

DOI:10.16781/j.0258-879x.2021.06.0585

· 论 著 ·

隆纳霉素引起三阴性乳腺癌细胞周期阻滞基因表达谱及关键通路的生物信息学分析

薛佳兴[△], 孔杰[△], 卢小玲, 孙筱, 于豪冰, 梁珍珍, 徐尧, 邓博文, 焦炳华^{*}
海军军医大学(第二军医大学)基础医学院生物化学与分子生物学教研室, 上海 200433

[摘要] **目的** 利用生物信息学方法分析隆纳霉素对人三阴性乳腺癌细胞系细胞周期阻滞的作用机制。**方法** 利用隆纳霉素处理人三阴性乳腺癌细胞系 MDA-MB-468 细胞 48、72 h, 计算 IC₅₀, 采用流式细胞术分析细胞周期变化。选取 0.8 μmol/L 隆纳霉素处理 48 h 的 MDA-MB-468 细胞和未加药处理的对照细胞进行转录组测序, 测序数据经质量控制过滤后, 使用 DESeq2 1.16.1 软件进行差异表达基因分析, 并进行基因本体和京都基因与基因组百科全书功能分析及基因集富集分析, 用 STRING 工具进行蛋白质相互作用网络分析及用 AutoDock 1.5.6 软件进行分子对接预测。**结果** 隆纳霉素干预 MDA-MB-468 细胞 48、72 h 的 IC₅₀ 分别为 2.733 和 0.866 μmol/L, 且流式细胞术分析显示隆纳霉素干预 48 h 有明显 G₂/M 阻滞作用。隆纳霉素干预后共筛选出 1 764 个差异表达基因, 主要定位于 NF-κB 介导的 TNF-α 通路和 P53 通路, G 蛋白 γ 亚基 7 (GNG7)、G 蛋白 γ 亚基 11 (GNG11)、C-X-C 基序趋化因子配体 8 (CXCL8)、腺苷酸环化酶 2 (ADCY2) 等基因可能是隆纳霉素作用基因网络中的关键节点。隆纳霉素与 DNA 分子、拓扑异构酶、基因表达调节蛋白 [20S 蛋白酶体 β 亚单位 5 (PSMB5) 和含有 SET 结构域 7 的组蛋白赖氨酸甲基转移酶 (SETD7)] 有较好的结合及相互作用。**结论** 隆纳霉素具有与蝴蝶霉素类似的功能, 其通过与 DNA- 拓扑异构酶 I 复合物中的 DNA 分子结合发挥抗肿瘤活性; 还可通过与基因表达调节蛋白 PSMB5 和 SETD7 结合影响 GNG7、GNG11、CXCL8、ADCY2 等基因的表达水平, 从而上调 NF-κB 介导的 TNF-α 通路和 P53 通路, 最终导致乳腺癌细胞发生 G₂/M 阻滞, 发挥抗肿瘤作用。

[关键词] 隆纳霉素; 蝴蝶霉素类似物; 三阴性乳腺癌; 细胞周期阻滞

[中图分类号] R 737.9; R 979.1

[文献标志码] A

[文章编号] 0258-879X(2021)06-0585-11

Bioinformatics analysis of gene expression profile and key pathways of loongmycin-induced cell cycle arrest in triple-negative breast cancer

XUE Jia-xing[△], KONG Jie[△], LU Xiao-ling, SUN Xiao, YU Hao-bing, LIANG Zhen-zhen, XU Yao, DENG Bo-wen, JIAO Bing-hua^{*}
Department of Biochemistry and Molecular Biology, College of Basic Medical Sciences, Naval Medical University (Second Military Medical University), Shanghai 200433, China

[Abstract] **Objective** To analyze the mechanism of loongmycin-induced cell cycle arrest in triple-negative breast cancer cell line by bioinformatics. **Methods** The triple-negative breast cancer cell line, MDA-MB-468, was treated with loongmycin for 48 and 72 h. The half inhibition concentration (IC₅₀) was calculated and the cell cycle was analyzed by flow cytometry. The MDA-MB-468 cells treated with 0.8 μmol/L loongmycin for 48 h and the control cells without drug treatment were selected for transcriptome sequencing. After the data were filtered by quality control, the differential gene expression was analyzed using DESeq2 1.16.1 software. Gene Ontology and Kyoto Encyclopedia of Genes and Genomes functional analyses and gene set enrichment analysis were performed. The protein-protein interaction network was analyzed by STRING and the docking prediction was performed using AutoDock 1.5.6 software. **Results** The IC₅₀ values of MDA-MB-468 cells intervened by loongmycin for 48 and 72 h were 2.733 and 0.866 μmol/L, respectively. The results of flow cytometry showed that loongmycin arrested MDA-MB-468 cell line in G₂/M phase after intervening for 48 h. A total of 1 764 differentially expressed genes were screened after intervention with loongmycin. The differentially expressed genes were mainly located in nuclear factor kappa B (NF-κB)-mediated tumor necrosis factor alpha (TNF-α) pathway and P53 pathway. G protein subunit gamma 7 (GNG7), G protein subunit gamma 11 (GNG11), C-X-C motif chemokine ligand 8 (CXCL8), adenylate cyclase

[收稿日期] 2021-02-26

[接受日期] 2021-04-30

[基金项目] 国家重点研发计划(2019YFC0312504). Supported by National Key Research and Development Program of China (2019YFC0312504).

[作者简介] 薛佳兴, 硕士生. E-mail: jx.xue@foxmail.com; 孔杰, 硕士, 助教. E-mail: kongjie1953@163.com

[△]共同第一作者 (Co-first authors).

^{*}通信作者 (Corresponding author). Tel: 021-81870965, E-mail: biojiao@163.com

2 (*ADCY2*) and other genes may be the key nodes in the gene network of drug action. Loongmycin and DNA molecules, topoisomerase, gene expression regulator (proteasome 20S subunit beta 5 [*PSMB5*] and SET domain containing histone lysine methyltransferase 7 [*SETD7*]) had good combination and interaction functions. **Conclusion** Loongmycin has similar functions to rebeccamycin; it exerts anti-tumor activity by binding with DNA molecules in DNA-topoisomerase I complex; and it can also affect the expression of *GNG7*, *GNG11*, *CXCL8*, *ADCY2* and other genes by binding gene expression regulatory proteins *PSMB5* and *SETD7*, and eventually up-regulate NF- κ B-mediated TNF- α pathway and P53 pathway, leading to G₂/M arrest in breast cancer cells, which plays an anti-tumor role.

[**Key words**] loongmycin; rebeccamycin analogs; triple-negative breast neoplasms; cell cycle arrest

[Acad J Sec Mil Med Univ, 2021, 42(6): 585-595]

海洋覆盖着地球表面积的 71%，海洋微生物所处的特殊环境导致了其物种的多样性，这种物种的多样性构成了抗肿瘤天然药物化学多样性的基础^[1]。随着提取分离技术和药物筛选手段的不断进步和发展，越来越多来自海洋微生物的代谢产物被发现具有体内外抗肿瘤活性。目前已经上市的海军药物中有 4 种具有抗肿瘤作用，分别为甲磺酸艾立布林 (eribulin mesylate, E7389; 商品名: 海乐卫)、 ω -3- 酸乙酯 (omega-3-acid ethyl esters)、曲贝替定 (trabectedin, ET-743) 和本妥昔单抗 (brentuximab vedotin, SGN-35)。海洋微生物来源的天然产物及其结构类似物因其特异的化学结构和良好的抗肿瘤活性成为抗肿瘤药物开发的重要来源^[2]，具有广阔的应用前景。

