

网络出版时间: 2019-1-3 17:25 网络出版地址: <http://kns.cnki.net/kcms/detail/34.1065.r.20181230.1809.002.html>

胃癌脂代谢通路基因表达的转录组学高通量分析

向丽娟¹, 汪圣毅², 包楚阳¹, 张焱³, 韩坤³, 刘虎^{1,4}

摘要 目的 探讨胃癌(GC)脂代谢(LM)通路的基因表达情况。方法 转录组测序筛选8例GC及4例癌旁组织之间的差异表达基因,京都基因与基因组百科全书(KEGG)富集分析GC脂代谢相关通路中的关键基因(KGS)。结果 高通量测序获得3 198个长度大于200 bp的差异表达基因。KEGG富集分析发现6个显著富集代谢通路,其中脂肪消化与吸收代谢通路有9个基因(MOGAT3、FABP2、FABP1、APOA4、NPC1L1、PLA2G2A、ABCG5、PLA2G4A、PLA2G12B)显著上调,2个基因(LIPF、PLA2G1B)显著下调。结论 GC脂代谢关键基因表达的异常状态为寻求胃癌诊断标志物提供了线索。

关键词 胃肿瘤; 转录组测序; 脂代谢; 关键基因

中图分类号 R 735.2

文献标志码 A **文章编号** 1000-1492(2019)01-0005-05

doi: 10.19405/j.cnki.issn1000-1492.2019.01.002

胃癌(gastric cancer, GC)是世界第五大恶性肿瘤和第三大常见癌症死亡原因^[1]。GC作为消化道肿瘤与营养代谢密切相关。随着脂质组学分析在癌症研究中取得的重要进展,脂代谢(lipid metabolism, LM)异常与癌症关系的研究日益受到关注^[2-5]。转录组是连接基因组遗传信息与生物功能的纽带,是揭示疾病基因突变规律、疾病发生发展重要机制以及发现致病基因调控关键靶点等领域的最

佳研究手段。该研究利用高通量测序对GC及癌旁组织的差异表达谱进行京都基因与基因组百科全书(Kyoto Encyclopedia of Genes and Genomes, KEGG)富集分析,以期探索GC中LM相关通路的关键基因(key genes, KGS),为GC诊断与治疗从转录组学和代谢组学上寻找新的线索。

1 材料与方法

1.1 病例资料 收集安徽医科大学第一附属医院2017年4月7日~19日住院GC患者术后GC组织标本8例及相应癌旁正常组织标本4例,均经病理学诊断证实。其中男3例,女5例,年龄25~77(60.1±15.3)岁。病理显示均为中-低分化腺癌,其中6例有淋巴结转移,但无远处脏器转移。TNM I B~II B期4例,III A~III C期4例。患者术前均未接受过化疗、放疗或生物治疗等干预措施,无其他肿瘤病史,无糖尿病、高血压、肾病、免疫系统疾病及上呼吸道感染等合并症。癌及癌旁两组性别、年龄差异均无统计学意义。标本于术后0.5 h内放入RN-Alater(美国Invitrogen公司)保存液中,−20℃保存,正常对照来自手术边界5 cm以上。所有患者签署知情同意书。该实验经安徽医科大学伦理委员会批准(伦理批准号:20140141)。

1.2 主要仪器与试剂 琼脂糖凝胶电泳系统购自美国Nanodrop公司; NanoPhotometer[®]分光光度计购自德国Implen公司;安捷伦2100生物分析仪购自美国Agilent公司; HiSeq Xten、TruSeq PE 簇生成试剂盒v3-cBot-HS 购自美国Illumina公司; NEBNext[®] Ultra[™] RNA文库制备试剂盒及Oligo d(T) 25磁珠购自美国NEB公司; Qubit[®] 2.0荧光计、Superscript II 逆转录试剂盒及TRIzol试剂购自美国Invitrogen公司; 核酸纯化试剂盒AMPure XP 购自美国Beck-

