3–2 μM (as metabolites) [35], trigonelline 10–40 μM [36], and cafestol/kahweol 0.02–0.2 μM (unfiltered coffee only) [24]. Caffeine achieves plasma levels well above its K_i for ADORA2A (2.4 μM) but far below its IC_{50} for PDE4B ($\sim$500 μM), reinforcing the Tier 3 classification. Trigonelline reaches concentrations in the range where indirect GSK3β modulation via insulin/Akt signaling is plausible, though direct kinase inhibition remains undemonstrated. The diterpenes cafestol and kahweol achieve the lowest plasma concentrations, suggesting their lipid-metabolic effects may require chronic accumulation or tissue-specific enrichment in hepatocytes, where first-pass metabolism concentrates these compounds [25]. Importantly, the elimination half-lives vary considerably: caffeine 3–7 h (CYP1A2-dependent, with wide interindividual variation) [31], chlorogenic acid metabolites 1–3 h [35], and trigonelline approximately 4–6 h [36]. These kinetic differences imply that the network is not engaged simultaneously at steady state but rather undergoes temporal evolution, with caffeine providing the most sustained target engagement throughout the day.

4.6 Testable Hypotheses from Network Analysis

Our network analysis generates several experimentally testable predictions that distinguish it from purely descriptive pathway mapping:

1. **Nrf2 synergy hypothesis:** Co-treatment of hepatocytes with kahweol (direct Nrf2 activator) and trigonelline (indirect GSK3β modulator) should produce greater ARE-luciferase reporter activation than either compound alone, due to simultaneous electrophilic activation and reduced phospho-dependent degradation of Nrf2. Predicted effect size: >1.5-fold

over additive expectation (Bliss independence model).

2. **Hub target essentiality:** CRISPR-mediated knockout of NFE2L2 in HepG2 cells should attenuate the anti-inflammatory transcriptional response to coffee extract by >50%, given Nrf2's role as a convergence node for three compound classes.
3. **Pharmacogenomic vulnerability:** Carriers of *CYP1A2*1F* slow-metabolizer alleles should show prolonged caffeine-mediated ADORA2A occupancy and enhanced neuroprotective benefit, as predicted by CYP1A2's bridging role between xenobiotic metabolism and neuroinflammation modules.
4. **Lipid module redundancy:** Cafestol's cholesterol-raising effect should be partially attenuated by co-administration of 5-CQA (which engages PPARG in the same lipid module), testable via a 2×2 factorial coffee-intervention design comparing filtered vs. unfiltered coffee with and without chlorogenic acid supplementation.

These predictions provide a roadmap for moving from computational network analysis to mechanistic validation, with each hypothesis specifying the expected direction, magnitude, and experimental system.

4.7 Clinical Relevance and Limitations

The network pharmacology approach employed here identifies plausible molecular mechanisms but does not directly measure pharmacological effects. Several limitations should be acknowledged. First, target identification relied on *in vitro* binding data that may not fully translate to *in vivo* conditions at physiologically achievable concentrations. Second, we did not model compound metabolism, which can generate bioactive metabolites (e.g., paraxanthine from caffeine, dihydrocaffeic acid from caffeic acid) with distinct target profiles [31]. Third, the PPI network represents a static snapshot that does not capture tissue-specific expression or temporal dynamics. Fourth, the STRING database relies heavily on text-mining as an evidence channel—indeed, text-mining contributed to all 19 PPI interactions in our network (Appendix B), while only 7 had direct experimental support. Text-mining scores can inflate interaction confidence for heavily co-studied protein pairs. However, we mitigate this concern by reporting individual evidence channel scores transparently and by demonstrating that the high-confidence subnetwork (≥ 0.700, retaining 7 interactions) preserves hub target rankings.

Despite these limitations, our analysis provides a testable framework for understanding coffee's multi-target pharmacology. The identified hub targets (Nrf2, COX-2, PPARγ) and bridging nodes (CYP1A2, GSK3β) represent priority candidates for mechanistic validation using CRISPR knockout studies, reporter assays, and targeted metabolomics in human intervention trials.

5 Conclusions

This integrative network pharmacology study systematically mapped the target landscape of six major coffee bioactives onto a compound–target–PPI network comprising 16 nodes and 36 edges. Three key findings emerged:

1. **Multi-target convergence:** Coffee bioactives collectively engage 10 validated human protein targets across five major biological pathways, with NFE2L2 (Nrf2), PTGS2 (COX-

2), and PPARγ as hub targets.

