



Integrating Genome-Wide Association Studies and Gene Expression Profiles With Chemical-Genes Interaction Networks to Identify Chemicals Associated With Colorectal Cancer

Xinyue Tan, Hanmin Tang, Liyun Gong, Lina Xie, Yutiantian Lei, Zhenzhen Luo, Chenchen He, Jinlu Ma and Suxia Han*

OPEN ACCESS

Edited by:

Ryan Matthew Layer,
University of Colorado Boulder,
United States

Reviewed by:

Shihao Shen,
University of California, Los Angeles,
United States

Lei Zhang,
Soochow University, China

*Correspondence:

Suxia Han
shan87@mail.xjtu.edu.cn

Specialty section:

This article was submitted to
Bioinformatics and Computational
Biology,
a section of the journal
Frontiers in Genetics

Received: 18 July 2019

Accepted: 27 March 2020

Published: 24 April 2020

Citation:

Tan X, Tang H, Gong L, Xie L,
Lei Y, Luo Z, He C, Ma J and Han S
(2020) Integrating Genome-Wide
Association Studies and Gene
Expression Profiles With
Chemical-Genes Interaction Networks
to Identify Chemicals Associated With
Colorectal Cancer.
Front. Genet. 11:385.
doi: 10.3389/fgene.2020.00385

Department of Oncology, The First Affiliated Hospital, Xi'an Jiaotong University, Xi'an, China

Colorectal cancer (CRC) is the third most common cancer and has the second highest mortality rate in global cancer. Exploring the associations between chemicals and CRC has great significance in prophylaxis and therapy of tumor diseases. This study aims to explore the relationships between CRC and environmental chemicals on genetic basis by bioinformatics analysis. The genome-wide association study (GWAS) datasets for CRC were obtained from the UK Biobank. The GWAS data for colon cancer (category C18) includes 2,581 individuals and 449,683 controls, while that of rectal cancer (category C20) includes 1,244 individuals and 451,020 controls. In addition, we derived CRC gene expression datasets from the NCBI-GEO (GSE106582). The chemicals related gene sets were acquired from the comparative toxicogenomics database (CTD). Transcriptome-wide association study (TWAS) analysis was applied to CRC GWAS summary data and calculated the expression association testing statistics by FUSION software. We performed chemicals related gene set enrichment analysis (GSEA) by integrating GWAS summary data, mRNA expression profiles of CRC and the CTD chemical-gene interaction networks to identify relationships between chemicals and genes of CRC. We observed several significant correlations between chemicals and CRC. Meanwhile, we also detected 5 common chemicals between colon and rectal cancer, including methylnitronitrosoguanidine, isoniazid, PD 0325901, sulindac sulfide, and importazole. Our study performed TWAS and GSEA analysis, linked prior knowledge to newly generated data and thereby helped identifying chemicals related to tumor genes, which provides new clues for revealing the associations between environmental chemicals and cancer.

Keywords: colorectal cancer, genome-wide association study, transcriptome-wide association study, gene set enrichment analysis, comparative toxicogenomics database

INTRODUCTION

Colorectal cancer (CRC) is the third most common cancer worldwide and has the second highest mortality rate in global cancer (Bray et al., 2018; Ferlay et al., 2019). In western countries, CRC accounts for about 10% of cancer deaths (Kuipers et al., 2015). The accepted view is that genetic, lifestyle, and environmental factors are closely related to CRC (Dekker et al., 2019). Current research shows that environmental chemicals play important roles in the etiology of CRC. Several chemicals have been suggested to promote the tumorigenesis and development of CRC. For instance, analysis of an Iowa Women's Health Study cohort suggested that exposure to TTHM in drinking water is associated with increased risk of rectal cancer (Jones et al., 2019). In addition, another case-control study observed that organochlorine and organophosphorus pesticides may induce CRC (Abolhassani et al., 2019). In contrast, numerous chemicals have been identified that inhibit CRC. Metastatic CRC (mCRC) often indicates a poor prognosis. The 5-year overall survival (OS) rate of patients with mCRC is less than 15% (Siegel et al., 2017; Bray et al., 2018), and the median OS of unresectable mCRC patients who received only supportive therapy was only 5 months (Lucas et al., 2011). However, the 5-years OS rate increased to 10% in such patients receiving 5-fluorouracil (5-FU)/leucovorin (LV) plus oxaliplatin (FOLFOX) (Gustavsson et al., 2015). Thus, FOLFOX chemotherapy regimen is still the standard first-line therapy for unresectable mCRC (Giacchetti et al., 2000; Goldberg et al., 2004; Kouhara et al., 2007; Bokemeyer et al., 2011). Recent studies have found that some non-chemotherapeutic chemicals also have an inhibitory effect on CRC, such as semisynthetic retinoid, lidocaine, and beta-carotene (Mattingly et al., 2003; Pham et al., 2013; Qu et al., 2018).

Therefore, it has great significance to clarify the relationship between chemicals in environmental and CRC for the treatment and prevention of diseases. But obtaining the entire life-time exposure of an individual is difficult and complex, for lacking sensitive methods to measure specific exposures. Although the exposure is known to have occurred, the transient character of the exposure indicators increases the difficulty of measuring the specific exposure (Messerlian et al., 2017). With the help of new technologies, such as genome-wide association research (GWAS), more convenient and efficient analyses have been produced to identify interactions between multiple environmental exposures with genes (Khoury et al., 2005). Studies of gene-environment interactions have been widely applied in psychological research, however, lack in the field of oncology (Manuck and McCaffery, 2014). The Comparative Toxicogenomics Database (CTD) is a public repository, aims to advance people's understanding of how environmental exposures affects human health (Mattingly et al., 2003). This database provides information regarding chemical-gene/protein interactions as well as chemical- and gene-disease relationships that is organized by individual genes, gene sets, organisms, chemicals, sequence type (DNA, mRNA, and protein), gene ontology annotations and sequences (Mattingly et al., 2006).

