

RESEARCH

Open Access



mGWAS identification of six novel single nucleotide polymorphism loci with strong correlation to gastric cancer

Shuangfeng Yang^{1,2}, Yuan-Liang Wang^{1,2}, Yanping Lyu^{1,2}, Yu Jiang^{1,2}, Jianjun Xiang^{1,2}, Shumi Ji^{1,2}, Shuling Kang^{1,2}, Xuejie Lyu^{1,2}, Chenzhou He^{1,2}, Peixin Li^{1,2}, Baoying Liu^{1,2*} and Chuancheng Wu^{1,2*} 

Abstract

Background: Metabolite genome-wide association studies (mGWAS) are key for understanding the genetic regulation of metabolites in complex diseases including cancers. Although mGWAS has revealed hundreds of metabolomics quantitative trait loci (mQTLs) in the general population, data relating to gastric cancer (GC) are still incomplete.

Methods: We identified mQTLs associated with GC by analyzing genome-wide and metabolome-wide datasets generated from 233 GC patients and 233 healthy controls.

Results: Twenty-two metabolites were statistically different between GC cases and healthy controls, and all of them were associated with the risk of gastric cancer. mGWAS analyses further revealed that 9 single nucleotide polymorphisms (SNPs) were significantly associated with 3 metabolites. Of these 9 SNPs, 6 loci were never reported in the previous mGWAS studies. Surprisingly, 4 of 9 SNPs were significantly enriched in genes involved in the T cell receptor signaling pathway.

Conclusions: Our study unveiled several novel GC metabolite and genetic biomarkers, which may be implicated in the prevention and diagnosis of gastric cancer.

Keywords: Metabolite, GWAS, Biomarkers, Gastric cancer

Introduction

Gastric cancer (GC) is one of the most common, multi-factorial malignancies mediated by environmental and genetic factors [1, 2]. Several genome-wide association studies (GWAS) have been performed to identify genetic biomarkers for GC [3–7]. Although multiple single nucleotide polymorphisms (SNPs) have been found to associate with GC, the effect size is small and the influence of genetic factors on the biological processes underlying GC remains elusive.

GWAS with intermediate phenotypes, like changes in metabolite and protein levels, are key in establishing functional links between genetic variants and disease end points [8]. Metabolomics is a rapidly developing discipline that focuses on the study and analysis of small endogenous molecules (MW < 1500 Da) [9]. As the end-point of the signaling cascade, it represents the final response of living systems to environmental, genetic, and disease factors [9, 10]. Advances in metabolomic tools provide a unique opportunity to reveal complex relationships between genotypes and phenotypes [10]. Therefore, metabolomics may be an ideal intermediate phenotype choice. Prior studies have shown that GC is associated with alterations in circulating metabolites,

* Correspondence: lby@fjmu.edu.cn; wcc@fjmu.edu.cn

¹School of Public Health, Fujian Medical University, Fuzhou, China
Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

such as nucleotides, lipids, and amino acids [11, 12]. Nevertheless, the relationship between metabolite levels and biological mechanisms of GC remains unclear.

GWAS for metabolic traits (metabolite genome-wide association studies, mGWAS) has revealed hundreds of metabolomic quantitative trait loci (mQTLs) in the general population [10, 13–15]. For GC, the identification of novel mQTLs may provide a valuable tool for discovering genetic biomarkers related to GC and provide new insights from the perspective of metabolism. Mapping GC-associated mQTLs could also shed light on the etiology and mechanisms of GC, which in turn is important for the prevention, early detection, diagnosis, and targeted treatment of this malignancy. However, to our knowledge, there are currently no studies of MGWAS associated with GC.

To identify the GC-associated mQTLs for the first time, we combined genome-wide and metabolome-wide datasets generated from 233 GC patients and 233 healthy controls. Patients were age-, sex-, smoking status-, alcohol consumption-, *H. pylori* infection-, and time of blood sample collection-matched to healthy controls. We hypothesized that by combining phenotypic, metabolomic, and genetic data, we could better identify individuals at risk of GC and uncover biological mechanisms related to GC.

Materials and methods

Study samples

This study was based on a population-based case-control study, in which GC patients and healthy individuals were consecutively enrolled in Xianyou County between March 2013 and December 2017. All cases were newly diagnosed histologically based on tissue specimens and had lived in Xianyou for at least 10 years. GC patients with other cancers, secondary or recurrent gastric cancer, gastritis, previously received neoadjuvant chemotherapy or chemoradiotherapy or radiotherapy, pregnant, metabolic diseases such as diabetes, gout, hyperlipidemia, systemic administration of corticosteroids, neurological and psychiatric diseases, severe hepatic and renal dysfunction, and severe respiratory disease requiring continuous oxygen treatment, etc., were excluded. For the healthy individuals, there was no abnormality of blood tests and endoscopic examinations.

The study initially included 244 GC cases and 244 healthy individuals. Ethical approval for the study was obtained from the Biomedical Research Ethics Committee of Fujian Medical University, China (No. 97,2014). Following an explanation of the study, written informed consent was obtained from all participants at study enrollment.

Metabolite profiling

Blood samples were collected after an overnight fast, immediately centrifuged, and stored at -80°C until assayed. Plasma metabolites were profiled as previously described using Agilent 1200 high-performance liquid chromatography combined with a 6520 accurate electrospray ionization/quadrupole-time-of-flight mass system (Agilent Technologies, CA, USA) [16]. A total of 225 compounds were detected. If a variable had a nonzero measurement value in at least 80% of the variables within one of the two subsets, the variable was included in the data set; otherwise, the variable was removed. This procedure will be referred to as the “80% rule” [17]. After being excluded based on extracted ion chromatogram (EIC) and “80% rule,” 60 metabolites have remained for subsequent analyses. The final metabolomics dataset contained 20 nucleotides, 19 lipids, 7 amino acids, 8 organoheterocyclic compounds/others, 3 peptides, and 3 unknown.

Genome-wide genotyping

Approximately 900,000 SNPs were genotyped using the Axiom™ Precision Medicine Research Array (Thermo Fisher Scientific, Waltham, MA, USA). Genotyping was performed as described according to the manufacturer's instructions. Briefly, genomic DNA was extracted using a Genomic DNA Isolation Kit (Biovision, CA, USA). Each sample was whole-genome amplified, fragmented, precipitated, dried, resuspended in appropriate hybridization buffer, chip cleaned, stained, and scanned. Participants were excluded if they satisfy any of the following items: (1) low call rate (overall rate $< 95\%$), (2) ambiguous gender, and (3) duplicates or familial relationship ($\text{PI_HAT} > 0.025$). SNPs were excluded if they (1) were not mapped to autosome chromosomes, (2) had a call rate $< 95\%$, (3) had minor allele frequency (MAF) < 0.05 in controls, and (4) were an excessive deviation from Hardy-Weinberg equilibrium in controls ($p < 1 \times 10^{-6}$). As a result, 233 pairs of subjects and 258,544 SNPs were left for subsequent analysis.

