



doi:10.7659/j.issn.1005-6947.2021.02.006
http://dx.doi.org/10.7659/j.issn.1005-6947.2021.02.006
Chinese Journal of General Surgery, 2021, 30(2):165-172.

· 基础研究 ·

基于生物信息学的胆囊癌差异表达谱中关键蛋白调控基因分析

韩伟光, 莘玮, 苏水霞, 王青

(空军军医大学第二附属医院 普通外科, 陕西 西安 710038)

摘要

背景与目的: 胆囊癌(GBC)的机制发生目前仍不清楚, 基因组与转录组层面的数据为研究GBC的分子生物学机制提供了基础数据资源。本研究运用生物信息学手段分析GBC组织和胆囊正常组织的差异表达基因以及GBC关键蛋白调控分子, 从而探讨GBC发生的潜在分子生物学机制。

方法: 基于GEO数据库中两个GBC转录组数据集筛选差异表达基因, 并对这些基因进行GO三种类型功能注释。利用STRING数据库构建蛋白质互作网络, 并进行模块挖掘寻找关键蛋白调控基因。评估关键蛋白调控基因的表达及预测效能。

结果: 共筛选到140个可重复GBC差异表达基因(上调基因20个, 下调基因120个), 主要与前脑发育和神经发生的正调控通路密切相关, 参与突触后膜和横小管组成。同时发现SFRP1这一关键调控蛋白基因在预测GBC的发生中具有一定作用。

结论: 本研究获得的GBC转录谱表达的信息可为GBC分子机制的研究提供框架和脉络, 关键蛋白调控基因SFRP1可能在GBC的发生与发展中发挥关键生物学作用。

关键词

胆囊肿瘤; 基因表达谱; 蛋白质相互作用图; 计算生物学

中图分类号: R735.8

Analysis of key protein regulatory genes in differential expression profile of gallbladder cancer based on bioinformatics approaches

HAN Weiguang, XIN Wei, SU Shuixia, WANG Qing

(Department of General Surgery, the Second Affiliated Hospital of Air Force Medical University, Xi'an 710038, China)

Abstract

Background and Aims: The underlying mechanism for the occurrence of gallbladder carcinoma (GBC) is still unclear at present. The available data at the genomic and transcriptomic levels provide the basic data source for investigation of the molecular biological mechanisms of GBC. Therefore, this study was conducted to analyze the differentially expressed genes in and normal gallbladder tissues and key protein regulatory molecules in GBC by bioinformatics approaches, so as to explore the potential molecular biological mechanism of GBC.

Methods: The differentially expressed genes were screened based on two GBC transcriptional datasets from GEO database, and the three GO functional annotations were performed on these genes. The STRING database was applied to construct a protein interaction network, and perform module mining in the investigation of key protein

收稿日期: 2020-12-05; 修订日期: 2021-01-23。

作者简介: 韩伟光, 空军军医大学第二附属医院主治医师, 主要从事肝胆胰疾病基础与临床方面的研究。

通信作者: 王青, Email: wangqing@fmmu.edu.cn

regulatory genes. Finally, the expression and predictive efficacy of the identified key protein regulatory genes were comprehensively evaluated.

Results: A total of 140 repeatable differentially expressed genes (20 up-regulated genes and 120 down-regulated genes) in GBC were screened, which are mainly related to the forebrain development and positive regulation of neurogenesis, and participate in the composition of the postsynaptic membrane and transverse tubules. Meanwhile, the SFRP1, a key protein regulatory molecule, had a certain ability in predicting the occurrence of GBC.

Conclusion: The information expressed by transcription spectrum of GBC obtained in this study can provide framework and thinking structure for studying the molecular mechanism of GBC. The key protein regulatory molecule SFRP1 probably plays a pivotal role in the occurrence and development of GBC.

Key words

Gallbladder Neoplasms; Gene Expression Profiling; Protein Interaction Maps; Computational Biology

CLC number: R735.8

胆囊癌 (gallbladder cancer, GBC) 作为胆道系统中最常见的恶性肿瘤, 具有高度侵袭性, 其病死率高且预后不良^[1-2]。GBC占胆道恶性肿瘤的80%~95%, 其总发生率为0.8%~1.2%^[3]。以往研究^[4]表明, 超过85%的GBC患者被诊断时患有胆结石, 证实胆结石的存在与GBC发生密切相关, 是GBC的一个重要危险因素。然而, 胆结石患者中少于3%会发展为GBC, 这表明需要额外的因素来驱动GBC的进展。因此, 解析GBC发生的潜在分子生物学机制对GBC的诊断及治疗具有重要的临床意义。

在组织学上, GBC主要分为腺癌和鳞癌两种, 分别占大约90%和1%~12%^[5-6]。手术切除是GBC患者可能治愈的一线选择, 但由于其在早期发病阶段通常无明显症状, 因此往往难以进行早期诊断^[7-8]。当出现黄疸和疼痛等明显症状时, 患者大多已经处于晚期, 并伴有浸润和转移, 普遍错过了接受根治性手术的机会^[9-10]。一些非手术方法, 如化学疗法、放射疗法和分子靶向疗法等可为GBC患者提供姑息性缓解, 减少肿瘤生长和抑制肿瘤转移^[11]。然而, 在过去的几十年中, GBC的总生存期并没有得到明显的改善^[12-14]。因此, 对GBC潜在发病机制的分子理解以及研究显得尤为重要, 同时发现特异的标记物用于GBC的早期诊