Genome-wide association studies (GWAS) analyze DNA sequence variations to provide associations for complex human

traits and diseases efficiently (Tam et al., 2019). Transcriptome-wide association studies (TWAS) is further developed on this basis, which can evaluate the association of each gene to diseases by integrating tissue-related gene expression measurements with GWAS summary data (Gong et al., 2018). Currently, TWAS has been proved with high efficiency in determining the genetic mechanism of complex diseases (Gusev et al., 2018; Wu et al., 2018). The Gene Expression Omnibus (GEO) is a worldwide resource which distributes a large number of high-throughput microarray and next-generation sequence functional genomic data sets (Barrett et al., 2013). Different from the traditional GWAS to explain the relationship between DNA and external phenotype, we simultaneously used the GEO to obtain the gene expression profile (mRNA expression profile chip data) of colorectal cancer, that is, a comprehensive analysis at the DNA and mRNA level. This is helpful to narrow the range of chemicals related candidate genes on the basis of traditional GWAS analysis.

Briefly, in this work, the CTD chemical-gene interaction networks, GWAS summary datasets and gene expression profiles were integrated. TWAS analysis was performed by FUSION software to evaluate the expression association testing statistics. The gene set enrichment analysis (GSEA) with the running-sum statistic and weighted Kolmogorov-Smirnov-like statistic were applied to detect the correlation between environmental chemicals and CRC (Charnpi and Ycart, 2015). Firstly, we obtained the empirical distributions of GSEA statistics for each chemical for statistical tests. Subsequently, the *P*-value of each chemical was conducted from the permuted empirical distribution of GSEA statistics. Finally, we summarized and analyzed the obtained chemicals associated with CRC.

MATERIALS AND METHODS

GWAS Summary Dataset for CRC

GeneATLAS¹, a huge resource storing the information of hundreds of traits and millions of related gene variants based on the UK Biobank cohort, provides a convenient way for researchers to acquire data from the UK Biobank (Lin et al., 2019). To be specific, it allows researchers to query genome-wide association results for 9,113,133 genetic variants and download over 30 million genetic variants (>23 billion phenotype-genotype pairs) for GWAS summary statistics (Canela-Xandri et al., 2018).

A large-scale GWAS summary data of colon cancer and rectal cancer in our study were downloaded from the GeneATLAS in UK Biobank. In the cancer register category, 103,470 data items are available from 84,726 participants. In brief, our GWAS summary data, which contained 5,899 available data items, were from categories C18 (malignant neoplasm of colon) and C20 (malignant neoplasm of rectum). Detailed information regarding the methods, process, and approaches were described in the previous studies (Hammerschlag et al., 2017).

¹<http://geneatlas.roslin.ed.ac.uk/>

Gene Expression Datasets of CRC

NCBI-GEO² is an international public repository with next-generation sequencing and microarray/gene profiles which was used in this study to obtain the mRNA expression profiles of mucosa and colorectal tumor tissues (GSE106582). CRC patients were recruited at the University Hospital of Heidelberg, from whom the gene expression profiles of 77 tumor and 117 mucosa tissues were obtained using an Illumina HumanHT-12 V4.0 expression beadchip. Using GEO2R, a web tool based on the GEO database, differential gene expression was assessed by comparing the expression of genes from colorectal tumor tissues to those of respective mucosa tissues.

Transcriptome-Wide Association Study (TWAS) Methodology

TWAS analysis utilizes disease GWAS summary statistics combining with pre-computed gene expression weights to calculate the association of every gene with known diseases (Gusev et al., 2018). In other words, TWAS can integrate the associations between GWAS and gene expression measurements to identify genes associated with traits. In this study, A TWAS for CRC was conducted using functional summary-based imputation (FUSION) software and the gene expression weight references of whole blood, rectum, and colon tissues were acquired from the FUSION website³. Specifically, the gene expression weights of whole blood were collected from 1,264 subjects of the Young Finns Study (Raitakari et al., 2008; Nuotio et al., 2014).

Firstly, based on FUSION software we performed prediction models to calculate the gene expression weights of different tissues (Huang et al., 2019). Then we conducted the correlation statistics between gene expressed level and CRC combining the gene expression weights and summary-level GWAS results. $Z_{TWAS} = w^*Z/(wLw)^{1/2}$ was used to calculate the association statistics. Z denotes the scores of CRC while w denotes the weights. L means the SNP-correlation linkage disequilibrium (LD) matrix. A TWAS p -value was calculated for each gene within whole blood, rectum and colon tissues, respectively (Qi et al., 2019). The genes with $p < 0.05$ were considered as significant. Detailed information can be found in the published study (Gusev et al., 2018).

Chemical-Gene Expression Interaction Database of the CTD Database

The Comparative Toxicogenomics Database (CTD)⁴ is a publicly and accessible database for toxicogenomic information (Zhang et al., 2019). The CTD currently includes more than 30.5 million toxicogenomic relationships associated with chemicals, proteins, etc. (Davis et al., 2017) and provides information regarding chemical, gene, phenotype, and disease relationships to advance our understanding of the effects of environmental toxin exposure on public health (Grondin et al., 2018). A unique and powerful feature of the CTD is knowledge transfer with

respect to any information that is directly annotated to chemicals, genes and diseases (Davis et al., 2013). This study download 11,190 chemicals related gene sets from the CTD. The process of retrieving information using CTD was described in the study previously (Mattingly et al., 2006).

Identification of Environmental Chemicals Elements Associated With Colorectal Cancer

The GSEA algorithm was originally used for microarray study and GWAS-based GSEA was developed subsequently (He et al., 2018). At present, it is utilized to identify abnormally expressed gene sets for target diseases, and has been applied in etiology researches of multiple diseases (Wang et al., 2007). Firstly, for the j th ($j = 1, 2, 3, \dots, N$) gene, the most significant GWAS association test statistics of the SNPs was assigned to j th gene according to the score r_j of the given gene. Secondly, all genes $G = (G_1^*, G_2^*, \dots, G_N^*)$ were ranked by their scores from the highest to the lowest (He et al., 2018), which was expressed as $U = (j_1^*, j_2^*, \dots, j_N^*)$. Thirdly, for a chemicals related gene set S , an enrichment score ES was calculated for CRC by the running sum statistic and weighted Kolmogorov-Smirnov-like statistic (Subramanian et al., 2005; Charmpi and Ycart, 2015). Gene set S independently derived from N_H genes. ES represents the overrepresentation of CRC associated genes in chemicals related gene set S . ES was calculated as:

$$ES(S) = \max_{1 \leq j \leq N} \left\{ \sum_{G_{j^*} \in S, j^* \leq j} \frac{|r_{j^*}|^p}{N_R} - \sum_{G_{j^*} \notin S, j^* \leq j} \frac{1}{N - N_H} \right\},$$

where

$$N_R = \sum_{G_j \in S} |r_{j^*}|^p.$$

Finally, after L time permutations, we can obtain the null distribution of $ES^{null} = (ES_1^{null}, ES_2^{null}, \dots, ES_L^{null})$. To control the effect of the gene sets with varying sizes, the observed $ES(S)$ is normalized by the average value and standard deviation of the permuted ES_S^{null} , defined by $NES^S = \frac{ES^S - \text{mean}(ES_S^{null})}{SD(ES_S^{null})}$. The P -values were finally calculated from the NES for each chemicals related gene set.