2. **Pathway complementarity:** Lipid metabolism showed the broadest compound coverage (3 compounds), followed by neuroinflammation (2 compounds), while oxidative stress defense exhibited the strongest convergence (3 compounds on 1 target).
3. **Synergistic potential:** The GSK3β–Nrf2 axis, connecting trigonelline to the antioxidant network, represents a predicted synergistic mechanism whereby GSK3β modulation potentiates Nrf2 activation by other coffee compounds.
4. **Network robustness:** Vulnerability analysis demonstrated that removal of any single hub target reduces compound–target edges by at most 17.6%, indicating that coffee's distributed target architecture provides inherent pharmacological redundancy against single-point failures.
5. **Quantitative benchmarking:** The Pathway Synergy Coverage Index (PSCI = 0.433) provides the first standardized metric for comparing multi-target dietary exposures, with coffee exceeding estimated values for green tea (0.28) and red wine (0.35).

These findings provide a systems pharmacology rationale for the cardiometabolic benefits observed in epidemiological studies of coffee consumption, introduce quantitative tools (PSCI, vulnerability analysis) for dietary network pharmacology, and generate four experimentally testable hypotheses for mechanistic validation.

From a practical perspective, our analysis highlights that coffee preparation method significantly modulates the bioactive network. Filtered coffee removes cafestol and kahweol, eliminating the lipid metabolism module (4 targets) but preserving the neuroinflammation, oxidative stress, and xenobiotic metabolism pathways. This is consistent with epidemiological data showing that filtered coffee is associated with cardiovascular benefit without the cholesterol-raising effects of unfiltered preparations [24, 1]. Our network framework thus provides a mechanistic basis for preparation-specific health recommendations. Table 3 summarizes these preparation-dependent differences.

Table 3: Preparation-dependent network engagement. Pathways active (+) or absent (−) in filtered versus unfiltered coffee, based on compound removal by paper filtration.

Pathway Module	**Filtered**	**Unfiltered**
Neuroinflammation (ADORA2A, PDE4B, GSK3B)	+	+
Oxidative stress (NFE2L2/Nrf2)	+	+
Xenobiotic metabolism (CYP1A2)	+	+
Inflammation (PTGS2/COX-2)	+	+
Lipid metabolism (PPARG, NR1H3, HMGCR, CYP7A1)	−	+
Active targets	6/10	10/10
Active compounds	4/6	6/6

References

References

[1] Poole R, Kennedy OJ, Roderick P, Fallowfield JA, Hayes PC, Parkes J. Coffee consumption and health: umbrella review of meta-analyses of multiple health outcomes. *BMJ*.

2017;359:j5024.

[2] Ding M, Bhupathiraju SN, Satija A, van Dam RM, Hu FB. Long-term coffee consumption and risk of cardiovascular disease: a systematic review and a dose-response meta-analysis of prospective cohort studies. *Circulation*. 2015;129(6):643–659.

[3] Loftfield E, Cornelis MC, Caporaso N, Yu K, Sinha R, Freedman N. Association of coffee drinking with mortality by genetic variation in caffeine metabolism: findings from the UK Biobank. *JAMA Intern Med*. 2022;182(10):1072–1083.

[4] Carlström M, Larsson SC. Coffee consumption and reduced risk of developing type 2 diabetes: a systematic review with meta-analysis. *Nutr Rev*. 2018;76(6):395–417.

[5] Chieng D, Canovas R, Segan L, Sugumar H, Voskoboinik A, Kistler PM. The impact of coffee subtypes on incident cardiovascular disease, arrhythmias, and mortality: long-term outcomes from the UK Biobank. *Eur J Prev Cardiol*. 2022;29(17):2240–2249.

[6] Santos C, Costa J, Santos J, Vaz-Carneiro A, Lunet N. Caffeine intake and dementia: systematic review and meta-analysis. *J Alzheimers Dis*. 2010;20(Suppl 1):S187–S204.

[7] Qi H, Li S. Dose-response meta-analysis on coffee, tea and caffeine consumption with risk of Parkinson's disease. *Geriatr Gerontol Int*. 2014;14(2):430–439.

[8] Farah A. Coffee constituents. In: Chu Y-F, editor. *Coffee: Emerging Health Effects and Disease Prevention*. Oxford: Wiley-Blackwell; 2012. p. 21–58.

[9] Ludwig IA, Clifford MN, Lean MEJ, Ashihara H, Crozier A. Coffee: biochemistry and potential impact on health. *Food Funct*. 2014;5(8):1695–1717.

[10] Hopkins AL. Network pharmacology: the next paradigm in drug discovery. *Nat Chem Biol*. 2008;4(11):682–690.

[11] Li S, Zhang B. Traditional Chinese medicine network pharmacology: theory, methodology and application. *Chin J Nat Med*. 2013;11(2):110–120.

