Network pharmacology of lycopene and Molecular Docking with Top Hub Proteins

Nisha Paudel\textsuperscript{1}, Umme Hani\textsuperscript{2}, Nagendra Prasad Awasthi\textsuperscript{3}, Manjunatha Hanumantappa\textsuperscript{1,2,*}, Rangaswamy Lakshminarayan\textsuperscript{1,*}

1. Bangalore University, Bengaluru 560056, Kanataka, India
2. kuvempu University, Shimigo, Karnataka, India
3. Deakin University, Victoria, Australia

Corresponding authors
Rangaswamy Lakshminarayan rlnarn21@gmail.com
Manjunatha Hanumantappa manjunatha75@gmail.com

Keywords: Network Pharmacology, Molecular docking, Lycopene
Abstract

Background
Lycopene is one of the potent antioxidants in the family of carotenoids that scavenges Reactive Oxygen Species (ROS) singlet oxygen which has been associated with various pathological consequences including atherosclerosis, myocardial infarction, and stroke and Sex hormone-induced cancers like breast cancer, endometrial cancer and prostate cancer. As multiple pathways are involved in the manifestation of aforementioned diseases initiated at the behest of ROS, it would be appropriate to understand the likely pathways triggered by the ROS and its eventual control by the action of lycopene through network pharmacology study, a robust paradigm for drug discovery via modulation of multiple targets.

Results
124 proteins were mined from CTD and STITCH databases, which showed some relationship with lycopene, among them strong association was found with TP53, STAT3 and CDK1 proteins. Lycopene showed a strong affinity with these proteins by hydrophobic interactions.

Conclusion
The topological analysis of a network created by the lycopene relevant genes showed its role as a potential therapeutic agent in cancer which further requires in vitro and in vivo studies to confirm these findings.
1. Background

Lycopene belongs to the carotenoids family, a class of compounds present in certain fruits and vegetables (1) which is of nutritional interest as its metabolism mediates various signaling effects (2). Lycopene undergoes extensive isomerization, a trans configuration of lycopene is found in fruits and vegetables (3,4) that gives the red hue whereas, it isomerizes in blood plasma and tissues after absorption into cis isomers (2,4). Lycopene is the most abundantly found carotenoid in the human body which is one of the most potent antioxidants that scavenges reactive oxygen species (ROS) singlet oxygen (5). Lycopene is a tetraterpene merged from eight isoprene units that are composed of carbon and hydrogen and it’s an open-polyene chain, without the ionone ring as in β-carotene, gives this molecule with an advantage of a very high antioxidant capacity, which is significantly higher than any other carotenoids, namely, more than twice that of β-carotene and about 10 times higher than that of α-tocopherol (6,7). Particularly, it is effective in the quenching of superoxide anion free radicals, (6). Therefore, its main protective mechanism is in its antioxidant properties, which have been well characterized in vitro (8).

Numerous studies and meta-analysis have been done in the association of lycopene with cardiovascular risks, including atherosclerosis myocardial infarction, and stroke (1,3,9-13) and Sex hormone-induced cancers like breast cancer, endometrial cancer and prostate cancer (14).

Various mechanisms have been postulated by which lycopene alleviates diseases such as prevention of oxidative DNA damage (15) and balancing
redox signaling (16), inhibition of androgen receptor signaling (17), inhibition of 5-lipoxygenase (18) and cell cycle regulatory proteins (19). These individual studies verify that instead of a single pathway, multiple mechanisms are involved (20). Oxidative stress is thought to be the main cause for the diseases and lycopene has reactive oxygen stress quenching capacity, as it stands out as a potential balanced redox state of a cell, however, the low bioavailability of lycopene is a major hindrance for its actions (21,22).

Network pharmacology is a new paradigm for drug discovery via modulation of multiple targets rather than a single target, encompassing system biology, network analysis, connectivity, redundancy, and pleiotropy (23). It is used to study compound-proteins/genes-disease pathways, which is capable of describing the complexities of biological systems, drugs, and diseases from a network perspective (24).

In the quest of finding possible approach of the working mechanism of lycopene for its various beneficial properties mentioned above, The present study is carried out, firstly, by extracting all the relevant genes/proteins from databases and then annotating by gene ontology, the network is constructed which showed the interaction in between genes. The network is analyzed and hub proteins are determined and docking of lycopene is done with those significant hub proteins.
2. Results

2.1 Data curation

Lycopene was found to be associated with 124 curated genes/proteins, Table 1, from the STRING database with a medium confidence level of 0.4 and was reviewed on the UniProt database to create a network.

