

# 基于生物信息学分析挖掘非小细胞肺癌中的预后基因

韩骐蔓, 朱静娟, 张传涛, 王力, 张晓春

青岛大学附属医院肿瘤精准医学中心, 山东 青岛  
Email: 15269757106@163.com, zxc9670@qdu.edu

收稿日期: 2021年3月15日; 录用日期: 2021年4月13日; 发布日期: 2021年4月20日

## 摘要

非小细胞肺癌(non-small cell lung cancer, NSCLC)仍是当今世界死亡率最高的恶性肿瘤, 但其发生和发展的分子机制的改变仍不清楚。本研究中, 我们通过基因表达综合(Gene Expression Omnibus, GEO)数据库下载基因表达微阵列数据GSE11830, 例用edgeR软件筛选肿瘤组织及周围正常组织之间上调和下调最明显的20个差异基因。为进一步了解差异基因的功能及机制, 我们通过基因本体论(Gene Ontology, GO)数据库及京都基因与基因组百科全书(Kyoto Encyclopedia of Gene and Genomes, KEGG)数据库对这40个差异基因的进行功能及通路富集进行分析。此外, 我们通过癌症基因组数据库(The Cancer Genome Atlas, TCGA)下载NSCLC患者的临床信息, 并对40个差异基因进行生存分析, 发现5个上调基因及3个下调基因与NSCLC生存期影响较为显著。利用cBioPortal可视化工具对这8个关键基因进行基因突变及DNA扩增频率的分析, 发现PKHD1L1、MME和IGSF10可能是NSCLC发生及影响预后的关键基因。

## 关键词

差异基因, 生信分析, 非小细胞肺癌, 预后

# Mining Prognostic Genes in Non-Small Cell Lung Cancer Based on Bioinformatics Analysis

Qiman Han, Jingjuan Zhu, Chuantao Zhang, Li Wang, Xiaochun Zhang

Oncology Precision Medicine Center, Affiliated Hospital of Qingdao University, Qingdao Shandong  
Email: 15269757106@163.com, zxc9670@qdu.edu

Received: Mar. 15<sup>th</sup>, 2021; accepted: Apr. 13<sup>th</sup>, 2021; published: Apr. 20<sup>th</sup>, 2021

## Abstract

Non-small cell lung cancer (NSCLC) is still the malignant tumor with the highest mortality rate in the world, but the molecular mechanism of occurrence and development of NSCLC is still unclear. In this study, we downloaded gene expression microarray data GSE11830 through the Gene Expression Omnibus (GEO) database, and used the edgeR software to screen the 20 most significantly up-regulated and down-regulated genes between tumor tissues and surrounding normal tissues. In order to further understand the functions and mechanisms of differential genes, we used the Gene Ontology (GO) database and the Kyoto Encyclopedia of Gene and Genomes (KEGG) database to perform enrichment analysis of these 40 differential genes. In addition, we downloaded the clinical information of NSCLC patients through The Cancer Genome Atlas (TCGA), and analyzed the survival of 40 differential genes, and found that 5 up-regulated genes and 3 down-regulated genes have a significant impact on NSCLC survival. In addition, we use cBioPortal which was a visualization tool to analyze the gene mutation and DNA amplification frequency of these 8 key genes, and it is found that *PKHD1L1*, *MME* and *IGSF10* may be the key genes for the occurrence and prognosis of NSCLC.

## Keywords

Differential Gene, Bioinformatics Analysis, Non-Small Cell Lung Cancer, Prognosis

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 背景

目前为止, 肺癌仍是中国发生率及死亡率最高的恶性肿瘤。根据世界卫生组织国际癌症研究机构(international agency for research on cancer, IARC)统计数据, 2020 年全球新发肺癌 220 万人, 其中中国新发肺癌 82 万人(<https://www.iarc.who.int/faq/latest-global-cancer-data-2020-qa/>)。非小细胞肺癌(non-small cell lung cancer, NSCLC)约占肺癌中的 80% [1], 包括腺癌、鳞状细胞癌及大细胞癌。相较于组织学分型, 基因分型在 NSCLC 的诊断、治疗、预后中发挥着越来越重要的作用。存在驱动基因突变的晚期 NSCLC 患者, 服用与之匹配的靶向药物比以铂类为基础的双联化疗方案可明显改善预后[2], 例如表皮生长因子受体(epidermal growth factor receptor, EGFR)突变和间变性淋巴瘤激酶(anaplastic lymphoma kinase, ALK)融合突变的患者可以服用小分子酪氨酸激酶抑制剂(tyrosine-kinase inhibitor, TKI)。但 EGFR 突变的患者占 NSCLC 患者的 10%~30% [3], ALK 融合突变患者仅占 3%~5% [4], NSCLC 的 5 年生存率也仅有 23% [5]。目前已知的 NSCLC 的驱动基因仍只是冰山一角, 与 NSCLC 发生发展过程相关的潜在基因仍需继续探索, 以研究 NSCLC 更有效的治疗策略。

随着微阵列技术的快速发展, 微阵列分析在获得和理解分子机制方面展现出巨大潜能[6], 我们可以通过一些高通量平台分析与肿瘤发生发展相关的差异性基因。在这项研究中, 我们从基因表达综合(Gene Expression Omnibus, GEO)数据库中获取 NSCLC 肿瘤组织与正常组织信使核糖核酸(message Ribonucleic Acid, mRNA)微阵列数据集, 通过生物信息学分析获取与 NSCLC 相关的关键基因, 寻找与 NSCLC 预后相关的潜在靶点。