断和靶向治疗也面临着巨大的挑战。

随着高通量测序技术的发展, 基因组与转录组层面的数据为研究GBC疾病的分子生物学机制提供了契机。目前已有研究^[15-16]表明基因组的不稳定性和变异性可能影响GBC疾病的发生发展。通过二代测序技术分析已鉴定出与GBC疾病发生有关的许多基因突变, 包括TP53和ERBB3^[17-18]。本研究通过转录组层面的GBC数据来筛选差异表达基因, 旨在发现GBC中异常表达的基因与相关生物学过程。此外对出现的关键蛋白调控基因进行评估, 为GBC患者的治疗提供新思路, 有助于对GBC疾病进行更深入认识。

1 资料与方法

1.1 材料

为研究GBC患者差异基因的表达特征, 本文在美国国家生物技术信息中心 (National Center for Biotechnology Information, NCBI) 的基因表达汇编 (Gene Expression Omnibus, GEO) 数据库中下载人类GBC相关的转录组公共数据, 数据集编号为GSE100363和GSE139682, 数据信息和样本分类见表1。

表1 GBC相关转录组数据信息和样本分类情况

Table 1 Data information of GBC-related transcriptome and sample classification

数据集编号	平台信息	样本来源	正常样本数	GBC 样本数	基因数
GSE100363	HiSeq X Ten (GPL20795)	组织样本	4	4	25343
GSE139682	HiSeq X Ten (GPL20795)	组织样本	10	10	41762

1.2 方法

1.2.1 转录组数据预处理 GSE100363 数据集的预处理过程: (1) 去除每段测序读段的接头序列并过滤低质量序列; (2) 使用 bowtie2 软件^[19] 将序列与参考基因组 hg19 进行比对; (3) 在 RSEM 软件^[20] 中计算基因表达的 FPKM (fragments per kilobase million mapped reads) 值。GSE139682 数据集的预处理过程: (1) 对序列进行过滤, 除去碱基质量低于 13 的 20% 的读段; (2) 使用 HISAT2.1.0 软件^[21] 将序列比对到参考基因组 hg38; (3) 根据 Chepelev 等^[22] 报道的方案计算基因表达的 RPKM (reads per kilobase per million mapped reads) 值。

1.2.2 筛选转录组数据差异表达基因 在 GBC 和正常样本中, 同时使用双总体 Student t 检验和差异倍数 FC (fold change) 法来识别差异表达基因, 本文定义符合阈值为 $P < 0.05$ 且 $|\log_2 FC| \geq 1$ 的基因为差异表达基因 (differentially expressed genes, DEGs)。

1.2.3 差异表达基因的 GO 注释 为深入了解 GBC 中差异表达基因的相关生物学过程, 本文使用 R 软件中 clusterProfiler、org.Hs.eg.db、DOSE 等程序包对差异表达基因进行 GO (gene ontology) 功能注释^[23], 包括基因参与的生物学过程 (biological process, BP)、基因所处的细胞组分 (cellular component, CC) 和基因执行的分子功能 (molecular function, MF)。在此采用多重检验校正的方法 Benjamini-Hochberg 对 P 值进行校正, 设置满足校正 $P < 0.05$ 的 GO 术语具有显著性。

1.2.4 差异表达基因的蛋白质互作网络构建与模块挖掘 随后采用 STRING^[24] 蛋白质互作数据库中的蛋白质互作信息对 GBC 差异表达基因构建蛋白质互作 (protein-protein interaction, PPI)

网络, 并用 Cytoscape 软件对蛋白质互作网络进行可视化, 设置最低互作分数为 0.4。同时使用 Cytoscape 中的图论聚类算法 MCODE (molecular complex detection) 在构建的蛋白互作网络中进行模块挖掘, 设置度的阈值为 2, 节点阈值为 0.2。

1.2.5 关键蛋白调控基因的表达分析和预测能力评估 采用 Graphpad Prism 8.3.0 软件对最终保留的关键蛋白调控基因表达量进行两独立样本 t 检验, $P < 0.05$ 为差异具有统计学意义。并且对关键蛋白调控基因绘制受试者工作特征曲线 (receiver operating characteristic curve, ROC)。ROC 曲线下面积 (area under curve, AUC) 越接近 1.0, 其诊断效果越好。