Statistical analysis

Before statistical analysis, each metabolic peak in all subject samples was normalized based on QC samples for removing the unwanted analytical variations occurring intra- and inter-batches. And the plasma abundant values of metabolites investigated were set to a log scale and auto-scaled (mean-centered and divided by the standard deviation of each variable) using MetaboAnalyst 4.0. An orthogonal partial least squares discriminant analysis (OPLS-DA) and univariate two-sides *t*-test were used for metabolic profile description and metabolic signature discovery between GC cases and healthy controls. The false discovery rate (FDR) method was used to

correct for multiple hypothesis testing and to reduce false positives. Those metabolic features with variable importance in the projection (VIP) values > 1.0 in the OPLS-DA model and FDR-adjusted p values < 0.05 in the t -test were considered to be significantly different between GC cases and healthy controls. Then, logistic regression was performed to test the association between discriminant metabolites and incident GC.

After that, we analyzed the associations between discriminant metabolite levels and genome chip variants using a generalized linear model adjusted for age, sex, family relationship, smoking status, alcohol consumption, pickled vegetable intake, and *H. pylori* infection analysis in TASSEL software (version 5.0). To control for false-positive error rates deriving from the large number of SNPs tested, a conservative Bonferroni-adjusted p -value of $p = 8.79 \times 10^{-9}$ ($= 0.05/(258544\text{SNPs} \times 22 \text{ metabolites})$) was applied for declaring genome-wide significance for the SNP-metabolite associations. The significant SNPs were annotated to the neighboring genes of 1000 Genomes Project (hg19/1000 Genomes ASN) downloaded from the University of California Santa Cruz (UCSC) genome browser (<http://hgdownload.cse.ucsc.edu>). Then, we conducted Gene Ontology (GO) enrichment analysis to investigate possible biological, molecular, or cellular processes associated with significant SNP-related genes using Metascape (<https://metascape.org/>) [18]. Furthermore, the interaction network between metabolites, genes, and GO terms were visualized using Cytoscape [19]. Finally,

RNA-seq data of gastric cancer tissue and adjacent normal tissue of six gastric cancer patients in Asia were downloaded from The Cancer Genome Atlas (TCGA; <https://www.cancer.gov/tcga>), and t test was performed to compare the candidate gene expression between two different groups shown in boxplot using R (R version 3.6.0).

Results

Study sample characteristics

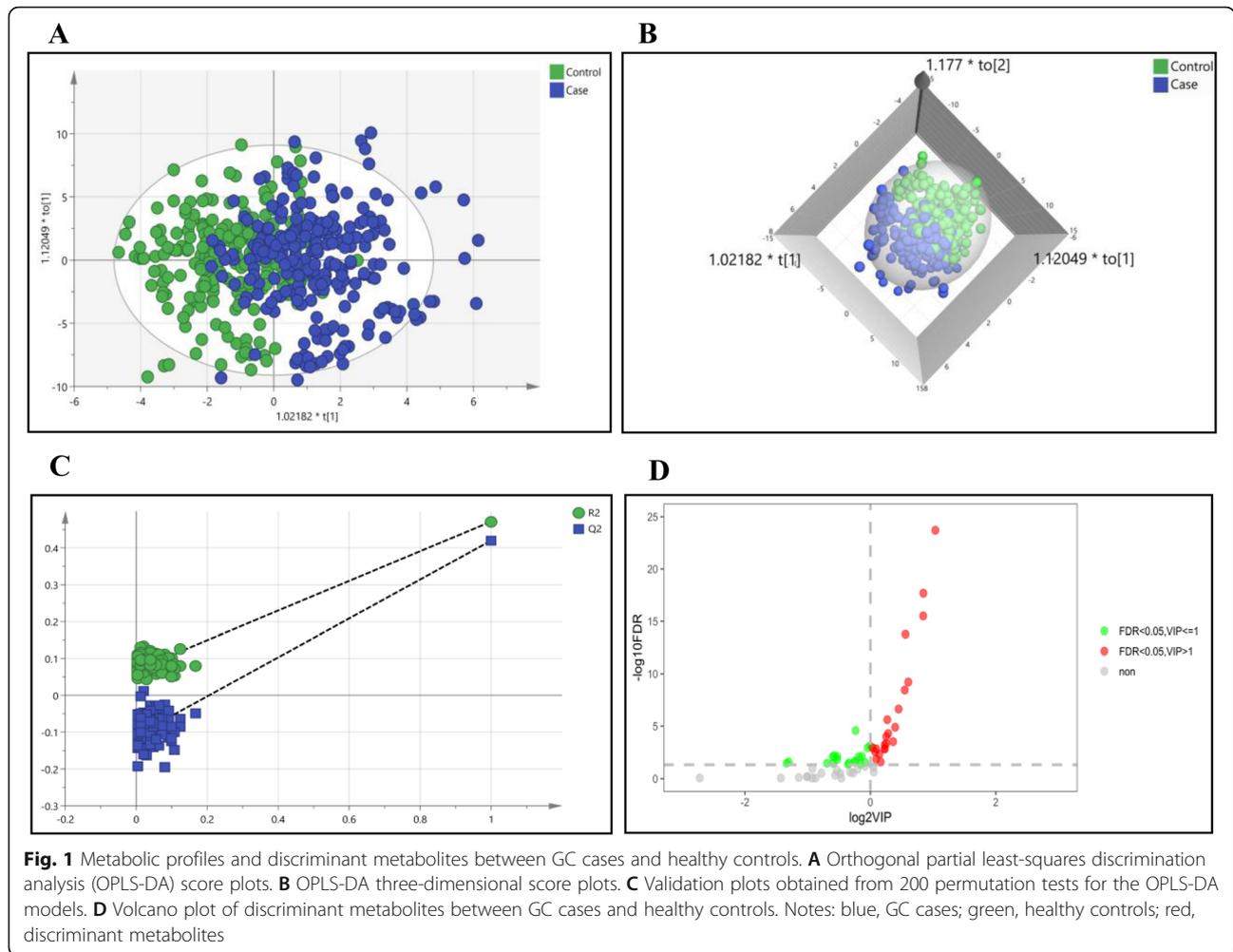
A total of 233 GC cases and 233 healthy controls were included, with the mean age of 64.80 ± 7.89 years and 65.16 ± 7.91 years, respectively. As shown in Table 1, except for pickled vegetable intake ($p < 0.05$), there was no statistically significant difference in gender, age, smoking status, alcohol consumption, and *H. pylori* infection, between GC cases and healthy controls ($p > 0.05$).

Metabolic profiles and discriminant metabolites between GC cases and healthy controls

To identify potential compounds that could be used to differentiate the metabolite profiles of GC cases and healthy controls, we established a supervised OPLS-DA model that focused on the actual class discriminating variation. As shown in Fig. 1A and B, we observed a clear separation between the two groups. Response permutation testing was performed to verify the reliability of the OPLS-DA model. As shown in Fig. 1C, the goodness of fit (R^2Y) and prediction ability of the model (Q^2) were 0.472 and 0.420 for differentiating GC cases and

Table 1 Demographic characteristics of study participants

	Control (n = 233)	Case (n = 233)	χ^2	p-value
Gender			0.000	1.00
Female	63	63		
Male	170	170		
Age			0.424	0.515
≤65y	123	130		
>65y	110	103		
Smoking status			0.009	0.926
No	109	108		
Yes	124	125		
Alcohol consumption			1.378	0.240
No	194	203		
Yes	39	30		
Pickled vegetable intake			17.003	<0.001
No	66	30		
Yes	167	203		
<i>H. pylori</i> infection			0.567	0.451
No	141	133		
Yes	92	100		



healthy controls, respectively, suggesting successful model construction.