2.2 Network construction and analysis

Based on the STRING database evidence, the network was constructed with lycopene associated proteins to build the protein-protein interaction using Cytoscape v 3.2. The resultant network had 928 nodes and 4390 edges, nodes represent proteins and the edges indicate their relations, Fig 1, represents a summary of simple parameters and ,Table 2, contains a list of genes analyzed by network analyzer based on degree. There significant proteins TP53, STAT3, and CDK1 were selected based on their degree 68, 63, and 45 and average shortest path length 2.9, 2.9, and 3.6 respectively. A radial network, Fig 2, was constructed with three hub proteins.
2.3 Gene Ontology analysis

To explore the functional characteristics of lycopene associated genes were subjected to GO analysis and enriched according to the available GO annotations with gene ID and gene accessions in the PANTHER database. In Fig 3 a, molecular function the genes associated with lycopene showed the highest activity in the catalytic activity. In Fig 3 b, the biological process, the associated genes had maximum involvement in the cellular process. In Fig 3 c, cellular component, genes are associated are found in Cell and Cell part. In Fig 3 d protein class, the associated genes had the highest activity in metabolite inter-conversion enzyme and in Fig 3 e, molecular pathway, the associated genes had a major role in CCKR signaling map and apoptosis signaling pathway.

2.4 Molecular docking analysis of lycopene

Top proteins from network analysis were used for docking. Molecular docking analysis of lycopene was performed using Auto Dock Vina Tool and their binding energy was noted. AutoDock Vina results were analyzed based on the interactions between hub proteins and ligand molecule, Table 3. The accuracies of the results were confirmed by considering the lowest binding free energy and hydrophobic binding between the receptor and ligand. The binding energy lycopene with proteins TP53 was $-6.3 \text{ kcal/mol}$, STAT3 was $-6.4\text{kcal/mol}$ and CDK1 was $-5.9 \text{ kcal/mol}$, Table 3.
No, any hydrogen bonds were found in all docking interactions. The hydrophobic interaction profile is shown in Fig 4. Based on the results, it can be inferred that all three proteins can bind with lycopene, which can be further validated by in vitro and in vivo studies.

3. Discussion

Network pharmacology is a relevantly new emerging technique for drug discovery approach, it’s an effort to address the side effects and toxicity of drug (Hopkins 2007). In the context of lycopene for the first time network pharmacology study has been conducted. Lycopene is a lipophilic, unsaturated carotenoid ranging from yellow to dark red color in various fruits and vegetables.

Numerous studies have been conducted on the beneficial roles of lycopene, such as reducing the risk of some cancers and cardiovascular diseases, however, one of the main obstacles, when considering the biological effects of lycopene, is its very low and variable bioavailability (21,25). The bioavailability depends on various factors including the nature of the double bonds, the geographic sociodemographic factors of an individual, health status, and the food ingestion context (22). These complex factors account for the variability of lycopene concentrations and biological effects in an individual. Furthermore, lycopene distribution throughout the human body is asymmetrical, being mostly found in the liver, adrenals, lungs, prostate, and skin (6). The diseases appear by
mediating various signaling effects and the cumulative effect can be studied with the help of network pharmacology.

The study began with data mining from the STITCH and CTD databases and was entered into the STRING database to get relevant genes. Medium confidence (0.400) level was chosen to obtain more number of interactions with affinity. The network was constructed using freeware, Cytoscape, due to the huge number of nodes and edges, which are 928 and 4390 edges, the network did not appear vivid so, a radial layout network was constructed only with the first neighbor nodes of three hub proteins. These three hub proteins were selected based on its topology parameters i.e. degree and average shortest path length; they are TP53, STAT3, and CDK1. All the top proteins obtained had a role in cancer therapeutics. Their structures were retrieved from the RSCB PDB database. 4MZR is a mutated structure of TP53 which no longer protects cells. TP53/P53 is easily mutable protein. So binding of small ligands may stabilize and reactivate the p53 protein activity. As it acts as a tumor suppressor in many tumor types; Involved in cell cycle regulation causes growth arrest or apoptosis depending on the physiological circumstances and cell type (26). 6NJS is STAT3 which is an attractive cancer therapeutic target. Induced degradation of STAT3 results in a strong suppression of its transcription network in leukemia and lymphoma cells degradation of STAT3 protein, therefore, is a promising cancer therapeutic strategy (27). 4YC6 is the structure CDK1 a family of PKs, which plays a key role in the regulation of the cell cycle, they depend on regulatory subunits, cyclin,
and their activities are regulated by CDKIPs. In some human cancers, CDKs are overexpressed or CDKIPs are either absent or mutated. Hence, CDKs are appealing therapeutic targets to prevent the unregulated proliferation of cancer cells. As a result, in the last few decades, selective CDK inhibitors have been noticed as effective chemotherapeutic agents (28).