2 结果

2.1 GBC 差异表达基因分析结果

在两个转录组数据中, 同时采用 t 检验和 FC 的差异基因分析方法, 设置 $P < 0.05$ 且 $|\log_2 FC| \geq 1$ 的阈值条件。GSE100363 数据集筛选出 644 个 (2.54%) 差异表达基因, 其中上调基因 282 个 (43.79%), 下调基因 362 个 (56.21%)。GSE139682 数据集筛选出 2 206 个 (5.28%) 差异表达基因, 上调基因 829 个 (37.58%), 下调基因 1 377 个 (62.42%)。两个数据集中可重复差异基因 144 个 (图 1A)。共同上调基因 20 个 (图 1B), 共同下调基因 120 个 (图 1C), 表 2 列举了在各自数据集中差异表达最大的 5 个共同上调基因和 5 个共同下调基因。由于这 140 个共同差异表达基因可能在 GBC 发生发展中发挥一定作用, 故随后的分析都基于此。

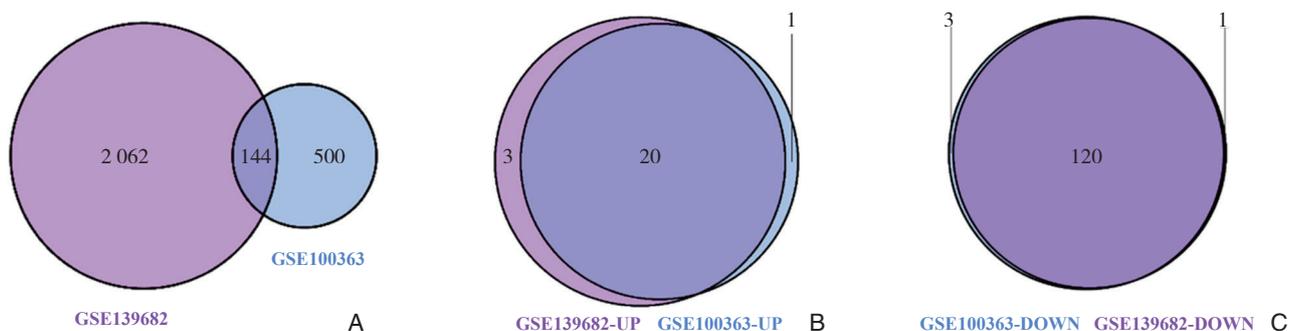


图 1 GBC 表达基因的 Venn 图 A: 两个数据集的共同差异表达基因 Venn 图; B: 两个数据集共同差异表达的上调基因 Venn 图; C: 两个数据集共同差异表达的下调基因 Venn 图

Figure 1 The Venn diagram of differentially expressed genes in GBC A: The Venn diagram of differentially expressed gene in GBC in both data sets; B: The Venn diagram of up-regulated genes in both data sets; C: The Venn diagram of down-regulated genes in both data sets

表2 GBC组织样本在各自数据集中差异表达最大的5个共同上调和5个共同下调基因

Table 2 The 5 common up-regulated and 5 common down-regulated genes with the largest differential expression in GBC tissue samples in each dataset

GSE100363	P	log ₂ FC	GSE139682	P	log ₂ FC
差异表达上调基因					
BLACAT1	0.002 8	6.774 8	KISS1	0.049 7	3.310 6
MISP	0.042 0	6.658 3	OIP5	0.002 9	1.856 2
LINC00673	0.018 4	4.257 6	TK1	0.000 7	1.825 8
LOC101927571	0.046 9	3.148 0	ACBD7	0.033 7	1.598 7
PLEKHN1	0.006 7	2.964 7	BLACAT1	0.009 5	1.584 4
差异表达下调基因					
ADCYAP1R1	0.021 3	-8.656 4	GRIK3	0.000 6	-3.131 8
SLITRK3	0.013 1	-8.422 1	NT5C1A	0.024 5	-2.640 5
SFRP1	0.008 7	-7.339 7	FHL1	0.000 1	-2.614 1
SMTNL2	0.019 2	-6.579 9	PCOLCE2	0.022 2	-2.476 3
ASTN1	0.024 4	-6.330 9	ADCYAP1R1	0.000 5	-2.439 9

2.2 差异表达基因的功能富集

为研究差异表达的140个基因的相关生物学功能,对这些基因进行GO注释,分别富集到2 249条生物学过程术语、247条细胞组分术语和295条分子功能术语。其中符合阈值校正 $P < 0.05$ 的生物学过程术语454条,从最显著的5条,可以看出这140个表达异常的基因主要影响前脑发育、端脑发育和神经发生的正调控等生物学过程,体现了GBC相关的通路出现了不同程度干扰。在这5条通路中,CDON (cell adhesion associated, oncogene regulatged)、FEZ1 (fasciculation and elongation protein zata 1)、GLI3 (GLI family zinc finger 3)、SLIT2 (slit homolog 2)、SOCS2 (suppressor of cytokine signaling 2)基因出现频率最高,表明这5个基因主要参与了GBC的生物学发生过程(图2)。符合阈值的2条细胞组分术语,其中9个基因参与突触后膜组成,4个基因参与横小管组成。另外在阈值的设定下没有符合的分子功能术语(图3)。

2.3 差异表达基因 PPI 网络分析及模块挖掘

对这140个差异表达基因在STRING数据库中筛选互作值 ≥ 0.4 的互作对,之后用Cytoscape软件对蛋白互作网络进行可视化(图4);用Cytoscape软件中的MCODE插件在PPI网络中进行模块挖掘(表3),PPI网络图中不同的颜色对应表3中不同的模块。结果表明SFRP1基因与GLI3、CCND2互作形成模块1;SEMA3D基因与SLIT2、NTN1互作形成模块2;ANK2基因与GNAO1、PTPRD互作形