Subsequently, to find discriminant metabolites between GC cases and healthy controls, a univariate two-sides *t*-test and the VIP scores from OPLS-DA models were carried out. Metabolites with FDR-adjusted *p*-values < 0.05 and VIP > 1 were selected as discriminants. Figure 1D shows that 22 metabolites were statistically different between GC cases and healthy controls, including 7 nucleotides, 9 lipids, and 6 others.

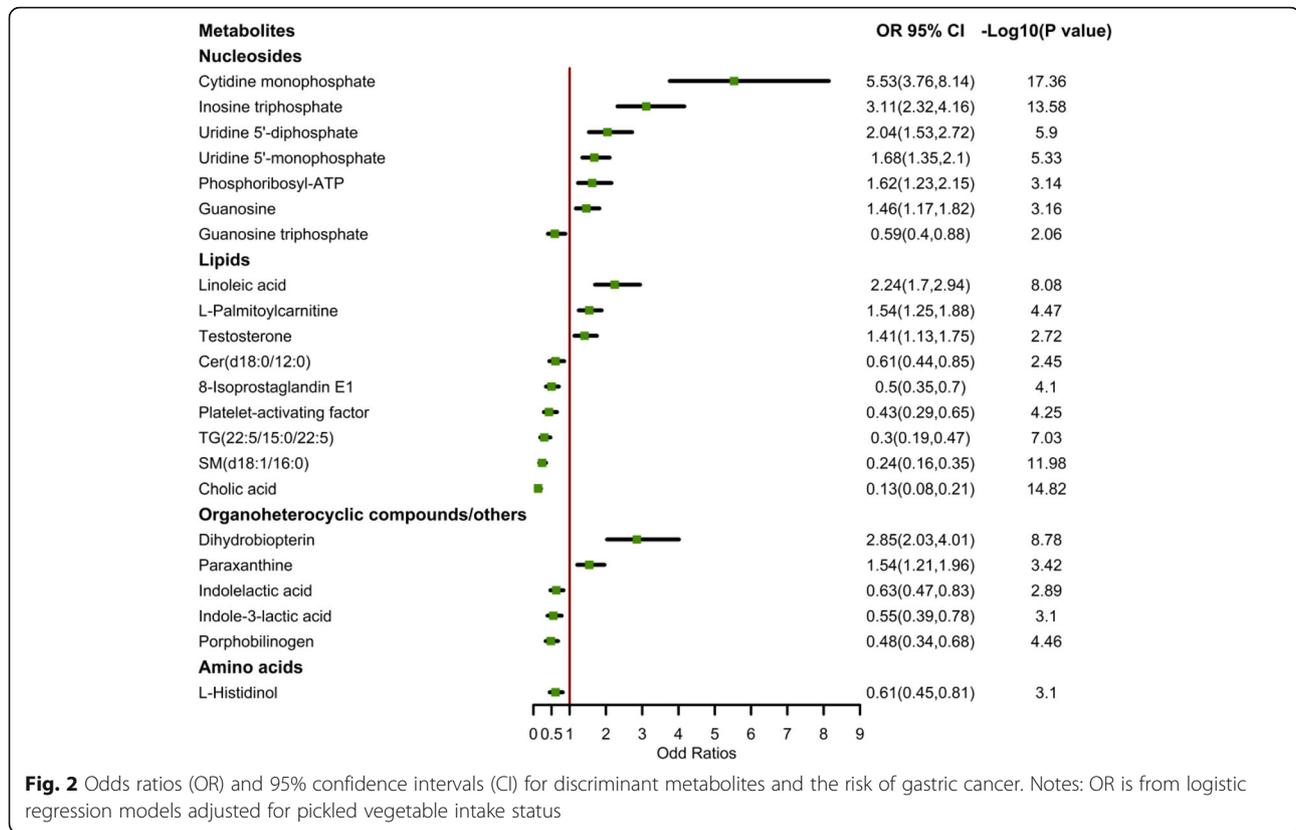
Association between discriminant metabolites and gastric cancer

Figure 2 denotes odds ratios for the association between 22 discriminant metabolites and GC risk, which was adjusted for pickled vegetable intake status. Eleven metabolites, including cytidine monophosphate (CMP), inosine triphosphate (ITP), uridine 5'-monophosphate (5'-UMP), uridine 5'-diphosphate (5'-UDP), guanosine, phosphoribosyl-ATP, linoleic acid, L-palmitoylcarnitine, testosterone, dihydrobiopterin, and paraxanthine, were

associated with increased risks of GC. Conversely, metabolites including guanosine triphosphate (GTP), Cer(d18:0/12:0), 8-Isoprostaglandin E1, platelet-activating factor, TG(22:5/15:0/22:5), SM(d18:1/16:0), cholic acid, indolelactic acid, indole-3-lactic acid, porphobilinogen, and L-histidinol were associated with decreased risk of GC. Receiver operating characteristic (ROC) curve analysis was used to assess the discriminative ability of the 22 metabolites. The area under the curve was 0.897 (95% CI 0.869 to 0.924), indicating a high diagnostic value for GC. In short, these metabolites are highly correlated with the occurrence of GC.

Genetic associations of discriminant metabolites

Next, we analyzed the association of these 22 metabolites with genome chip variants. As listed in Table 2, 9 SNPs were significantly associated with 3 metabolite concentrations at a stringent genome-wide cutoff of 8.79×10^{-9} . Of these, 4 loci are in the enhancer regions, 2 are in the 3'UTR. Besides, 6 loci have not been reported before, while 3 loci were reported in the previous



mGWAS study in the general population [20], including rs12204145, rs12208390, and rs3927423. The last two loci were in strong linkage disequilibrium (LD) ($r^2 = 1$) with the variant rs12202419. And markers R^2 for 3 metabolites ranged from 7.23 to 8.46% and the total phenotypic variance explained by 9 SNPs was 70.46%.

Gene and metabolite interaction network

To investigate which biological pathways were associated with these SNPs and the corresponding metabolites, we carried out a GO enrichment analysis. The most significantly enriched gene-based pathways were the T cell

receptor signaling pathway (GO:0002376) with five genes (i.e., *PDE4D*, *BTN3A3*, *BTN2A2*, *BTN3A1*, *BTN2A3P*) (Fig. 3). Figure 3 also denotes the relationships between metabolites, genes, and GO terms.