In the simple parameters, obtained after network analysis, the clustering coefficient was found to be 0.795, which signifies at a higher degree, nodes tend to cluster together. The largest distance between the two nodes is 12 which is the diameter of the network whereas the shortest distance between any two nodes is 1 which is the radius of the network. Network centralization is 0.063 which means that the network is more uniformly connected. The nodes have an average shortest path length of 4.639 and at an average number single node has 9.461 neighbors. Network heterogeneity is the tendency of the network to contain the hub nodes which is 0.717.

In the molecular function, the proteins associated with lycopene showed the highest activity in the catalytic activity followed by binding and transporter activity whereas in the biological process the associated proteins had maximum involvement on the cellular process and metabolic process. In the cellular component, these proteins/genes are associated with both cellular and extracellular region which includes plasma membrane and external encapsulating structures. In the molecular pathway, genes showed the highest activity on Protein class CCKR
signaling map and Apoptosis signaling pathway, in protein class, the target genes were enriched in metabolite interconversion enzyme.

When the docking was done all the three proteins showed a good affinity with the proteins due to its lipophilic nature with acyclic structure (29), it showed good hydrophobic interactions. In the wet lab studies, using real-time PCR assay, lycopene promoted upregulation of TP53 and Bax transcript expression and also downregulation of Bcl-2 expression in PCa (30). Lycopene treatment inhibits the activation of Jak1/stat3 signaling in gastric disease (31). Lycopene increases the expression of CDK inhibitors including p21 and p27, as well as the tumor suppressor gene p53, and decreases the expression of Skp2 (32).

It is important to note that the conflicting results about the lycopene also have been published may be related to the wide variety of experimental protocols used to discover any association between lycopene consumption and cardiovascular disease. Furthermore, some have even argued that the sole lycopene is not effective as compared to consuming whole tomato products (25,33).
4. Conclusions

The topological analysis of a network created by the lycopene relevant genes showed its role as a therapeutic agent in cancer. Furthermore, the docking results also supported a good affinity binding with TP53, STAT3, and CDK1. Since the bioavailability of lycopene is less and it is converted to a metabolite that may show different affinity with these proteins. Hence, encapsulated lycopene can be used to confirm these in silico findings in vitro and in vivo studies.

5. Methods

All the programs were run on a 4.00 GB RAM Intel® Core i3 2.5GHz with a 64-bit Windows 10 Operating System in the HP notebook. The research approach was executed by implementing the following methods:

5.1 Data acquisition

For the identification of target genes/proteins associated with lycopene relevant to Homo sapiens, data mining was carried from CTD (34) which contains manually curated information about chemical–gene/protein interactions, chemical–disease and gene–disease relationships (35) and STITCH (36) which integrates information about interactions from metabolic pathways, crystal structures, binding experiments and drug-target relationships (37). Genes were curated with a medium confidence score (0.400) to find maximum interactions with affinity. The target genes
were screened manually in UniProt (38) and only reviewed one was selected.

Genes obtained from CTD and STITCH databases were enriched in the online tool STRING (39), which aims to provide critical predicted information on protein-protein interactions inferred from genomic resources, experimental shreds of evidence, text mining, co-occurrence and co-expression (40). The data in STRING were weighted and integrated by assigning a medium confidence score of 0.400, which indicates the probability that the interaction is maximum. Such networks can be used for filtering and examining the functional genomics data and for showing an intuitive design for annotating structural, functional and evolutionary aspects of proteins, and also exploring the predicted interaction networks suggests new directions for future experimental research (41).