成模块3。在不同基因表达的蛋白互作下,3个种子基因对于GBC的生物学过程可能显得更为重要。

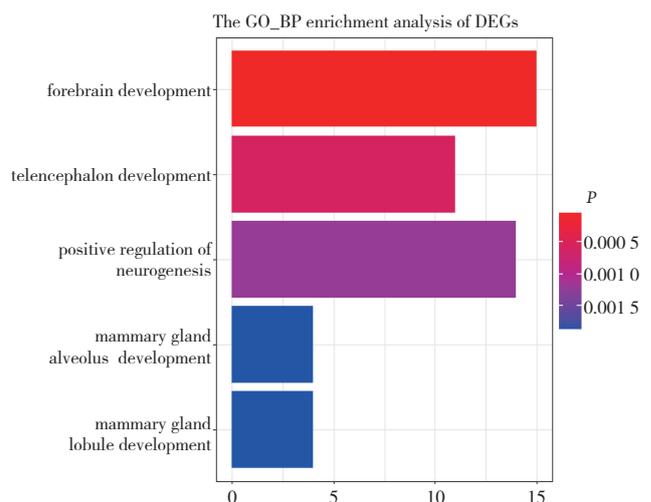


图2 差异表达基因的GO生物学过程注释(纵坐标表示富集到的不同GO生物学过程,横坐标表示注释到某条术语的基因数目;不同色阶表示为不同校正P值,颜色越红表示校正P值越小,其差异越显著,越蓝表示校正P值越大越不显著)

Figure 2 Annotation of GO biological process of differentially expressed genes (the vertical axis showing enrichment of different biological processes, the horizontal axis showing the number of genes annotated to a certain term; different color gradations standing for different adjusted P values, namely the darker the red color, the smaller the adjusted P value and the more significant the difference, and the darker the blue color, the bigger the adjusted P value and the less significant the difference)

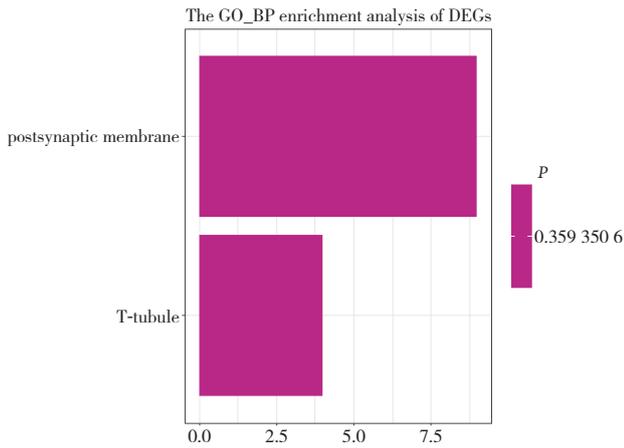


图 3 差异表达基因的 GO 细胞组分注释 (纵坐标表示富集到的不同 GO 细胞组分, 横坐标表示注释到某条术语的基因数目; 图中展示的两条术语校正 P 值在 0.036 附近)

Figure 3 Annotation of GO cellular component of differentially expressed genes (the vertical axis showing enrichment of different cellular components, the horizontal axis showing the number of genes annotated to a certain term; the adjusted P values of the two terms displayed in above picture are around 0.036)

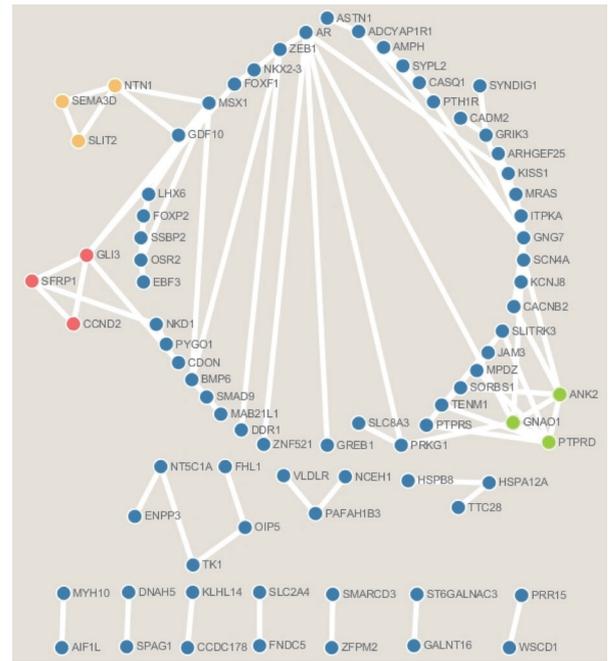


图 4 差异表达基因的 PPI 网络图

Figure 4 PPI network diagram of differentially expressed genes

表 3 PPI 网络中对应的模块信息

Table 3 The corresponding module information in PPI network

模块分类	基因名称	节点类型	网络度	得分
■ 1	SFRP1	种子节点	3	2
■ 1	GLI3	聚类节点	5	2
■ 1	CCND2	聚类节点	2	2
■ 2	SEMA3D	种子节点	2	2
■ 2	SLIT2	聚类节点	2	2
■ 2	NTN1	聚类节点	4	2
■ 3	ANK2	种子节点	5	1.2
■ 3	GNAO1	聚类节点	5	1.2
■ 3	PTPRD	聚类节点	6	1.2