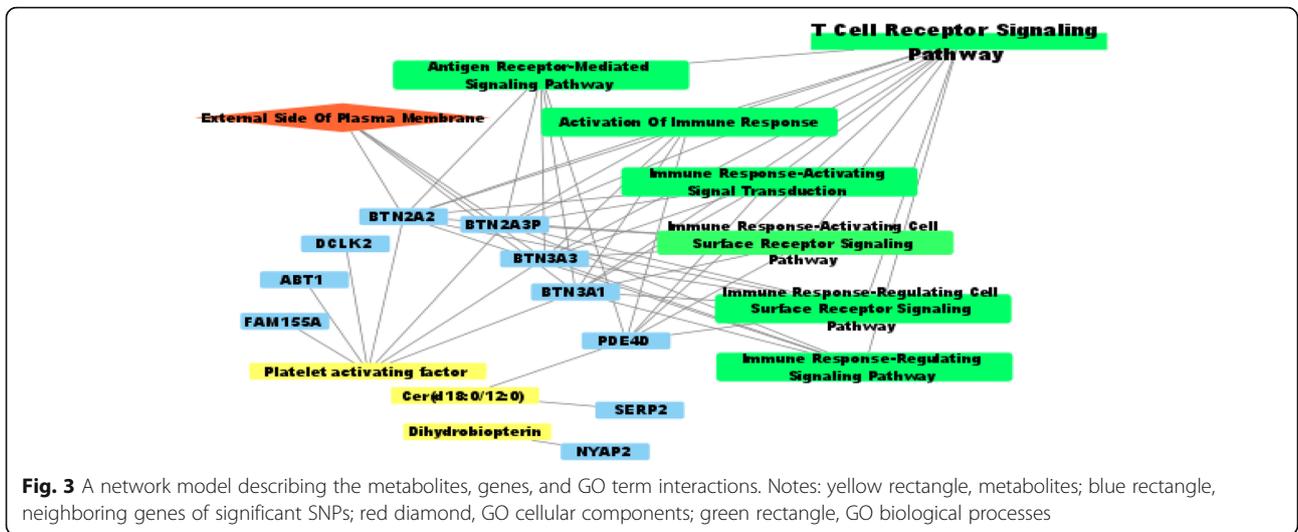
Expression of 10 candidate genes between gastric cancer tissue and adjacent normal tissue

Finally, GWAS was performed by plink to investigate whether the genotype distribution of the 9 SNPs is different between GC cases and healthy controls. Unfortunately, no statistically significant association was found. Then, we downloaded RNA-seq data of gastric cancer

Table 2 SNP associations of discriminant metabolites with $p < 8.79 \times 10^{-9}$

Metabolites	SNP	Chr	Pos	Alt	Ref	p value	R ²	Candidate gene	Function
Dihydrobiopterin	rs140991639	2	226294958	A	G	6.62E-09	7.23%	NYAP2	Intron, Enh60914
Platelet-activating factor	rs12208390	6	26437106	T	C	1.69E-09	8.10%	BTN3A3,BTN2A3P	Upstream
	rs3927423	6	26393487	T	C	2.73E-09	7.93%	BTN2A2	UTR-3
	rs6912853	6	26401438	T	C	3.80E-09	7.82%	BTN3A1	Upstream
	rs140044870	4	150707364	A	C	5.88E-09	7.62%	DCLK2	Upstream, Enh90264
	rs12204145	6	26600156	A	G	7.73E-09	7.51%	ABT1	UTR-3
	rs957788	13	108257220	C	T	8.66E-09	7.51%	FAM155A	Intron
Cer(d18:0/12:0)	rs1183416	13	44873710	T	C	1.10E-09	8.46%	SERP2	Upstream, Enh30302
	rs295977	5	58893286	A	C	1.70E-09	8.28%	PDE4D	Intron, Enh55272

SNP, single nucleotide polymorphism; Chr, chromosome; Pos, position; Alt, alternative allele; Ref, reference allele

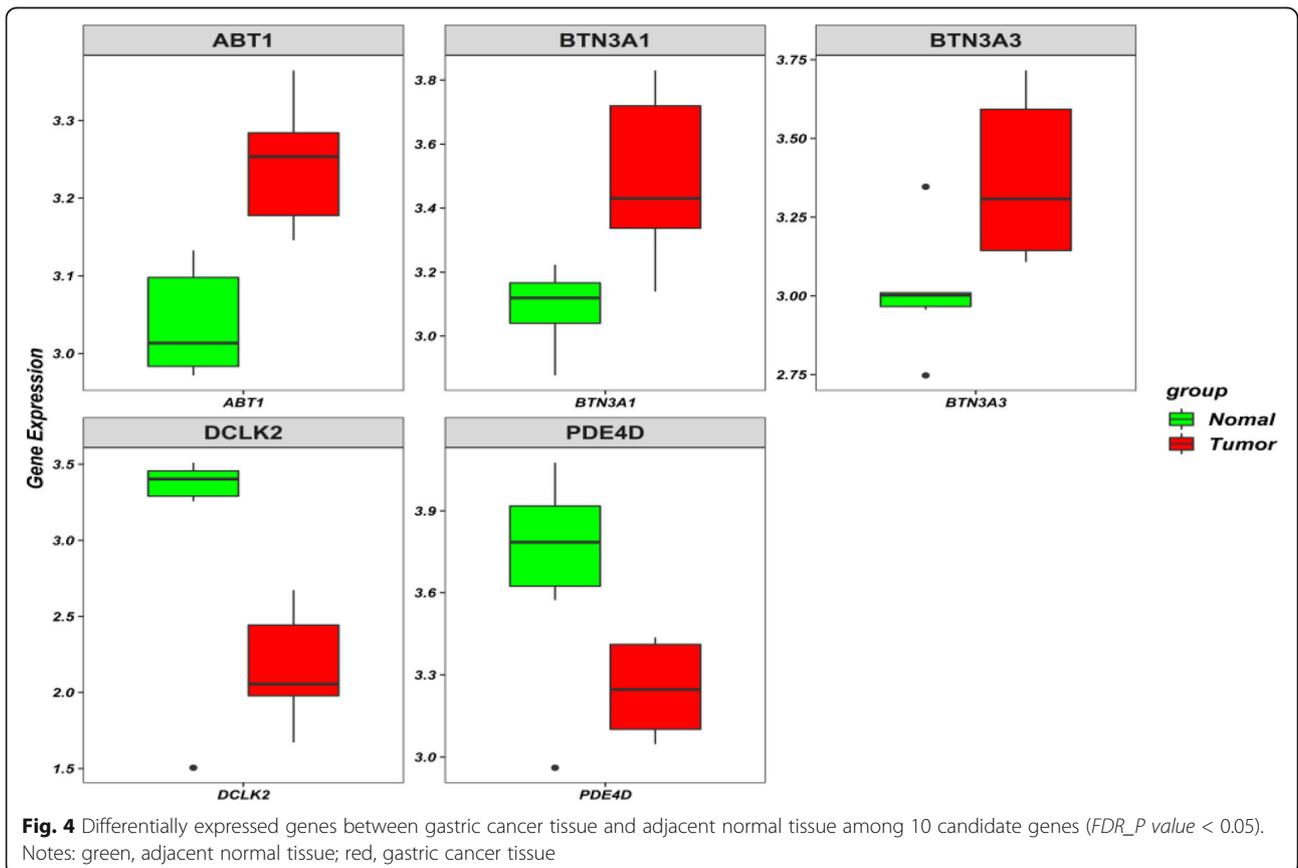


tissue and adjacent normal tissue from TCGA and compare the expression of 10 candidate genes between GC tissue and adjacent normal tissue. As shown in Fig. 4, there were statistically significant differences in the expression of 5 genes between the two groups (FDR_P value < 0.05), which indirectly verifies the reliability of this study.

Discussion

Discriminant metabolites between GC cases and healthy controls

In this study, 22 metabolites were found to be associated with the occurrence of GC, including 7 nucleotides, 9 lipids, 5 organoheterocyclic compounds, and 1 amino acid.