5.2 Gene Ontology analysis

To know the functional association of obtained genes from databases with lycopene, GO enrichment analysis using PANTHER (42) was performed. It is a tool for the classification of genes based on their evolutionary history and functions (43,44). Five aspects of GO determined are; (a) Molecular function where the molecular-level activities performed by gene products, the (b) cellular component represents locations relative to cellular structures in which a gene product performs a function, (c) biological process is for larger processes, or ‘biological programs’ accomplished by multiple molecular activities (45), (d) protein class and (e) pathway.
5.3 Network construction

The Protein-Protein Interaction (PPI) network of genes associated with the lycopene molecule was constructed and visualized using open-source Cytoscape v 3.2.2 software. This software provides integrated models for biomolecular interactome networks displaying nodes as proteins and edges as the physical relationship between them (46).

The genes curated were merged using a tool ‘Advanced Network Merge’ and further analyzed by ‘Network Analyser’ to find the topological parameters of the target genes. Top hub proteins were selected and the network was created using the top three hub proteins’ first neighboring proteins. The various parameters of the network were analyzed; like the average shortest path length which is the expected route between the nodes to connect it, betweenness centrality is the number of shortest paths going via a certain node, node degree is the number of edges linked to it, shared neighbor distribution is the maximum number of times, a node shares its neighbor with other nodes, node degree distribution, average clustering coefficient distribution, a topological coefficient is a relative measure for the extent to which a node shares neighbors with other nodes (47). Gene list in descending order with a degree was created along with various parameters of gene-like Neighborhood Connectivity, Average Shortest PathLength, Betweenness Centrality, Topological Coefficient, and Clustering Coefficient. Based on these, three proteins, namely, TP53, STAT3, and CDK1 were selected for docking.
5.4 Molecular docking study of the top hub node

The top three hub proteins, namely, TP53, STAT3, and CDK1, identified through network analysis were subjected to molecular docking study. The X-ray crystallographic structure of TP53 (PDB ID: 4mzr, resolution: 2.9 Å), STAT3 (PDB ID: 1WZY, resolution: 2.7 Å) and CDK1 (PDB ID, resolution: 2.6Å) was retrieved from RCSB PDB (48). The energy minimized structures of these compounds were obtained by Swiss Pdb-Viewer v 4.1.0 (49). The active sites were chosen based on software CASTp (50) and finally confirmed by visualizing them proteins in PyMol v 2.3. Active sites of TP53 are GLY334 ARG337 PHE338 GLU 339 GLU340 PHE 341 ASN 345, STAT3 are SER611 GLU612 SER613 TRP623 PRO629 SER636 VAL637 GLU638 PRO639 TYR640, and CDK1 ILE49 52ILE 53SER 56LYS 69VAL 79LEU. AutoDock Vina program was used over Autodock v 4 to dock ligand molecules with a receptor, as it calculates its grid maps automatically and ranks the results without saving its intermediate results which makes it fast and less space-consuming along with this Vina comes with a new knowledge-based, a statistical scoring function that replaces the semiempirical force field of AutoDock (51). The heteroatoms and water molecules in the protein structure were removed, then hydrogen atoms and, Gasteiger charges were added and torsion degrees of freedom and Solvation parameters were assigned and saved in PDBQT format for further analysis (52).
A target-based docking method was used in this study to prepare the grid maps. The coordinate of the grid box was set at x=30, y=30 and z=30, and the grid center was set at X=-74.881 Y= 23.427 and Z= 6.486 for TP53, X=13.453 Y= 53.543 and Z= 2.754 for STAT3 and X=-38.074 Y= -14.226 and Z= -12.556 for CDK1. Docking simulation was executed with several modes set to 10 to get more accurate results. The pose with the highest negative Gibbs free energy (ΔG) value was considered as the best conformation and selected for further analysis.

The post-docking analysis was performed using LigPlot v.4.5.3 which displays the protein-ligand interactions mediated by the hydrophobic contacts and hydrogen bonds. In Fig 4, 5 the hydrophobic contacts are represented by an arc with spokes radiating towards the ligand atoms they contact. The contacted atoms are depicted with spokes radiating back. The final result was viewed as the PyMol v.2.3 is a 3D structure visualization tool.
References

1. Mordente A, Guantario B, Meucci E, Silvestrini A, Lombardi E, E. Martorana G, et al. Lycopene and Cardiovascular Diseases: An Update. Curr Med Chem. 2011;18(8):1146–63.