## 2. 方法

### 2.1. 基因表达谱数据采集

在 GEO (<http://www.ncbi.nlm.nih.gov/geo>) 数据库中检索“NSCLC”，按以下标准进行筛选：所有组织样本均为人类 NSCLC 肿瘤组织及周围正常组织；微阵列或 RNA 测序数据应未 mRNA；组织样本由权威机构测试获得美国食品药品监督管理局(FDA)认可。选择“表达谱芯片”、“智人”，下载发布日期较新且肿瘤组织与正常组织样品数目相匹配的微阵列数据集 GSE11830，其微阵列数据包含 6 例肿瘤组织样本和 6 例周围正常组织样本，所有数据均通过 Affymetrix U33Plu2.0 平台进行检测。

### 2.2. 数据预处理与寻找差异基因

首先下载 GSE11830 的平台注释文件，将基因名称与探针名称一一对应并删除不能相互匹配的探针，当多个探针对应同一个基因时取中位数作为基因的最终表达率。edgeR 软件(版本：4.0.2, <http://www.bioconductor.org/packages/release/bioc/html/edgeR.html>) 的 limma 包(版本：3.10.3, <http://www.bioconductor.org/packages/2.9/bioc/html/limma.html>)通过经典贝叶斯方法(topological map-matching, TMM)对微阵列数据进行差异性分析[7]。使用 Benjamini & Hochberg 方法执行获得校正后的差异显著性(adj.P.Value)，选择差异显著性小于 0.05 且选择绝对对数(|logFC|)变化大于 1 作为阈值进行下面的分析。然后利用 edgeR 安装 pheatmap 包对微阵列数据进行差异分析，并选取上调及下调最显著的差异基因各 20 个绘制热图[8]。

### 2.3. 功能及通路富集

基因本体论(Gene Ontology, GO)数据库包括生物学过程(Biological, BP)、细胞组分(Cellular Component, CC)、分子功能(Molecular Function, MF)，京都基因与基因组百科全书(Kyoto Encyclopedia of Gene and Genomes, KEGG)数据库整合基因组学、生物化学及系统功能组学信息。通过 Cluster Profiler 软件包(版本：3.1, <http://bioconductor.org/packages/3.1/bioc/html/clusterProfiler.html>)将差异显著性小于 0.05 的 40 个差异基因的进行功能富集分析及通路富集分析[9]。

### 2.4. 差异基因的生存分析及其肺癌中基因组变化情况

入癌症基因组数据库(The Cancer Genome Atlas, TCGA)数据库，下载肺恶性肿瘤转录组分析基因表 RNA 测序数据(HTSeq-FPKM)，manifest 以及样品数据，metadata 数据及 clinical 数据中的 JSON 文件下载。例用下载的 NSCLC 患者的临床信息分析对 40 个差异基因进行生存分析，安装 edgeR 软件 surviva 包(版本 3.2-7, <https://cran.r-project.org/web/packages/survival/index.html>)及 survminer (版本：0.4.8, <https://cran.r-project.org/web/packages/survminer/index.html>)包对差异基因进行生存分析并选取绘制生存曲线[10]，Kaplan-meier 生存曲线用 log-rank 进行验证，将病人分为基因的高表达组与低表达组，认为 P 值小于 0.05 有统计学意义寻找关键基因。癌症基因组学的可视化工具 cBioPortal (<http://cbioportal.org>)提供可视化数据和多维癌症基因数据分析[11]。本项研究中，我们从 6531 个样本中分析关键基因突变及 DNA 拷贝数目的发生频率。

## 3. 结果

### 3.1. 差异基因分析

从 GSE11830 数据集中，NSCLC 肿瘤组织与周围正常组织之间差异最明显的 20 个上调基因为 *SERTM1*、*FLJ34503*、*LOC101928161*、*IGSF10*、*CLDN18*、*FCN3*、*CADM3-AS1*、*SGCG*、*LOC400568*、