[12] Boezio B, Audouze K, Ducrot P, Taboureau O. Network-based approaches in pharmacology. *Mol Inform*. 2017;36(10):1700048.

[13] Szklarczyk D, Kirsch R, Koutrouli M, et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res*. 2023;51(D1):D599–D606.

[14] Wishart DS, Feunang YD, Guo AC, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res*. 2018;46(D1):D1074–D1082.

[15] Zdrazil B, Felix E, Hunter F, et al. The ChEMBL Database in 2023: a drug discovery platform spanning genomics, chemical biology and clinical data. *Nucleic Acids Res*. 2024;52(D1):D1180–D1192.

[16] Daina A, Michielin O, Zoete V. SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. *Nucleic Acids Res.* 2019;47(W1):W357–W364.

[17] Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* 2023;51(D1):D587–D592.

[18] Hagberg AA, Schult DA, Swart PJ. Exploring network structure, dynamics, and function using NetworkX. In: Varoquaux G, Vaught T, Millman J, editors. *Proceedings of the 7th Python in Science Conference (SciPy2008).* Pasadena, CA; 2008. p. 11–15.

[19] Boettler U, Sommerfeld K, Volz N, et al. Coffee constituents as modulators of Nrf2 nuclear translocation and ARE (EpRE)-dependent gene expression. *J Nutr Biochem.* 2011;22(5):426–440.

[20] Liang N, Kitts DD. Role of chlorogenic acids in controlling oxidative and inflammatory stress conditions. *Nutrients.* 2016;8(1):16.

[21] Ruggiero E, Di Castelnuovo A, Costanzo S, et al. Daily coffee drinking is associated with lower risks of cardiovascular and total mortality in a general Italian population: results from the Moli-sani study. *J Nutr.* 2021;151(2):395–404.

[22] Ong KW, Hsu A, Tan BKH. Chlorogenic acid stimulates glucose transport in skeletal muscle via AMPK activation: a contributor to the beneficial effects of coffee on diabetes. *PLoS One.* 2012;7(3):e32718.

[23] Meng S, Cao J, Feng Q, Peng J, Hu Y. Roles of chlorogenic acid on regulating glucose and lipids metabolism: a review. *Evid Based Complement Alternat Med.* 2013;2013:801457.

[24] Urgert R, Katan MB. The cholesterol-raising factor from coffee beans. *Annu Rev Nutr.* 1997;17:305–324.

[25] Ricketts ML, Boekschoten MV, Kreber AJ, et al. The cholesterol-raising factor from coffee beans, cafestol, as an agonist ligand for the farnesoid and pregnane X receptors. *Mol Endocrinol.* 2007;21(7):1603–1616.

[26] Dinkova-Kostova AT, Holtzclaw WD, Cole RN, et al. Direct evidence that sulfhydryl groups of Keap1 are the sensors regulating induction of phase 2 enzymes that protect against carcinogens and oxidants. *Proc Natl Acad Sci USA.* 2002;99(18):11908–11913.

[27] Salazar M, Rojo AI, Velasco D, de Sagarra RM, Cuadrado A. Glycogen synthase kinase-3β inhibits the xenobiotic and antioxidant cell response by direct phosphorylation and nuclear exclusion of the transcription factor Nrf2. *J Biol Chem.* 2006;281(21):14841–14851.

[28] Rada P, Rojo AI, Chowdhry S, McMahon M, Hayes JD, Cuadrado A. SCF/β-TrCP promotes glycogen synthase kinase 3-dependent degradation of the Nrf2 transcription factor in a Keap1-independent manner. *Mol Cell Biol.* 2011;31(6):1121–1133.

[29] Zhou J, Chan L, Zhou S. Trigonelline: a plant alkaloid with therapeutic potential for diabetes and central nervous system disease. *Curr Med Chem.* 2012;19(21):3523–3531.

[30] Nehlig A. Is caffeine a cognitive enhancer? *J Alzheimers Dis.* 2010;20(Suppl 1):S85–S94.

[31] Nehlig A. Interindividual differences in caffeine metabolism and factors driving caffeine consumption. *Pharmacol Rev.* 2018;70(2):384–411.

[32] Hooper C, Killick R, Lovestone S. The GSK3 hypothesis of Alzheimer's disease. *J Neurochem.* 2008;104(6):1433–1439.

[33] Calkin AC, Tontonoz P. Transcriptional integration of metabolism by the nuclear sterol-activated receptors LXR and FXR. *Nat Rev Mol Cell Biol.* 2012;13(4):213–224.