2. Ross AB, Vuong LT, Ruckle J, Synal HA, Schulze-König T, Wertz K, et al. Lycopene bioavailability and metabolism in humans: An accelerator mass spectrometry study. Am J Clin Nutr. 2011;93(6):1263–73.

3. Story EN, Kopec RE, Schwartz SJ, Harris GK. An Update on the Health Effects of Tomato Lycopene. Annu Rev Food Sci Technol. 2010;1(1):189–210.

4. Shi J, Le Maguer M. Lycopene in tomatoes: chemical and physical properties affected by food processing. Vol. 40, Critical Reviews in Food Science and Nutrition. 2000. 1–42 p.

5. Fiedor J, Burda K. Potential role of carotenoids as antioxidants in human health and disease. Nutrients. 2014;6(2):466–88.

6. Moran NE, Clinton SK, Erdman JW. Differential Bioavailability, Clearance, and Tissue Distribution of the Acyclic Tomato Carotenoids Lycopene and Phytoene in Mongolian Gerbils. J Nutr. 2013;143(12):1920–6.

7. Milani A, Basirnejad M, Shahbazi S, Bolhassani A. Carotenoids: biochemistry, pharmacology and treatment. Br J Pharmacol. 2017;174(11):1290–324.

8. Kelkel M, Schumacher M, Dicato M, Diederich M. Antioxidant and anti-proliferative properties of lycopene. Free Radic Res. 2011;45(8):925–40.

9. Biddle MJ, Lennie TA, Bricker G V., Kopec RE, Schwartz SJ, Moser DK. Lycopene Dietary Intervention: A pilot study in patients with heart failure. J Cardiovasc Nurs. 2015;30(3):205–12.

10. Böhm V. Lycopene and heart health. Mol Nutr Food Res. 2012;56(2):296–303.

11. Cheng HM, Koutsidis G, Lodge JK, Ashor A, Siervo M, Lara J. Tomato and lycopene supplementation and cardiovascular risk factors: A systematic review and meta-analysis. Atherosclerosis. 2017;257:100–8.

12. Giovannucci E. Tomatoes, Tomato-Based Products, Lycopene, and Cancer: Review of the Epidemiologic Literature. 1999;91(4):317–31.

13. Viuda-Martos M, Sanchez-Zapata E, Sayas-Barberá E, Sendra E, Pérez-Álvarez JA, Fernández-López J. Tomato and Tomato
Byproducts. Human Health Benefits of Lycopene and Its Application to Meat Products: A Review. Crit Rev Food Sci Nutr. 2014;54(8):1032–49.

14. Sharoni Y, Linnewiel-Hermoni K, Zango G, Khanin M, Salman H, Veprik A, et al. The role of lycopene and its derivatives in the regulation of transcription systems: Implications for cancer prevention. Am J Clin Nutr. 2012;96(5).

15. Rao AV, Fleshner N, Agarwal S. Serum and tissue lycopene and biomarkers of oxidation in prostate cancer patients: A case-control study. Nutr Cancer. 1999;33(2):159–64.

16. Palozza P, Parrone N, Simone R, Catalano A. Role of Lycopene in the Control of ROS-Mediated Cell Growth: Implications in Cancer Prevention. Curr Med Chem. 2011;18(12):1846–60.

17. Siler U, Barella L, Spitzer V, Schnorr J, Lein M, Goralczyk R, et al. Lycopene and Vitamin E interfere with autocrine/paracrine loops in the Dunning prostate cancer model. FASEB J. 2004;18(9):1019–21.

18. Hazai E, Bikádi Z, Zsila F, Lockwood SF. Molecular modeling of the non-covalent binding of the dietary tomato carotenoids lycopene and lycophyll, and selected oxidative metabolites with 5-lipoxygenase. Bioorganic Med Chem. 2006;14(20):6859–67.

19. Levy J, Bosin E, Feldman B, Giat Y, Miinster A, Danilenko M, et al. Lycopene Is A More Potent Inhibitor Of Human Cancer Cell Proliferation Than Either A-Carotene Or β-Carotene. Nutr Cancer. 1995;24(3):257–66.