2018-09-12 接收

基金项目: 国家自然科学基金(编号:81472750); 安徽省自然科学基金(编号:1508085MH171)

作者单位: 安徽医科大学第一附属医院¹ 肿瘤内科、² 普外科, 合肥 230022

³ 北京诺禾致源科技股份有限公司, 北京 100080

⁴ 中国科学技术大学附属第一医院西区(安徽省肿瘤医院), 合肥 230031

作者简介: 向丽娟, 女, 硕士研究生;

刘虎, 男, 博士, 副教授, 副主任医师, 硕士生导师, 责任作者, E-mail: drliuhu@yahoo.com

heart weight/tibia length, myocardial cell cross-sectional area, myocardial fibrosis and ROS in myocardium were significantly increased in IH group. Intervention with SAHA could significantly improve myocardial injury induced by intermittent hypoxia and ROS increase. **Conclusion** SAHA, a histone deacetylase inhibitor, can improve intermittent hypoxia-induced myocardial injury in mice by inhibiting oxidative stress.

Key words intermittent hypoxia; histone deacetylase inhibitor; myocardial injury; oxidative stress

man Coulter 公司。

1.3 cDNA 文库构建及质检 ① 按 TRIzol 试剂说明书进行总 RNA 提取, 随后质检(样品无污染时, RNA 质量 $\geq 1 \mu\text{g}$, RNA 完整性(RNA integrity number, RIN) ≥ 6.8 ; 样品有轻微污染时, RNA 质量 $\geq 2 \mu\text{g}$, RIN ≥ 6.8); ② 通过 Oligo(dT) 磁珠富集 mRNA, 在 NEBNext 碎片缓冲区用二价阳离子将 mRNA 随机打断; ③ 将片段化 mRNA 为模板合成 cDNA 第一条链, 随后用核糖核酸酶 H 降解 RNA 链, 以 dNTPs 为原料合成 cDNA 第二条链; ④ 末端修复、加 A 尾并连接测序接头, 用 AMPure XP 筛选 150 ~ 200 bp 的 cDNA 片段, 进行 PCR 扩增并再次使用 AMPure XP 进行纯化, 最终获得 cDNA 文库; ⑤ 使用 Qubit 2.0 荧光计对文库进行初步定量, 安捷伦 2100 生物分析仪对文库的插入片段大小进行检测, 随后 qRT-PCR 对文库有效浓度进行准确定量(文库有效浓度高于 $2 \times 10^{-9} \text{ mol/L}$), 以保证文库质量。从而得到最终的 cDNA 文库。

1.4 上机测序 根据制造商的说明, 利用 TruSeq PE 簇生成试剂盒 v3-cBot-HIS 生成簇, 然后在 HiSeq Xten 平台上对文库进行双端 150 bp 测序。

1.5 统计学处理 经过测序错误率检查、GC 含量分布检查及原始数据(raw read)过滤, 获得后续待分析数据(clean reads), 采用 STAR(v2.5.1b) 软件对获得的 clean reads 进行比对分析, 使用 HTSeq(v0.6.0) 进行基因水平定量分析, RSEM(v1.2.28) 进行转录本水平的定量分析。采用 edgeR(v3.12.1) 软件进行表达差异显著性分析, 将 q (多重假设检验校正的 P 值) < 0.05 为差异有统计学意义。采用 clusterProfiler(v3.0.5) 软件对差异基因集进行 KEGG 通路富集分析, 将 $q < 0.05$ 为富集有统计学意义, 筛选出与胃癌相关的脂代谢通路。

2 结果

2.1 mRNA 转录组文库的构建 采用 TRIzol 法提取的 GC 和癌旁正常组织的总 RNA 的各项质量指标都基本达标, 然后用 Oligo 磁珠分离出 mRNA, 合成双链 cDNA 文库, 文库质检合格后, 上机测序。