Nucleotide synthesis is key for cell proliferation as it is required for DNA replication, gene transcription, and ribosome biogenesis [21]. Tumor cells are in a state of such rapid proliferation and differentiation that frequent nucleotide synthesis and metabolism are upregulated significantly. Our study found that the relative abundance of CMP, 5'-UMP, 5'-UDP, ITP, guanosine, and phosphoribosyl-ATP were increased in the GC cases compared with healthy controls, while GTP was decreased. Previous research showed that guanosine was also increased in GC samples, but no noticeable difference of ATP and GTP was observed between normal and GC tissues [22]. This discrepancy may be due to the difference in experimental materials. Nucleotides in the blood are either passively released from stressed or dying cells due to cell lysis or actively released through membrane channels by exocytosis of vesicles or as part of exosomes [23]. The increase of nucleotides in plasma in our study may indicate an increase of dead cells in the blood of GC patients. On the one hand, it may come from the death of normal cells due to undernutrition [24, 25]. Alternatively, it may come from the increase of programmed cell death process triggered by cellular stress, DNA damage, and immune surveillance, to resist tumor proliferation [26].

Lipids, including phospholipids, fatty acids, triglycerides, sphingolipids, cholesterol, and cholesteryl esters, are key constituents of all biological membrane structures [27, 28]. Moreover, lipids could function as second messengers to transduce signals within cells, and serve as important energy sources when nutrients are limited [29]. As a result, proliferating cells need to acquire sufficient lipids to support membrane growth and integrity. Accumulating evidence has demonstrated that lipid metabolism is substantially reprogrammed in cancers including GC [11, 30]. In this study, linoleic acid, L-palmitoylcarnitine, testosterone, Cer(d18:0/12:0), 8-Isoprostaglandin E1, platelet-activating factor, TG (22:5/15:0/22:5), SM (d18:1/16:0), and cholic acid were found to be associated with GC. Yu et al. demonstrated that there was no difference in plasma linoleic acid concentration between gastric cancer patients and healthy individuals [31], but some studies also found that the increase of linoleic acid in plasma or diet may increase the risk of hepatocellular carcinoma, colorectal cancer, and prostate cancer [32, 33]. Linoleic acid can be metabolized to arachidonic acid, which plays an important role in inflammatory processes, as the latter can serve as a substrate for the production of some pro-inflammatory eicosanoids, leading to the production of inflammatory mediators such as tumor necrosis factor- α (TNF- α) and interleukin-1 (IL-1) [34]. Higher L-palmitoylcarnitine levels were found in mice with colon cancer and patients with acute pneumonia, indicating an increase in

inflammation and fatty acid beta-oxidation [35, 36]. Findings of our study suggest that linoleic acid and L-palmitoylcarnitine may also play a pro-inflammatory role in gastric cancer.

Cer(d18:0/12:0) and SM(d18:1/16:0) are members of sphingolipids (SPs), which are structural molecules of cell membranes with important roles in maintaining barrier function and fluidity [35, 37]. Sphingolipids also regulate various biological processes such as growth, proliferation, migration, invasion, and/or metastasis by controlling signaling functions within the cancer cell signal transduction network [38]. Changes in cellular ceramide levels are followed by the activation of downstream effectors, which results in cell cycle arrest, senescence, or programmed cell death [39]. In this study, the decrease of Cer(d18:0/12:0) in plasma may indicate the disorder of the cell cycle and programmed cell death process, resulting in the exuberant proliferation of cancer cells and the continuous production of the cell membrane.

Platelet-activating factor, also known as PAF or PC(O-16:0/2:0), is a ubiquitous, potent phospholipid activator and mediator of inflammation that plays an important role in the pathogenesis of inflammatory disorders [40]. In vitro and animal studies suggest that PAF can act on the growth of various human tumor cell lines, increasing the adhesiveness of tumor cells to vascular endothelia, enhancing oncogene expression, and contributing to tumor development through enhancement of cell motility and stimulation of angiogenic response [41–44]. PAF receptor expression is increased in patients with gastric adenocarcinoma, and it is closely related to tumor proliferation ability and tumor size [45], which indicates that gastric adenocarcinoma tissue needs to consume more PAF. Transcripts of PAF and PAF receptors were also significantly increased in hepatocellular carcinoma specimens compared with non-cancer specimens [46]. In conclusion, PAF can change local angiogenesis and cytokine networks and is essential to suppress the immune system and promote metastasis and tumor growth [47]. In this study, the decrease of the platelet-activating factor in plasma may indicate the increase of the platelet-activating factor in the tumor microenvironment. However, these hypotheses need to be effectively verified by subsequent experiments.

The relative abundance of dihydrobiopterin (BH₂), the precursor of tetrahydrobiopterin (BH₄) of folate biosynthesis, was increased in gastric cancer patients compared with the healthy controls. Tetrahydrobiopterin (BH₄) is essential for the synthesis of nitric oxide (NO) [48], and it is recognized as an essential step in initiating neoplastic transformation [49]. A previous study demonstrated that the BH₄:BH₂ ratio is lower in tumor tissues. As a consequence, nitric oxide synthase activity generates

more peroxydinitrite and superoxide anion than nitric oxide, resulting in tumor growth and anti-apoptotic signaling [50].

Genetic associations of discriminant metabolites

Although we found some metabolites associated with GC, the biological mechanisms remain unclear. Hence, it is essential to combine metabolomics with other -omics methods to get a more integrated understanding of gastric carcinogenesis. mGWAS is an excellent method in quantifying metabolic data and uncovering genetic variants affecting metabolite levels [13]. In this study, the area under the ROC of 22 discriminant metabolites was 0.897 (95% CI 1.12 to 1.62), indicative of an ideal intermediate phenotype choice of GC. Therefore, we proceeded to identify mQTLs associated with GC by testing the genetic associations of 22 discriminant metabolites with genome-wide data. We found that 9 SNPs were significantly associated with 3 metabolites. Of them, 3 loci were reported in the previous mGWAS [20], and 6 loci were novel. Surprisingly, 4 of these 9 SNPs were significantly enriched in five genes (i.e., PDE4D, BTN3A3, BTN2A2, BTN3A1, BTN2A3P) involved in the T cell receptor signaling pathway (GO: 0002376).

Besides, 6 SNPs were found to be associated with PAF, which influences the occurrence of gastric cancer. They included rs12208390, rs3927423, and rs6912853 in butyrophilins (BTN) gene; rs140044870 in DCLK2; rs12204145 in ABT1; and rs957788 in FAM155A. The BTN gene belongs to the immunoglobulin superfamily. There are seven human BTN genes in three related phylogenetic subfamilies. BTN1 subfamily contains only the BTN1A1 gene, while BTN2 contains BTN2A1, BTN2A2, and BTN2A3 pseudogenes, and BTN3 contains BTN3A1, BTN3A2, and BTN3A3 genes [51]. Studies have shown that overexpression of BTN3A and dominant expression of BTN3A2 subtypes are closely related to the poor prognosis of pancreatic ductal adenocarcinoma and gastric cancer [52, 53]. Overexpression of the BTN3A2 gene is associated with increased proliferation and invasion of gastric cancer cells [52]. Furthermore, Sarterk et al. found that BTN2A2 is a co-inhibitory molecule that modulates T cell-mediated immune responses [54]. Also, Lebrero-Fernandez et al. found that BTN2A2 gene expression was significantly increased in the tissues of colon cancer patients compared with healthy controls [55]. Importantly, we found that the relative abundance of PAF in the plasma of the gastric cancer group and the control group was different, and PAF-related genes were mainly enriched in the T cell receptor signaling pathway. Therefore, we speculate that PAF-related SNPs may regulate PAF metabolism by regulating the expression of gastric cancer BTN family

genes, mediate the T cell receptor signaling pathway, promote angiogenesis, and promote the growth and metastasis of gastric cancer cells.