20. Qiu X, Yuan Y, Vaishnav A, Tessel MA, Nonn L, Van Breemen RB. Effects of lycopene on protein expression in human primary prostatic epithelial cells. Cancer Prev Res. 2013;6(5):419–27.

21. Borel P, Desmarchelier C, Nowicki M, Bott R. Lycopene bioavailability is associated with a combination of genetic variants. Free Radic Biol Med [Internet]. 2015;83:238–44.

22. Petyaev IM. Lycopene Deficiency in Ageing and Cardiovascular Disease. Oxid Med Cell Longev. 2016;2016.

23. Hopkins AL. Network pharmacology: The next paradigm in drug discovery. Nat Chem Biol. 2008;4(11):682–90.

24. Zhang R, Zhu X, Bai H, Ning K. Network pharmacology databases for traditional Chinese medicine: Review and assessment. Front Pharmacol. 2019;10(February):1–14.

25. Costa-Rodrigues J, Pinho O, Monteiro PRR. Can lycopene be considered an effective protection against cardiovascular disease? Food Chem [Internet]. 2018;245:1148–53.

26. Emamzadah S, Tropia L, Vincenti I, Falquet B, Halazonetis TD.
27. Bai L, Zhou H, Xu R, Zhao Y, Chinnaswamy K, McEachern D, et al. A Potent and Selective Small-Molecule Degrader of STAT3 Achieves Complete Tumor Regression In Vivo. Cancer Cell [Internet]. 2019;36(5):498-511.e17.

28. Brown NR, Korolchuk S, Martin MP, Stanley WA, Moukhametzianov R, Noble MEM, et al. CDK1 structures reveal conserved and unique features of the essential cell cycle CDK. Nat Commun [Internet]. 2015;6:1–12.

29. Kong KW, Khoo HE, Prasad KN, Ismail A, Tan CP, Rajab NF. Revealing the power of the natural red pigment lycopene. Molecules. 2010;15(2):959–87.

30. Soares N da CP, Machado CL, Trindade BB, Lima IC do C, Gimba ERP, Teodoro AJ, et al. Lycopene extracts from different tomato-based food products induce apoptosis in cultured human primary prostate cancer cells and regulate TP53, Bax and Bcl-2 transcript expression. Asian Pacific J Cancer Prev. 2017;18(2):339–45.

31. Park B, Lim JW, Kim H. ScienceDirect Lycopene treatment inhibits activation of Jak1 / Stat3 and Wnt / β -catenin signaling and attenuates hyperproliferation in gastric epithelial cells . Nutr Res [Internet]. 2018;1–12.

32. Trejo-Solís C, Pedraza-Chaverrí J, Torres-Ramos M, Jiménez-Farfán D, Cruz Salgado A, Serrano-García N, et al. Multiple molecular and cellular mechanisms of action of lycopene in cancer inhibition. Evidence-based Complement Altern Med. 2013;2013(I).

33. Friedman M. Anticarcinogenic, cardioprotective, and other health benefits of tomato compounds lycopene, α-tomatine, and tomatidine in pure form and in fresh and processed tomatoes. J Agric Food Chem. 2013;61(40):9534–50.

34. CTD (Comparative toxicogenomics Database (http://ctdbase.org/) accessed at 17 February 2020.

35. Davis AP, Grondin CJ, Johnson RJ, Sciaky D, McMorran R, Wiegers J, et al. The Comparative Toxicogenomics Database: Update 2019. Nucleic Acids Res. 2019;47(D1):D948–54.

36. STITCH (Search Tool for Interacting Chemicals) Database v. 5.0 (http://stitch.embl.de/ ) accessed at 15 February 2020.

37. Kuhn M, von Mering C, Campillos M, Jensen LJ, Bork P. STITCH: Interaction networks of chemicals and proteins. Nucleic Acids Res. 2008;36(SUPPL. 1):684–8.

38. UniProt (https://www.uniprot.org/) accessed at 20 February 2020.
39. STRING v.11 database (https://string-db.org/) accessed at 28 February 2020.

40. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607–13.

41. Schwartz AS, Yu J, Gardenour KR, Finley RL, Ideker T. Cost-effective strategies for completing the interactome. Nat Methods. 2009;6(1):55–61.

42. PANTHER v 15.0 (Protein Analysis Through Evolutionary Relationships) (http://www.pantherdb.org/) accessed at 4 March 2020.

43. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res. 2019;47(D1):D419–26.