2.2 RNA-seq 数据处理 将构建文库用 HiSeq Xten 进行转录组的双端测序, 为了保证信息分析数据质量, 将 12 个样本的测序数据进行过滤和比对分析。clean reads 均占相应 raw reads 的 90% 以上, $Q20 \geq 94\%$, $Q30 \geq 88\%$ (Q 为 20 和 30 分别代表碱基测序错误率为 0.01 和 0.001, $Q20$ 和 $Q30$ 分别表

示质量值大于 20 和 30 的碱基占总碱基的百分比)。质量控制合格, 后续分析都基于高质量的 clean reads。然后对 RNA-seq 测序数据进行比对分析、基因水平定量及标准化处理。有 91% ~ 95% 的 reads 可以比对到人类参考基因组上, 两组的唯一一对比率 66% ~ 93%。

2.3 差异表达分析 根据以上判断标准, 利用 edgeR 软件筛选差异基因。将 $q < 0.05$ 作为 GC 和癌旁正常组织差异表达基因筛选标准。一共检测出差异表达基因 3 198 个。与正常胃黏膜相比, GC 中表达上调基因共 1 776 个, 下调基因共 1 422 个。火山图表明了 GC 癌组织和正常组织间差异有统计学意义的表达基因分布情况, 见图 1, 上调基因用红点表示, 下调基因用绿点表示, PC 表示癌旁组, DEG 表示差异表达基因。差异倍数(fold change)表示基因在癌组织与正常组织间表达倍数变化, 根据差异分析软件中的收缩模型计算得到。

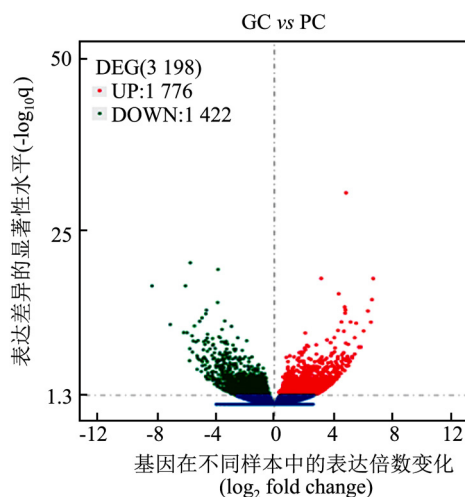


图1 胃癌与正常胃黏膜差异表达基因火山图

2.4 差异表达 KEGG 分析 进一步行 KEGG 富集分析, 评估差异表达基因涉及的生物途径。结果表明, 共映射到 215 种不同的代谢通路上。其中显著富集的途径是蛋白质的消化与吸收, 细胞外基质受体相互作用, 精氨酸和脯氨酸代谢, 粘着斑, 脂肪的消化与吸收, 胃酸分泌, 见表 1。

2.5 脂肪消化与吸收代谢通路的关键基因 脂肪的消化与吸收代谢通路中有 18 个基因显著差异表达, 其中 14 个上调和 4 个下调, 见表 2。代谢通路图见图 2。将基因差异倍数 $|\text{fold change}| > 2$ 作为进一步筛选标准, 得到 9 个上调基因和 2 个下调基因, 可能为脂代谢相关通路的潜在 KGS。

表1 GC和正常组织差异表达基因的KEGG显著富集通路($q < 0.05$)