Previous studies revealed that the Neuronal tyrosine-phosphorylated Adaptor for The PI3-kinase (NYAPs) could activate the PI3K signaling pathway [56]. Moreover, activation of the PI3K signaling pathway promotes cell growth and survival and is a significant cancer development regulator [57]. Consistently, this study found that the rs140991639 loci in NYAP2 were associated with BH2, which may be related to the occurrence of gastric cancer. Other targets found in this study that may be related to the occurrence of gastric cancer include DCLK2 (associated with reduced survival in cancer patients [58]), ABT1 (encodes proteins needed by ribosomes and contains genetic modifiers responsible for promoting nerve cell survival [59]), and FAM155A (associated with tumor invasion phenotype and early distant metastasis in patients with surgically treated renal clear cell carcinoma [60]). It is worth noting that the rs957788 in FAM155A was found to be associated with anorexia nervosa in a GWAS study [61]. Furthermore, we found rs295977 in PDE4D and rs1183416 in SERP2 were related to Cer(d18:0/12:0). To our knowledge, both PDE4D (phosphodiesterase 4D) and SERP2 (stress-associated endoplasmic reticulum protein family member 2) are associated with tumors [62, 63]. Among them, PDE4D has been reported to act as a proliferation promoter in several different tumors [64–66], and mutations in SERP2 are associated with leukemia [67]. In our study, both the 2 loci were connected to Cer(d18:0/12:0), but further studies are needed to explore the specific relationship between the occurrence of gastric cancer.

Conclusions

In summary, we found 22 metabolites that were statistically different between GC cases and healthy controls. All of them were associated with the risk of gastric cancer. The area under the ROC of 22 discriminant metabolites was 0.897 (95% CI 0.869 to 0.924), indicating a high diagnostic value for GC and an ideal intermediate phenotype choice of GC. More importantly, we identified 6 new mQTLs in GC cases that may be a valuable tool for discovering genetic biomarkers related to GC. Therefore, this paper provides new insights into the prevention, early detection, diagnosis, and targeted treatment of gastric cancer. However, a validation cohort and in vivo and in vitro experiments are required to confirm these findings and fully explore the implications of this study.

Acknowledgements

Not applicable.

Authors' contributions

S.Y.: formal analysis, writing a original draft preparation; Y.W.: investigation; Y.L.: methodology; Y.J.: software and project administration; J.X.: software and project administration; S.J. and SK: project administration; X.L.: project administration; C.H.: project administration; P.L.: project administration; B.L.: supervision; C.W.: writing, review, and editing and supervision. All authors have read and agreed to the published version of the manuscript. The author(s) read and approved the final manuscript.

Funding

This study was supported by the funds from the Xianyou County in Fujian gastrointestinal cancer etiology, epidemiology research, Xianyou County government [grant number 2013B008 085094]; Fujian Natural Science Fund [grant number 2017J01811; 2015J01673]; and Fujian Medical Innovation Fund [grant number 2016-CX-41].

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding authors on reasonable request.

Declarations**Ethics approval and consent to participate**

Ethical approval for the study was obtained from the Biomedical Research Ethics Committee of Fujian Medical University, China (No. 97,2014).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹School of Public Health, Fujian Medical University, Fuzhou, China. ²Fujian Provincial Key Laboratory of Environment Factors and Cancer, Fuzhou, China.