乳腺癌是女性发病率最高、最常见的恶性肿瘤。WHO 国际癌症研究机构发布了 2018 年全球癌症负担数据，全球女性有超过 208 万乳腺癌新发病例，居女性恶性肿瘤发病率的首位；死于乳腺癌的女性患者约有 62 万例，也位于女性恶性肿瘤病死率的首位^[3]。乳腺癌的分子分型有 Luminal A 型、Luminal B 型、人表皮生长因子受体 2 (human epidermal growth factor receptor 2, HER2) 过表达型、基底样型/三阴性 (雌激素受体、孕激素受体

和 HER2 均为阴性) 及特殊类型，其中三阴性乳腺癌占比约为 12.7%^[4]。三阴性乳腺癌的治疗进展一直落后于其他分子分型，迄今为止主要局限于化学治疗^[5]。虽然目前有几种免疫疗法和靶向疗法正在研究中，但由于三阴性乳腺癌包含一系列异质性的分子结构，这些疗法只对部分患者有效^[6]。所以积极寻找三阴性乳腺癌新的作用靶点及新的化学治疗药物仍然十分必要。

隆纳霉素 (loongmycin) 是从海洋放线菌 NA01583 中分离得到的新的次级代谢产物，属于蝴蝶霉素 (图 1) 类似物，在前期的研究中，隆纳霉素对人乳腺癌细胞 MDA-MB-468、MDA-MB-231 表现出良好的细胞毒性。蝴蝶霉素是一种吡啶啉类的新型天然抗生素，其类似物能够结合在 DNA 与拓扑异构酶 I 形成的复合物上抑制 DNA 链切口处的重连接^[7]，引起 DNA 单链和双链断裂，从而产生较强的抗肿瘤活性。但蝴蝶霉素水溶性较差，影响其进一步的体内药理学活性评价，目前仅有巴特卡令 (becatecarin)^[8]、edotecacin^[9-10] 和 NSC#655649^[11] 3 种蝴蝶霉素类似物药物进入临床研究。隆纳霉素较蝴蝶霉素多了一个稀有糖基，增加了其水溶性，具有更好的成药潜力。

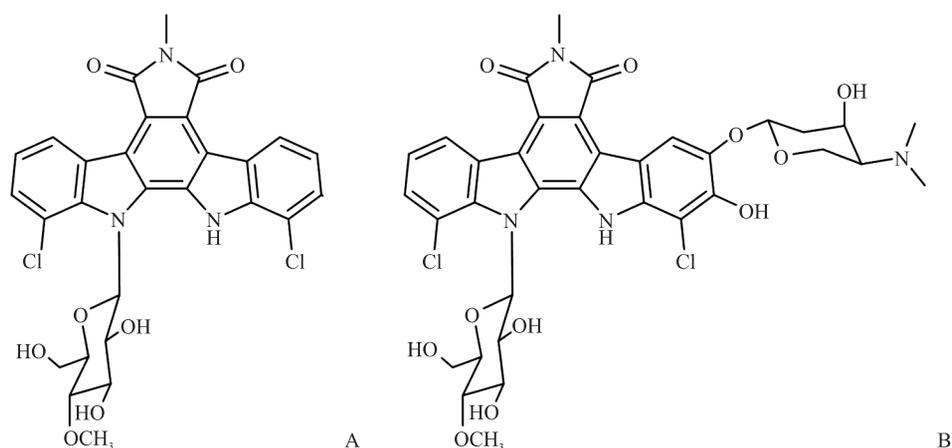


图 1 蝴蝶霉素 (A) 与隆纳霉素 (B) 的化学结构

Fig 1 Structures of rebeccamycin (A) and loongmycin (B)

转录组测序技术 (RNA-seq) 能够反映生物体在某一特定时空条件下基因的表达情况。通过分析测序数据的差异表达基因、功能富集等定位肿瘤发生和发展、药物处理等情况下的细胞生物学通路, 或许能筛选出具有生物学意义的靶基因。本研究利用转录组测序技术初步筛选在隆纳霉素抑制三阴性乳腺癌细胞系增殖过程中起关键作用的通路和基因, 为进一步研究其治疗靶作用提供参考。

1 材料和方法

1.1 细胞、试剂与仪器

1.1.1 细胞系 人三阴性乳腺癌细胞系 MDA-MB-468 (货号 TCHu136) 购至上海生物化学与细胞生物学细胞库。

1.1.2 药品和试剂 隆纳霉素由南京大学戈惠明课题组提供。链霉素/青霉素混合液、0.25% 胰酶+0.2% EDTA 溶液、FBS 均购自美国 Gibco 公司。RPMI-1640 培养基购自美国 HyClone 公司。膜联蛋白 (annexin) V -FITC/PI 凋亡检测试剂盒、DMSO 均购自美国 Origen 公司。BCA 蛋白定量试剂盒购自英国 Abcam 公司。甲醇溶液 (分析纯) 购自中国国药集团。PBS、十二烷基硫酸钠上样缓冲液购自生工生物工程 (上海) 股份有限公司。CCK-8 试剂 [2-(2-甲氧基-4-硝基苯基)-3-(4-硝基苯基)-5-(2,4-二磺酸苯)-2H-四唑单钠盐] 购自美国 Sigma 公司。

转录组测序工作由上海吉凯基因医学科技股份有限公司完成, 其中 M-MuLV 反转录酶反应体系、核糖核酸酶 H、DNA 聚合酶 I 均购自生工生物工程 (上海) 股份有限公司, AMPure XP 磁珠购自美国 Beckman 公司, 测序文库构建使用试剂盒为美国 Illumina 公司的第二代测序 RNA 文库制备试剂盒 (NEBNext® Ultra™ RNA Library Prep Kit for Illumina®)。

1.1.3 主要仪器及耗材 超净工作台 (深圳市九鼎净化设备有限公司), CKX41 型倒置显微镜 (日本 Olympus 公司), HERACELL150i 型细胞培养箱 (美国 Thermo 公司), 5417R 型小型台式冷冻离心机 (德国 Eppendorf 公司), Invitrogen Countess 型细胞计数仪 (赛默飞世尔科技中国有限公司), FACSCanto II 型流式细胞仪 (美国 BD 公司), 超高灵敏度化学发光成像系统 (美国 Bio-

Rad 公司), HiSeq X ten PE150 NovaSeq 6000 型第二代测序分析仪 (美国 Illumina 公司)。

75 cm² 一次性细胞培养瓶、25 cm² 一次性细胞培养瓶、96 孔板、6 孔板均购自美国 ThermoFisher 公司。各规格离心管购自美国 Corning 公司。细胞计数板购自上海拜力生物科技有限公司。

1.2 细胞培养 MDA-MB-468 细胞用含 10% FBS、100 U/mL 青霉素及 100 mg/L 链霉素的 RPMI-1640 培养基, 在 37 °C、5% CO₂、饱和湿度培养箱中培养至贴壁。

1.3 CCK-8 法检测隆纳霉素对细胞生长曲线的影响 取贴壁 MDA-MB-468 细胞, 用含 EDTA 的胰酶消化后, 培养基洗涤稀释至 5×10⁵/mL, 每孔 100 μL 接种于 2 块 96 孔板内, 分别标记为 48 和 72 h 处理组。置于饱和湿度细胞培养箱 (37 °C、5% CO₂) 培养过夜。将培养基换为加药培养基 100 μL, 根据药物终浓度设阴性对照组 (含体积分数为 0.06% DMSO 的 PBS) 及不同浓度 (0.062 5、0.125、0.25、0.5、1、2、4 μmol/L) 的隆纳霉素处理组, 每组设置 3 个复孔。分别培养 48 和 72 h 后加入 10 μL CCK-8 试剂, 反应 1 h 后置于摇床轻度震荡 5 min, 然后将 96 孔板置于酶标仪中, 测定 450 nm 波长处光密度值 (D_{450}), 计算各组肿瘤细胞的存活率: 细胞存活率 (%) = (隆纳霉素处理组 D_{450} - 空白对照组 D_{450}) / (阴性对照组 D_{450} - 空白对照组 D_{450}) × 100%, 其中空白对照组为加入 CCK-8 试剂、不含肿瘤细胞株和隆纳霉素的培养基。计算各肿瘤细胞株的 IC₅₀, 实验重复 3 次。此处以 IC₅₀ 作为药物作用效果的评价指标, 其含义为在该处理时长下细胞增殖抑制率达到 50% 时所用药物的浓度。

