Bioinformatic Prediction of Possible Targets and Mechanisms of Action of the Green Tea Compound Epigallocatechin-3-Gallate Against Breast Cancer

Xinqiang Song [1,2]*, Mu Zhang [3], Lei Chen [1] and Qingsong Lin [2]*

[1] Department of Biological Sciences, Xinyang Normal University, Henan, China, [2] Department of Biological Sciences, National University of Singapore, Singapore, Singapore, [3] Hospital Attached to Xinyang Normal University, Xinyang, China

Epigallocatechin-3-gallate (EGCG), a bioactive compound in green tea, is the most abundant and biologically active catechin, and it exerts multiple effects in humans through mechanisms that remain to be clarified. The present study used bioinformatics to identify possible mechanisms by which EGCG reduces risk of breast cancer. Possible human protein targets of EGCG were identified in the PubChem database, possible human gene targets were identified in the NCBI database, and then both sets of targets were analyzed using Ingenuity Pathway Analysis to predict molecular networks affected by EGCG in breast cancer. The results suggest that signaling proteins affected by EGCG in breast cancer, which include JUN, FADD, NFKB1, Bcl-2, GNAO1, and MMP14, are involved primarily in cell death and survival; DNA replication, recombination and repair; and the cell cycle. The main networks affected by EGCG are predicted to involve the cell cycle; cellular assembly and organization; DNA replication, recombination and repair; and cell death and survival. These results identify several specific proteins and pathways that may be affected by EGCG in breast cancer, and they illustrate the power of integrative bioinformatics and chemical fragment analysis for focusing mechanistic studies.

Keywords: epigallocatechin-3-gallate, breast cancer, ingenuity pathway analysis, bioinformatics analysis, targets

INTRODUCTION

Breast cancer is frequent in women, and incidence continues to increase, even though improvements in prevention, mammography-based screening, and treatments (e.g., endocrine therapy) have reduced patient mortality. Approximately one in eight women in the US will develop invasive breast cancer during their lifetime, and breast cancer incidence among women in the UK has increased by 6% in recent years (Harbeck and Gnant, 2017). Incidence is even increasing in countries where it has been relatively low, such as Japan and China (Xiang et al., 2016).

In 1997, epidemiological work in Japan suggested that drinking green tea, which is prepared from *Camellia sinensis* leaves, could reduce risk of breast cancer among women, especially those drinking more than 10 cups of green tea per day (Imai et al., 1997). Since then, numerous cohort studies and case-control studies in China, USA, and Singapore have confirmed an association

between green tea consumption and reduction of breast cancer risk (Wu et al., 2003; Yuan et al., 2005; Zhang et al., 2007, 2009; Ganmaa et al., 2008; Inoue et al., 2008; Kumar et al., 2009; Shrubsole et al., 2009; Chen et al., 2010; Dai et al., 2010; Li et al., 2016). For example, one study involving 5082 women in USA showed that women who drank at least three cups of green tea per day had 37% lower risk of breast cancer than women who did not drink any tea (Kumar et al., 2009).

The ability of green tea to protect against breast cancer appears to be mediated by catechins, which are polyphenols accounting for 30–40% of the dry weight of brewed green tea. The four major catechins in green tea are (–)-epigallocatechin-3-gallate (EGCG), (–)-epigallocatechin, (–)-epicatechin gallate, and (–)-epicatechin, and EGCG is both the most abundant and most biologically active (Du et al., 2012; Kanwar et al., 2012). EGCG and other compounds extracted from green tea have been shown to suppress carcinogen-induced production of reactive oxygen species (ROS) and DNA damage, as well as alter cell signaling pathways (Ruch et al., 1989; Kaur et al., 2007; Rathore et al., 2012). Some researchers (Jones and Takai, 2001; Fang et al., 2003; Lee et al., 2005) found that EGCG can suppress cancer by inhibiting DNA methylation, anti-proliferation and inducing cancer cell apoptosis. One or several of these mechanisms may help explain how green tea extracts reduce risk of breast cancer.

The present study explored possible downstream proteins and pathways that EGCG may affect, in an effort to guide more detailed mechanistic studies to elucidate how EGCG reduces risk of breast cancer. This study used integrative bioinformatics analysis to bring together predictions of protein and pathway targets, followed by Ingenuity Pathway Analysis to build these predicted targets into a network model of interacting molecules that may help explain the presumably complex effects that green tea exerts in breast cancer.