Received: 4 March 2021 Accepted: 8 September 2021

Published online: 26 September 2021

References

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424. <https://doi.org/10.3322/caac.21492>.
- Machlowska J, Baj J, Sitarz M, et al. Gastric cancer: epidemiology, risk factors, classification, genomic characteristics and treatment strategies. *Int J Mol Sci*. 2020, 21(11):1611–8. <https://doi.org/10.1136/gutjnl-2015-309340>.
- Hu N, Wang Z, Song X, Wei L, Kim BS, Freedman ND, et al. Genome-wide association study of gastric adenocarcinoma in Asia: a comparison of associations between cardia and non-cardia tumours. *Gut*. 2016;65(10):1611–8. <https://doi.org/10.1136/gutjnl-2015-309340>.
- Tanikawa C, Kamatani Y, Toyoshima O, Sakamoto H, Ito H, Takahashi A, et al. Genome-wide association study identifies gastric cancer susceptibility loci at 12q24.11-12 and 20q11.21. *Cancer Sci*. 2018;109(12):4015–24. <https://doi.org/10.1111/cas.13815>.
- Shi Y, Hu Z, Wu C, Dai J, Li H, Dong J, et al. A genome-wide association study identifies new susceptibility loci for non-cardia gastric cancer at 3q13.31 and 5p13.1. *Nature genetics*. 2011;43(12):1215–8. <https://doi.org/10.1038/ng.978>.
- Wang Z, Dai J, Hu N, Miao X, Abnet CC, Yang M, et al. Identification of new susceptibility loci for gastric non-cardia adenocarcinoma: pooled results from two Chinese genome-wide association studies. *Gut*. 2017;66(4):581–7. <https://doi.org/10.1136/gutjnl-2015-310612>.
- Yan C, Zhu M, Ding Y, Yang M, Wang M, Li G, et al. Meta-analysis of genome-wide association studies and functional assays decipher susceptibility genes for gastric cancer in Chinese populations. *Gut*. 2020; 69(4):641–51. <https://doi.org/10.1136/gutjnl-2019-318760>.
- Suhre K, Arnold M, Bhagwat A M, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. *Nature Communications*. 2017; 8:14357.
- Nielsen J. Systems biology of metabolism: a driver for developing personalized and precision medicine. *Cell Metab*. 2017;25(3):572–9. <https://doi.org/10.1016/j.cmet.2017.02.002>.
- Shin S-Y, Fauman EB, Petersen A-K, et al. An atlas of genetic influences on human blood metabolites. *Nature genetics*. 2014;46(6):543–50. <https://doi.org/10.1038/ng.2982>.
- Xiao S, Zhou L. Gastric cancer: metabolic and metabolomics perspectives (Review). *Int J Oncol*, 2017, 51(1); Gastric cancer: Metabolic and metabolomics perspectives (Review), 51, 1, 51(17, DOI: <https://doi.org/10.3892/ijo.2017.4000>
- Huang S, Guo Y, Li Z, Zhang Y, Zhou T, You W, et al. A systematic review of metabolomic profiling of gastric cancer and esophageal cancer. *Cancer Biol Med*. 2020;17(1):181–98. <https://doi.org/10.20892/j.issn.2095-3941.2019.0348>.
- Long T, Hicks M, Yu H-C, Biggs WH, Kirkness EF, Menni C, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nature genetics*. 2017;49(4):568–78. <https://doi.org/10.1038/ng.3809>.
- Rhee EP, Ho JE, Chen M-H, Shen D, Cheng S, Larson MG, et al. A genome-wide association study of the human metabolome in a community-based cohort. *Cell Metab*. 2013;18(1):130–43. <https://doi.org/10.1016/j.cmet.2013.06.013>.
- Köttgen A, Albrecht E, Teumer A, et al. Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. *Nature genetics*. 2013;45(2):145–54. <https://doi.org/10.1038/ng.2500>.
- Yang S, Lv Y, Wu C, et al. Pickled vegetables intake impacts the metabolites for gastric cancer. *Cancer Manag Res*. 2020; 12:8263-73.
- Bijlsma S, Bobeldijk I, Verheij ER, Ramaker R, Kochhar S, Macdonald IA, et al. Large-scale human metabolomics studies: a strategy for data (pre-) processing and validation. *Anal Chem*. 2006;78(2):567–74. <https://doi.org/10.1021/ac051495j>.
- Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10(1):1523. <https://doi.org/10.1038/s41467-019-09234-6>.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–504. <https://doi.org/10.1101/gr.1239303>.
- Suhre K, Shin SY, Petersen AK, et al. Human metabolic individuality in biomedical and pharmaceutical research. *Nature*. 2011;477(7362):54–60. <https://doi.org/10.1038/nature10354>.
- Zhu J, Thompson CB. Metabolic regulation of cell growth and proliferation. *Nat Rev Mol Cell Biol*. 2019;20(7):436–50. <https://doi.org/10.1038/s41580-019-0123-5>.
- Hirayama A, Kami K, Sugimoto M, Sugawara M, Toki N, Onozuka H, et al. Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry. *Cancer research*. 2009;69(11):4918–25. <https://doi.org/10.1158/0008-5472.CAN-08-4806>.
- Riera-Domingo C, Audign A, Granja S, et al. Immunity, hypoxia, and metabolism—the MC, Audign A, Grancancer: implications for immunotherapy. *Physiological reviews*, 2020, 100(1); Immunity, Hypoxia, and Metabolism—the Ménage à Trois of Cancer: Implications for Immunotherapy, 100, 1, 100102, DOI: <https://doi.org/10.1152/physrev.00018.2019>
- Campbell SL, Wellen KE. Metabolic signaling to the nucleus in cancer. *Mol Cell*. 2018;71(3):398–408. <https://doi.org/10.1016/j.molcel.2018.07.015>.
- Pavlova NN, Thompson CB. The emerging hallmarks of cancer metabolism. *Cell Metab*. 2016;23(1):27–47. <https://doi.org/10.1016/j.cmet.2015.12.006>.
- Carneiro BA, El-Deiry WS. Targeting apoptosis in cancer therapy. *Nat Rev Clin Oncol*. 2020;17(7):395–417. <https://doi.org/10.1038/s41571-020-0341-y>.
- Holthuis JC, Menon AK. Lipid landscapes and pipelines in membrane homeostasis. *Nature*. 2014;510(7503):48–57. <https://doi.org/10.1038/nature13474>.
- Maxfield FR. Plasma membrane microdomains. *Curr Opin Cell Biol*. 2002; 14(4):483–7. [https://doi.org/10.1016/S0955-0674\(02\)00351-4](https://doi.org/10.1016/S0955-0674(02)00351-4).
- Cheng C, Geng F, Cheng X, Guo D. Lipid metabolism reprogramming and its potential targets in cancer. *Cancer Commun (Lond)*. 2018;38(1):27. <https://doi.org/10.1186/s40880-018-0301-4>.
- Santos CR, Schulze A. Lipid metabolism in cancer. *FEBS J*. 2012;279(15):2610–23. <https://doi.org/10.1111/j.1742-4658.2012.08644.x>.