KEGG 通路编号	通路编号的 功能描述	注释到通路 编号上的基因数	q 值	注释到通路编号上的基因
hsa04 974	Protein digestion and absorption	31	0	PGA5/ATP1B3/COL4A1/PGA3/COL1A1/KCNQ1/COL3A1/ COL1A2/COL6A3/PGA4/COL5A2/COL12A1/COL4A2/SLC7A9/ COL5A1/CPA2/SLC16A10/COL11A1/PRSS3/MEP1A/COL18A1/ DPP4/SLC3A1/SLC9A3/MEP1B/ATP1A1/COL2A1/SLC6A19/ COL6A2/COL5A3/COL15A1
hsa04 512	ECM-receptor interaction	31	0	COL4A1/THBS2/COL1A1/THBS4/COL3A1/COL1A2/COL6A3/ LAMC2/SPPI/COL5A2/COL4A2/COMP/COL5A1/ITGB1/ITGB5/ COL11A1/ITGA2/ITGB8/ITGA5/AGRN/FN1/TNN/LAMC1/THBS1/ COL2A1/HMMR/COL6A2/ITGA11/COL5A3/ITGB7/LAMA4
hsa00 330	Arginine And proline me- tabolism	21	0.003	CKMT2/GLUL/CKM/CPS1/NOS2/CKB/AOC1/AGMAT/GOT1/SMS/ GAMT/ALDH4A1/P4HA3/ALDH1B1/ODC1/ACY1/ARG1/ASS1/ NAGS/CKMT1B/CKMT1A
hsa04 510	Focal adhesion	53	0.003	COL4A1/THBS2/COL1A1/THBS4/COL3A1/COL1A2/COL6A3/ LAMC2/SPPI/COL5A2/RAF1/COL4A2/COMP/COL5A1/ITGB1/IT- GB5/COL11A1/VAV3/PDGFB/ITGA2/ITGB8/PCF/PDGFD/ITGA5/ PAK3/PDGFRB/FN1/ELK1/TNN/RAC1/LAMC1/THBS1/ACTN4/ COL2A1/ACTB/PRKCG/JUN/MYL12B/SOS2/ACTN1/COL6A2/ SOS1/ITGA11/VAV2/FIGF/ZYX/VEGFC/COL5A3/BRAF/CTNNB1/ SHC4/ITGB7/LAMA4
hsa04 975	Fat digestion and absorption	18	0.005	LIPF/MOGAT3/JMJD7-PLA2G4B/FABP2/FABP1/APOA4/AGPAT2/ NPC1L1/PPAP2C/PLA2G2A/ABCG5/PLA2G4A/PLA2G12A/ SLC27A4/PLA2G1B/PLA2G12B/SCARB1/DGAT2
hsa04 971	Gastric acid se- cretion	24	0.009	CCKBR/SLC26A7/ATP1B3/HRH2/KCNJ10/KCNQ1/CALM2/ SLC4A2/CHRM3/KCNE2/PRKACB/ATP4A/SSTR2/KCNJ16/ ATP4B/KCNK2/GNAI3/ATP1A1/SLC9A4/PLCB3/KCNJ15/ACTB/ PRKCG/ADCY3

表2 脂肪消化与吸收代谢通路显著上调或下调基因

基因名称	up/down	q 值	$\log_2$ fold change	基因名称	up/down	q 值	$\log_2$ fold change
MOGAT3	up	0	4.356	PLA2G4A	up	0.007 8	2.257
FABP2	up	0.000 1	4.117	SLC27A4	up	0.014 2	0.843
FABP1	up	0.001 0	3.724	PLA2G12B	up	0.030 4	2.857
APOA4	up	0.001 4	3.786	SCARB1	up	0.034 8	0.919
AGPAT2	up	0.003 2	1.516	DGAT2	up	0.035 2	0.997
NPC1L1	up	0.003 5	2.880	LIPF	down	0	-7.049
PPAP2C	up	0.004 4	1.296	JMJD7-PLA2G4B	down	0	-1.056
PLA2G2A	up	0.006 9	3.440	PLA2G12A	down	0.010 4	-0.802
ABCG5	up	0.007 2	2.674	PLA2G1B	down	0.020 8	-2.383

3 讨论

自从癌症被提出是一种代谢异常性疾病以来,已有诸多研究^[6-8]探索对肿瘤发生发展起关键作用的特定生化代谢途径。脂质具有多种重要的生物学功能,脂质谱系变化及LM异常在癌症的发生发展过程中扮演着十分重要的作用^[4,9]。