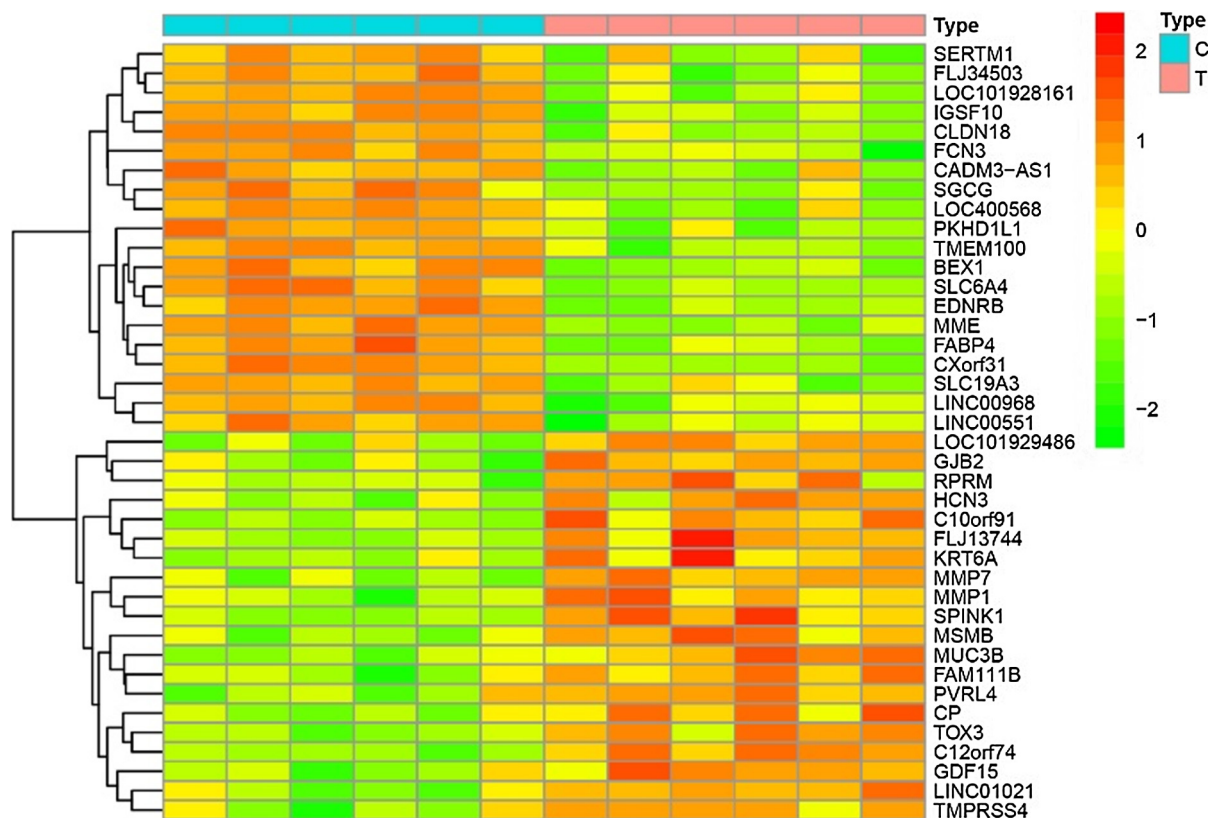
*PKHD1L1*、*TMEM100*、*BEX1*、*SLC6A4*、*EDNRB*、*MME*、*FABP4*、*CXorf31*、*SLC19A3*、*LINC00968*、*LINC00551*; 差异最明显的 20 个下调基因分别为 *LOC101929486*、*GJB2*、*RPRM*、*HCN3*、*C10orf91*、*FLJ13744*、*KRT6A*、*MMP7*、*MMP1*、*SPINK1*、*MSMB*、*MUC3B*、*FAM111B*、*PVRL4*、*CP*、*TOX3*、*C12orf74*、*GDF15*、*LINC01021*、*TMPRSS4* (见图 1)。

### 3.2. 富集分析

为进一步了解差异基因的功能及机制, 我们例用 ClusterProfiler 软件包对其进行富集分析。差异基因的 GO 富集分析显示, BP 主要与老化、痛觉、多细胞生物过程、负生长调节相关; CC 分析可见差异基因主要与膜筏、投影终端、神经细胞体、突触前相关; MF 方面主要与激素结合、金属肽酶抑制剂活性、氯离子跨膜转运蛋白活性、蛋白酶结合相关(见图 2(a)~(c))。KEGG 通路富集分析显示四种重要的途径, 包括血管平滑肌收缩, 心肌细胞的肾上腺素信号传导, cGMP-PKG 信号通路, 肾素-血管紧张素系统(见图 2(d))。

### 3.3. 生存分析

为探索差异基因与 NSCLC 总生存期(overall survival, OS)之间的关系, 我们通过绘制生存曲线进行统计学分析。我们发现表达上调基因 *GDF15* (log-rank,  $p = 0.032$ )、*IGSF10* (log-rank,  $p = 0.009$ )、*MME* (log-rank,  $p = 0.021$ )、*PKHD1L1* (log-rank,  $p = 0.050$ )、*SGCG* (log-rank,  $p = 0.034$ )会降低 NSCLC 的 OS, 而表达下调差异基因 *FAM111B* (log-rank,  $p = 0.002$ )、*KRT6A* (log-rank,  $p = 0.026$ )、*RPRM* (log-rank,  $p = 0.045$ )则会改善 OS (见图 3(a)~(h))。



**Figure 1.** The GSE11830 data set has the most significant difference between up-regulated and down-regulated genes. Red: high expression; green: ground expression