This study conducted a total of 5,000 permutations to calculate the empirical distributions of GSEA statistics of each chemical. And the chemicals related gene sets with $P < 0.05$ are considered statistically significant. Previous research provides the detailed descriptions regarding this approach (Zhao et al., 2018). Similarly, all mRNA expression profile from GEO were analyzed using the same approach (Weng et al., 2011).

RESULTS

Environmental Chemicals Associated With Colorectal Cancer

From the CRC GWAS summary datasets, we identified 175 chemicals that were significantly associated with colon cancer

²<https://www.ncbi.nlm.nih.gov/gds/>

³<http://gusevlab.org/projects/fusion/>

⁴<http://ctdbase.org/>

(including 34 for colon tissue and 141 for whole blood) as well as 103 chemicals significantly associated with rectal cancer (including 20 for rectal tissue and 83 for whole blood) ($P < 0.05$; **Supplementary Tables S1, S2**). For the expression profile of CRC, we identified 1,198 significant chemicals ($P < 0.05$; **Supplementary Table S3**).

After a comparative analysis of the TWAS and mRNA expression profile GSEA results, we significantly detected several chemicals associated with the colon cancer and rectal cancer ($P < 0.05$). For colon cancer, 104 common chemicals were detected, including 83 in colon tissue and 24 in blood tissue, and 3 in both tissues (**Supplementary Table S4**), such as Antirheumatic Agents (P -value1 = 0.0244, P -value2 = 0.0230), Chenodeoxycholic Acid (P -value1 = 0.0002, P -value2 = 0.0002) and Trientine (P -value1 = 0.0464, P -value2 = 0.0314; Supplementary note: In this paragraph, P -value1 is P -value in GWAS dataset and P -value2 is P -value in mRNA expression profile). For rectal cancer, 51 common chemicals were discovered, including 12 in rectum tissue and 39 in blood tissue (**Supplementary Table S5**). **Tables 1, 2** summarized the top 10 chemicals identified for the colon cancer and rectal cancer separately.

TABLE 1 | List of top ten chemicals identified for colon cancer after a comparative of GWAS and mRNA GSEA results.

Chemical Name	P -value1 ^a	P -value2 ^b
Antirheumatic Agents	0.0002	0.0002
LG 100815	0.0004	0.0004
Zinc Acetate	0.0010	0.0004
Aerosols	0.0016	0.0002
Titanium dioxide	0.0026	0.0002
Motexafin gadolinium	0.0046	0.0006
Clofibrac Acid	0.0052	0.0002
Vitallium	0.0052	0.0002
Raloxifene Hydrochloride	0.0066	0.0002
Soman	0.0094	0.0002

^a P -value1: P -value in GWAS dataset. ^b P -value2: P -value in mRNA expression profile.

TABLE 2 | List of top ten chemicals identified for rectal cancer after a comparative of GWAS and mRNA GSEA results.

Chemical name	P -value1 ^a	P -value2 ^b
NAD	0.0020	0.0004
Sulindac sulfide	0.0052	0.0002
Casticin	0.0086	0.0002
Benz(a)anthracene	0.0124	0.0002
Methylnitronitrosoguanidine	0.0132	0.0002
Afimoxifene	0.0134	0.0002
4-phenylbutyric acid	0.0150	0.0004
Nickel	0.0178	0.0002
Ochratoxin A	0.0180	0.0002
Promethazine	0.0196	0.0006

^a P -value1: P -value in GWAS dataset. ^b P -value2: P -value in mRNA expression profile.

TABLE 3 | The common significant chemicals between colon cancer and rectal cancer GSEA results.