KEGG是一个用于从基因组和分子水平了解生物系统高级功能和效用的数据库资源,它通过图形来表示细胞内代谢、膜运输、信号传导和细胞生长周期等生物学过程。该研究对GC差异表达基因行KEGG富集分析,发现LM通路中脂肪的消化与吸收代谢通路显著富集。

脂质的消化吸收分为胆汁酸盐协助脂质消化酶消化脂质和吸收的脂质经再合成进入血液循环两部分。脂质消化产生的长链脂肪酸、2-甘油一酯、胆固醇和溶血磷脂等,在小肠进入肠黏膜细胞后需要单酰甘油-O-酰基转移酶3(monoacylglycerol O-acyl-transferase 3,MOGAT3)、脂肪酸结合蛋白2(fatty acid binding protein 2,FABP2)、载脂蛋白A4(apolipoprotein A4,APOA4)、NPC1样细胞内胆固醇转运蛋白1(NPC1 like intracellular cholesterol transporter 1,NPC1L1)等协助合成乳糜微粒,进而被肠黏膜细胞分泌,经淋巴系统进入血液循环。而该研究中合成这些蛋白的基因高表达,提示吸收的脂质再合成途径未影响GC患者脂质的消化与吸收,而主要是

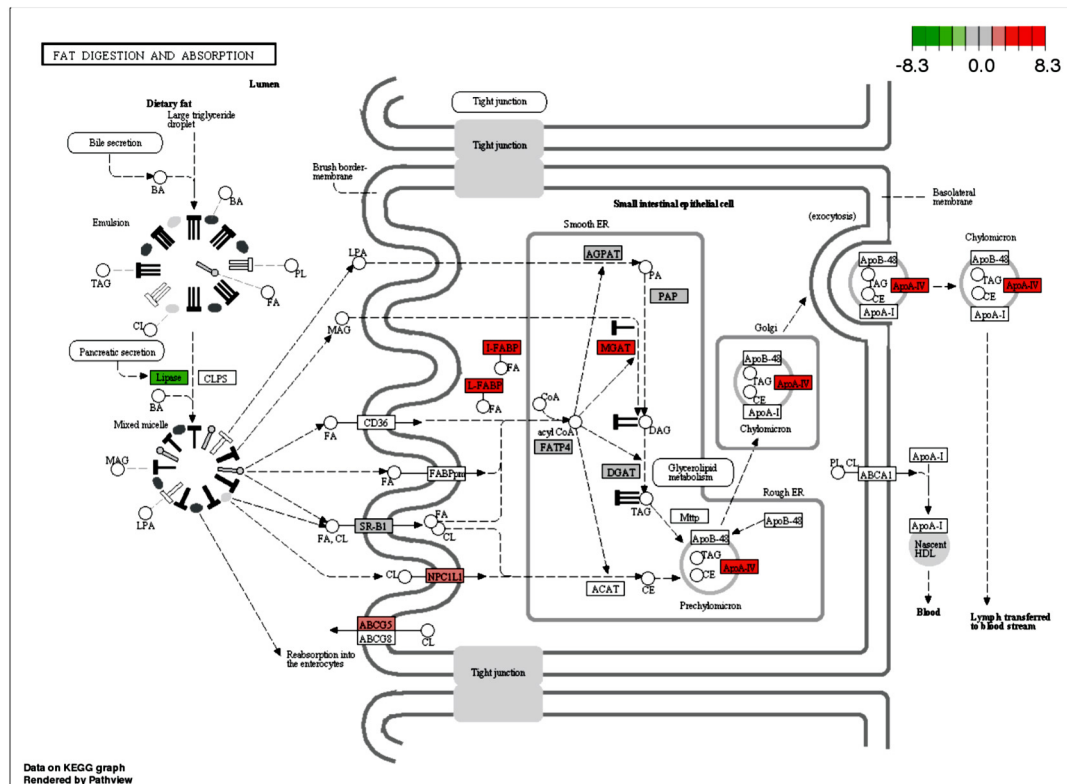


图2 脂肪消化与吸收的代谢通路图

红色表示基因上调,绿色表示基因下调;其中 Lipase 中包括 PLA2G1B、PLA2G2A、PLA2G12B、PLA2G12A、LIPF

脂质消化酶的低表达起了主导作用。

脂肪酶 F_{胃型} (lipase F_{gastric type}, LIPF) 是最显著下调的基因,它编码的胃脂肪酶参与胃肠道中膳食甘油三酯的消化;其表达下调可严重影响机体对脂肪的存储。而脂肪存储障碍是引起癌症恶病质的重要因素之一。因此, LIPF 可能与癌症恶病质的发生紧密相关。