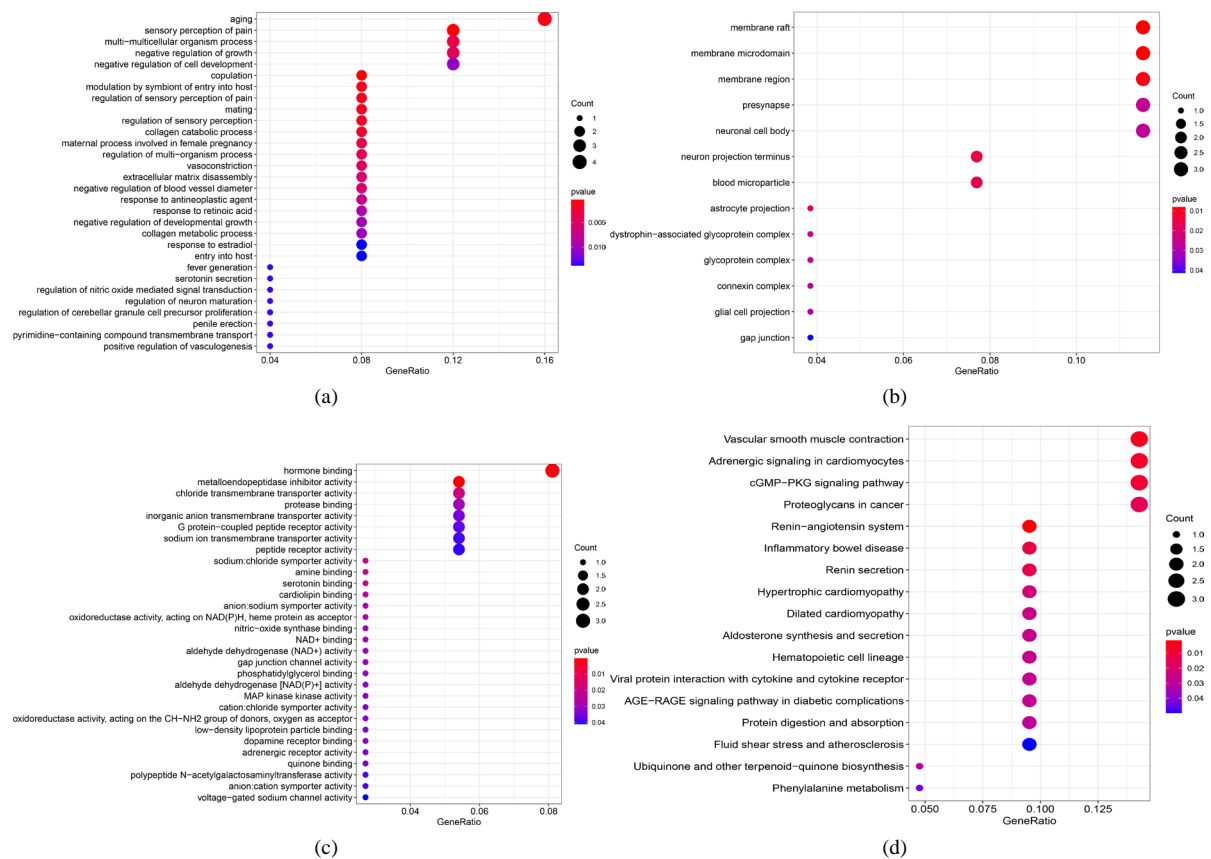
**图 1.** GSE11830 数据集差异最显著的上调基因及下调基因。红色: 高表达; 绿色: 地表达

### 3.4. 预后相关差异基因在肺癌中的发生频率

根据癌症基因组学 cBioPortal 获得的数据, 在 6531 个样本中对 8 个与预后相关的差异基因进行基因突变类型及 DNA 扩增的分析, 如图 4 所示。可见上调基因中, *GDF15*、*IGSF10*、*MME*、*PKHD1L1*、*SGCGFAM* 基因发生突变或 DNA 扩增的频率分别为 1.2%、12%、12%、15%、1.5%, 下调基因 *FAM111B*、*KRT6A*、*RPRM* 发生突变或 DNA 扩增的频率为 1.2%、1.9%、0.7%。