[34] Macheiner L, Schmidt J, Graziadio C, et al. Chlorogenic acid inhibits CYP1A2 *in vitro* and *in vivo*: implications for coffee–drug interactions. *Food Chem Toxicol.* 2022;167:113254.

[35] Stalmach A, Mullen W, Barron D, et al. Metabolite profiling of hydroxycinnamate derivatives in plasma and urine after the ingestion of coffee by humans: identification of biomarkers of coffee consumption. *Drug Metab Dispos.* 2009;37(8):1749–1758.

[36] Lang R, Dieminger N, Beusch A, et al. Bioappearance and pharmacokinetics of bioactives upon coffee consumption. *Anal Bioanal Chem.* 2013;405(26):8487–8503.

[37] Gan RY, Li HB, Sui ZQ, Corke H. Absorption, metabolism, anti-cancer effect and molecular targets of epigallocatechin gallate (EGCG): an updated review. *Crit Rev Food Sci Nutr.* 2018;58(6):924–941.

A Appendix: Complete Compound–Target Interaction Table

Table 4: Complete mapping of coffee bioactives to validated protein targets with evidence sources, primary pharmacological actions, and supporting references.

Compound	Target (Gene)	Evidence	Action	Reference
Caffeine	ADORA2A	ChEMBL, DrugBank	Antagonist ($K_i \approx 2.4\,\mu$M)	[14, 15]
Caffeine	PDE4B	ChEMBL (Tier 3)	Inhibitor ($IC_{50} \approx 0.5$ mM)[§]	[15]
Caffeine	CYP1A2	DrugBank	Substrate/inhibitor	[14]
Cafestol	NR1H3	Molecular docking	LXR-α ligand ($\Delta G = -6.67$ kcal/mol)[‡]	Docking (t
Cafestol	HMGCR	Literature	Upregulates expression	[24]
Cafestol	CYP7A1	Literature	Modulates bile acid synthesis	[25]
Kahweol	NR1H3	Molecular docking	LXR-α ligand ($\Delta G = -6.63$ kcal/mol)[‡]	Docking (t
Kahweol	PTGS2	PubChem BioAssay	COX-2 inhibitor	PubChem
Kahweol	NFE2L2	Literature	Nrf2 activator	[19]
5-CQA	PPARG	SwissTargetPred	PPARγ partial agonist	[16]
5-CQA	NFE2L2	Literature	Nrf2 activator	[20]
5-CQA	CYP1A2	Literature	Inhibitor	[34]
Trigonelline	GSK3B	Literature (Tier 3)	GSK-3β modulator (indirect[†])[§]	[29]
Trigonelline	ADORA2A	SwissTargetPred	Weak A_{2A} modulator	[16]
Caffeic acid	NFE2L2	ChEMBL	Nrf2 activator	[15]
Caffeic acid	PTGS2	ChEMBL	COX-2 inhibitor	[15]
Caffeic acid	CYP1A2	Literature	Inhibitor	[34]

[†]No direct GSK-3β inhibition demonstrated; indirect pathway modulation via Wnt/β-catenin signaling.
[‡]Ricketts et al. (2007) demonstrated cafestol activation of FXR (NR1H4) and PXR (NR1I2), but did **not** find LXR-α (NR1H3) activation. The LXR interaction is based on molecular docking predictions from the companion paper in this series (cafestol: $\Delta G = -6.67$ kcal/mol; kahweol: $\Delta G = -6.63$ kcal/mol against PDB 5HJP, 2.6 Å). Note: PDB 5HJP is an LXR-β (NR1H2) structure; results are informative for the NR1H3 network node given ∼77% LBD sequence identity between LXR subtypes, but the receptor subtype distinction should be noted.
[§]Tier 3 interaction: retained for topological completeness but does not meet the primary $< 10\,\mu$M affinity threshold or lacks direct binding evidence. See Methods for tier classification criteria.

STRING Query Parameters

- **Database version:** STRING v12.0
- **Species:** *Homo sapiens* (NCBI Taxonomy 9606)
- **Confidence threshold:** Medium (combined score ≥ 0.400)
- **Query proteins:** PPARG, NR1H3, HMGCR, CYP7A1, ADORA2A, CYP1A2, PDE4B, GSK3B, NFE2L2, PTGS2
- **Interactions returned:** 19

- **Score range:** 0.421–0.973
- **Query date:** March 2026

B Appendix: Complete STRING PPI Data

Table 5 presents all 19 protein–protein interactions retrieved from the STRING database (v12.0) for the 10 query proteins, including individual evidence channel scores.

Table 5: Complete STRING protein–protein interaction data for coffee bioactive targets. All 19 interactions with combined score ≥ 0.400. Individual channel scores: Exp = experimental, DB = database, TM = textmining, Coex = co-expression.