1.4 流式细胞术检测隆纳霉素对细胞周期的影响 应用流式细胞仪, 采用膜联蛋白 V -FITC/PI 双标法检测隆纳霉素对 MDA-MB-468 细胞周期的影响。取培养 24 h 的 MDA-MB-468 细胞接种至 6 孔板, 每孔 1×10⁶ 个细胞。待细胞贴壁后弃上清液, 实验组加入含浓度为 0.4 和 0.8 μmol/L 隆纳霉素的培养基, 孵育 48 h 后收集 5×10⁵ 个细胞。然后加入 500 μL 的结合缓冲液充分混匀, 重悬细胞后加入 2.5 μL 膜联蛋白 V -FITC 溶液充分混匀, 再加入 2.5 μL PI 混匀。放置于室温、避光环境中反应 15 min, 立即用流式细胞仪检测细胞周期。应用

SPSS 22.0 软件,采用单因素方差分析进行数据比较,检验水准(α)为0.05。

1.5 送测序样本制备 将处于对数生长期的细胞用0.25%的胰酶消化,接种至6孔板。每孔接种约 1×10^5 个细胞。培养12 h后换成含浓度为0.8 $\mu\text{mol/L}$ 隆纳霉素的培养基,继续培养48 h后,与不加药对照组各3个复孔的细胞,经TRIzol试剂破碎和裂解后,交由上海吉凯基因医学科技股份有限公司进行RNA提取、质量控制、建立测序文库和转录组测序。

1.6 RNA的提取、检测、测序文库构建及转录组测序 将细胞在液氮中研磨成粉末,使用TRIzol试剂提取总RNA,并用DNA酶I去除基因组DNA。使用2100型生物分析仪(安捷伦科技有限公司上海分公司)检测RNA完整性。

利用第二代测序RNA文库制备试剂盒,按照说明书方法构建测序文库。以mRNA片段化产物为模板,添加随机寡核苷酸引物,在M-MuLV反转录酶反应体系中得到cDNA第1条链。用核糖核酸酶H降解杂合双链中的RNA链,并在DNA聚合酶I体系下,以dNTP为原料合成cDNA第2条链。纯化后的双链cDNA经过末端修复、加A尾并连接测序接头,用AMPure XP磁珠筛选长度为250~300 bp的cDNA,以此为模板进行PCR扩增。再次使用AMPure XP磁珠纯化PCR产物获得测序文库。

对获得的测序文库先使用Qubit 2.0 荧光计进行初步定量,根据定量结果,将测序文库稀释至1.5 ng/ μL ,然后使用2100型生物分析仪检测文库的有效测序片段长度,符合250~300 bp,正态分布标准化后,用qRT-PCR定量文库的有效浓度,确保文库有效浓度高于2 nmol/L。文库检验合格后,上机测序。

1.7 质量控制、数据组装及基因功能注释 对原始数据进行过滤,主要包括去除测序片段的接头序列、去除含无法确定的碱基信息的测序片段、去除低质量测序片段。Illumina®平台测序结果的碱基质量值用Qphred表示,其计算公式为 $Q_{\text{phred}} = -10 \lg e$,其中 e 表示碱基测序错误率。将 $Q_{\text{phred}} \leq 20$ 碱基数占全长50%的片段判为低质量,同时统计Q20所占比例。之后,对过滤后数据进行Q30和GC含量计算,进一步确认测序数据质量。

后续所有分析均基于此处过滤得到的数据进行。

应用HISAT2 2.0.5软件将配对末端测序片段与参照基因组(Species: Humangenus GRCh38.p12)比对。采用StringTie 1.3.3b软件进行新基因预测。使用featureCounts 1.5.0-p3软件计算映射到每个基因的读数,该部分处理均采用软件推荐的默认参数。然后根据基因长度计算每个基因的每千个碱基的转录每百万映射读取的片段数(fragments per kilobase of exon model per million mapped fragments, FPKM),基因表达的阈值定为1,计算映射到该基因的读数,得到表达值矩阵。

1.8 差异表达基因分析与功能富集 使用DESeq2 1.16.1软件^[12]进行组间差异表达基因分析(每组设3个生物学重复)。标准化方法为DESeq标准算法,使用基于负二项式分布的 P 值计算模型确定基因表达数据中的差异表达,筛选标准定为错误发现率(false discovery rate, FDR) ≤ 0.05 且 $\log_2|FC| \geq 1$ 或 ≤ -1 ,其中FC为基因表达变化的差异倍数(fold change)。使用Benjamini-Hochberg方法校正 P 值以控制FDR。通过clusterProfiler 3.4.4软件实现差异表达基因的基因本体(Gene Ontology, GO)分析^[13-14]和京都基因与基因组百科全书(Kyoto Encyclopedia of Genes and Genomes, KEGG)^[15]通路富集分析,其中修正了基因长度偏差。以上分析基于R 3.6.0软件实现。考虑校正的 P 值 < 0.05 的通路差异表达基因显著富集,基于基因集做排列检验(permutation test)并进行基因集富集分析(gene set enrichment analysis, GSEA)^[16]。

1.9 蛋白质-蛋白质相互作用(protein-protein interacton, PPI)网络构建及枢纽基因(hub-gene)筛选 将差异表达基因输入到STRING在线工具(<https://string-db.org/>)^[17]进行PPI分析,将节点信息导入Cytoscape 3.7.2软件,采用其中的分子复合物检测(molecular complex detection, MCODE)扩展包2.0.0^[17]进行枢纽基因分析。

1.10 多数据库交叉比较 取表达差异倍数较大的前400个基因,通过查找MalaCards(<https://www.malacards.org>)、GeneCards(<https://www.genecards.org>)、癌症基因组图谱(The Cancer Genome Atlas, TCGA; <https://portal.gdc.cancer.gov>)数据库,鉴定是否已有文献报道这些基因与乳腺癌相关。

1.11 药物-靶蛋白对接分析 使用PyMOL 2.4.1软件^[18]和AutoDock 1.5.6软件^[19]进行分子对接预测。使用ChemDraw 19软件绘制并导出隆纳霉素分子数据,从ZINC数据库(<http://zinc.docking.org>)下载蝴蝶霉素分子数据,登记号为ZINC28520217。从RCSB蛋白数据库(<https://www.rcsb.org>)获取20S蛋白酶体 β 亚单位5(proteasome 20S subunit beta 5, PSMB5; DOI: 10.2210/pdb6RGQ/pdb)和含有SET结构域7的组蛋白赖氨酸甲基转移酶(SET domain containing 7 histone lysine methyltransferase, SETD7; DOI: 10.2210/pdb4JLG/pdb)蛋白及DNA分子结构(DOI: 10.2210/pdb1SC7/pdb)。