## 4. 讨论

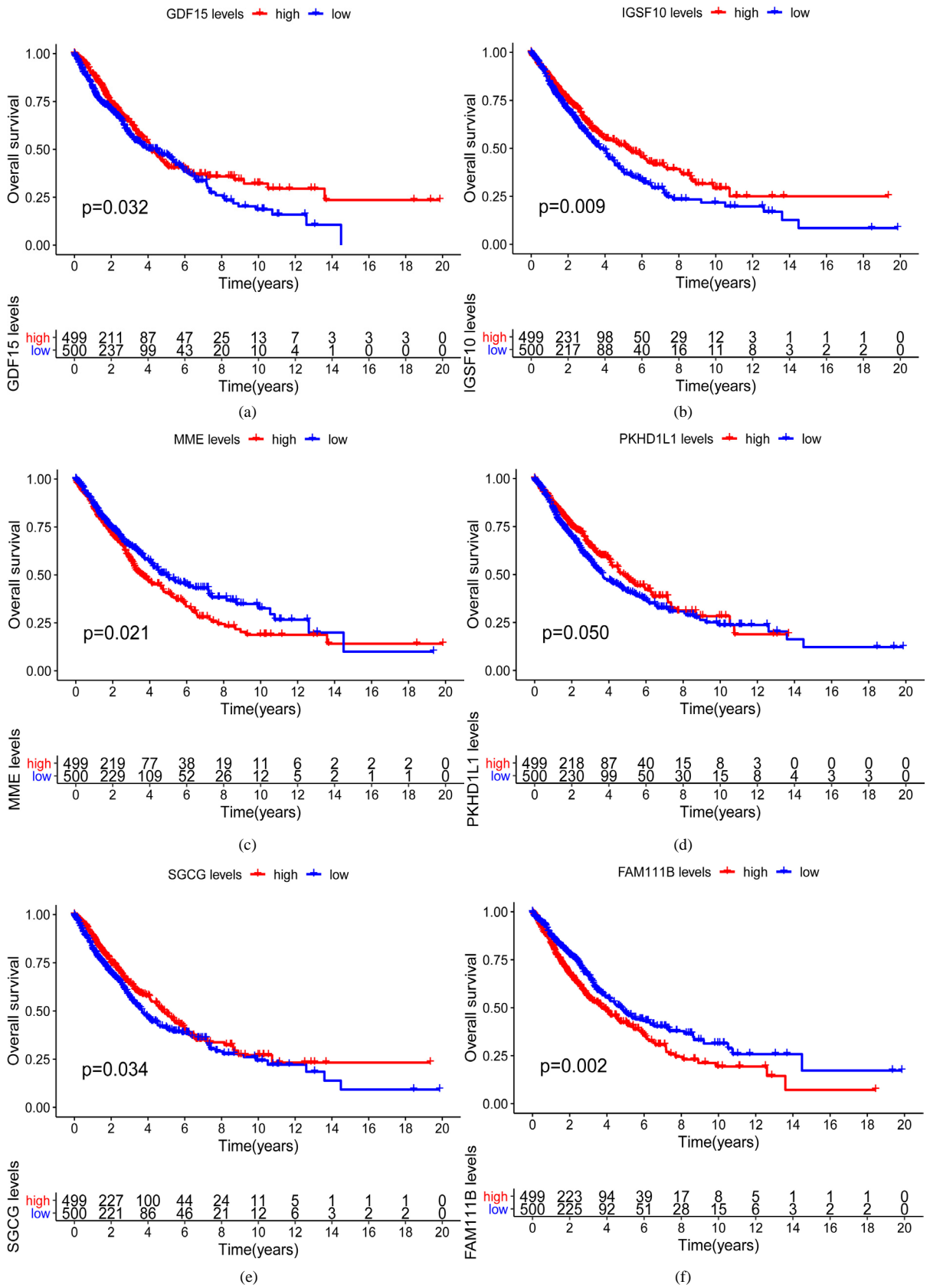
尽管近年来晚期 NSCLC 在放疗、化疗、靶向治疗、免疫治疗方面取得了一些突破性的进展, 但其 5 年生存率仍旧很低。微阵列技术目前已经广泛应用于肿瘤的诊断、治疗及预后评估。在本项研究中, 我们进行了一系列的生物信息学分析, 在 NSCLC 肿瘤组织与正常组织之间筛选出差异常最明显的 20 个上调基因和 20 个下调基因。通过富集分析发现差异基因通路富集主要位于血管平滑肌收缩, 心肌细胞的肾上腺素信号传导, cGMP-PKG 信号通路, 肾素-血管紧张素系统。主要功能富集于细胞老化、多细胞生物过程、负生长调节等生物学过程, 激素结合、金属肽酶抑制剂活性、氯离子跨膜转运蛋白活性、蛋白酶结合等分子功能, 这些均显示差异基因与肿瘤发生或转移密切相关。此外, 我们利用 TCGA 数据库下载 NSCLC 患者的临床信息, 对 40 个差异基因进行生存分析, 筛选出对 NSCLC 预后有明显影响的关键基因, 其中包括 5 个上调基因和 3 个下调基因, 表明这些基因与肺癌显著相关, 可能在肺癌的发生发展中

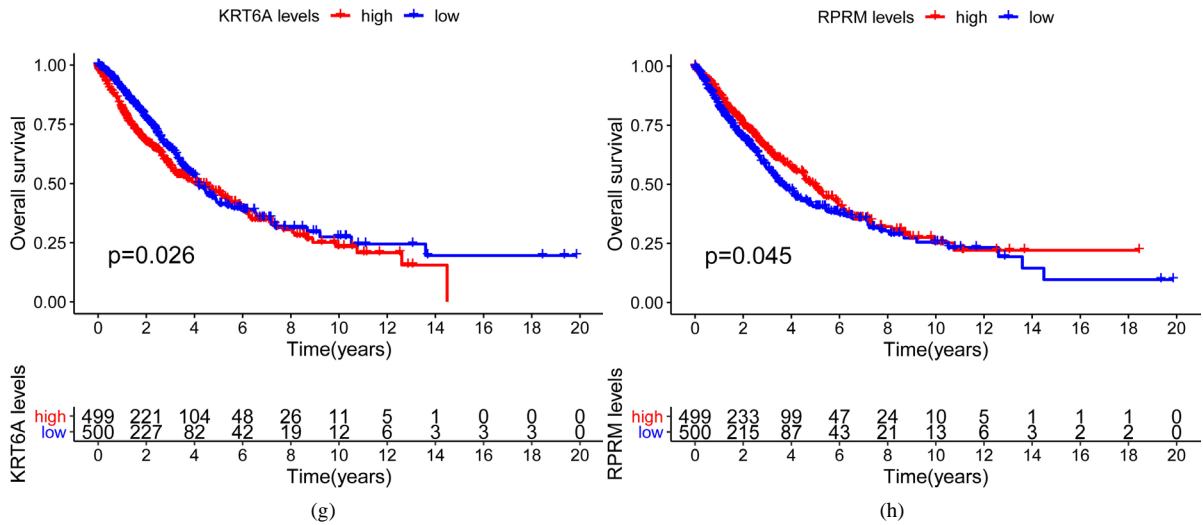


**Figure 2.** Differential gene enrichment analysis of function and pathway. (a) BP enrichment analysis. (b) CC enrichment analysis. (c) MF enrichment analysis. (d) KEGG enrichment analysis.