31. Yu L, Aa J, Xu J, Sun M, Qian S, Cheng L, et al. Metabolomic phenotype of gastric cancer and precancerous stages based on gas chromatography time-of-flight mass spectrometry. *J Gastroenterol Hepatol*. 2011;26(8):1290–7. <https://doi.org/10.1111/j.1440-1746.2011.06724.x>.
32. Jee SH, Kim M, Kim M, Yoo HJ, Kim H, Jung KJ, et al. Metabolomics profiles of hepatocellular carcinoma in a Korean prospective cohort: the Korean Cancer Prevention Study-II. *Cancer Prev Res (Phila)*. 2018;11(5):303–12. <https://doi.org/10.1158/1940-6207.CAPR-17-0249>.
33. Edin ML, Duval C, Zhang G, Zeldin DC. Role of linoleic acid-derived oxylipins in cancer. *Cancer Metastasis Rev*. 2020;39(3):581–2. <https://doi.org/10.1007/s10555-020-09904-8>.
34. Burns J L, Nakamura M T, Ma D W L. Differentiating the biological effects of linoleic acid from arachidonic acid in health and disease. *Prostaglandins Leukot Essent Fatty Acids*. 2018; 135:1-4.
35. Sun R, Gu J, Chang X, et al. Metabonomics study on orthotopic transplantation mice model of colon cancer treated with Astragalus membranaceus-Curcuma wenyujin in different proportions via UPLC-Q-TOF/MS. *J Pharm Biomed Anal*, 2021, 193(113708).
36. To K K W, Lee K-C, Wong S S Y, et al. Lipid metabolites as potential diagnostic and prognostic biomarkers for acute community acquired pneumonia. *Diagn Microbiol Infect Dis*. 2016;85(2):249–54. <https://doi.org/10.1016/j.diagmicrobio.2016.03.012>.
37. Hannun YA, Obeid LM. Principles of bioactive lipid signalling: lessons from sphingolipids. *Nat Rev Mol Cell Biol*. 2008;9(2):139–50. <https://doi.org/10.1038/nrm2329>.
38. Ogretmen B. Sphingolipid metabolism in cancer signalling and therapy. *Nat Rev Cancer*. 2018;18(1):33–50. <https://doi.org/10.1038/nrc.2017.96>.
39. Jeffries K A, Krupenko N I. Ceramide signaling and p53 pathways. *Adv Cancer Res*, 2018, 140: 191-215.
40. Prescott S M, Zimmerman G A, Stafforini D M, et al. Platelet-activating factor and related lipid mediators. *Annu Rev Biochem*, 2000, 69: 419-45.
41. Bussolati B, Biancone L, Cassoni P, Russo S, Rola-Pleszczynski M, Montrucchio G, et al. PAF produced by human breast cancer cells promotes migration and proliferation of tumor cells and neo-angiogenesis. *Am J Pathol*. 2000;157(5):1713–25. [https://doi.org/10.1016/S0002-9440\(10\)64808-0](https://doi.org/10.1016/S0002-9440(10)64808-0).
42. Maggi M, Bonaccorsi L, Finetti G, Carloni V, Muratori M, Laffi G, et al. Platelet-activating factor mediates an autocrine proliferative loop in the endometrial adenocarcinoma cell line HEC-1A. *Cancer Res*. 1994;54(17): 4777–84.
43. Mannori G, Barletta E, Mugnai G, Ruggieri S. Interaction of tumor cells with vascular endothelia: role of platelet-activating factor. *Clin Exp Metastasis*. 2000;18(1):89–96. <https://doi.org/10.1023/A:1026548700247>.
44. Montrucchio G, Lupia E, Battaglia E, del Sorbo L, Boccellino M, Biancone L, et al. Platelet-activating factor enhances vascular endothelial growth factor-induced endothelial cell motility and neoangiogenesis in a murine matrigel model. *Arterioscler Thromb Vasc Biol*. 2000;20(1):80–8. <https://doi.org/10.1161/01.ATV.20.1.80>.
45. Giaginis C, Kourou E, Giagini A, Goutas N, Patsouris E, Kouraklis G, et al. Platelet-activating factor (PAF) receptor expression is associated with histopathological stage and grade and patients' survival in gastric adenocarcinoma. *Neoplasma*. 2014;61(3):309–17. https://doi.org/10.4149/neo_2014_040.
46. Mathonnet M, Descottes B, Valleix D, Truffinet V, Labrousse F, Denizot Y. Platelet-activating factor in cirrhotic liver and hepatocellular carcinoma. *World J Gastroenterol*. 2006;12(17):2773–8. <https://doi.org/10.3748/wjg.v12i17.2773>.
47. Lordan R, Tsoupras A, Zabetakis I. The potential role of dietary platelet-activating factor inhibitors in cancer prevention and treatment. *Adv Nutrition (Bethesda, Md)*, 2019, 10(1): 148-64.
48. Rabender CS, Alam A, Sundaresan G, Cardnell RJ, Yakovlev VA, Mukhopadhyay ND, et al. The role of nitric oxide synthase uncoupling in tumor progression. *Mol Cancer Res*. 2015;13(6):1034–43. <https://doi.org/10.1158/1541-7786.MCR-15-0057-T>.
49. Pacher P, Beckman JS, Liaudet L. Nitric oxide and peroxynitrite in health and disease. *Physiol Rev*. 2007;87(1):315–424. <https://doi.org/10.1152/physrev.00029.2006>.
50. Chittiboyina S, Chen Z, Chiorean EG, Kamendulis LM, Hocevar BA. The role of the folate pathway in pancreatic cancer risk. *Plos One*. 2018;13(2): e0193298. <https://doi.org/10.1371/journal.pone.0193298>.
51. Yokoyama K, Tezuka T, Kotani M, Nakazawa T, Hoshina N, Shimoda Y, et al. NYAP: a phosphoprotein family that links PI3K to WAVE1 signalling in neurons. *EMBO J*. 2011;30(23):4739–54. <https://doi.org/10.1038/emboj.2011.348>.
52. Zhu M, Yan C, Ren C, Huang X, Zhu X, Gu H, et al. Exome array analysis identifies variants in SPOCD1 and BTN3A2 that affect risk for gastric cancer. *Gastroenterology*. 2017;152(8):2011–21. <https://doi.org/10.1053/j.gastro.2017.02.017>.
53. Benyamine A, Loncle C, Foucher E, et al. BTN3A is a prognosis marker and a promising target for Vt for Vor V for Vor V93298.cer risk. *Plos One*, 2018, 13(2): e0193298.298.s, 19Oncoimmunology, 2017, 7(1): e1372080.
54. Sarter K, Leimgruber E, Gobet F, Agrawal V, Dunand-Sauthier I, Barras E, et al. Btn2a2, a T cell immunomodulatory molecule coregulated with MHC class II genes. *J Exp Med*. 2016;213(2):177–87. <https://doi.org/10.1084/jem.20150435>.
55. Lebrero-Fernandez C, Wenzel UA, Akeus P, et al. Altered expression of Butyrophilin (BTN) and BTN-like (BTNL) genes in intestinal inflammation and colon cancer. *Immun Inflamm Dis*. 2016;4(2):191–200. <https://doi.org/10.1002/iid3.105>.
56. Zocchi MR, Costa D, VenR R, et al. Zoledronate can induce colorectal cancer microenvironment expressing BTN3A1 to stimulate effector tordition and antitumor activity. *Oncoimmunology*. 2017;6(3):e1278099. <https://doi.org/10.1080/2162402X.2016.1278099>.
57. Marcu-Malina V, Garelick D, Peshes-Yeloz N, et al. Peripheral blood-derived, gamma9delta2 t cell-enriched cell lines from glioblastoma multiforme patients exert anti-tumoral effects in vitro. *J Biol Regul Homeost Agents*. 2016;30(1):17–30.
58. Barrow TM, Wong Doo N, Milne RL, Giles GG, Willmore E, Strathdee G, et al. Analysis of retrotransposon subfamily DNA methylation reveals novel early epigenetic changes in chronic lymphocytic leukemia. *Haematologica*. 2021; 106(1):98–110. <https://doi.org/10.3324/haematol.2019.228478>.
59. de Planell-Saguer M, Schroeder DG, Rodicio MC, Cox GA, Mourelatos Z. Biochemical and genetic evidence for a role of IGHMBP2 in the translational machinery. *Hum Mol Genet*. 2009;18(12):2115–26. <https://doi.org/10.1093/hmg/ddp134>.
60. Kang HW, Park H, Seo SP, Byun YJ, Piao XM, Kim SM, et al. Methylation signature for prediction of progression free survival in surgically treated clear cell renal cell carcinoma. *J Korean Med Sci*. 2019;34(19):e144. <https://doi.org/10.3346/jkms.2019.34.e144>.
61. Wang K, Zhang H, Bloss CS, Duvvuri V, Kaye W, Schork NJ, et al. A genome-wide association study on common SNPs and rare CNVs in anorexia nervosa. *Mol Psychiatry*. 2011, Sep;16(9):949–59. <https://doi.org/10.1038/mp.2010.107>.
62. Lundin C, Hjorth L, Behrendtz M, Nordgren A, Palmqvist L, Andersen MK, et al. High frequency of BTG1 deletions in acute lymphoblastic leukemia in children with down syndrome. *Genes Chromosomes Cancer*. 2012;51(2): 196–206. <https://doi.org/10.1002/gcc.20944>.
63. Chen L, Gao H, Liang J, Qiao J, Duan J, Shi H, et al. miR-203a-3p promotes colorectal cancer proliferation and migration by targeting PDE4D. *Am J Cancer Res*. 2018;8(12):2387–401.
64. Rahrmann EP, Collier LS, Knutson TP, Doyal ME, Kuslak SL, Green LE, et al. Identification of PDE4D as a proliferation promoting factor in prostate cancer using a Sleeping Beauty transposon-based somatic mutagenesis screen. *Cancer research*. 2009;69(10):4388–97. <https://doi.org/10.1158/0008-5472.CAN-08-3901>.
65. Pullamsetti SS, Banat GA, Schmall A, Szibor M, Pomagurk D, Hänze J, et al. Phosphodiesterase-4 promotes proliferation and angiogenesis of lung cancer by crosstalk with HIF. *Oncogene*. 2013;32(9):1121–34. <https://doi.org/10.1038/onc.2012.136>.
66. Xu T, Wu S, Yuan Y, et al. Knockdown of phosphodiesterase 4D inhibits nasopharyngeal carcinoma proliferation via the epidermal growth factor receptor signaling pathway. *Oncol Lett*. 2014;8(5):2110–6. <https://doi.org/10.3892/ol.2014.2422>.
67. Safavi S, Hansson M, Karlsson K, Biloglav A, Johansson B, Paulsson K. Novel gene targets detected by genomic profiling in a consecutive series of 126 adults with acute lymphoblastic leukemia. *Haematologica*. 2015;100(1):55–61. <https://doi.org/10.3324/haematol.2014.112912>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.