MATERIALS AND METHODS
Dataset of Breast Cancer-Related Genes
The National Center for Biotechnology Information (NCBI) Gene Database (http://www.ncbi.nlm.nih.gov/gene; up to 15 December, 2016), which integrates information from a wide variety of species, was searched for genes related to breast cancer using the search term "breast cancer." Search hits were filtered to retain only *Homo sapiens* genes (Supplementary Table S1).

Dataset of EGCG-Targeted Proteins
The PubChem database of small molecules (http://pubchem. ncbi.nlm.nih.gov; up to 15 December, 2016), including the Compound, Bioassay, and Substance sub-databases, was searched for proteins shown in bioassays to be affected by EGCG (CID:65064) or predicted to be affected by EGCG based on similarity with known binders. Search hits were limited to *Homo sapiens* proteins (Supplementary Table S2).

Prediction of Interaction Networks Affected by EGCG
A network of interacting molecules was built using on-line Ingenuity Pathway Analysis (IPA, www.ingenuity.com) based on the dataset of human genes related to breast cancer and the dataset of EGCG-targeted human proteins ("focus molecules"). Based on the functions of these focus molecules, Pathway Analysis generated a set of networks likely to be affected by EGCG. Molecules were represented as nodes with different shapes depending on their function; and lines were drawn between nodes shown to be biologically related in at least one reference from the literature, a textbook, or other canonical information stored in the Ingenuity Knowledge Base.

The networks generated by IPA were scored according to the significance of the molecules in the network, then the "Compare" module within IPA was used to determine the significance of the association between focus molecules and canonical pathways, based on Fisher's exact test. Finally, we overlaid the two networks to discover the most likely targets of EGCG in breast cancer (**Figure 1**).
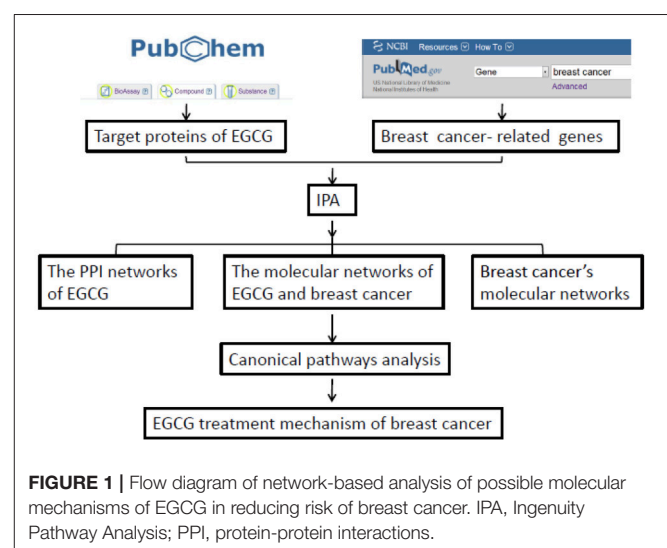
RESULTS
Breast Cancer-Related Gene Networks and Their Functions
A total of 3,237 human genes linked to breast cancer were identified in the GenBank database (Supplementary Table S1), and the encoded proteins were assembled into a set of 25 networks using IPA. These pathways involve primarily cell death and survival, cellular growth and proliferation, inflammatory response, cell-to-cell signaling and interaction, metabolic disease, as well as cellular assembly and organization (Supplementary Table S3, Supplementary Figure S1).

EGCG-Targeted Protein Networks and Their Functions
A total of 65 human proteins targeted by EGCG were identified from the PubChem database (Supplementary Table S2) and their GenInfo Identifier numbers were imported into IPA, which generated protein-protein interaction networks

FIGURE 1 | Flow diagram of network-based analysis of possible molecular mechanisms of EGCG in reducing risk of breast cancer. IPA, Ingenuity Pathway Analysis; PPI, protein-protein interactions.

(Supplementary Table S4, Supplementary Figure S2). Proteins targeted by EGCG participate primarily in the cell cycle; cellular assembly and organization; DNA replication, recombination and repair; cell death and survival; gastrointestinal disease; hepatic system disease; cell morphology; nervous system development and function; organ morphology; and carbohydrate metabolism.