**图 2.** 差异基因的功能和通路富集分析。(a) BP 富集分析。(b) CC 富集分析。(c) MF 富集分析。(d) KEGG 富集分析

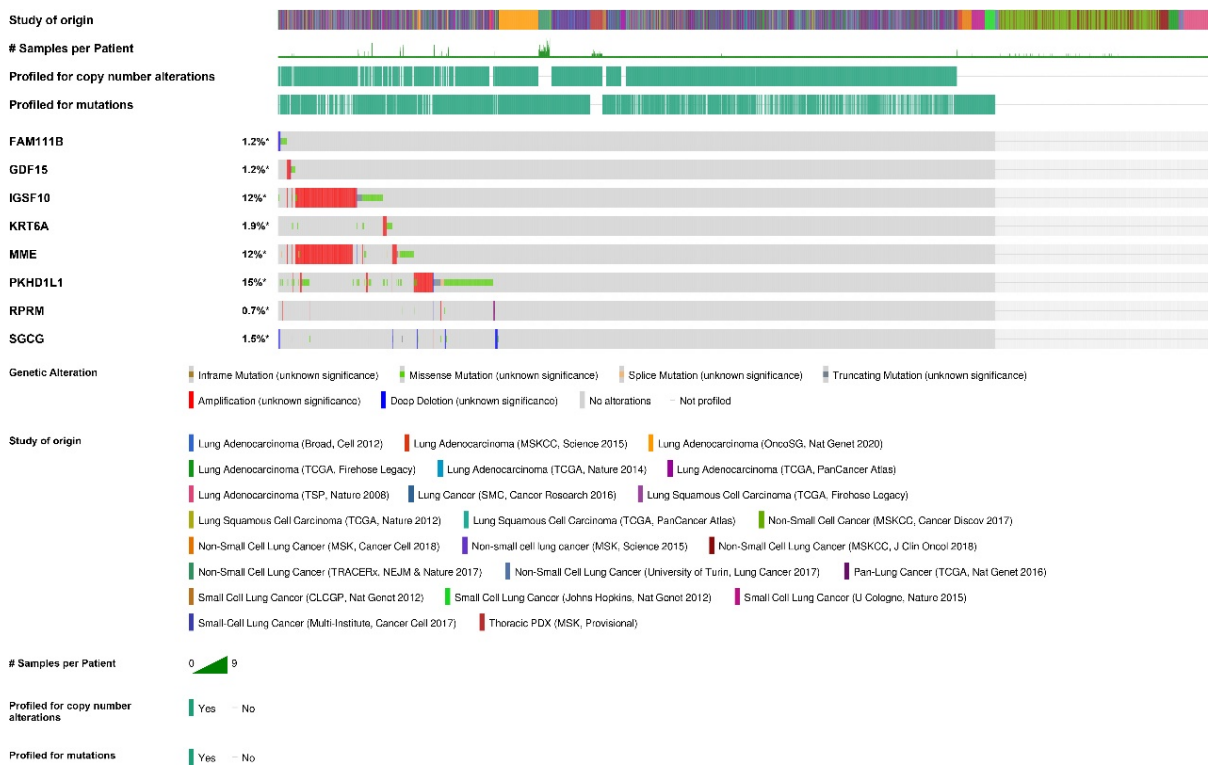






**Figure 3.** Prognostic Survival Analysis of Differential Genes in NSCLC. (a) *GDF15*; (b) *IGSF10*; (c) *MME*; (d) *PKHD1L1*; (e) *SGCG*; (f) *FAM111B*; (g) *KRT6A*; (h) *RPRM*

**图 3.** 差异基因在 NSCLC 中的预后生存分析。(a) *GDF15*; (b) *IGSF10*; (c) *MME*; (d) *PKHD1L1*; (e) *SGCG*; (f) *FAM111B*; (g) *KRT6A*; (h) *RPRM*



**Figure 4.** Visual analysis of the frequency of prognosis-related gene mutations and DNA amplification

**图 4.** 预后相关基因的基因突变及 DNA 扩增频率的可视化分析。

发挥了重要作用。例用 cBioPortal 工具研究这 8 个关键基因的突变或 DNA 扩增频率，其中 *PKHD1L1*、*MME* 和 *IGSF10* 的频率最高，分别为 15%、12%、12%。

多囊肾和肝病 1-L (polycystic kidney and hepatic disease-like1, *PKHDIL-1*) 是一种蛋白质编码基因，它编码的纤维囊蛋白-L 在细胞免疫中发挥重要作用，在活化的 CD4+ T 细胞及 CD8+ T 细胞上调，

被用作 T 淋巴细胞激活信号的启动子[12] [13], 具有信号受体活性的功能。Zheng 等证明 *PKHD1L-1* 在甲状腺癌组织中的表达明显低于与之匹配的正常组织, 可能是与甲状腺癌细胞的相关的抑癌基因, 与肿瘤的增值、集落形成和细胞侵袭相关[14]。与 *PKHD1L-1* 有关的疾病还包括大颗粒 T 淋巴细胞白血病和子宫内膜癌[15] [16] [17]。我们的结果提示此基因可能与 NSCLC 的预后相关, 可能是其潜在治疗靶点。

膜金属肽内切酶(membrane metalloendopeptidase, *MME*)是一种蛋白质编码基因, 该基因编码的蛋白质是 II 型跨膜糖蛋白和急性淋巴细胞性白血病抗原, 也是一种中性内肽酶, 可在疏水残基的氨基侧切割肽, 并使几种肽激素失活, 包括胰高血糖素, 脑啡肽, P 物质, 神经降压素, 催产素和缓激肽。*MME* 具有金属内肽活性、外肽酶活性等功能[18] [19], 同时参与神经肽加工、蛋白质加工及水解等过程[18] [20]。Benjamin 等证明 *MME* 是雄激素调控基因, 存在转移的前列腺癌的原发组织中表达下调[21]。*MME* 通过 *PTEN* 结合抑制前列腺癌细胞生长和转移[22], Osman 等提出 *MME* 的缺失与前列腺癌的复发相关[23]。我们的研究结果提示, *MME* 基因在肺癌中存在错译突变及 DNA 扩增, 可能在未来肺癌的诊断和靶向治疗中发挥重要作用。