2.4 关键蛋白调控基因表达分析情况

将GBC组织和正常胆囊组织相比, 采用两独立样本t检验, 发现GBC组织的关键蛋白调控基因SFRP1、SEMA3D、ANK2的表达均明显下降, 其

差异均具有统计学意义 (均 $P < 0.05$) (图5)。这一结果表明调控这3种基因的高表达可能降低GBC的发生。SFRP1和SEMA3D基因的差异情况略优于ANK2基因。

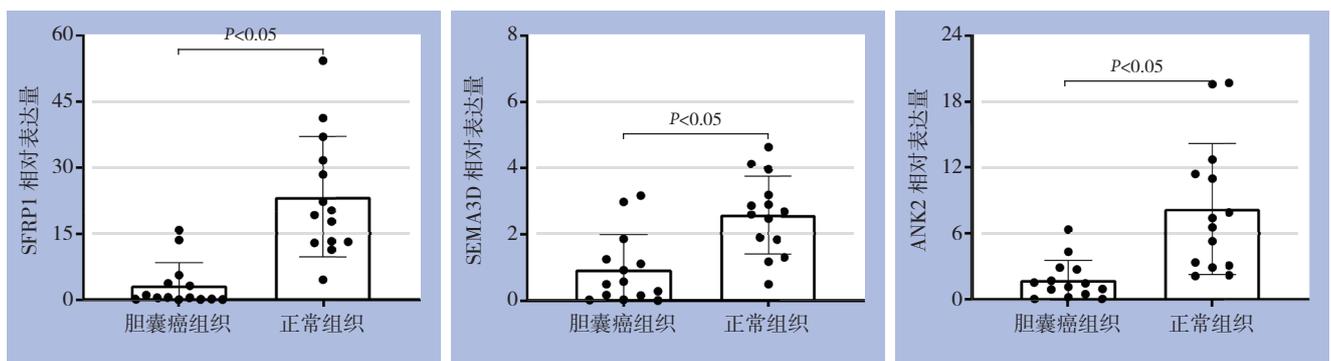


图 5 关键蛋白调控基因表达情况分析

Figure 5 Analysis of expressions of the key protein regulating genes

2.5 评估关键蛋白调控基因对GBC的预测能力

通过绘制ROC曲线来评估关键蛋白调控基因诊断和预测GBC的效果。结果发现SFRP1和ANK2基因对GBC具有一定的预测能力，其

AUC值均>0.9。SFRP1基因的预测效果略优。SFRP1、ANK23、SEMA3D联合应用的AUC值最大，为0.9643（图6）。

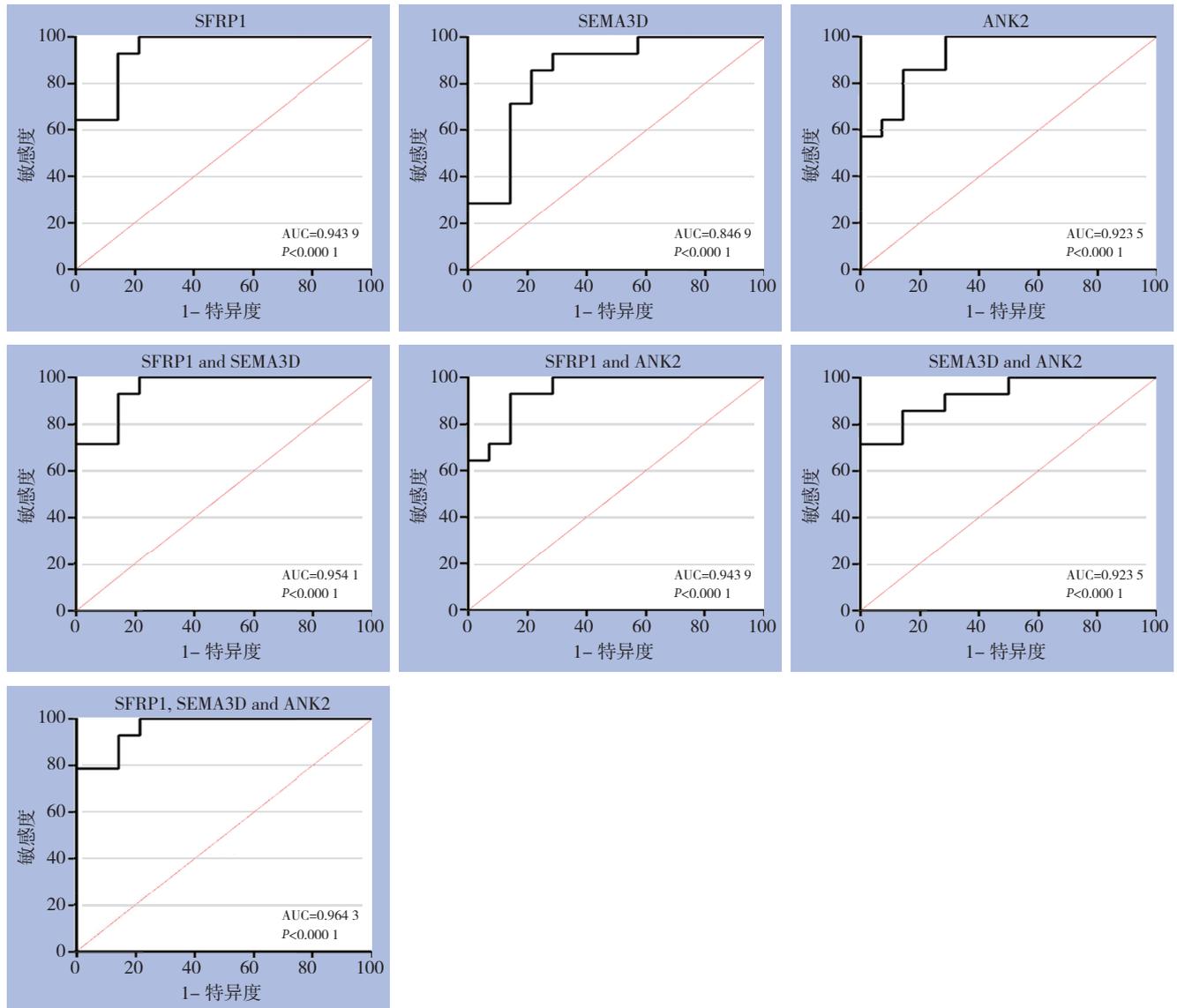


图6 关键蛋白调控基因诊断能力分析

Figure 6 Analysis of diagnostic ability of the key protein regulatory genes

3 讨论

GBC是最常见的胆道恶性肿瘤，由于缺乏特异性症状和可靠敏感的疾病标志，大多数GBC患者被诊断即为晚期，因此迫切面临对分子机制的准确理解与研究。本研究旨在利用转录组层面的高通量数据，使用T检验和倍数变换法筛选差异表达基因，结果发现在两个数据集中符合阈值为 $P<0.05$ 且 $|\log_2FC| \geq 1$ 的共同出现的差异基因140个

（上调20个，下调120个）。通过GO中3种类型的功能注释，在阈值为校正 $P<0.05$ 的条件下，发现差异基因主要与前脑发育、端脑发育和神经发生的正调控等生物学过程有关，与突触后膜和横小管这两种细胞组密切相关，而没有发现与分子功能显著相关的注释信息。随即对差异表达基因构建蛋白互作网络以发现存在明显相互作用的基因。值得注意的是3个种子节点的基因SFRP1、SEMA3D、ANK2，它们与正常胆囊组织相比，在