44. Carbon S, Dietze H, Lewis SE, Mungall CJ, Munoz-Torres MC, Basu S, et al. Expansion of the gene ontology knowledgebase and resources: The gene ontology consortium. Nucleic Acids Res. 2017;45(D1):D331–8.

45. Gaudet P, Livstone MS, Lewis SE, Thomas PD. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. Brief Bioinform. 2011;12(5):449–62.

46. Paul Shannon 1, Andrew Markiel 1, Owen Ozier, 2 Nitin S. Baliga, 1 Jonathan T. Wang, 2 Daniel Ramage 2, Nada Amin 2, Benno Schwikowski, 1, 5 and Trey Ideker2, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Res [Internet]. 2003;13(2498–2504):426.

47. Doncheva NT, Assenov Y, Domingues FS, Albrecht M. Topological analysis and interactive visualization of biological networks and protein structures. Nat Protoc [Internet]. 2012;7(4):670–85.

48. RCSB (Research Collaboratory for Structural Bioinformatics ) PDB (Protein Databank) https://www.rcsb.org accessed at 8 -10 March 2020.

49. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis. 1997;18(15):2714–23.

50. Tian W, Chen C, Lei X, Zhao J, Liang J. CASTp 3.0: Computed atlas of surface topography of proteins. Nucleic Acids Res. 2018;46(W1):W363–7.

51. Trott, oleg oslan A. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient
List of Abbreviations

ROS  Reactive Oxygen Species
CTD  Comparative Toxicogenomics Database
STITCH  Search Tool for Interacting Chemicals
STRING  Search Tool for the Retrieval of Interacting Genes/Proteins
GO  Gene Ontology
PASS  Prediction of Activity Spectra for Substances
PPI  Protein-Protein Interaction
RCSB  Research Collaboratory for Structural Bioinformatics
PDB  Protein Databank
CASTp  Computed Atlas of Surface Topology of Proteins
ΔG  Gibbs Free Energy
Pa  Probabilities of being Active
Pi  Probabilities of being Inactive
CCKR  Cholecystokinin Receptor
Acknowledgements

The authors are grateful to Bangalore University Biotechnology Department faculty members.

Funding

No funding received.

Contributions

NP curated data designed the study, NP, NPA and UH cytoscape and docking for multiple times to make it reproducible, NP, HM and RL analysed data, NP, NPA, UH, MH and RL drafted the manuscript.

Corresponding author

Correspondence to Nisha Paudel (paudelbella@gmail.com)
Ethics declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
All other authors declare that they have no competing interests.

Supplementary information

Supplementary files
https://github.com/Nishapaudel/Network-Pharmacolgy
Fig 1: Network Parameters containing the basic details of network

| Parameter                  | Value          |
|----------------------------|----------------|
| Clustering coefficient     | 0.795          |
| Connected components       | 15             |
| Network diameter           | 12             |
| Network radius             | 1              |
| Network centralization     | 0.063          |
| Shortest paths             | 584982 (68%)   |
| Characteristic path length | 4.639          |
| Avg. number of neighbors   | 9.461          |
| Number of nodes            | 928            |
| Network density            | 0.010          |
| Network heterogeneity      | 0.717          |
| Isolated nodes             | 1              |
| Number of self-loops       | 0              |
| Multi-edge node pairs      | 0              |
| Analysis time (sec)        | 21.236         |

Fig 2: Radial network of genes associated with lycopene constructed by the first neighbor
Fig 3: Gene Ontology results a. Molecular function b. Biological Process c. Cellular component d. Protein Class e. Pathway
Fig: 4 a) Molecular docking TP53-lycopene dock complex ligplot diagram and PyMol presentation respectively
Fig: 4 b) Molecular docking STAT3-lycopene dock complex ligplot diagram and PyMol presentation respectively
Fig: 4 c) Molecular docking CDK1-lycopene dock complex), ligplot diagram, and PyMol presentation respectively
Table 1: List of Lycopene associated Genes