免疫球蛋白超家族成员 10 (immunoglobulin superfamily member 10, *IGSF10*)同样是一种蛋白质编码基因, 与细胞分化、多细胞生物发育、神经元迁移的调节等过程相关, 被证实影响大鼠骨肉瘤的发生发展[24]。Ling 等证实敲除 *IGSF10* 可激活整合素 *b1/FAK* 通路, 从而影响肺癌细胞的增值和黏附[25]。Wu 等验证了 *IGSF10* 在乳腺癌组织中显著下调[26], Thutkawkorapin 等提出 *IGSF10* 可能与直肠癌和胃癌的发生发展相关, 但仍需要进一步验证[27]。

本研究从 GEO 网站中下载 NSCLC 肿瘤组织及与之匹配的正常组织的微阵列数据集, 通过绘制热图对差异基因进行可视化分析, 同时对差异基因的功能和通路进行富集。接着, 我们 TCGA 网站的临床数据对差异基因进行生存分析从而筛选出与预后相关的关键基因。最后, 我们通过 cBioPortal 中的数据库对关键基因的突变及 DNA 扩增频率进行统计。本研究也存在一定的一些局限性, 我们无法评价 GEO、TCGA 及 cBioPortal 数据库的数据质量, 同时需要更多的临床研究来证实我们的结果。

## 5. 结论

综上所述, 本研究通过生物信息学分析, 得出 *PKHD1L1*、*MME* 和 *IGSF10* 可能是 NSCLC 发生及影响预后的关键基因。然而, 我们还需要进一步的实验和临床数据来证实我们的研究结果, 指导探索 NSCLC 新的诊断及治疗途径。

## 致 谢

感谢本文各位作者为本文所需的软件提供技术支持。

## 参考文献

- [1] Zarogoulidis, K., Zarogoulidis, P., Darwiche, K., Boutsikou, E. and Machairiotis, N. (2013) Treatment of Non-Small Cell Lung Cancer (NSCLC). *Journal of Thoracic Disease*, **5**, S389-S396.
- [2] Mitsudomi, T., Morita, S., Yatabe, Y., Negoro, S., Okamoto, I., Tsurutani, J., *et al.* (2010) Gefitinib versus Cisplatin plus Docetaxel in Patients with Non-Small-Cell Lung Cancer Harboring Mutations of the Epidermal Growth Factor Receptor (WJTOG3405): An Open Label, Randomised Phase 3 Trial. *The Lancet Oncology*, **11**, 121-128. [https://doi.org/10.1016/S1470-2045\(09\)70364-X](https://doi.org/10.1016/S1470-2045(09)70364-X)
- [3] Hirsch, F.R. and Bunn, P.A. (2009) EGFR Testing in Lung Cancer Is Ready for Prime Time. *The Lancet Oncology*, **10**, 432-433. [https://doi.org/10.1016/S1470-2045\(09\)70110-X](https://doi.org/10.1016/S1470-2045(09)70110-X)
- [4] Soda, M., Choi, Y.L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., *et al.* (2007) Identification of the Trans-