2 结果

2.1 隆纳霉素对细胞活性及细胞周期的影响 隆纳霉素对人三阴性乳腺癌细胞株MDA-MB-468细胞作用48、72 h,均表现出较高的细胞毒性,且具有浓度依赖性(图2A)。隆纳霉素处理48、72 h后,拟计算该细胞系的 IC_{50} ,分别为2.733和0.866 $\mu\text{mol/L}$ 。考虑到误差及药物代谢情况,选取用药时长48 h进行实验,设置0.4和0.8 $\mu\text{mol/L}$ 2个浓度进行流式细胞术检测,结果显示隆纳霉素对MDA-MB-468细胞有周期阻滞作用,且在0.8 $\mu\text{mol/L}$ 浓度下与对照组相比差异有统计学意义($P<0.05$,图2B),故选取0.8 $\mu\text{mol/L}$ 隆纳霉素处理48 h的细胞进行转录组测序。

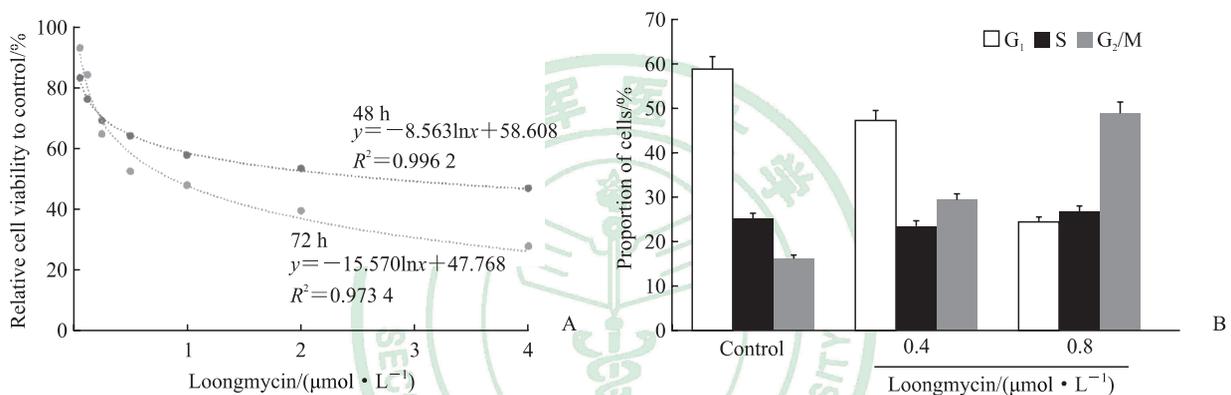


图2 隆纳霉素对人三阴性乳腺癌MDA-MB-468细胞活性及细胞周期的影响

Fig 2 Effect of loongmycin on activity and cell cycle of human triple-negative breast cancer cell line MDA-MB-468 cells
 A: Cell counting kit 8 assay results revealed that the cytotoxic activity showed a concentration dependent manner after 48 and 72 h of loongmycin treatment; B: Flow cytometry results revealed that loongmycin arrested MDA-MB-468 cells in G₂/M phase, with significant difference at 0.8 $\mu\text{mol/L}$ ($P<0.05$). $n=3$, $\bar{x} \pm s$.

2.2 差异表达基因筛选 按照 $FDR \leq 0.05$ 和 $\log_2|FC| \geq 1$ 或 ≤ -1 的标准共筛选出1764个差异表达基因(图3A),与对照组相比,隆纳霉素处理组上调基因737个、下调基因1027个。差异结果热图及聚类(图3B)显示,样本分组聚类符合实验设计,组间差异大于组内差异,其中表达上调的前10个基因是IL-6、海马钙蛋白样蛋白4(hippocalcin like 4, HPCAL4)、角蛋白相关蛋白3-1(keratin-associated protein 3-1, KRTAP3-1)、胃泌素(gastrin, GAST)、嗅觉受体家族2亚家族H成员2(olfactory receptor family 2 subfamily H member 2, OR2H2)、新转录本AC117394.2和AC002480.1、上池蛋白(suprabasin, SBSN)、新转录本AL133444.1、胶原VI型 $\alpha 2$ 链蛋白(collagen type VI alpha 2 chain,

COL6A2),表达下调的前10个基因是肌聚糖蛋白 δ (sarcoglycan delta, SGCD)、新转录本AC027312.1、甘氨酸受体 $\alpha 3$ (glycine receptor alpha 3, GLRA3)、Nanog同源异形假基因7(Nanog homeobox pseudogene 7, NANOGP7)、钙黏着蛋白17(cadherin 17, CDH17)、长链非编码RNA 1630(long intergenic non-protein coding RNA 1630, LINC01630)及新转录本AC095055.1、AL606534.1、FP236383.2、AP001381.1。

2.3 差异表达基因功能分析 GO分析(图4)提示,表达下调的基因在突触后膜明显富集,主要参与神经突触信号转导,在分子功能方面主要与肌动蛋白结合及离子门控通道有关;表达上调的基因主要在细胞外基质相关条目富集,多参与信号受体活

性调节、内皮细胞增殖调控、细胞运动等生物学过程,在分子功能方面主要富集于受体-配体活性调节相关条目。KEGG 分析结果(图5)显示,下调表达的基因主要集中于神经相关功能通路及 Ras 相

关蛋白1(Ras-related protein 1, RAP1)信号通路,上调表达的基因主要集中于细胞因子-受体相互作用过程。

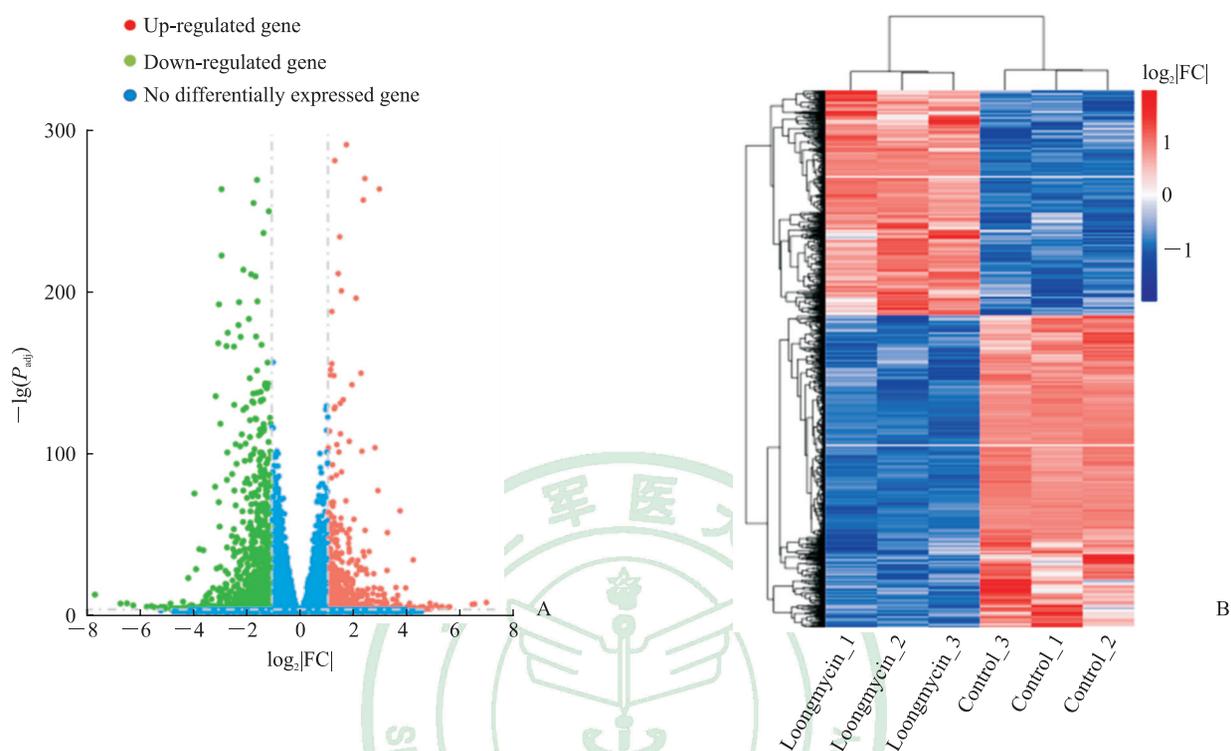


图3 隆纳霉素处理组与对照组人三阴性乳腺癌 MDA-MB-468 细胞差异表达基因筛选

Fig 3 Differentially expressed genes of human triple-negative breast cancer cell line MDA-MB-468 cells in loongmycin group compared with control group

A: Compared with the control group, there were 737 up-regulated genes and 1 027 down-regulated genes in the loongmycin group (the vertical dashed lines indicate $\log_2|FC| = -1$ or 1 , and horizontal dashed line indicates $P_{adj} = 0.05$); B: Heat map and clustering of differentially expressed genes showed that the sample clustering conformed to the experimental design, and the differences between groups are greater than the differences within the groups (_1, _2 and _3 represent different samples). P_{adj} : P value adjusted for multiple testing using Benjamini-Hochberg to estimate the false discovery rate; FC: Fold change.