GBC组织中表达量均为明显下降趋势,可能在GBC发生中起着尤为重要的作用。其中,SFRP1基因相比于其他两个基因的差异表达量最显著且诊断效果也最明显。

既往研究证实由SFRP1基因所编码的SFRP1蛋白作为Wnt信号蛋白的拮抗剂,在Wnt/catenin信号通路的调控中发挥重要作用。此外,SFRP1基因在胆道系统中也被报道是高度相关的甲基化基因,基于其甲基化的分子标记物已成功地用作常规、预后及预测指标,以更好地管理各种癌症患者,一些研究还描述了SFRP1作为癌症的预后和预测生物标志物的价值^[25-26]。在人类中,高度相关甲基化基因的生物标记,允许从相应的正常组织中区分出肿瘤组织,这些标志物还与独特的癌症表型有关,通过进一步优化分析SFRP1基因可能会增加对GBC早期的发现与诊断,并在人体液体中表现相对可靠的监测潜力^[27-28]。

同时,针对目前发现的在差异最显著的5条GO生物学过程中出现的频率最多的基因也有研究证实它们的发生与GBC有关。有研究表明,GBC的许多抑癌基因中已经描述了杂合子丢失,包括染色体1p34-36(p73),5q21(APC),8p21-23(PRLTS和FEZ1)和17p13(p53基因)等^[29]。最新报道^[30]发现,参与GBC发生发展的长链非编码RNA(long-chain non-coding RNA, lncRNA)SNHG6的沉默和miR-26b-5p的上调可促进细胞凋亡、抑制细胞生长和上皮间质转化,并促进GLI3和E-钙黏着蛋白表达的上调,进而影响GBC的生物学过程。另外转录因子GLI3是Hedgehog(Hh/HH)信号通路的成员,可以全长(Gli3-FL/GLI3-FL)或阻遏物(Gli3-R/GLI3-R)的形式存在。GLI3调节各种对于癌细胞生长和发展至关重要的生物过程,锚定非依赖性生长、增殖和迁移的肿瘤细胞系^[31]。另外SLIT2基因的甲基化在多种肿瘤中也有不同程度的研究^[32-33]。