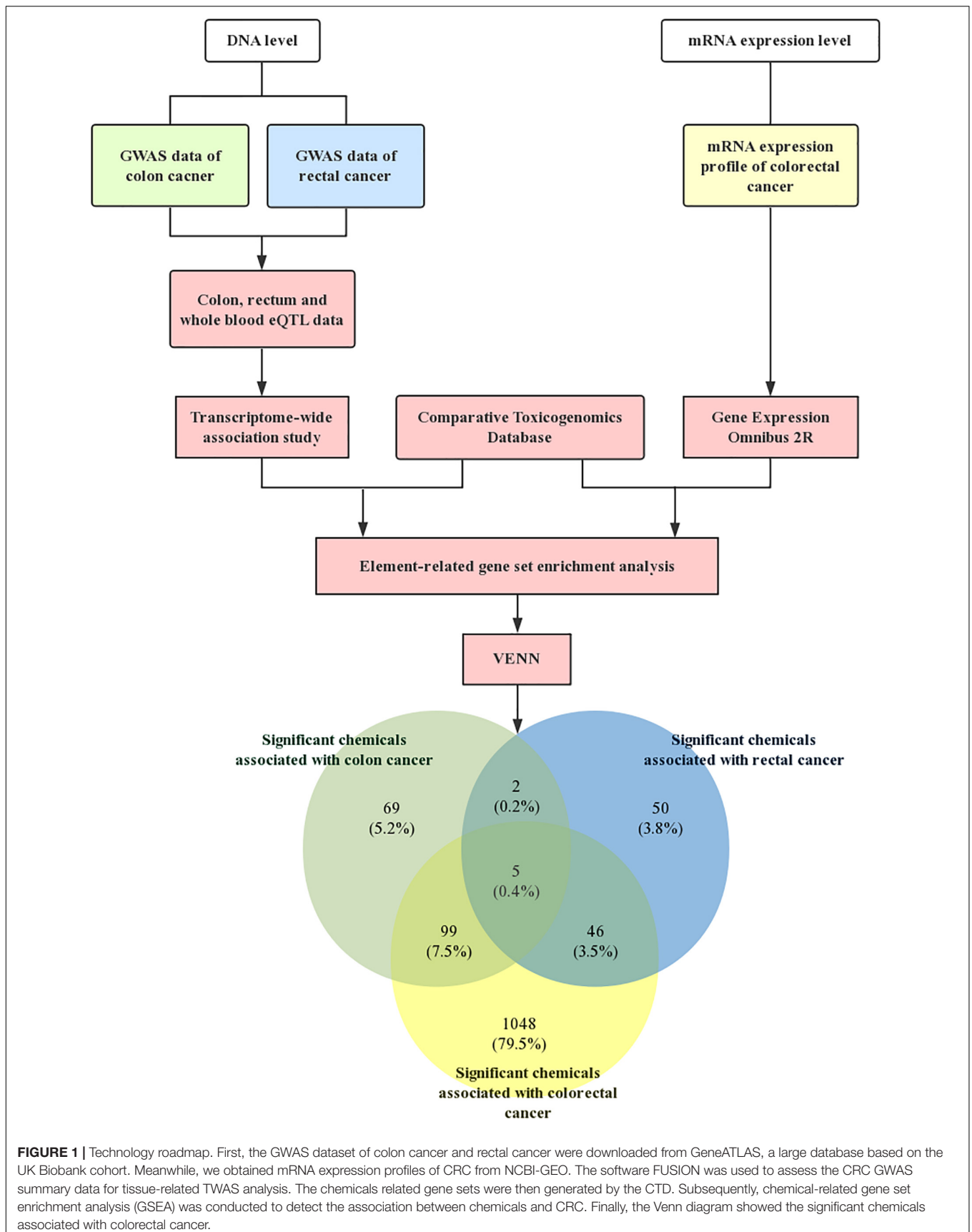
Chemical name	P -value1 ^a	P -value2 ^b	P -value3 ^c
Methylnitronitrosoguanidine	0.0394	0.0132	0.0002
Isoniazid	0.0164	0.0262	0.0068
PD 0325901	0.0348	0.0406	0.0012
Sulindac sulfide	0.0374	0.0052	0.0002
Importazole	0.0378	0.0450	0.0224

^a P -value1: P -value in colon GWAS dataset. ^b P -value2: P -value in rectal GWAS dataset. ^c P -value3: P -value in mRNA expression profile.

Meanwhile, **Table 3** shows the common significant environmental chemicals between colon cancer and rectal cancer. We detected 5 chemicals, including methylnitronitrosoguanidine (P -value1 = 0.0394, P -value2 = 0.0132, P -value3 = 0.0002), isoniazid (P -value1 = 0.0164, P -value2 = 0.0262, P -value3 = 0.0068), PD 0325901 (P -value1 = 0.0348, P -value2 = 0.0406, P -value3 = 0.0012), sulindac sulfide (P -value1 = 0.0374, P -value2 = 0.0052, P -value2 = 0.0002), importazole (P -value1 = 0.0378, P -value2 = 0.0450, P -value3 = 0.0224; Supplementary note: In this paragraph, P -value1 is P -value in colon GWAS dataset, P -value2 is P -value in rectal GWAS dataset, P -value3 is P -value in mRNA expression profile). The specific technology roadmap and Venn diagram are shown in **Figure 1**.

DISCUSSION

CRC is the fourth deadliest cancer lead to 900,000 deaths worldwide annually (Dekker et al., 2019). It has become a global public health problem due to its high morbidity and mortality worldwide. Both genetic and environmental factors play significant roles in the etiology of colorectal cancer. Cancer risk factors include biological agents (infection), exposure to synthetic chemicals, and lifestyle factors, which together contribute to the development of 70–95% of cancers (Wu et al., 2016). Several chemicals have been reported promote the tumorigenesis and tumor development of CRC (Abolhassani et al., 2019; Cernigliaro et al., 2019; Jones et al., 2019). This provides a new clue for us to prevent the occurrence of colorectal cancer. Meanwhile, except for the standard treatment, many chemicals have been reported to inhibit CRC in recent years. For example, the anti-colorectal cancer effect of awsonaringenin (LSG), a flavonoid compound, has been demonstrated in previous research (Anwar et al., 2018). Environmental chemicals are related to various malignant tumors besides CRC. For example, acrylamide, benzo(a)pyrene and polychlorinated biphenyls can induce carcinogenesis for cytotoxicity and DNA damage to hepatic cells (Erkekoglu et al., 2017). The discovery for the active substance in chemicals related cancer is of great significance for the treatment to tumor patients. Since the chemicals environmental exposure is usually complex and accurately measuring exposure levels *in vivo* is still with many objective problems, we try to explore the relationships between chemicals and cancer in an easier way.



In this study, we extended the classical GSEA approach to detect associations between chemicals and CRC using TWAS data and gene expression datasets. We identified several chemicals showing genetic correlation evidence with the CRC.

We identified several significant chemicals for the colon cancer, such as aspirin and titanium dioxide, which have been reported by previous study. Aspirin, a well-known antirheumatic drug, is proved that can prolong the survival of patients with colorectal cancer and activate T cell-mediated antitumor immunity (Hamada et al., 2017). Bettini, Boutet-Robinet et al. has reported that daily oral food-grade titanium dioxide (TiO₂) intake is related to an chronic intestinal inflammation and will increase the risk of carcinogenesis (Bettini et al., 2017).

NAD and Nickel are two remarkable chemicals associated with rectal cancer. A recent study revealed that increased nicotinamide adenine dinucleotide pool suppressed reactive oxygen species level to promote progression of colon cancer (Hong et al., 2019). In a previous study, trace elements in normal and cancerous tissue which obtained from 18 patients suffering from colon and rectum cancer were quantitatively determined by X-ray fluorescence, and the result showed that Nickel elevated in cancerous tissues (Gregoriadis et al., 1983).