Protein A	**Protein B**	**Combined**	**Exp**	**DB**	**TM**	**Coex**
CYP7A1	CYP1A2	0.973	0.000	0.900	0.526	0.000
GSK3B	NFE2L2	0.953	0.547	0.750	0.620	0.000
HMGCR	CYP7A1	0.925	0.000	0.000	0.909	0.212
CYP7A1	NR1H3	0.879	0.056	0.000	0.867	0.111
PPARG	NR1H3	0.865	0.405	0.000	0.763	0.123
PPARG	NFE2L2	0.831	0.000	0.000	0.830	0.042
PPARG	PTGS2	0.735	0.000	0.000	0.729	0.063
PTGS2	NFE2L2	0.698	0.000	0.000	0.674	0.109
PPARG	HMGCR	0.661	0.000	0.000	0.660	0.044
GSK3B	PDE4B	0.634	0.000	0.000	0.634	0.000
PPARG	GSK3B	0.634	0.071	0.000	0.622	0.000
PPARG	CYP7A1	0.616	0.056	0.000	0.585	0.099
HMGCR	GSK3B	0.601	0.049	0.000	0.586	0.069
HMGCR	NR1H3	0.596	0.000	0.000	0.596	0.000
GSK3B	PTGS2	0.573	0.000	0.000	0.573	0.000
NFE2L2	NR1H3	0.558	0.000	0.000	0.558	0.000
HMGCR	PTGS2	0.463	0.000	0.000	0.461	0.044
PDE4B	PTGS2	0.446	0.000	0.000	0.292	0.249
CYP1A2	ADORA2A	0.421	0.000	0.000	0.421	0.000

Textmining is the dominant evidence channel, contributing to all 19 interactions. Experimental evidence supports 7 interactions, with GSK3B–NFE2L2 (escore = 0.547) and PPARG–NR1H3 (escore = 0.405) showing the strongest experimental support. Database evidence is present for 2 interactions: CYP7A1–CYP1A2 (dscore = 0.900, cytochrome P450 family annotation) and GSK3B–NFE2L2 (dscore = 0.750, kinase–substrate relationship).

C Appendix: NetworkX Analysis Code

The following Python code implements the network construction and topology analysis described in Section 2.3. The full source code is available in the supplementary file `network_analysis.py`.

Listing 1: Network construction and centrality analysis using NetworkX. STRING PPI data is loaded from the REST API JSON response.

```
import json
import networkx as nx

# Compound-target mapping (literature-curated)
COMPOUND_TARGETS = {
    'Caffeine':     ['ADORA2A', 'PDE4B', 'CYP1A2'],
    'Cafestol':     ['NR1H3', 'HMGCR', 'CYP7A1'],
    'Kahweol':      ['NR1H3', 'PTGS2', 'NFE2L2'],
    '5-CQA':        ['PPARG', 'NFE2L2', 'CYP1A2'],
    'Trigonelline': ['GSK3B', 'ADORA2A'],
    'Caffeic acid': ['NFE2L2', 'PTGS2', 'CYP1A2'],
}

def build_bipartite_network(string_data):
    """Build compound-target-PPI heterogeneous network."""
    G = nx.Graph()

    # Add compound nodes and compound-target edges
    for compound, targets in COMPOUND_TARGETS.items():
        G.add_node(compound, node_type='compound')
        for target in targets:
            G.add_node(target, node_type='target')
            G.add_edge(compound, target,
                       edge_type='compound-target')

    # Add STRING PPI edges (score >= 0.400)
    for interaction in string_data:
        a = interaction['preferredName_A']
        b = interaction['preferredName_B']
        score = interaction['score']
        if a in G.nodes and b in G.nodes:
            G.add_edge(a, b, edge_type='ppi',
                       score=score)
    return G

# Load STRING data and build network
with open('string_results.json') as f:
    string_data = json.load(f)

G = build_bipartite_network(string_data)

# Compute centrality metrics
```

```
targets = [n for n, d in G.nodes(data=True)
           if d.get('node_type') == 'target']
degrees = {n: G.degree(n) for n in targets}
betweenness = nx.betweenness_centrality(G)

# Network density
density = nx.density(G)
print(f"Nodes: {G.number_of_nodes()}, "
      f"Edges: {G.number_of_edges()}, "
      f"Density: {density:.3f}")

# Print target centrality rankings
for t in sorted(degrees, key=degrees.get,
                reverse=True):
    print(f"{t:10s}: degree={degrees[t]}, "
          f"betweenness={betweenness[t]:.4f}")
```