

生物信息学分析：COL1A1可能是幽门螺旋杆菌通过细胞毒素相关蛋白A参与胃腺癌进程的关键靶点

阚笑^{1,2}, 王云¹, 宋晓萍^{2*}

¹青岛大学基础医学院病原微生物学系, 山东 青岛

²青岛市中医医院(青岛大学附属青岛市海慈医院)检验科, 山东 青岛

收稿日期: 2025年5月19日; 录用日期: 2025年6月13日; 发布日期: 2025年6月23日

摘要

目的: 幽门螺旋杆菌细胞毒素相关蛋白A是一种已知的可诱导胃癌发展的重要毒力因子。本研究旨在通过生物信息学进一步确定具体在这一过程中发挥作用的生物学靶点。材料方法: 首先, 从基因表达综合数据库中筛选mRNA表达谱数据集GSE29998, 并借助已有的研究成果, 使用R软件(4.0.0版)分析幽门螺旋杆菌通过细胞毒素相关蛋白A参与胃腺癌进程的潜在差异表达基因, 再借助富集分析进一步阐述其意义。并最终根据蛋白质互作网络的hub基因的排名, 筛选出可能的潜在靶向基因。结果: 我们通过差异性分析确定了540个可能的与胃腺癌相关的潜在差异基因。再结合陈等人的研究成果所提供的基因数据库, 我们进一步确认18个与潜在的幽门螺旋杆菌细胞毒素相关蛋白A相关的胃腺癌差异表达基因。蛋白质-蛋白质相互作用网络提示只有COL1A1处于枢纽位置。因此我们认定COL1A1是与幽门螺旋杆菌细胞毒素相关蛋白A相关的胃腺癌差异表达基因。结论: 我们的研究鉴定了在胃腺癌中18个潜在的幽门螺旋杆菌细胞毒素相关蛋白A相关的差异表达基因。幽门螺旋杆菌细胞毒素相关蛋白A可能借助COL1A1参与胃腺癌的发生发展过程中。本研究结果首次借助细胞实验相关数据筛选出临床样本中幽门螺旋杆菌可能影响胃腺癌发生的潜在生物学靶点, 这可能有助于我们进一步理解这一疾病。

关键词

胃腺癌, 幽门螺旋杆菌, 幽门螺旋杆菌细胞毒素相关蛋白A, COL1A1

*通讯作者。

Bioinformatics Analysis: COL1A1 May Be a Key Target for Helicobacter Pylori to Participate in the Process of Gastric Adenocarcinoma through Cytotoxin-Associated Protein A

Xiao Kan^{1,2}, Yun Wang¹, Xiaoping Song^{2*}

¹Department of Pathogenic Biology, School of Basic Medicine, Qingdao University, Qingdao Shandong

²Clinical Laboratory, Qingdao Traditional Chinese Medicine Hospital (Qingdao Hiser Hospital Affiliated of Qingdao University), Qingdao Shandong

Received: May 19th, 2025; accepted: Jun. 13th, 2025; published: Jun. 23rd, 2025

Abstract

Introduction: Helicobacter pylori cytotoxin-associated protein A is an important virulence factor known to induce the development of gastric cancer. This research aims to further determine the specific biological targets that play a role in this process through bioinformatics. **Material method:** Firstly, the mRNA expression profile dataset GSE29998 was screened from the comprehensive database of gene expression. With the help of existing research results, the potential differentially expressed genes of Helicobacter pylori involved in the process of gastric adenocarcinoma through cytotoxin-associated protein A were analyzed using R software (version 4.0.0), and its significance was further elaborated with the help of enrichment analysis. And finally, based on the ranking of hub genes in the protein-protein interaction network, possible potential targeted genes are screened out. **Result:** Through differential analysis, we identified 540 possible potentials differentially expressed genes related to gastric adenocarcinoma. Combined with the gene database provided by the research results of Chen *et al.*, we further confirmed 18 differentially expressed genes of gastric adenocarcinoma related to the potential Helicobacter pylori cytotoxin-associated protein A. The protein-protein interaction network suggests that only COL1A1 is in the hub position. Therefore, we determined that COL1A1 is A differentially expressed gene in gastric adenocarcinoma associated with Helicobacter pylori cytotoxin-associated protein A. **Conclusion:** Our research identified 18 potentials differentially expressed genes related to Helicobacter pylori cytotoxin-associated protein A in gastric adenocarcinoma. Helicobacter pylori cytotoxin-associated protein A may be involved in the occurrence and development of gastric adenocarcinoma with the help of COL1A1. The results of this study are the first to screen out potential biological targets that Helicobacter pylori may affect the occurrence of gastric adenocarcinoma in clinical samples with the help of relevant data from cell experiments, which may help us further understand this disease.