Five overlapped chemicals have been identified associated with CRC, including the carcinogens methylnitrosoguanidine, isoniazid. And PD 0325901, sulindac sulfide and importazole have the ability to inhibit the carcinogenesis and development of cancer.

Methylnitrosoguanidine (MNNG) is anticipated to be declared a human carcinogen based on sufficient evidence of its carcinogenicity from investigations involving animal models. MNNG caused tumors at different tissue sites in several animal model species by several different exposure routes. Research indicated that the intrarectal infusion of MNNG into large intestine of rats can cause tumors (Tsukamoto et al., 2015; U.S. Department of Health and Human Services, 2016).

Isoniazid (INH) is an irreversible inhibitor of Monoamine oxidase A (MAOA) that is widely regarded as a major anti-tuberculosis drug (Zareifopoulos and Panayiotakopoulos, 2017). MAOA is a mitochondrial-bound enzyme. It was confirmed that MAOA may promote the progression of prostate cancer by mediating EMT (Wu et al., 2014; Lv et al., 2018). However, because conflicting results have been reported for the importance of MAOA in HCC and cholangiocarcinoma (Huang et al., 2012; Li et al., 2014), the role of MAOA may vary across cancer types. Lee et al. demonstrated that Monoamine Oxidase Inhibitors (MAOIs) are associated with increased colorectal cancer risk (adjusted OR = 1.22, 95% CI = 1.06-1.41; Lee et al., 2017).

PD 0325901 is an MEK inhibitor. Interestingly, Roper et al. (2014) have shown PI3K/MEK inhibition combined with NVP-BKM120 and PD-0325901 treatment can induce tumor progression in a wild-type PIK3CA mouse model, KRAS mutant CRC, based on the inhibition of mTORC1 and MCL-1 and the activation of BIM. Moreover, PD0325901 was reported to inhibit oxaliplatin-induced neuropathy and enhance oxaliplatin efficacy (Tsubaki et al., 2015).

Liggett et al. observed that the non-steroidal anti-inflammatory drug sulindac sulfide inhibits the expression of the potential oncogene structural protein nesprin-2 in CRC cells (Liggett et al., 2014). The results of another study suggested the inhibition of sulindac sulfide on the growth of colon cancer cells and down-regulation of specific transcription factors (Li et al., 2015). Furthermore, the inhibitory effects of 5-fluorouracil and oxaliplatin on human CRC cell survival were demonstrated to be synergistically enhanced by sulindac sulfide (Flis and Splwinski, 2009).

Importazole is a small molecule inhibitor of the transport receptor importin- β (Soderholm et al., 2011) that can inhibit the proliferation and induce apoptosis of multiple myeloma cells by blocking the NF- κ B signaling pathway (Yan et al., 2015). Moreover, intravenous administration of the specific KPNB1 inhibitor importazole was effective in reducing the volume and weight of prostate cancer tumor in mice inoculated with PC3 PCA cells (Yang et al., 2019). Thus, the results of the above studies show that importazole can inhibit tumors.

We conducted a large scale correlation study between colorectal cancer and environmental chemicals and explored the associations between chemicals and colorectal cancer systematically. Our analysis approach has two advantages. Firstly, we identified interaction between chemicals and genes directly. From the perspective of genome, the result is more stable to overcome the shortcomings of traditional exposure measurement methods. From the perspective of benefit, genome-wide summary data usually can be obtained online conveniently. Secondly, our research analyzed summaries of TWAS and mRNA expression profiles, in other words, we made a comprehensive analysis in the DNA and mRNA expression levels. This is helpful to narrow the range of chemicals related candidate genes on the basis of traditional GWAS analysis and make the results more reliable. Current research shows that chemicals in environmental factors have great significance in the etiology of multiple cancers (Thompson et al., 2015). However, we only researched the colon cancer and rectal cancer. As cancer sequencing gene data sets increasing, we will apply our method to large-scale studies of cancer gene-environment interactions.

In summary, we conducted an integrative analysis of GWAS summary data, mRNA expression profiles and chemical-gene interaction networks. Tools such as TWAS and GSEA helped linking these datasets and identifying several chemicals associated with CRC. The results of our study evaluate the associations between CRC and chemicals systematically, and provide new clues for revealing the association between chemicals and genes and their effects on cancer. Furthermore, our method can be used to analyze other chemicals and complex malignant disease, which is helpful for assessing the relationship between environmental exposure and cancer.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: <https://www.ncbi.nlm.nih.gov/gds/>

(GSE106582), <http://geneatlas.roslin.ed.ac.uk/> (categories C18, categories C20), <http://gusevlab.org/projects/fusion/>, <http://ctdbase.org/>.

AUTHOR CONTRIBUTIONS

XT and SH designed experiments. XT, HT, ZL, and YL reviewed and downloaded the original data. XT, LG, and LX processed and analyzed the data. XT, CH, and JM analyzed experimental results. XT, HT, and SH wrote the manuscript.