2.4 GSEA 结果 以标志基因集为探索集的 GSEA 结果(图 6A、6B)显示,隆纳霉素处理组 NF- κ B 介导的 TNF- α 信号通路及 P53 信号通路存在上调。以调控靶基因集为探索集的 GSEA 结果(图 6C、6D)显示,隆纳霉素处理组基因的表达差异可能主要受 PSMB5 和 SETD7 调控。

2.5 PPI 网络及枢纽基因获取 通过 MCODE 筛选出 MCODE 得分排前 3 位的 3 个枢纽模块,提取各模块得分排前 5 位的基因,分别为 G 蛋白 γ 亚基(G protein subunit gamma, *GNG*) 11、*GNG7*、C-X-C 基序趋化因子配体(C-X-C motif chemokine ligand, *CXCL*) 8、腺苷酸环化酶 2(adenylate cyclase 2, *ADCY2*)、

CXCL1、骨形态发生蛋白 4(bone morphogenetic protein 4, *BMP4*)、载脂蛋白 E(apolipoprotein E, *APOE*)、富半胱氨酸血管生成诱导因子 61(cysteine rich angiogenic inducer 61, *CYR61*)、潜在转化生长因子 β 结合蛋白 1(latent transforming growth factor beta binding protein 1, *LTBP1*)、纤维蛋白原 α 链(fibrinogen alpha chain, *FGA*),以及核苷酸外焦磷酸酶/磷酸二酯酶(ectonucleotide pyrophosphatase/phosphodiesterase, *ENPP*) 3、*ENPP1*、腺苷酸激酶 9(adenylate kinase 9, *AK9*)、磷酸二酯酶(phosphodiesterase, *PDE*) 1C、*PDE1A*。

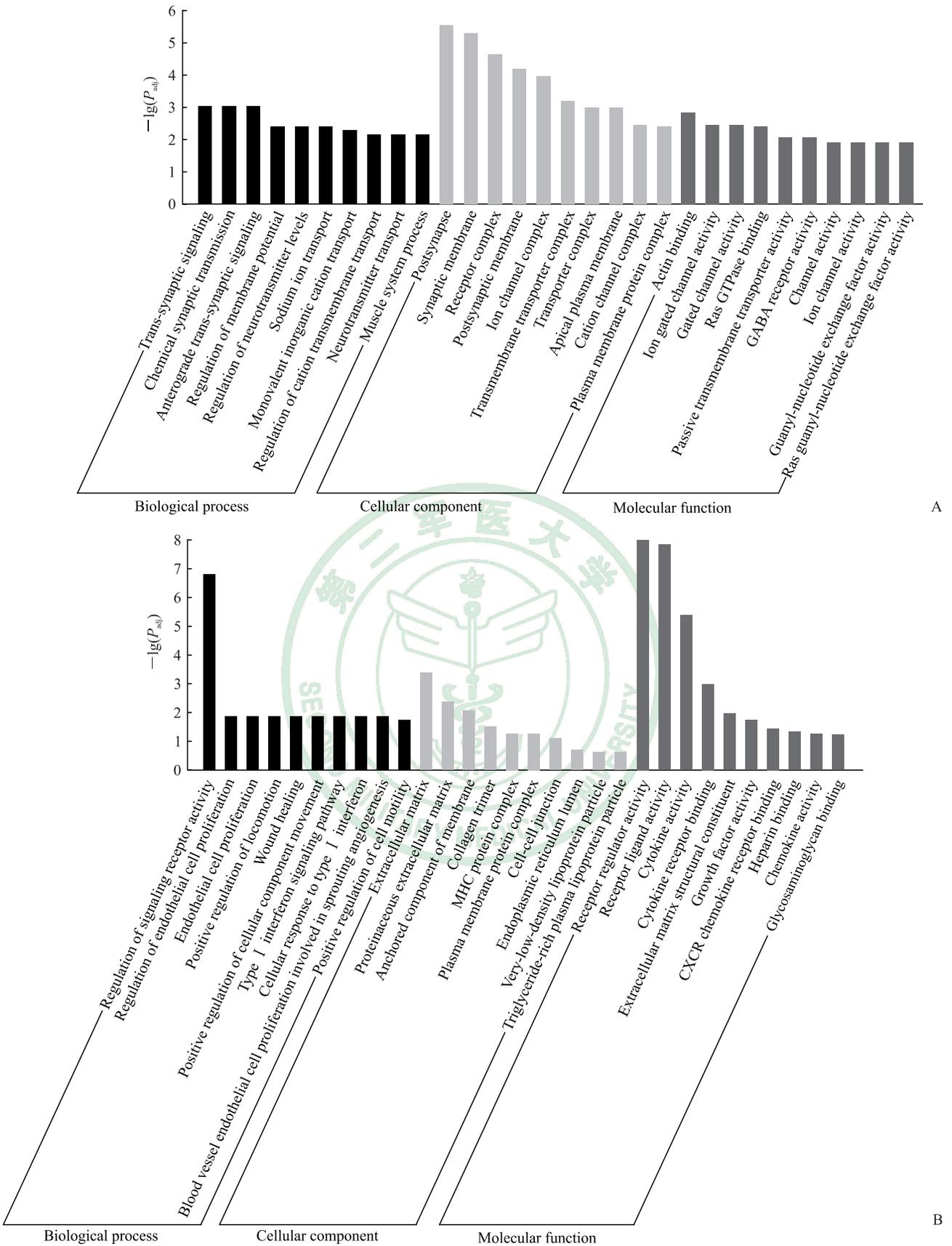


图 4 差异表达基因的基因本体功能分析

Fig 4 Gene Ontology analysis of differentially expressed genes

A: Down-regulated genes; B: Up-regulated genes. P_{adj} : P value adjusted for multiple testing using Benjamini-Hochberg to estimate the false discovery rate; GTPase: Guanosine triphosphatase; GABA: Gamma-aminobutyric acid; MHC: Major histocompatibility complex; CXCR: C-X-C motif chemokine receptor.