此外,本文也存在如下不足之处。首先,由于是基于生物信息学方法进行的研究,所以部分结果还需要进一步的功能实验进行验证,因而可以为之后的临床应用提供更可靠的依据。其次,由于GBC疾病本身发生机制的异质性和研究样本数量的有限性,快速发展的高通量单细胞层面测序技术可能会解决这一问题。最后,因为GBC疾病的发生是受多方面因素影响的,只从转录组层面看存在一定的片面性,本文尚未结合基因组、

蛋白质组、代谢组等多组学数据信息,多组学数据之间的相关性有待研究。同时环境的因素也可能是GBC发生与进展的重要原因。因此,本研究结果可能适用于部分GBC患者,并非适用于所有GBC患者。

综上,本研究通过分析两个转录组层面数据集取交集的相关差异基因的GO生物学过程和细胞组分,关键蛋白调控基因SFRP1可能在GBC中发挥不可忽视的作用。研究结果具有一定的可靠性,可以为更深刻的认识GBC分子机制提供新的见解与方向。

参考文献

- [1] Miller KD, Siegel RL, Lin CC, et al. Cancer treatment and survivorship statistics, 2016[J]. CA Cancer J Clin, 2016, 66(4):271-289. doi: 10.3322/caac.21349.
- [2] Hundal R, Shaffer EA. Gallbladder cancer: epidemiology and outcome[J]. Clin Epidemiol, 2014, 6:99-109. doi: 10.2147/CLEP.S37357.
- [3] Siegel RL, Fedewa SA, Miller KD, et al. Cancer statistics for hispanics/latinos, 2015 [J]. CA Cancer J Clin, 2015, 65(6):457-480. doi: 10.3322/caac.21314.
- [4] Misra S, Chaturvedi A, Misra NC, et al. Carcinoma of the gallbladder[J]. Lancet Oncol, 2003, 4(3):167-176. doi: 10.1016/s1470-2045(03)01021-0.
- [5] Samuel S, Mukherjee S, Ammannagari N, et al. Clinicopathological characteristics and outcomes of rare histologic subtypes of gallbladder cancer over two decades: A population-based study[J]. PloS one, 2018, 13(6):e0198809. doi: 10.1371/journal.pone.0198809.
- [6] Roa JC, Tapia O, Cakir A, et al. Squamous cell and adenosquamous carcinomas of the gallbladder: Clinicopathological analysis of 34 cases identified in 606 carcinomas[J]. Mod Pathol, 2011, 24(8):1069-1078. doi: 10.1038/modpathol.2011.68.
- [7] Reid KM, Ramos-De la Medina A, Donohue JH. Diagnosis and surgical management of gallbladder cancer: A review[J]. J Gastrointest Surg, 2007, 11(5):671-681. doi: 10.1007/s11605-006-0075-x.
- [8] Henley SJ, Weir HK, Jim MA, et al. Gallbladder cancer incidence and mortality, united states 1999-2011 [J]. Cancer Epidemiol Biomarkers Prev, 2015, 24(9):1319-1326. doi: 10.1158/1055-9965.EPI-15-0199.
- [9] Kakaei F, Beheshtirouy S, Nejatollahi SM, et al. Surgical treatment of gallbladder carcinoma: a critical review [J]. Updates Surg, 2015,