Network Overlap to Predict Pathways Affected by EGCG in Breast Cancer

The "Canonical Pathway" module of IPA identified 485 signaling pathways linked to breast cancer and 246 targeted by EGCG, with 235 signaling pathways shared between the two sets. These overlapping pathways primarily involve molecular mechanisms of cancer, inflammation, glucocorticoid receptor signaling, and cytokine signaling (**Figure 2**).

The "Networks" module of IPA identified 25 networks linked to breast cancer and 6 targeted by EGCG, with 5 networks shared between the two sets (**Figure 3**, Supplementary Table S5). These networks are primarily involved in cell death and survival; DNA replication, recombination and repair; cell cycle; cellular assembly and organization; post-translational modifications related to development; post-transcriptional modifications; and protein synthesis.

Prediction of Specific EGCG Target Proteins in Breast Cancer

Proteins linked to breast cancer and targeted by EGCG participate in several canonical pathways underlying a range of biological activities. To demonstrate the ability of this integrative bioinformatics approach to propose specific protein targets for detailed mechanistic studies, we selected one pathway in the IPA category "molecular mechanisms of cancer" that was linked to breast cancer and targeted by EGCG. Several nodes in this pathway emerged as potential direct targets of EGCG in breast cancer: JUN, FADD, NFKB1, Bcl-2, GNAO1, and MMP14 (**Figure 4**, Supplementary Table S6). These potential
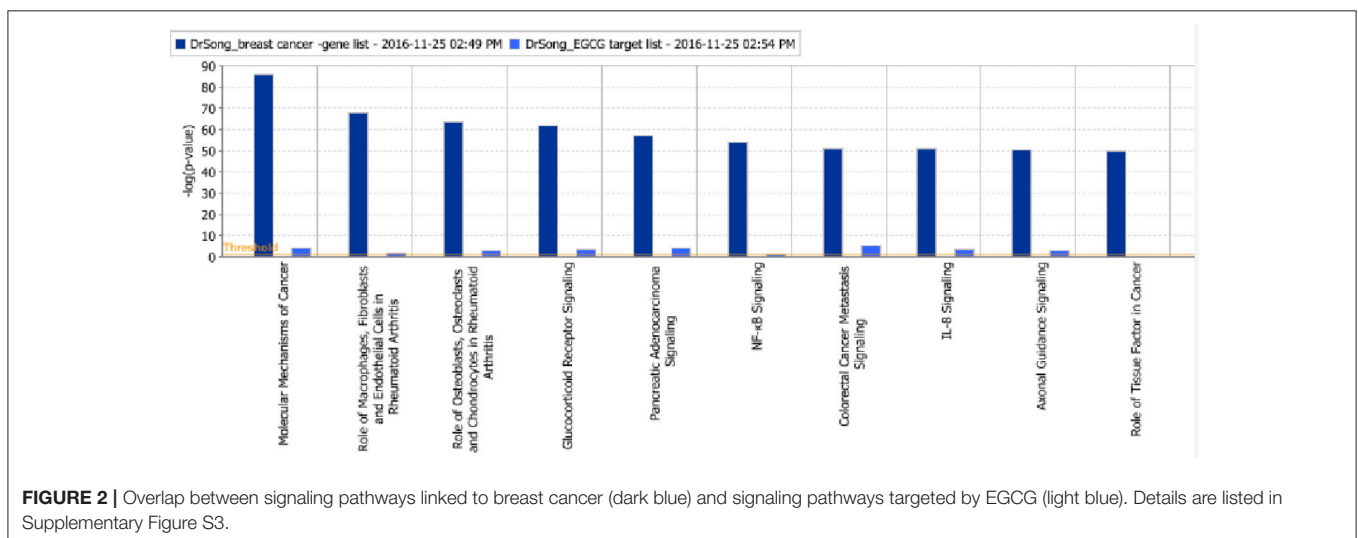
target molecules also appeared in other canonical pathways linked to breast cancer and targeted by EGCG (Supplementary Figures S4, S5).

DISCUSSION

Here we applied an integrative bioinformatics approach drawing on free, publicly available databases of genes linked to breast cancer and of proteins or protein fragments known or predicted to be affected by EGCG. Drawing on the large size of both databases, we were able to identify numerous signaling pathways and networks potentially linked to breast cancer and potentially regulated by EGCG. Then we were able to predict several specific proteins likely to be affected by EGCG in breast cancer. These are strong leads for detailed mechanistic studies, illustrating the power of this bioinformatics-based "panning" or "screening" to guide studies of how EGCG may help reduce risk of breast cancer.