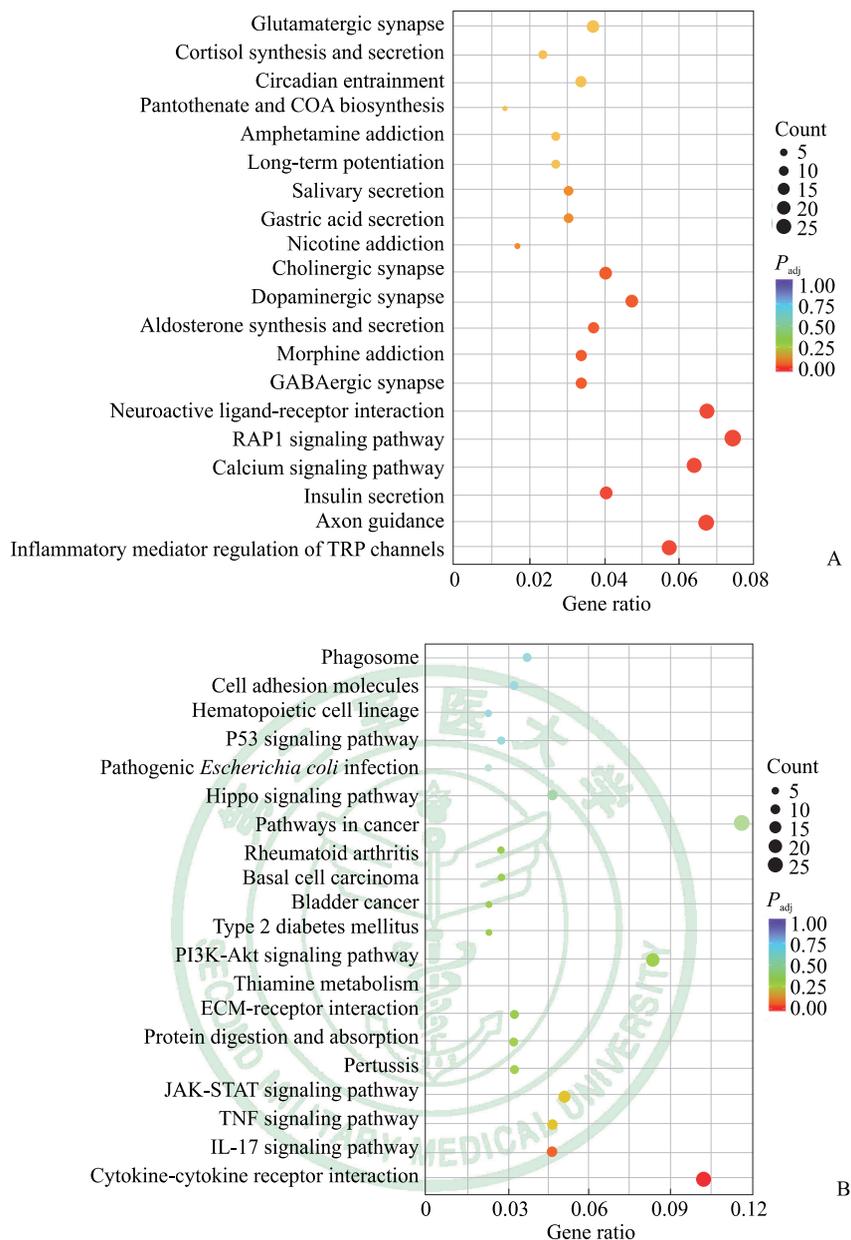


图5 差异表达基因 KEGG 富集分析

Fig 5 KEGG enrichment analysis of differentially expressed genes

A: Down-regulated genes; B: Up-regulated genes. KEGG: Kyoto Encyclopedia of Genes and Genomes; P_{adj} : P value adjusted for multiple testing using Benjamini-Hochberg to estimate the false discovery rate; COA: Coenzyme A; GABA: Gamma-aminobutyric acid; RAP1: Ras-related protein 1; TRP: Transient receptor potential; PI3K: Phosphatidylinositol 3-kinase; Akt: Protein kinase B; ECM: Extracellular matrix; JAK: Janus kinase; STAT: Signal transducers and activators of transcription; TNF: Tumor necrosis factor; IL-17: Interleukin 17.

2.6 基于多数据库的乳腺癌相关基因筛选 通过查阅文献并参考人类疾病数据库 MalaCards, 人工筛选出 27 个乳腺癌相关差异表达基因, 分别为神经元五聚蛋白 1 (neuronal pentraxin 1, *NPTX1*)、小核仁 RNA C/D 盒蛋白 104 (small nucleolar RNA C/D box 104, *SNORD104*)、LY6/PLAUR 结构域蛋白 1 (LY6/PLAUR domain containing 1, *LYPD1*)、*CXCL5*、Ras 相关地塞米松诱导蛋白 1 (Ras related dexamethasone induced 1, *RASDI*)、*SBSN*、维甲

酸代谢早期转录本 1L (retinoic acid early transcript 1L, *RAETIL*)、亨廷顿蛋白相关蛋白 1 (huntingtin-associated protein 1, *HAPI*)、肌动蛋白集束蛋白 1 (fascin actin-bundling protein 1, *FSCN1*)、生殖细胞碱性磷酸酶 (alkaline phosphatase germ cell, *ALPG*)、神经生长因子受体 (nerve growth factor receptor, *NGFR*)、组织因子途径抑制蛋白 2 (tissue factor pathway inhibitor 2, *TFPI2*)、G 蛋白调节诱导神经突增生蛋白 3 (GPRIN family member 3,

GPRIN3)、清道夫受体 A5 (scavenger receptor class A member 5, *SCARA5*)、肝细胞黏附分子 2 (HEPACAM family member 2, *HEPACAM2*)、奇跳相关转录因子 1 (odd-skipped related transcription factor 1, *OSR1*)、雌激素受体 γ (estrogen-related receptor gamma, *ESRRG*)、核受体亚家族 2F1 反义 RNA1 (NR2F1 antisense RNA 1, *NR2F1-AS1*)、*SGCD*、kit 原癌基因受体酪氨酸蛋白激酶 (kit proto-oncogene receptor tyrosine kinase, *KIT*)、脆性组氨酸三联体 (fragile histidine triad, *FHIT*)、

γ -氨基丁酸 A 受体 π 亚基 (gamma-aminobutyric acid type A receptor pi subunit, *GABRP*)、*ENPP3*、G 蛋白偶联受体 55 (G protein-coupled receptor 55, *GPRP55*)、*PDE1C*、含三联基序 63 蛋白 (tripartite motif containing 63, *TRIM63*)、三磷酸肌醇相互作用结构域 1 蛋白 (ITPR interacting domain containing 1, *ITPRID1*)。在隆纳霉素处理组 *HAP1* ($\log_2|FC|=3.560$) 及 *FSCN1* ($\log_2|FC|=4.029$) 表达显著增加, *KIT* ($\log_2|FC|=-3.573$) 与 *GABRP* ($\log_2|FC|=-3.746$) 表达显著下降。