REFERENCES

- Abolhassani, M., Asadikaram, G., Paydar, P., Fallah, H., Aghaee-Afshar, M., Moazed, V., et al. (2019). Organochlorine and organophosphorous pesticides may induce colorectal cancer: A case-control study. *Ecotoxicol. Environ. Saf.* 178, 168–177. doi: 10.1016/j.ecoenv.2019.04.030
- Anwar, A., Uddin, N., Siddiqui, B. S., Siddiqui, R. A., Begum, S., and Choudhary, M. I. (2018). A natural flavonoid lawsonaringenin induces cell cycle arrest and apoptosis in HT-29 colorectal cancer cells by targeting multiple signalling pathways. *Mol. Biol. Rep.* 45, 1339–1348. doi: 10.1007/s11033-018-4294-5
- Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomaszewski, M., et al. (2013). NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* 41, D991–D995. doi: 10.1093/nar/gks1193
- Bettini, S., Boutet-Robinet, E., Cartier, C., Comera, C., Gaultier, E., Dupuy, J., et al. (2017). Food-grade TiO₂ impairs intestinal and systemic immune homeostasis, initiates preneoplastic lesions and promotes aberrant crypt development in the rat colon. *Sci. Rep.* 7:40373. doi: 10.1038/srep40373
- Bokemeyer, C., Bondarenko, I., Hartmann, J., De Braud, F., Schuch, G., Zube, A., et al. (2011). Efficacy according to biomarker status of cetuximab plus FOLFOX-4 as first-line treatment for metastatic colorectal cancer: the OPUS study. *Ann. Oncol.* 22, 1535–1546. doi: 10.1093/annonc/mdq632
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492
- Canela-Xandri, O., Rawlik, K., and Tenesa, A. (2018). An atlas of genetic associations in UK Biobank. *Nat. Genet.* 50, 1593–1599. doi: 10.1038/s41588-018-0248-z
- Cernigliaro, C., D'Anneo, A., Carlisi, D., Giuliano, M., Marino Gammazza, A., Barone, R., et al. (2019). Ethanol-mediated stress promotes autophagic survival and aggressiveness of colon cancer cells via activation of Nrf2/HO-1 pathway. *Cancers* 11:505. doi: 10.3390/cancers11040505
- Charnpi, K., and Ycart, B. (2015). Weighted kolmogorov smirnov testing: an alternative for gene set enrichment analysis. *Stat. Appl. Genet. Mol. Biol.* 14, 279–293. doi: 10.1515/sagmb-2014-0077
- Davis, A. P., Grondin, C. J., Johnson, R. J., Sciaky, D., King, B. L., McMorran, R., et al. (2017). The comparative toxicogenomics database: update 2017. *Nucleic Acids Res.* 45, D972–D978. doi: 10.1093/nar/gkw838
- Davis, A. P., Murphy, C. G., Johnson, R., Lay, J. M., Lennon-Hopkins, K., Saraceni-Richards, C., et al. (2013). The comparative toxicogenomics database: update 2013. *Nucleic Acids Res.* 41, D1104–D1114. doi: 10.1093/nar/gks994
- Dekker, E., Tanis, P. J., Vleugels, J. L. A., Kasi, P. M., and Wallace, M. B. (2019). Colorectal cancer. *Lancet* 394, 1467–1480. doi: 10.1016/S0140-6736(19)32319-0
- Erkekoglu, P., Oral, D., Chao, M. W., and Kocer-Gumusel, B. (2017). Hepatocellular carcinoma and possible chemical and biological causes: a review. *J. Environ. Pathol. Toxicol. Oncol.* 36, 171–190. doi: 10.1615/JEnvironPatholToxicolOncol.2017020927
- Ferlay, J., Colombet, M., Soerjomataram, I., Mathers, C., Parkin, D. M., Pineros, M., et al. (2019). Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int. J. Cancer* 144, 1941–1953. doi: 10.1002/ijc.31937

FUNDING

This work was supported by grants from the Innovation Capacity Support Plan of Shaanxi Province 2018TD–002.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.00385/full#supplementary-material>

- Flis, S., and Splwinski, J. (2009). Inhibitory effects of 5-fluorouracil and oxaliplatin on human colorectal cancer cell survival are synergistically enhanced by sulindac sulfide. *Anticancer Res.* 29, 435–441.
- Giacchetti, S. P. B., Zidani, R., Le Bail, N., Faggiuolo, R., and Focan, C. (2000). Phase III multicenter randomized trial of oxaliplatin added to chronomodulated fluorouracil–leucovorin as first-line treatment of metastatic colorectal cancer. *J. Clin. Oncol.* 18, 136–147. doi: 10.1200/JCO.2000.18.1.136
- Goldberg, R. M., Sargent, D. J., Morton, R. F., Fuchs, C. S., Ramanathan, R. K., Williamson, S. K., et al. (2004). A randomized controlled trial of fluorouracil plus leucovorin, irinotecan, and oxaliplatin combinations in patients with previously untreated metastatic colorectal cancer. *J. Clin. Oncol.* 22, 23–30. doi: 10.1200/JCO.2004.09.046
- Gong, L., Zhang, D., Lei, Y., Qian, Y., Tan, X., and Han, S. (2018). Transcriptome-wide association study identifies multiple genes and pathways associated with pancreatic cancer. *Cancer Med.* 7, 5727–5732. doi: 10.1002/cam4.1836
- Gregoriadis, G. C., Apostolidis, N. S., Romanos, A. N., and Paradellis, T. P. (1983). A comparative study of trace elements in normal and cancerous colorectal tissues. *Cancer* 52, 508–519. doi: 10.1002/1097-0142(19830801)52:3<508::AID-CNCR2820520322>3.0.CO;2-8
- Grondin, C. J., Davis, A. P., Wieggers, J. A., and Mattingly, C. J. (2018). Accessing an expanded exposure science module at the comparative toxicogenomics database. *Environ. Health Perspect.* 126:014501. doi: 10.1289/EHP2873
- Gusev, A., Mancuso, N., Won, H., Kousi, M., Finucane, H. K., Reshef, Y., et al. (2018). Transcriptome-wide association study of schizophrenia and chromatin activity yields mechanistic disease insights. *Nat. Genet.* 50, 538–548. doi: 10.1038/s41588-018-0092-1
- Gustavsson, B., Carlsson, G., Machover, D., Petrelli, N., Roth, A., Schmoll, H. J., et al. (2015). A review of the evolution of systemic chemotherapy in the management of colorectal cancer. *Clin. Colorectal Cancer* 14, 1–10. doi: 10.1016/j.clcc.2014.11.002
- Hamada, T., Cao, Y., Qian, Z. R., Masugi, Y., Nowak, J. A., Yang, J., et al. (2017). Aspirin use and colorectal cancer survival according to tumor CD274 (Programmed Cell Death 1 Ligand 1) expression status. *J. Clin. Oncol.* 35, 1836–1844. doi: 10.1200/JCO.2016.70.7547
- Hammerschlag, A. R., Stringer, S., de Leeuw, C. A., Sniekers, S., Taskesen, E., Watanabe, K., et al. (2017). Genome-wide association analysis of insomnia complaints identifies risk genes and genetic overlap with psychiatric and metabolic traits. *Nat. Genet.* 49, 1584–1592. doi: 10.1038/ng.3888
- He, A., Wang, W., Prakash, N. T., Tinkov, A. A., Skalny, A. V., Wen, Y., et al. (2018). Integrating genome-wide association study summaries and element-gene interaction datasets identified multiple associations between elements and complex diseases. *Genet. Epidemiol.* 42, 168–173. doi: 10.1002/gepi.22106
- Hong, S. M., Hwang, S. W., Wang, T., Park, C. W., Ryu, Y. M., Jung, J. H., et al. (2019). Increased nicotinamide adenine dinucleotide pool promotes colon cancer progression by suppressing reactive oxygen species level. *Cancer Sci.* 110, 629–638. doi: 10.1111/cas.13886
- Huang, H., Cheng, S., Ding, M., Wen, Y., Ma, M., Zhang, L., et al. (2019). Integrative analysis of transcriptome-wide association study and mRNA expression profiles identifies candidate genes associated with autism spectrum disorders. *Autism Res.* 12, 33–38. doi: 10.1002/aur.2048