Our network analysis implicated several pathways by which EGCG may reduce breast cancer risk, involving cell death and survival; DNA replication, recombination and repair; cell cycle; cellular assembly and organization. These results are consistent with several studies *in vitro* and *in vivo* suggesting that EGCG exerts anti-carcinogenic activity by protecting DNA from ROS-induced damage and generally alleviating ROS stress (Ruch et al., 1989; Kaur et al., 2007), as well as by inhibiting DNA cleavage (Rathore et al., 2012). Our results are also consistent with work suggesting that EGCG suppresses proliferation and induces apoptosis by down-regulating anti-apoptotic factors such as B cell lymphoma 2 (Bcl-2), Bcl-xL, and vimentin (Leone et al., 2003; Ermakova et al., 2005), as well as by inhibiting NF-κB, JAK/STAT, and PI3K pathways (Surh et al., 2001; Lambert et al., 2010; Van Aller et al., 2011; Senggunprai et al., 2014).

Our analysis of canonical pathways in IPA suggests that EGCG may reduce breast cancer risk by altering pathways involved in molecular mechanisms of cancer, inflammatory
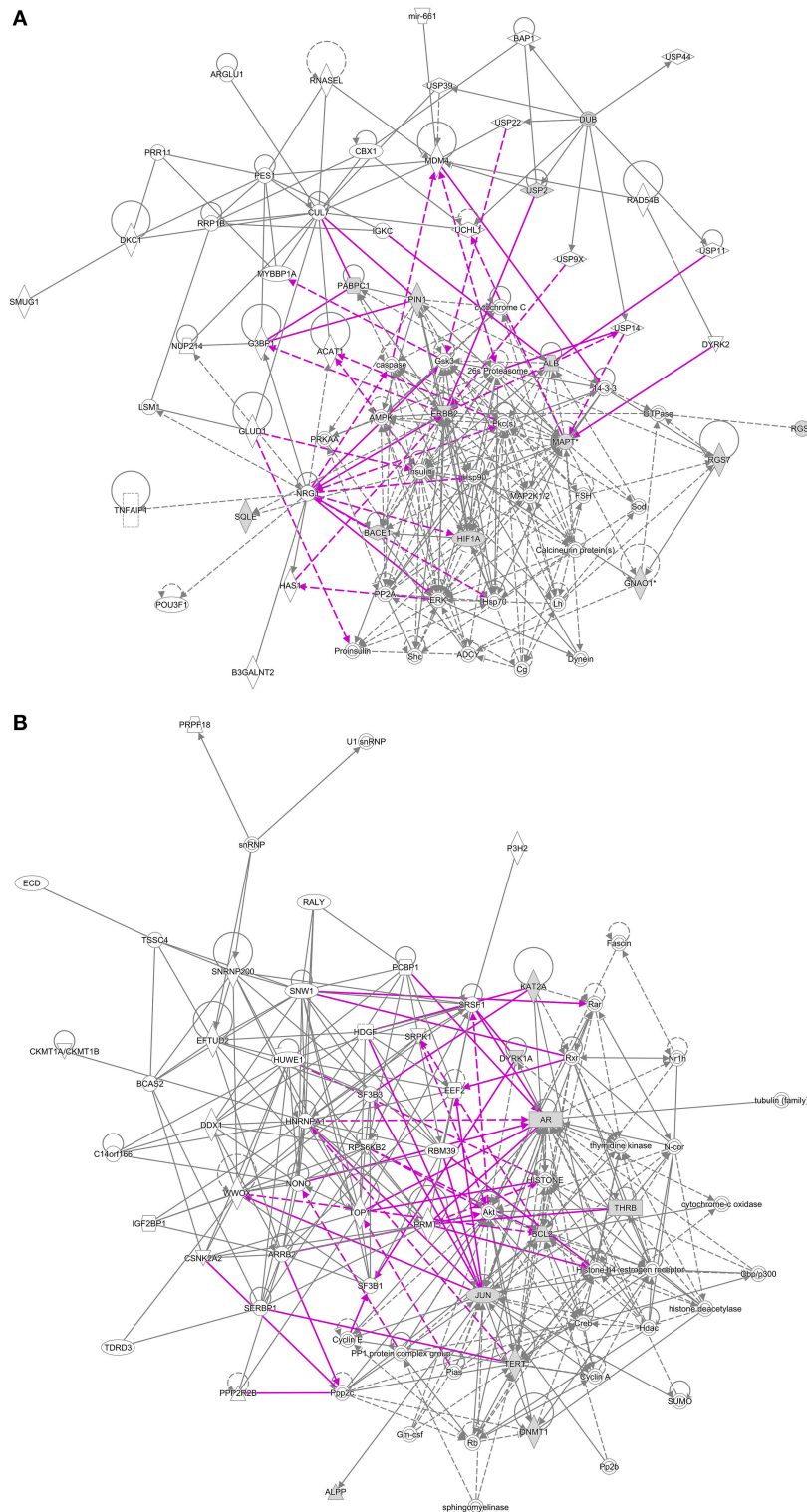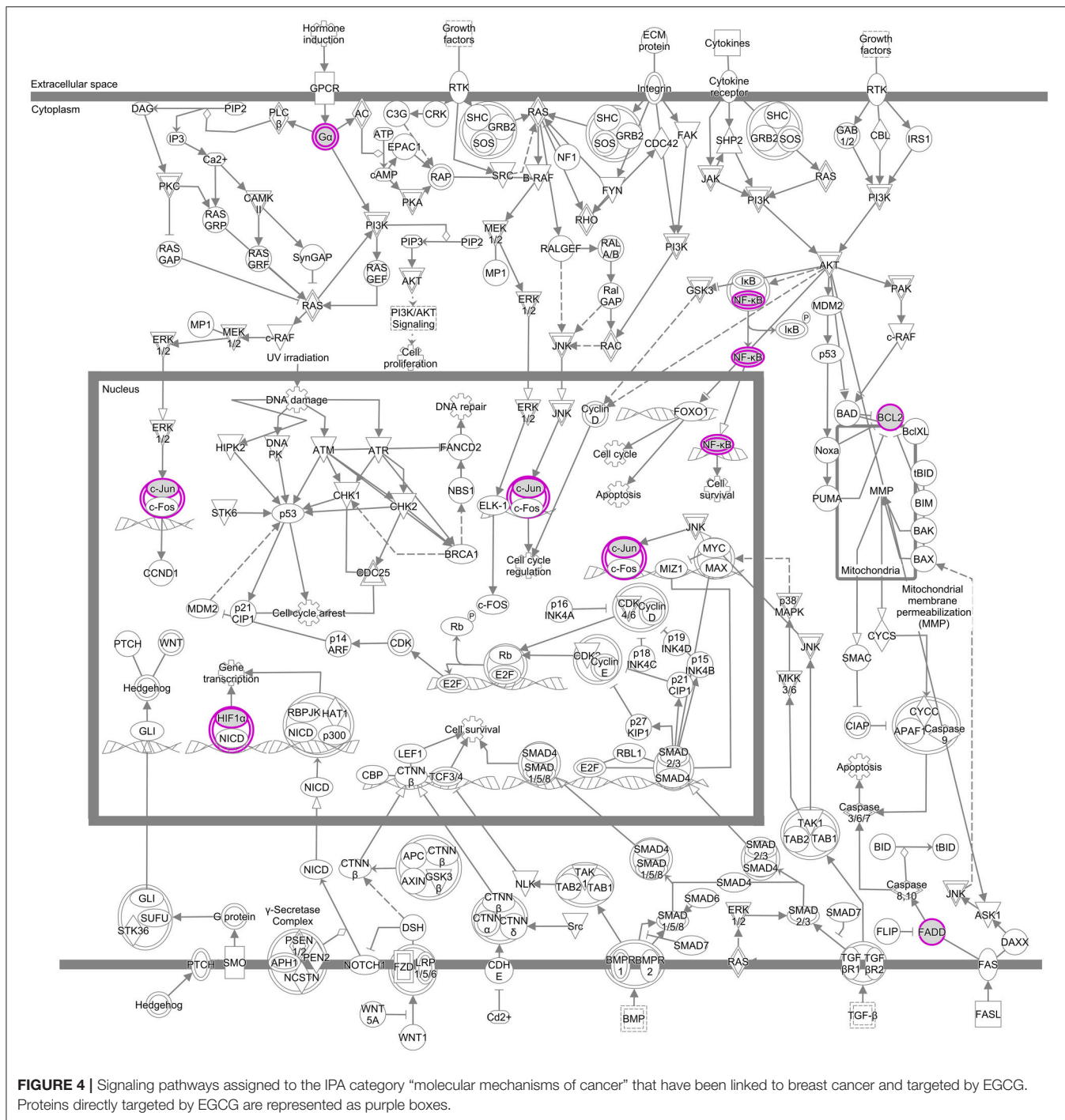