由磷脂酶 A2 组 IIA (phospholipase A2 group IIA, PLA2G2A) 编码的蛋白质是磷脂酶 a2 家族的成员,它催化磷酸甘油酯的 sn-2 脂肪酸酰基酯键的水解,并且被认为参与调节生物膜中的磷脂代谢。有研究^[10]显示 PLA2G2A 在 GC 细胞质中显著表达,其表达程度与肿瘤大小、分化程度及分期相关,是 GC 患者生存的独立预测因子。体外实验显示它的敲除能增加 GC 对 5-氟尿嘧啶的敏感性^[11]。该研究显示 PLA2G2A 在 GC 中上调,这与以前的研究结果一致,提示 PLA2G2A 可能成为 GC 诊断、治疗及预后的分子靶点。另外, PLA2G2A 的过表达与直肠癌、肺癌的预后也相关^[4, 12],并且可能参与前列腺癌的进展^[13]。

PLA2G4A 为花生四烯酸生成途径的限速酶,而花生四烯酸代谢途径与胃肠癌炎症相关肿瘤的发生

相关。这提示 PLA2G4A 在胃肠癌的产生中发挥关键作用。Zhang et al.^[14]发现与非癌组织相比,GC 中 PLA2G4A 的表达显著降低,它的降低与 GC 患者的不利生存相关。而该研究却显示 PLA2G4A 在 GC 组高表达。有研究^[5]显示 PLA2G4A 在乳腺癌中也高表达,与其不良预后相关。此外, PLA2G4A 过表达可以促进癌细胞迁移和侵袭^[2]。各种癌症中都报道了 FABP1 的过表达,并能促进肿瘤血管生成和迁移^[9]。ATP 结合转运蛋白 G 超家族成员 5 (ATP binding cassette subfamily G member 5, ABCG5) 编码的蛋白质作为半转运体可以限制肠吸收并促进胆固醇的胆汁排泄,其高表达可能与 GC 患者的吸收功能缺陷有关,尚需进一步研究。

综上所述,对脂肪的消化与吸收代谢通路的分析显示 LIPF、PLA2G2A、PLA2G4A、FABP1 及 ABCG5 可作为 GC 诊断与治疗的候选生物标志物。本研究为探明 GC 脂质代谢改变机制提供了线索,有利于监控代谢异常和及早干预恶病质的发生,延长患者生存期和提高生活质量。