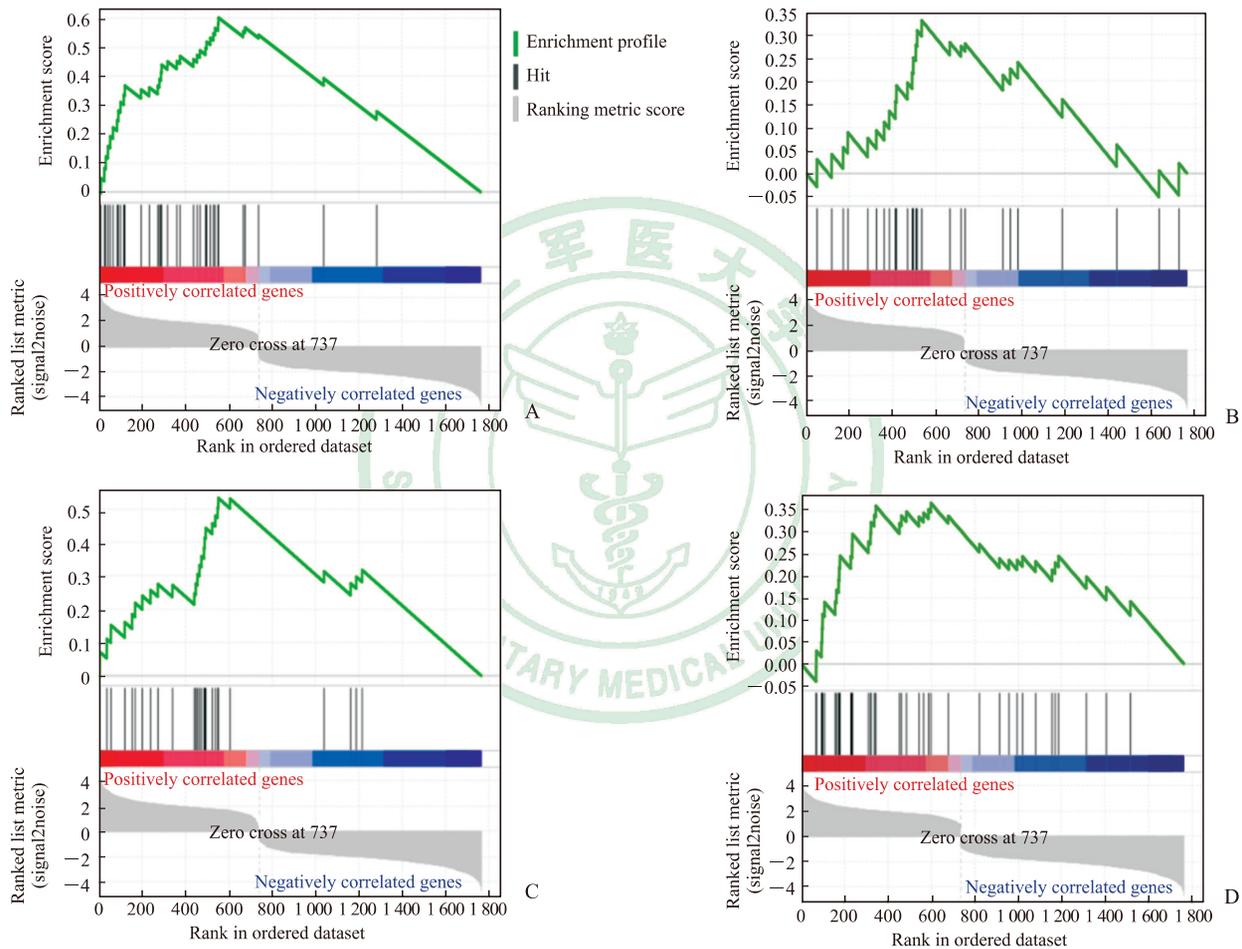


图 6 差异表达基因 GSEA 结果

Fig 6 GSEA results of differentially expressed genes

A: Enrichment plot of hall mark TNF- α signaling via NF- κ B; B: Enrichment plot of hall mark P53 pathway; C: Enrichment plot of PSMB5 target genes; D: Enrichment plot of SETD7 target genes. GSEA: Gene set enrichment analysis; TNF- α : Tumor necrosis factor alpha; NF- κ B: Nuclear factor kappa B; PSMB5: Proteasome 20S subunit beta 5; SETD7: SET domain containing 7 histone lysine methyltransferase.

2.7 分子对接分析 隆纳霉素和蝴蝶霉素对 DNA 分子的结合能分别为 -5.19 和 -3.52 kcal/mol (1 kcal=4.186 kJ)。隆纳霉素与 PSMB5 第 81 位的丙氨酸和第 84 位的酪氨酸作用, 结合能为 -7.22 kcal/mol; 与 SETD7 第 340 位的丝氨酸作用,

结合能为 -6.43 kcal/mol。蝴蝶霉素与 PSMB5 第 128 位的天冬氨酸和第 125 位的酪氨酸作用, 结合能为 -3.58 kcal/mol; 与 SETD7 的结合能为 -2.77 kcal/mol, 无明显作用位点 (图 7)。

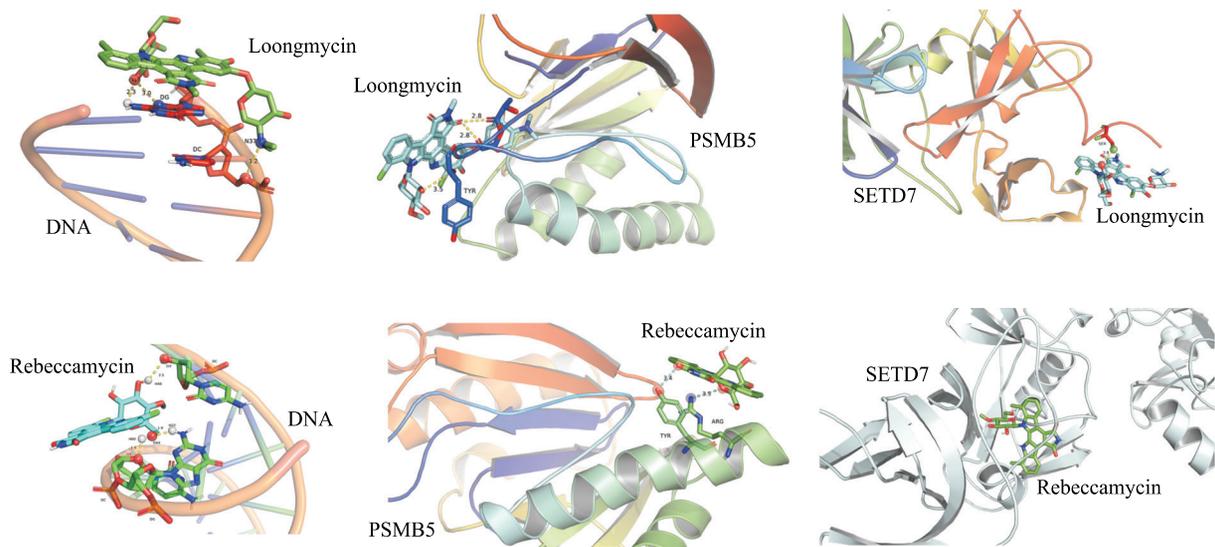


图7 隆纳霉素及蝴蝶霉素分别与DNA、PSMB5、SETD7分子对接示意图

Fig 7 Rigid molecular docking of loongmycin or rebeccamycin with DNA, PSMB5 and SETD7

PSMB5: Proteasome 20S subunit beta 5; SETD7: SET domain containing 7 histone lysine methyltransferase.

3 讨论

本实验分析结果提示NF- κ B介导的TNF- α 信号通路及P53信号通路可能是隆纳霉素作用通路,参与调节的上游蛋白为PSMB5和SETD7。分子对接结果显示,相比蝴蝶霉素,隆纳霉素分子能够更好地插入DNA分子内,起到抑制拓扑异构酶功能的作用,且隆纳霉素与DNA片段及差异表达基因潜在的关键调节蛋白PSMB5和SETD7具有相互作用。

TNF是一种重要的细胞因子,可诱导多种细胞内信号通路,包括细胞凋亡、细胞存活及炎症和免疫。活化的TNF被组装到一个同型三聚体上,并与其受体(TNF受体1、2)结合,TNF受体1可诱导许多基因的激活,主要受2个不同的通路调控,其中NF- κ B通路导致凋亡、MAPK级联反应引起坏死。TNF受体2激活NF- κ B通路,包括PI3K依赖的NF- κ B通路和JNK通路,促进细胞维持生存状态。隆纳霉素作用于人三阴性乳腺癌MDA-MB-468细胞后,IL-6等炎症因子及CXCL1等趋化因子的表达增加,符合TNF受体1介导的信号调节结果。