FIGURE 2 | Overlap between signaling pathways linked to breast cancer (dark blue) and signaling pathways targeted by EGCG (light blue). Details are listed in Supplementary Figure S3.

FIGURE 3 | Overlap between networks linked to breast cancer and networks targeted by EGCG. Proteins linked to breast cancer are represented as black lines; EGCG-targeted proteins, as purple lines. **(A)** Shared networks involved in post-translational modification, protein synthesis, cell morphology, and nervous system development. **(B)** Shared networks involved in post-transcriptional modification, cell death and survival, cellular development, cell cycle, cancer, and endocrine disorders.

FIGURE 4 | Signaling pathways assigned to the IPA category "molecular mechanisms of cancer" that have been linked to breast cancer and targeted by EGCG. Proteins directly targeted by EGCG are represented as purple boxes.

signaling, glucocorticoid receptor signaling, and cytokine signaling. In particular, we identified JUN, FADD, NFKB1, Bcl-2, GNAO1, and MMP14 as potential targets of EGCG in breast cancer.

It was reported that p21-activated protein kinase 1 induced the invasion of gastric cancer cells through c-Jun NH2-terminal kinase-mediated activation of matrix metalloproteinase-2, and FADD protected pancreatic cancer cells from drug-induced apoptosis (Li et al., 2017; Zhang et al., 2017).

Researchers Pei et al. (2013) and Stawowczyk et al. (2017) had shown that down-regulation of GNAO1 increased cell proliferation and MMP14 promoted lung cancer by cleavage of heparin-binding EGF-like growth factor. How EGCG may work against these proteins needs more experiments to verify it. In agreement with our results, EGCG has been shown to bind with high affinity to Bcl-2, as well as to down-regulate Bcl-2 and NF-κB. The remaining targets that we identified have not previously been reported in the literature, to the best of our knowledge.

This means that they may be novel potential targets that merit validation and, if positive, detailed mechanistic studies.

CONCLUSION

Integrative bioinformatics analysis and chemical fragment analysis is a promising method for *in silico* "panning" or "screening" of proteins and pathways that EGCG may affect and thereby reduce risk of breast cancer. This approach may be suitable for analyzing the mechanism of action of other bioactive compounds. Our network analysis allowed us to identify several pathways by which EGCG may reduce cancer risk; these pathways are involved mainly in cell death and survival; DNA replication, recombination and repair; cell cycle; cellular assembly and organization; post-translational modification related to development; post-transcriptional modification; and protein synthesis. Our network analysis also allowed us to identify several specific proteins that EGCG may help regulate in breast cancer, including JUN, FADD, NFKB1, Bcl-2, GNAO1, and MMP14.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: XS. Analyzed the data: MZ, LC, QL. Wrote the paper: XS.

INFORMED CONSENT

Informed consent was obtained from all individual participants included in the study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmolb.2017.00043/full#supplementary-material

REFERENCES

Chen, X., Lu, W., Zheng, Y., Gu, K., Chen, Z., Zheng, W., et al. (2010). Exercise, tea consumption, and depression among breast cancer survivors. *J. Clin. Oncol.* 28, 991–998. doi: 10.1200/JCO.2009.23.0565

Dai, Q., Shu, X. O., Li, H., Yang, G., Shrubsole, M. J., Cai, H., et al. (2010). Is green tea drinking associated with a later onset of breast cancer? *Ann. Epidemiol.* 20, 74–81. doi: 10.1016/j.annepidem.2009.09.005

Du, G. J., Zhang, Z., Wen, X. D., Yu, C., Calway, T., Yuan, C. S., et al. (2012). Epigallocatechin Gallate (EGCG) is the most effective cancer chemopreventive polyphenol in green tea. *Nutrients* 4, 1679–1691. doi: 10.3390/nu4111679

Ermakova, S., Choi, B. Y., Choi, H. S., Kang, B. S., Bode, A. M., and Dong, Z. (2005). The intermediate filament protein vimentin is a new target for epigallocatechin gallate. *J. Biol. Chem.* 280, 16882–16890. doi: 10.1074/jbc.M414185200

Fang, M. Z., Wang, Y., Ai, N., Hou, Z., Sun, Y., Lu, H., et al. (2003). Tea polyphenol (-)-epigallocatechin-3-gallate inhibits DNA methyltransferase and reactivates methylation-silenced genes in cancer cell lines. *Cancer Res.* 63, 7563–7570.