- Huang, L., Frampton, G., Rao, A., Zhang, K. S., Chen, W., Lai, J. M., et al. (2012). Monoamine oxidase A expression is suppressed in human cholangiocarcinoma via coordinated epigenetic and IL-6-driven events. *Lab. Invest.* 92, 1451–1460. doi: 10.1038/labinvest.2012.110
- Jones, R. R., DellaValle, C. T., Weyer, P. J., Robien, K., Cantor, K. P., Krasner, S., et al. (2019). Ingested nitrate, disinfection by-products, and risk of colon and rectal cancers in the Iowa Women's Health Study cohort. *Environ. Int.* 126, 242–251. doi: 10.1016/j.envint.2019.02.010
- Khoury, M. J., Davis, R., Gwinn, M., Lindegren, M. L., and Yoon, P. (2005). Do we need genomic research for the prevention of common diseases with environmental causes? *Am. J. Epidemiol.* 161, 799–805. doi: 10.1093/aje/kwi113
- Kouhara, J., Yoshida, T., Nakata, S., Horinaka, M., Wakada, M., Ueda, Y., et al. (2007). Fenretinide up-regulates DR5/TRAIL-R2 expression via the induction of the transcription factor CHOP and combined treatment with fenretinide and TRAIL induces synergistic apoptosis in colon cancer cell lines. *Int. J. Oncol.* 30, 679–687. doi: 10.3892/ijo.30.3.679
- Kuipers, E. J., Grady, W. M., Lieberman, D., Seufferlein, T., Sung, J. J., Boelens, P. G., et al. (2015). Colorectal cancer. *Nat. Rev. Dis. Primers* 1:15065. doi: 10.1038/nrdp.2015.65
- Lee, H. C., Chiu, W. C., Wang, T. N., Liao, Y. T., Chien, I. C., Lee, Y., et al. (2017). Antidepressants and colorectal cancer: A population-based nested case-control study. *J. Affect. Disord.* 207, 353–358. doi: 10.1016/j.jad.2016.09.057
- Li, J., Yang, X. M., Wang, Y. H., Feng, M. X., Liu, X. J., Zhang, Y. L., et al. (2014). Monoamine oxidase A suppresses hepatocellular carcinoma metastasis by inhibiting the adrenergic system and its transactivation of EGFR signaling. *J. Hepatol.* 60, 1225–1234. doi: 10.1016/j.jhep.2014.02.025
- Li, X., Pathi, S. S., and Safe, S. (2015). Sulindac sulfide inhibits colon cancer cell growth and downregulates specificity protein transcription factors. *BMC Cancer* 15:974. doi: 10.1186/s12885-015-1956-8
- Liggett, J. L., Choi, C. K., Donnell, R. L., Kihm, K. D., Kim, J. S., Min, K. W., et al. (2014). Nonsteroidal anti-inflammatory drug sulindac sulfide suppresses structural protein Nesprin-2 expression in colorectal cancer cells. *Biochim. Biophys. Acta* 1840, 322–331. doi: 10.1016/j.bbagen.2013.09.032
- Lin, B. M., Nadkarni, G. N., Tao, R., Graff, M., Fornage, M., Buyske, S., et al. (2019). Genetics of chronic kidney disease stages across ancestries: the PAGE study. *Front. Genet.* 10:494. doi: 10.3389/fgene.2019.00494
- Lucas, A. S., O'Neil, B. H., and Goldberg, R. M. (2011). A decade of advances in cytotoxic chemotherapy for metastatic colorectal cancer. *Clin. Colorectal Cancer* 10, 238–244. doi: 10.1016/j.clcc.2011.06.012
- Lv, Q., Wang, D., Yang, Z., Yang, J., Zhang, R., Yang, X., et al. (2018). Repurposing antitubercular agent isoniazid for treatment of prostate cancer. *Biomater. Sci.* 7, 296–306. doi: 10.1039/C8BM01189C
- Manuck, S. B., and McCaffery, J. M. (2014). Gene-environment interaction. *Annu. Rev. Psychol.* 65, 41–70. doi: 10.1146/annurev-psych-010213-115100
- Mattingly, C. J., Colby, G. T., Forrest, J. N., and Boyer, J. L. (2003). The comparative toxicogenomics database (CTD). *Environ. Health Perspect.* 111, 793–795. doi: 10.1289/ehp.6028
- Mattingly, C. J., Rosenstein, M. C., Colby, G. T., Forrest, J. N. Jr., and Boyer, J. L. (2006). The Comparative Toxicogenomics Database (CTD): a resource for comparative toxicological studies. *J. Exp. Zool. A Comp. Exp. Biol.* 305, 689–692. doi: 10.1002/jez.a.307
- Messerlian, C., Martinez, R. M., Hauser, R., and Baccarelli, A. A. (2017). 'Omics' and endocrine-disrupting chemicals - new paths forward. *Nat. Rev. Endocrinol.* 13, 740–748. doi: 10.1038/nrendo.2017.81
- Nuotio, J., Oikonen, M., Magnussen, C. G., Jokinen, E., Laitinen, T., Hutri-Kahonen, N., et al. (2014). Cardiovascular risk factors in 2011 and secular trends since 2007: the cardiovascular risk in young finns study. *Scand. J. Public Health* 42, 563–571. doi: 10.1177/1403494814541597
- Pham, D. N., Leclerc, D., Levesque, N., Deng, L., and Rozen, R. (2013). beta, beta-carotene 15,15'-monooxygenase and its substrate beta-carotene modulate migration and invasion in colorectal carcinoma cells. *Am. J. Clin. Nutr.* 98, 413–422. doi: 10.3945/ajcn.113.060996
- Qi, X., Wang, S., Zhang, L., Liu, L., Wen, Y., Ma, M., et al. (2019). An integrative analysis of transcriptome-wide association study and mRNA expression profile identified candidate genes for attention-deficit/hyperactivity disorder. *Psychiatry Res.* 282:112639. doi: 10.1016/j.psychres.2019.112639