参考文献

- [1] Torrel A, Bray F, Siegel R L, et al. Global cancer statistics,

- 2012[J]. *CA Cancer J Clin*, 2015, 65(2): 87–108.
- [2] Fu H, He Y, Qi L, et al. cPLA2alpha activates PI3K/AKT and inhibits Smad2/3 during epithelial-mesenchymal transition of hepatocellular carcinoma cells[J]. *Cancer Lett*, 2017, 403: 260–70.
- [3] Kato I, Land S, Majumdar A P, et al. Functional polymorphisms to modulate luminal lipid exposure and risk of colorectal cancer[J]. *Cancer Epidemiol*, 2010, 34(3): 291–7.
- [4] He H L, Lee Y E, Shiue Y L, et al. PLA2G2A overexpression is associated with poor therapeutic response and inferior outcome in rectal cancer patients receiving neoadjuvant concurrent chemoradiotherapy[J]. *Histopathology*, 2015, 66(7): 991–1002.
- [5] Chen L, Fu H, Luo Y, et al. cPLA2alpha mediates TGF-beta-induced epithelial-mesenchymal transition in breast cancer through PI3k/Akt signaling[J]. *Cell Death Dis*, 2017, 8(4): e2728.
- [6] DeBerardinis R J, Chandel N S. Fundamentals of cancer metabolism[J]. *Sci Adv*, 2016, 2(5): e1600200.
- [7] Vander Heiden M G, DeBerardinis R J. Understanding the intersections between metabolism and cancer biology[J]. *Cell*, 2017, 168(4): 657–69.
- [8] Borroughs L K, DeBerardinis R J. Metabolic pathways promoting cancer cell survival and growth[J]. *Nat Cell Biol*, 2015, 17(4): 351–9.
- [9] Ku C Y, Liu Y H, Lin H Y, et al. Liver fatty acid-binding protein (L-FABP) promotes cellular angiogenesis and migration in hepatocellular carcinoma[J]. *Oncotarget*, 2016, 7(14): 18229–46.
- [10] Wang X, Huang C J, Yu G Z, et al. Expression of group IIA phospholipase A2 is an independent predictor of favorable outcome for patients with gastric cancer[J]. *Hum Pathol*, 2013, 44(10): 2020–7.
- [11] Chen D, Jiao X L, Liu Z K, et al. Niu M. knockdown of PLA2G2A sensitizes gastric cancer cells to 5-FU *in vitro* [J]. *Eur Rev Med and Pharmacol Sci*, 2013, 17(13): 1703–8.
- [12] Yu J A, Mauchley D, Li H, et al. Knockdown of secretory phospholipase A2 IIA reduces lung cancer growth *in vitro* and *in vivo* [J]. *J Thorac Cardiovasc Surg*, 2012, 144(5): 1185–91.
- [13] Dong Z, Liu Y, Scott K F, et al. Secretory phospholipase A2-IIa is involved in prostate cancer progression and may potentially serve as a biomarker for prostate cancer[J]. *Carcinogenesis*, 2010, 31(11): 1948–55.
- [14] Zhang X, Wu Q, Gan L, et al. Reduced group IVA phospholipase A2 expression is associated with unfavorable outcome for patients with gastric cancer[J]. *Med Oncol*, 2013, 30(1): 454.

Transcriptomics high-throughput analysis of gene expression in lipid metabolic pathways in gastric cancer

Xiang Lijuan¹, Wang Shengyi², Bao Chuyang¹, et al

(¹Dept of Oncology, ²Dept of General Surgery, The First Affiliated Hospital of Anhui Medical University, Hefei 230022)

Abstract Objective To investigate the gene expression of gastric cancer (GC) lipid metabolism (LM) pathways. **Methods** Transcriptome sequencing was used to screen the differentially expressed genes in 8 cases of gastric cancer and 4 cases of adjacent tissues and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was performed to explore the key genes in lipid metabolism related pathways in gastric cancer. **Results** High-throughput sequencing obtained 3198 differentially expressed genes with a length greater than 200 bp. KEGG enrichment analysis found six significantly enriched metabolic pathways, of which nine genes (MOGAT3, FABP2, FABP1, APOA4, NPC1L1, PLA2G2A, ABCG5, PLA2G4A, PLA2G12B) were significantly up-regulated in fat digestion and absorption metabolic pathway, and 2 genes (LIPF, PLA2G1B) was significantly downregulated. **Conclusion** Abnormal expression of key genes in GC lipid metabolism provides clues to the diagnosis markers of gastric cancer.

Key words stomach neoplasms; transcriptome sequencing; lipid metabolism; key genes