Ganmaa, D., Willett, W. C., Li, T. Y., Feskanich, D., van Dam, R. M., Lopez-Garcia, E., et al. (2008). Coffee, tea, caffeine and risk of breast cancer: a 22-year follow-up. *Int. J. Cancer* 122, 2071–2076. doi: 10.1002/ijc.23336

Harbeck, N., and Gnant, M. (2017). Breast cancer. *Lancet.* 389, 1134–1150. doi: 10.1016/S0140-6736(16)31891-8

Imai, K., Suga, K., and Nakachi, K. (1997). Cancer-preventive effects of drinking green tea among a Japanese Population (Paris). *Prev Med.* 26, 769–775.

Inoue, M., Robien, K., Wang, R., Van Den Berg, D. J., Koh, W. P., and Yu, M. C. (2008). Green tea intake, MTHFR/TYMS genotype and breast cancer risk: the Singapore Chinese Health Study. *Carcinogenesis* 29, 1967–1972. doi: 10.1093/carcin/bgn177

Jones, P. A., and Takai, D. (2001). The role of DNA methylation in mammalian epigenetics. *Science* 293, 1068–1070. doi: 10.1126/science.1063852

Kanwar, J., Taskeen, M., Mohammad, I., Huo, C., Chan, T. H., and Dou, Q. P. (2012). Recent advances on tea polyphenols. *Front. Biosci.* 4, 111–131. doi: 10.2741/e363

Kaur, S., Greaves, P., Cooke, D. N., Edwards, R., Steward, W. P., Gescher, A. J., et al. (2007). Breast cancer prevention by green tea catechins and black tea theaflavins in the C3(1) SV40 T,t antigen transgenic mouse model is accompanied by increased apoptosis and a decrease in oxidative DNA adducts. *J. Agric. Food Chem.* 55, 3378–3385. doi: 10.1021/jf0633342

Kumar, N., Titus-Ernstoff, L., Newcomb, P. A., Trentham-Dietz, A., Anic, G., and Egan, K. M. (2009). Tea consumption and risk of breast cancer. *Cancer Epidemiol. Biomarkers Prev.* 18, 341–345. doi: 10.1158/1055-9965.EPI-08-0819

Lambert, J. D., Sang, S., Hong, J., and Yang, C. S. (2010). Anticancer and anti-inflammatory effects of cysteine metabolites of the green tea polyphenol, (-)-epigallocatechin-3-gallate. *J. Agric. Food Chem.* 58, 10016–10019. doi: 10.1021/jf102311t

Lee, W. J., Shim, J. Y., and Zhu, B. T. (2005). Mechanisms for the inhibition of DNA methyltransferases by tea catechins and bioflavonoids. *Mol. Pharmacol.* 68, 1018–1030. doi: 10.1124/mol.104.008367

Leone, M., Zhai, D., Sareth, S., Kitada, S., Reed, J. C., and Pellecchia, M. (2003). Cancer prevention by tea polyphenols is linked to their direct inhibition of antiapoptotic Bcl-2-family proteins. *Cancer Res.* 63, 8118–8121.

Li, L. H., Wu, G. Y., Lu, Y. Z., Chen, X. H., Liu, B. Y., Zheng, M. H., et al. (2017). p21-activated protein kinase 1 induces the invasion of gastric cancer cells through c-Jun NH2-terminal kinase-mediated activation of matrix metalloproteinase-2. *Oncol. Rep.* 38, 193–200. doi: 10.3892/or.2017.5643

Li, M., Tse, L. A., Chan, W. C., Kwok, C. H., Leung, S. L., Wu, C., et al. (2016). Evaluation of breast cancer risk associated with tea consumption by menopausal and estrogen receptor status among Chinese women in Hong Kong. *Cancer Epidemiol.* 40, 73–78. doi: 10.1016/j.canep.2015.11.013