- 67(4):339–351. doi: 10.1007/s13304-015-0328-x.
- [10] Lazcano-Ponce EC, Miquel JF, Muñoz N, et al. Epidemiology and molecular pathology of gallbladder cancer[J]. *CA Cancer J Clin*, 2001, 51(6):349–364. doi: 10.3322/canjclin.51.6.349.
- [11] Lee DG, Lee SH, Kim JS, et al. Loss of *ndrg2* promotes epithelial-mesenchymal transition of gallbladder carcinoma cells through *mmp-19*-mediated slug expression [J]. *J Hepatol*, 2015, 63(6):1429–1439. doi: 10.1016/j.jhep.2015.08.007.
- [12] Kanthan R, Senger JL, Ahmed S, et al. Gallbladder cancer in the 21st century[J]. *J Oncol*, 2015, 2015:967472. doi: 10.1155/2015/967472.
- [13] Caldwell Pilgrim CH, Groeschl RT, Quebbeman EJ, et al. Recent advances in systemic therapies and radiotherapy for gallbladder cancer[J]. *Surg Oncol*, 2013, 22(1):61–67. doi: 10.1016/j.suronc.2012.12.001.
- [14] Horgan AM, Amir E, Walter T, et al. Adjuvant therapy in the treatment of biliary tract cancer: A systematic review and meta-analysis[J]. *J Clin Oncol*, 2012, 30(16):1934–1940. doi: 10.1200/JCO.2011.40.5381.
- [15] Razumilava N, Gores GJ. Building a staircase to precision medicine for biliary tract cancer [J]. *Nat Genet*, 2015, 47(9):967–968. doi: 10.1038/ng.3386.
- [16] Mhatre S, Wang Z, Nagrani R, et al. Common genetic variation and risk of gallbladder cancer in india: A case-control genome-wide association study[J]. *Lancet Oncol*, 2017, 18(4):535–544. doi: 10.1016/S1470-2045(17)30167-5.
- [17] Jiao Y, Pawlik TM, Anders RA, et al. Exome sequencing identifies frequent inactivating mutations in *bap1*, *arid1a* and *pbrm1* in intrahepatic cholangiocarcinomas[J]. *Nat Genet*, 2013, 45(12):1470–1473. doi: 10.1038/ng.2813.
- [18] Nakamura H, Arai Y, Totoki Y, et al. Genomic spectra of biliary tract cancer[J]. *Nat Genet*, 2015, 47(9):1003–1010. doi: 10.1038/ng.3375.
- [19] Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2 [J]. *Nat Methods*, 2012, 9(4):357–359. doi: 10.1038/nmeth.1923.
- [20] Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome[J]. *BMC Bioinformatics*, 2011, 12:323. doi: 10.1186/1471-2105-12-323.
- [21] Kim D, Paggi JM, Park C, et al. Graph-based genome alignment and genotyping with hisat2 and hisat-genotype [J]. *Nat Biotechnol*, 2019, 37(8):907–915. doi: 10.1038/s41587-019-0201-4.
- [22] Chepelev I, Wei G, Tang Q, et al. Detection of single nucleotide variations in expressed exons of the human genome using rna-seq[J]. *Nucleic Acids Res*, 2009, 37(16):e106. doi: 10.1093/nar/gkp507.
- [23] Gene Ontology Consortium. Gene Ontology Consortium: going forward[J]. *Nucleic Acids Res*, 2015, 43(Database issue):D1049–1056. doi: 10.1093/nar/gku1179.
- [24] von Mering C, Huynen M, Jaeggi D, et al. String: A database of predicted functional associations between proteins[J]. *Nucleic Acids Res*, 2003, 31(1):258–261. doi: 10.1093/nar/gkg034.
- [25] Leygo C, Williams M, Jin HC, et al. DNA methylation as a noninvasive epigenetic biomarker for the detection of cancer[J]. *Dis Markers*, 2017, 2017:3726595. doi: 10.1155/2017/3726595.
- [26] Zheng L, Sun D, Fan W, et al. Diagnostic value of *sfrp1* as a favorable predictive and prognostic biomarker in patients with prostate cancer[J]. *PLoS One*, 2015, 10(2):e0118276. doi: 10.1371/journal.pone.0118276.
- [27] Amornpisutt R, Prongvitaya S, Jearanaikoon P, et al. DNA methylation level of *opcml* and *sfrp1*: A potential diagnostic biomarker of cholangiocarcinoma [J]. *Tumour Biol*, 2015, 36(7):4973–4978. doi: 10.1007/s13277-015-3147-2.
- [28] Yamashita K, Hosoda K, Nishizawa N, et al. Epigenetic biomarkers of promoter DNA methylation in the new era of cancer treatment[J]. *Cancer Sci*, 2018, 109(12):3695–3706. doi: 10.1111/cas.13812.
- [29] Kuroki T, Tajima Y, Matsuo K, et al. Genetic alterations in gallbladder carcinoma[J]. *Surg Today*, 2005, 35(2):101–105. doi: 10.1007/s00595-004-2906-2.
- [30] Liu XF, Wang K, Du HC. Lncrna *snhg6* regulating hedgehog signaling pathway and affecting the biological function of gallbladder carcinoma cells through targeting *mir-26b-5p*[J]. *Eur Rev Med Pharmacol Sci*, 2020, 24(14):7598–7611. doi: 10.26355/eurev_202007_22258.
- [31] Matissek SJ, Elsawa SF. GLI3: a mediator of genetic diseases, development and cancer[J]. *Cell Commun Signal*, 2020, 18(1):54. doi: 10.1186/s12964-020-00540-x.
- [32] Chen H, Zhang T, Sheng Y, et al. Methylation profiling of multiple tumor suppressor genes in hepatocellular carcinoma and the epigenetic mechanism of *3ost2* regulation[J]. *J Cancer*, 2015, 6(8):740–749. doi: 10.7150/jca.11691.
- [33] Suzuki M, Shiraishi K, Eguchi A, et al. Aberrant methylation of *line-1*, *slit2*, *mal* and *igfbp7* in non-small cell lung cancer[J]. *Oncol Rep*, 2013, 29(4):1308–1314. doi: 10.3892/or.2013.2266.

(本文编辑 姜晖)

本文引用格式: 韩伟光, 莘玮, 苏水霞, 等. 基于生物信息学的胆囊癌差异表达谱中关键蛋白调控基因分析[J]. 中国普通外科杂志, 2021, 30(2):165–172. doi:10.7659/j.issn.1005-6947.2021.02.006

Cite this article as: Han WG, Xin W, Su SX, et al. Analysis of key protein regulatory genes in differential expression profile of gallbladder cancer based on bioinformatics approaches[J]. *Chin J Gen Surg*, 2021, 30(2):165–172. doi:10.7659/j.issn.1005-6947.2021.02.006