| Gene 1 | Gene 2 | Gene 3 | Gene 4 | Gene 5 |
|--------|--------|--------|--------|--------|
| ACHE   | C4B    | CYP7B1 | JUN    | RB1    |
| ADD1   | CASP3  | CYP11A1| KPNA2  | RDX    |
| AGER   | CASP8  | DDB2   | LDLR   | RPS6KA3|
| AHR    | CASP9  | ESR1   | LIPA   | RRM2   |
| AIF1   | CAT    | FANCC  | LPO    | SBP    |
| AKT1   | CCND1  | FASLG  | MAP2K1 | SCGB1D4|
| AR     | CCNE1  | FOS    | MAP2K6 | SELE   |
| ATP1A1 | CDK1   | GET1   | MAPK8  | SIRT1  |
| ATP1A2 | CDK4   | GFAP   | MGAT2  | SKAP2  |
| ATP1A3 | CDK5   | GJA1   | MICA   | SMAD3  |
| ATP1A4 | CDK7   | GJA4   | MLH1   | SMC1A  |
| ATP1B1 | CDK9   | GPX4   | MMP2   | SRD5A1 |
| ATP1B2 | CDKN1A | GSR    | MSH6   | STAR   |
| ATP1B3 | CHGA   | HMGB1  | NFE2L2 | STAT3  |
| ATP1B4 | CHGB   | HMOX1  | NFkB1  | TFF1   |
| ATP2C1 | CRB1   | HSPA4  | NOS1   | TFP12  |
| ATP2C2 | CTSS   | ICAM1  | NQO1   | TGFA   |
| BACH1  | CXCL14 | IGF1   | NUP214 | TLR2   |
| BAP1   | CYB5A  | IGFR   | PAH    | TLR4   |
| BAX    | CYP1A1 | IL10   | PARP1  | TNF    |
| BCL2   | CYP1A2 | IL1B   | PCNA   | TP53   |
| BCO1   | CYP1B1 | IL2    | PLA2G2A| VCAM1  |
| BDNF   | CYP2A4 | IL6    | POMC   | VCP    |
| BRAF   | CYP2E1 | IVL    | PTGS2  | ZNF138 |
| C1QA   | CYP3A4 | JAK1   | RASSF7 |
Table 2: List of genes based on the degree

| Name  | Degree | Neighborhood Connectivity | Average Shortest Path Length | Betweenness Centrality | Topological Coefficient | Clustering Coefficient |
|-------|--------|---------------------------|-------------------------------|------------------------|-------------------------|------------------------|
| TP53  | 68     | 15.6                      | 2.9                           | 0.19                   | 0.05                    | 0.11                   |
| STAT3 | 63     | 18.63                     | 2.9                           | 0.15                   | 0.06                    | 0.15                   |
| JUN   | 55     | 15.27                     | 3                             | 0.17                   | 0.05                    | 0.11                   |
| CDK1  | 45     | 13.26                     | 3.6                           | 0.06                   | 0.08                    | 0.17                   |
| MAPK  | 1      | 45                        | 17.08                         | 3.1                    | 0.07                    | 0.07                   | 0.19                   |
| TNF   | 45     | 14.17                     | 3.6                           | 0.02                   | 0.09                    | 0.20                   |
| AKT1  | 41     | 19.09                     | 3.1                           | 0.11                   | 0.07                    | 0.19                   |
| MAPK  | 3      | 40                        | 15.55                         | 3.3                    | 0.04                    | 0.08                   | 0.21                   |
| IGF1  | 38     | 13.10                     | 3.6                           | 0.04                   | 0.10                    | 0.20                   |
| CCND1 | 37     | 20.51                     | 3                             | 0.04                   | 0.07                    | 0.27                   |

Table 3: Docking score and hydrophobic bonds of docked complexes

| Hub proteins | Affinity (Kcal/mol) | Hydrophobic Interactions |
|--------------|---------------------|--------------------------|
| TP53         | -6.3                | Lys321 Ile332 Glu332     |
|              |                     | Thr322 Phe328            |
|              |                     | Leu330 Asn345            |
|              |                     | Glu326 Asp324            |
| STAT3        | -6.4                | Thr641 Gln644 Gln635     |
|              |                     | Ser614 Pro639            |
|              |                     | Ile659 Glu638            |
|              |                     | Trp623 Asp627            |
|              |                     | Ala662                   |
| CDK1         | -5.9                | Arg59 Leu58 Leu55        |
|              |                     | His60 Glu57 Arg122       |
|              |                     | Val124 Leu54            |
|              |                     | Arg151 Thr47            |
|              |                     | Ala628                 |
|              |                     | Ala662                 |