Pei, X., Zhang, J., Wu, L., Lu, B., Zhang, X., Yang, D., et al. (2013). The down-regulation of GNAO1 and its promoting role in hepatocellular carcinoma. *Biosci. Rep.* 33:e00069. doi: 10.1042/BSR20130001

Rathore, K., Choudhary, S., Odoi, A., and Wang, H. C. (2012). Green tea catechin intervention of reactive oxygen species-mediated ERK pathway activation and chronically induced breast cell carcinogenesis. *Carcinogenesis* 33, 174–183. doi: 10.1093/carcin/bgr244

Ruch, R. J., Cheng, S. J., and Klaunig, J. E. (1989). Prevention of cytotoxicity and inhibition of intercellular communication by antioxidant catechins isolated from Chinese green tea. *Carcinogenesis* 10, 1003–1008. doi: 10.1093/carcin/10.6.1003

Senggunprai, L., Kukongviriyapan, V., Prawan, A., and Kukongviriyapan, U. (2014). Quercetin and EGCG exhibit chemopreventive effects in cholangiocarcinoma cells via suppression of JAK/STAT signaling pathway. *Phytother. Res.* 28, 841–848. doi: 10.1002/ptr.5061

Shrubsole, M. J., Lu, W., Chen, Z., Shu, X. O., Zheng, Y., Dai, Q., et al. (2009). Drinking green tea modestly reduces breast cancer risk. *J. Nutr.* 139, 310–316. doi: 10.3945/jn.108.098699

Stawowczyk, M., Wellenstein, M. D., Lee, S. B., Yomtoubian, S., Durrans, A., Choi, H., et al. (2017). Matrix Metalloproteinase 14 promotes lung cancer by cleavage of Heparin-Binding EGF-like Growth Factor. *Neoplasia* 19, 55–64. doi: 10.1016/j.neo.2016.11.005

Surh, Y. J., Chun, K. S., Cha, H. H., Han, S. S., Keum, Y. S., Park, K. K., et al. (2001). Molecular mechanisms underlying chemopreventive activities of anti-inflammatory phytochemicals: down-regulation of COX-2 and iNOS through suppression of NF-kappa B activation. *Mutat. Res.* 480–481, 243–268. doi: 10.1016/S0027-5107(01)00183-X

Van Aller, G. S., Carson, J. D., Tang, W., Peng, H., Zhao, L., Copeland, R. A., et al. (2011). Epigallocatechin gallate (EGCG), a major component of green tea, is a dual phosphoinositide-3-kinase/mTOR inhibitor. *Biochem. Biophys. Res. Commun.* 406, 194–199. doi: 10.1016/j.bbrc.2011.02.010

Wu, A. H., Yu, M. C., Tseng, C. C., Hankin, J., and Pike, M. C. (2003). Green tea and risk of breast cancer in Asian Americans. *Int. J. Cancer* 106, 574–579. doi: 10.1002/ijc.11259

Xiang, L. P., Wang, A., Ye, J. H., Zheng, X. Q., Polito, C. A., Lu, J. L., et al. (2016). Suppressive effects of tea catechins on breast cancer. *Nutrients* 8:E458. doi: 10.3390/nu8080458

Yuan, J. M., Koh, W. P., Sun, C. L., Lee, H. P., and Yu, M. C. (2005). Green tea intake, ACE gene polymorphism and breast cancer risk among Chinese women in Singapore. *Carcinogenesis* 26, 1389–1394. doi: 10.1093/carcin/bgi080

Zhang, M., Holman, C. D., Huang, J. P., and Xie, X. (2007). Green tea and the prevention of breast cancer: a case-control study in Southeast China. *Carcinogenesis* 28, 1074–1078. doi: 10.1093/carcin/bgl252

Zhang, M., Huang, J., Xie, X., and Holman, C. D. (2009). Dietary intakes of mushrooms and green tea combine to reduce the risk of breast cancer in Chinese women. *Int. J. Cancer* 124, 1404–1408. doi: 10.1002/ijc.24047

Zhang, R., Liu, Y., Hammache, K., He, L., Zhu, B., Cheng, W., et al. (2017). The role of FADD in pancreatic cancer cell proliferation and drug resistance. *Oncol. Lett.* 13, 1899–1904. doi: 10.3892/ol.2017.5636