- forming EML4-ALK Fusion Gene in Non-Small-Cell Lung Cancer. *Nature*, **448**, 561-566. <https://doi.org/10.1038/nature05945>
- [5] Miller, K.D., Nogueira, L., Mariotto, A.B., Rowland, J.H. and Siegel, R.L. (2019) Cancer Treatment and Survivorship Statistics, 2019. *CA: A Cancer Journal for Clinicians*, **69**, 363-385. <https://doi.org/10.3322/caac.21565>
- [6] Ding, Y., Yang, D.Z., Zhai, Y.N., Xue, K. and Wang, S.M. (2017) Microarray Expression Profiling of Long Non-Coding RNAs in Epithelial Ovarian Cancer. *Oncology Letters*, **14**, 2523-2530. <https://doi.org/10.3892/ol.2017.6448>
- [7] Smyth, G.K. (2005) Limma: Linear Models for Microarray Data. In: *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, Springer, New York, 397-420. [https://doi.org/10.1007/0-387-29362-0\\_23](https://doi.org/10.1007/0-387-29362-0_23)
- [8] Sun, X. Li, J. (2013) Pairheatmap: Comparing Expression Profiles of Gene Groups in Heatmaps. *Computer Methods and Programs in Biomedicine*, **112**, 599-606. <https://doi.org/10.1016/j.cmpb.2013.07.010>
- [9] Mccarthy, D.J., Yunshun, C. and Smyth, G.K. (2012) Differential Expression Analysis of Multifactor RNA-Seq Experiments with Respect to Biological Variation. *Nucleic Acids Research*, **40**, 4288-4297. <https://doi.org/10.1093/nar/gks042>
- [10] Fox, J. and Carvalho, M.S. (2012) The RcmdrPlugin.survival Package: Extending the R Commander Interface to Survival Analysis. *Journal of Statistical Software*, **49**, 1-32. <https://doi.org/10.18637/jss.v049.i07>
- [11] Tang, Z., et al. (2017) GEPIA: A Web Server for Cancer and Normal Gene Expression Profiling and Interactive Analyses. *Nucleic Acids Research*, **45**, W98-W102. <https://doi.org/10.1093/nar/gkx247>
- [12] Lian, P.W., Fu, Y.L., Li, A., Dai, B.Z. and Wu, G.Q. (2011) Preparation and Characterization of a Polyclonal Antibody against Human Fibrocytin-L. *Chinese Journal of Cellular & Molecular Immunology*, **27**, 78-81.
- [13] Hogan, M.C., Griffin, M.D., Sandro, R., Torres, V.E., Ward, C.J. and Harris, P.C. (2003) PKHD1L1, a Homolog of the Autosomal Recessive Polycystic Kidney Disease Gene, Encodes a Receptor with Inducible T Lymphocyte Expression. *Human Molecular Genetics*, **12**, 685-698. <https://doi.org/10.1093/hmg/ddg068>
- [14] Zheng, C., Quan, R., Xia, E.J., Bhandari, A. and Zhang, X. (2019) Original Tumour Suppressor Gene Polycystic Kidney and Hepatic Disease 1-Like 1 Is Associated with Thyroid Cancer Cell Progression. *Oncology Letters*, **18**, 3227-3235. <https://doi.org/10.3892/ol.2019.10632>
- [15] Makrogkikas, S. (2017) Molecular and Cellular Mechanism of Function of the PKHD1L1 Gene in Vertebrates. *Mechanisms of Development*, **145**, S64. <https://doi.org/10.1016/j.mod.2017.04.140>
- [16] Suzuki, A., Fukushige, S., et al. (1997) Frequent Gains on Chromosome Arms 1q and/or 8q in Human Endometrial Cancer. *Human Genetics*, **100**, 629-636. <https://doi.org/10.1007/s004390050565>
- [17] Izykowska, K., et al. (2014) Submicroscopic Genomic Rearrangements Change Gene Expression in T-Cell Large Granular Lymphocyte Leukemia. *European Journal of Haematology*, **93**, 143-149. <https://doi.org/10.1111/ejh.12318>
- [18] Pascale, G., Livstone, M.S., Lewis, S.E. and Thomas, P.D. (2011) Phylogenetic-Based Propagation of Functional Annotations within the Gene Ontology Consortium. *Briefings in Bioinformatics*, **12**, 449-462. <https://doi.org/10.1093/bib/bbr042>
- [19] Skidgel, R.A., et al. (1984) Hydrolysis of Substance P and Neurotensin by Converting Enzyme and Neutral Endopeptidase. *Peptides*, **5**, 769-776. [https://doi.org/10.1016/0196-9781\(84\)90020-2](https://doi.org/10.1016/0196-9781(84)90020-2)
- [20] Spencer, B., Verma, I., Desplats, P., Morvinski, D., Rockenstein, E., Adame, A., et al. (2014) A Neuroprotective Brain-Penetrating Endopeptidase Fusion Protein Ameliorates Alzheimer Disease Pathology and Restores Neurogenesis. *Journal of Biological Chemistry*, **289**, 17917-17931. <https://doi.org/10.1074/jbc.M114.557439>
- [21] Thomas, B.C., Kay, J.D., Menon, S., Vowler, S.L., Dawson, S.N., Bucklow, L.J., et al. (2016) Whole Blood mRNA in Prostate Cancer Reveals a Four-Gene Androgen Regulated Panel. *Endocrine Related Cancer*, **23**, 797-812. <https://doi.org/10.1530/ERC-16-0287>
- [22] Sumitomo, M., Iwase, A., Zheng, R., Navarro, D. and Nanus, D.M. (2004) Synergy in Tumor Suppression by Direct Interaction of Neutral Endopeptidase with PTEN. *Cancer Cell*, **5**, 67-78. [https://doi.org/10.1016/S1535-6108\(03\)00331-3](https://doi.org/10.1016/S1535-6108(03)00331-3)
- [23] Osman, I. (2004) Neutral Endopeptidase Protein Expression and Prognosis in Localized Prostate Cancer. *Clinical Cancer Research*, **10**, 4096-4100. <https://doi.org/10.1158/1078-0432.CCR-04-0120>
- [24] Daino, K., Ugolin, N., Altmeyer-Morel, S., Guilly, M.-N. and Chevillard, S. (2009) Gene Expression Profiling of Alpha-Radiation-Induced Rat Osteosarcomas: Identification of Dysregulated Genes Involved in Radiation-Induced Tumorigenesis of Bone. *International Journal of Cancer*, **125**, 612-620. <https://doi.org/10.1002/ijc.24392>
- [25] Ling, B., Liao, X., Huang, Y., Liang, L., Jiang, Y., Pang, Y., et al. (2020) Identification of Prognostic Markers of Lung Cancer through Bioinformatics Analysis and *in Vitro* Experiments. *International Journal of Oncology*, **56**, 193-205.

- [26] Wu, M., Li, Q. and Wang, H. (2021) Identification of Novel Biomarkers Associated with the Prognosis and Potential Pathogenesis of Breast Cancer via Integrated Bioinformatics Analysis. *Technology in Cancer Research & Treatment*, **20**, 1533033821992081. <https://doi.org/10.1177/1533033821992081>
- [27] Thutkawkorapin, J., Picelli, S., Kontham, V., Liu, T., Nilsson, D. and Lindblom, A. (2016) Exome Sequencing in One Family with Gastric- and Rectal Cancer. *BMC Genetics*, **17**, 41. <https://doi.org/10.1186/s12863-016-0351-z>