DNA损伤、氧化应激和激活的致癌基因等应激信号能够激活P53。P53蛋白被用作P53调控基因的转录激活因子,能够导致细胞周期阻滞、衰老或凋亡。推测隆纳霉素能够促进P53上调从而导致癌细胞G₂/M阻滞。

PSMB5是一个多催化功能的蛋白酶体复合物,催化非溶酶体依赖性蛋白质裂解,其在乳腺

癌组织中高表达且与较差的预后相关^[20]。研究表明,PSMB5对三阴性乳腺癌细胞的增殖及抗药性有促进作用^[21]。结合本实验结果,推测PSMB5是隆纳霉素发挥作用的关键靶点之一,隆纳霉素能够通过与其结合抑制人三阴性乳腺癌细胞系MDA-MB-468细胞的增殖。

SETD7是一种甲基转移酶,催化组蛋白H3的第4位赖氨酸单甲基化,是表观遗传转录激活的特殊标志。SETD7也具有对P53等非组蛋白的甲基转移酶活性,在终末分化的细胞重新进入细胞周期后表达上调^[22]。研究表明SETD7低表达乳腺癌患者的中位生存时间(18.1年)约为SETD7高表达患者(9.5年)的2倍^[23]。以上结果提示,隆纳霉素可能通过与SETD7结合抑制后者发挥功能。

另外,基于多数据库的基因筛选结果提示,乳腺癌患者的HAPI、FSCN1、KIT和GABRP等基因表达的变化也应予以关注。

根据分析结果推测:(1)隆纳霉素具有与蝴蝶霉素类似的生物学功能,其通过与DNA-拓扑异构酶I复合物中的DNA分子结合发挥抗肿瘤活性。

(2)隆纳霉素可通过与基因表达调节蛋白PSMB5和SETD7结合,影响GNG7、GNG11、CXCL8、ADCY2等基因的表达水平,从而上调NF- κ B介导的TNF- α 通路和P53通路,最终导致MDA-MB-468细胞G₂/M阻滞,发挥抗肿瘤作用。

[参考文献]

[1] SITHRANGA BOOPATHY N, KATHIRESAN K.

- Anticancer drugs from marine flora: an overview[J/OL]. *J Oncol*, 2010, 2010: 214186. DOI: 10.1155/2010/214186.
- [2] 文雯,周博,陈胜发. 海洋微生物代谢产物中抗肿瘤活性物质的研究进展[J]. *科技视界*, 2019(3): 114-116.
- [3] BRAY F, FERLAY J, SOERJOMATARAM I, SIEGEL R L, TORRE L A, JEMAL A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA Cancer J Clin*, 2018, 68: 394-424.
- [4] GARCÍA FERNÁNDEZ A, CHABRERA C, GARCÍA FONT M, FRAILE M, LAIN J M, GÓNZALEZ S, et al. Differential patterns of recurrence and specific survival between luminal A and luminal B breast cancer according to recent changes in the 2013 St Gallen immunohistochemical classification[J]. *Clin Transl Oncol*, 2015, 17: 238-246.
- [5] SCHMID P, ADAMS S, RUGO H S, SCHNEEWEISS A, BARRIOS C H, IWATA H, et al. Atezolizumab and nab-paclitaxel in advanced triple-negative breast cancer[J]. *N Engl J Med*, 2018, 379: 2108-2121.
- [6] COLLIGNON J, LOUSBERG L, SCHROEDER H, JERUSALEM G. Triple-negative breast cancer: treatment challenges and solutions[J]. *Breast Cancer (Dove Med Press)*, 2016, 8: 93-107.
- [7] BAILLY C, COLSON P, HOUSIER C, RODRIGUES-PEREIRA E, PRUDHOMME M, WARING M J. Recognition of specific sequences in DNA by a topoisomerase I inhibitor derived from the antitumor drug rebeccamycin[J]. *Mol Pharmacol*, 1998, 53: 77-87.
- [8] SCHWANDT A, MEKHAIL T, HALMOS B, O'BRIEN T, MA P C, FU P, et al. Phase- II trial of rebeccamycin analog, a dual topoisomerase- I and - II inhibitor, in relapsed "sensitive" small cell lung cancer[J]. *J Thorac Oncol*, 2012, 7: 751-754.
- [9] SAIF M W, SELLERS S, DIASIO R B, DOUILLARD J Y. A phase I dose-escalation study of edotecarin (J-107088) combined with infusional 5-fluorouracil and leucovorin in patients with advanced/metastatic solid tumors[J]. *Anticancer Drugs*, 2010, 21: 716-723.
- [10] MAMIDALA R, MAJUMDAR P, JHA K K, BATHULA C, AGARWAL R, CHARY M T, et al. Identification of *Leishmania donovani* topoisomerase I inhibitors via intuitive scaffold hopping and bioisosteric modification of known Top 1 inhibitors[J/OL]. *Sci Rep*, 2016, 6: 26603. DOI: 10.1038/srep26603.
- [11] LANGEVIN A M, BERNSTEIN M, KUHN J G, BLANEY S M, IVY P, SUN J, et al; Children's Oncology Group. A phase II trial of rebeccamycin analogue (NSC #655649) in children with solid tumors: a Children's Oncology Group study[J]. *Pediatr Blood Cancer*, 2008, 50: 577-580.
- [12] YU G, WANG L G, HAN Y, HE Q Y. clusterProfiler: an R package for comparing biological themes among gene clusters[J]. *OMICS*, 2012, 16: 284-287.
- [13] The Gene Ontology Consortium, ASHBURNER M, BALL C A, BLAKE J A, BOTSTEIN D, BUTLER H, et al. Gene Ontology: tool for the unification of biology[J]. *Nat Genet*, 2000, 25: 25-29.
- [14] Gene Ontology Consortium. The Gene Ontology resource: enriching a Gold mine[J]. *Nucleic Acids Res*, 2021, 49: D325-D334.
- [15] KANEHISA M, FURUMICHI M, SATO Y, ISHIGURO-WATANABE M, TANABE M. KEGG: integrating viruses and cellular organisms[J]. *Nucleic Acids Res*, 2021, 49: D545-D551.
- [16] MOOTHA V K, LINDGREN C M, ERIKSSON K F, SUBRAMANIAN A, SIHAG S, LEHAR J, et al. PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes[J]. *Nat Genet*, 2003, 34: 267-273.
- [17] BADER G D, HOGUE C W. An automated method for finding molecular complexes in large protein interaction networks[J]. *BMC Bioinform*, 2003, 4: 1-27.
- [18] JANSON G, PAIARDINI A. PyMod 3: a complete suite for structural bioinformatics in PyMOL[J/OL]. *Bioinformatics*, 2020: btaa849. DOI: 10.1093/bioinformatics/btaa849.
- [19] TROTT O, OLSON A J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading[J]. *J Comput Chem*, 2010, 31: 455-461.
- [20] WANG C Y, LI C Y, HSU H P, CHO C Y, YEN M C, WENG T Y, et al. PSMB5 plays a dual role in cancer development and immunosuppression[J]. *Am J Cancer Res*, 2017, 7: 2103-2120.
- [21] WEI W, ZOU Y, JIANG Q, ZHOU Z, DING H, YAN L, et al. PSMB5 is associated with proliferation and drug resistance in triple-negative breast cancer[J]. *Int J Biol Markers*, 2018, 33: 102-108.
- [22] SOSHNIKOVA N. Functions of SETD7 during development, homeostasis and cancer[J/OL]. *Stem Cell Investig*, 2019, 6: 26. DOI: 10.21037/sci.2019.06.10.
- [23] HUANG R, LI X, YU Y, MA L, LIU S, ZONG X, et al. SETD7 is a prognosis predicting factor of breast cancer and regulates redox homeostasis[J]. *Oncotarget*, 2017, 8: 94080-94090.