- Qu, X., Yang, L., Shi, Q., Wang, X., Wang, D., and Wu, G. (2018). Lidocaine inhibits proliferation and induces apoptosis in colorectal cancer cells by upregulating mir-520a-3p and targeting EGFR. *Pathol. Res. Pract.* 214, 1974–1979. doi: 10.1016/j.prp.2018.09.012
- Raitakari, O. T., Juonala, M., Ronnema, T., Keltikangas-Jarvinen, L., Rasanen, L., Pietikainen, M., et al. (2008). Cohort profile: the cardiovascular risk in Young Finns Study. *Int. J. Epidemiol.* 37, 1220–1226. doi: 10.1093/ije/dym225
- Roper, J., Sinnamon, M. J., Coffee, E. M., Belmont, P., Keung, L., Georgeon-Richard, L., et al. (2014). Combination PI3K/MEK inhibition promotes tumor apoptosis and regression in PIK3CA wild-type, KRAS mutant colorectal cancer. *Cancer Lett.* 347, 204–211. doi: 10.1016/j.canlet.2014.02.018
- U.S. Department of Health and Human Services (2016). *14th Report on Carcinogens*. Washington, DC: U.S. Department of health and human services.
- Siegel, R. L., Miller, K. D., and Jemal, A. (2017). Cancer statistics, 2017. *CA Cancer J. Clin.* 67, 7–30. doi: 10.3322/caac.21387
- Soderholm, J. F., Bird, S. L., Kalab, P., Sampathkumar, Y., Hasegawa, K., Uehara-Bingen, M., et al. (2011). Importazole, a small molecule inhibitor of the transport receptor importin-beta. *ACS Chem. Biol.* 6, 700–708. doi: 10.1021/cb2000296
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–15550. doi: 10.1073/pnas.0506580102
- Tam, V., Patel, N., Turcotte, M., Bosse, Y., Pare, G., and Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* 20, 467–484. doi: 10.1038/s41576-019-0127-1
- Thompson, P. A., Khatami, M., Baglole, C. J., Sun, J., Harris, S. A., Moon, E. Y., et al. (2015). Environmental immune disruptors, inflammation and cancer risk. *Carcinogenesis* 36(Suppl. 1), S232–S253. doi: 10.1093/carcin/bgv038
- Tsubaki, M., Takeda, T., Tani, T., Shimaoka, H., Suzuyama, N., Sakamoto, K., et al. (2015). PKC/MEK inhibitors suppress oxaliplatin-induced neuropathy and potentiate the antitumor effects. *Int. J. Cancer* 137, 243–250. doi: 10.1002/ijc.29367
- Tsukamoto, H., Mizoshita, T., Katano, T., Hayashi, N., Ozeki, K., Ebi, M., et al. (2015). Preventive effect of rebamipide on N-methyl-N'-nitro-N-nitrosoguanidine-induced gastric carcinogenesis in rats. *Exp. Toxicol. Pathol.* 67, 271–277. doi: 10.1016/j.etp.2015.01.003
- Wang, K., Li, M., and Bucan, M. (2007). Pathway-based approaches for analysis of genomewide association studies. *Am. J. Hum. Genet.* 81, 1278–1283. doi: 10.1086/522374
- Weng, L., Macciardi, F., Subramanian, A., Guffanti, G., Potkin, S. G., Yu, Z., et al. (2011). SNP-based pathway enrichment analysis for genome-wide association studies. *BMC Bioinformatics* 12:99. doi: 10.1186/1471-2105-12-99
- Wu, J. B., Shao, C., Li, X., Li, Q., Hu, P., Shi, C., et al. (2014). Monoamine oxidase A mediates prostate tumorigenesis and cancer metastasis. *J. Clin. Invest.* 124, 2891–2908. doi: 10.1172/JCI70982
- Wu, L., Shi, W., Long, J., Guo, X., Michailidou, K., Beesley, J., et al. (2018). A transcriptome-wide association study of 229,000 women identifies new candidate susceptibility genes for breast cancer. *Nat. Genet.* 50, 968–978. doi: 10.1038/s41588-018-0132-x
- Wu, S., Powers, S., Zhu, W., and Hannun, Y. A. (2016). Substantial contribution of extrinsic risk factors to cancer development. *Nature* 529, 43–47. doi: 10.1038/nature16166
- Yan, W., Li, R., He, J., Du, J., and Hou, J. (2015). Importin beta1 mediates nuclear factor-kappaB signal transduction into the nuclei of myeloma cells and affects their proliferation and apoptosis. *Cell. Signal.* 27, 851–859. doi: 10.1016/j.cellsig.2015.01.013
- Yang, J., Guo, Y., Lu, C., Zhang, R., Wang, Y., Luo, L., et al. (2019). Inhibition of Karyopherin beta 1 suppresses prostate cancer growth. *Oncogene* 38, 4700–4714. doi: 10.1038/s41388-019-0745-2

- Zareifopoulos, N., and Panayiotakopoulos, G. (2017). Neuropsychiatric effects of antimicrobial agents. *Clin. Drug Investig.* 37, 423–437. doi: 10.1007/s40261-017-0498-z
- Zhang, Y. F., Huang, Y., Ni, Y. H., and Xu, Z. M. (2019). Systematic elucidation of the mechanism of geraniol via network pharmacology. *Drug Des. Devel. Ther.* 13, 1069–1075. doi: 10.2147/DDDT.S189088
- Zhao, Y., He, A., Zhu, F., Ding, M., Hao, J., Fan, Q., et al. (2018). Integrating genome-wide association study and expression quantitative trait locus study identifies multiple genes and gene sets associated with schizophrenia. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 81, 50–54. doi: 10.1016/j.pnpbp.2017.10.003

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Tan, Tang, Gong, Xie, Lei, Luo, He, Ma and Han. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.