Keywords

Gastric Adenocarcinoma, Helicobacter Pylori, Helicobacter Pylori Cytotoxin-Associated Protein A, COL1A1



1. 介绍

胃癌作为一种恶性肿瘤发病率仅次于肺癌、乳腺癌、结直肠癌和前列腺癌，同时也是第三大最常见的癌症死亡原因[1]。胃腺癌占有所有胃恶性肿瘤的 90%以上[2]。目前临床上关于胃腺癌的治疗主要包含全身化疗、放疗、手术、免疫治疗和靶向治疗[3]。有研究指出胃癌根治性手术后病灶可能存在微小残留病灶的标志物，目前正在试图确定包括靶向治疗在内的额外的辅助治疗是否可以清除这些标志物[3]。由此可见继续探索胃腺癌相关的生物学标志至关重要。

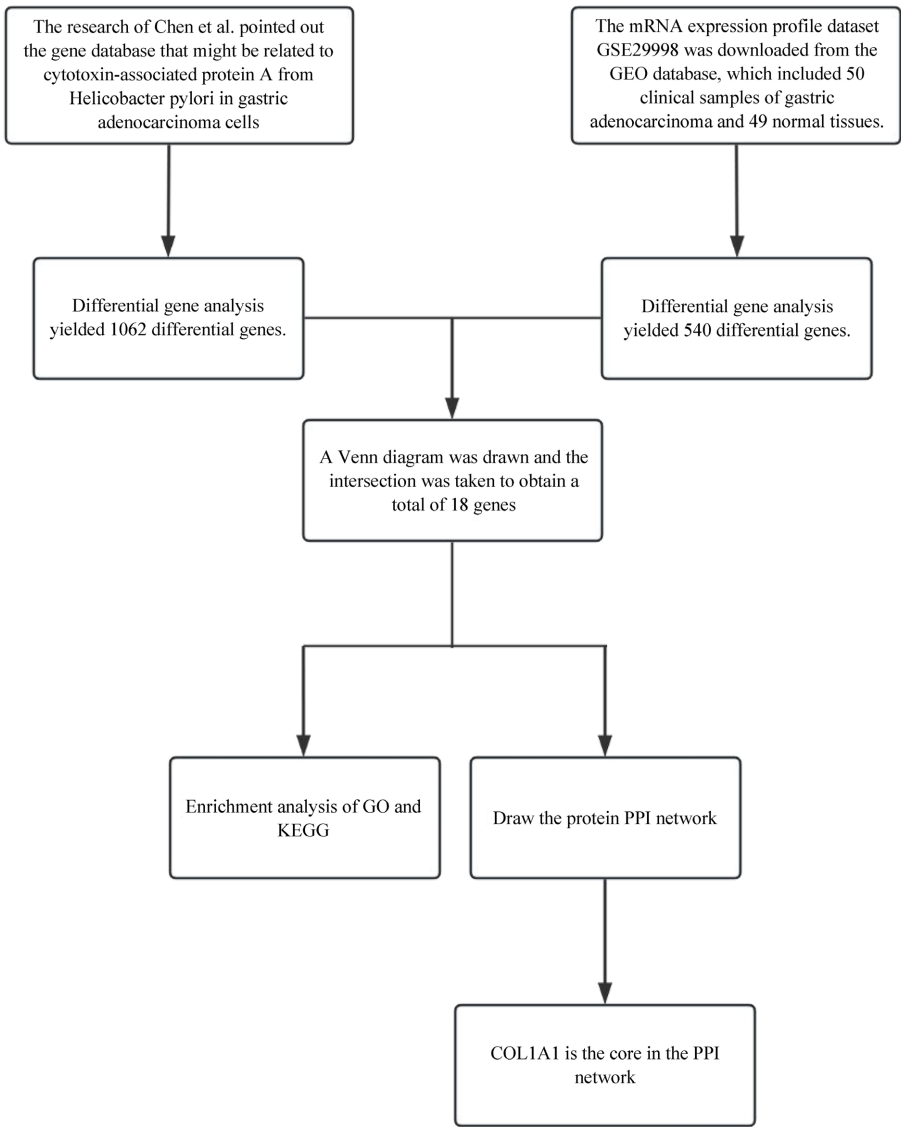


Figure 1. Flowchart
图 1. 流程图

早在 1994 年国际癌症研究机构(IARC)便提出“有足够的证据表明幽门螺旋杆菌感染对人类具有致癌性”[4]。幽门螺旋杆菌感染是已知最重要的胃癌危险因素[5]-[7]。有证据指出胃粘膜中的幽门螺旋杆菌相关的长期炎症反应可能导致持续的组织损伤,从而导致远期胃癌的发生,El-Omar 等人的研究提出 IL1B 和 IL1RN 基因(IL1B 编码白细胞介素(IL)-1 β 和 IL1RN 编码其天然存在的受体拮抗剂)的多态性与幽门螺旋杆菌相关的胃癌风险升高有关[8][9]。早期研究发现,并非所有幽门螺旋杆菌菌株感染都导致剧烈的粘膜炎症反应[10]。进一步的研究提示这是因为一部分幽门螺旋杆菌菌株表达细胞毒素相关基因 A [10]。后者被认为是一种 120 kDa 大小的癌蛋白,通过病原体的 IV 型分泌系统转位到胃上皮细胞中,诱导多个信号级联反应[11][12]。因此讨论细胞毒素相关基因 A 与胃腺癌的关系非常重要,而寻找其潜在的生物学靶点可能是突破口之一。

公开的基因表达数据集 GSE29998 来自于基因表达综合数据库(GEO)。该数据集包括 50 个胃腺癌样本和 49 个匹配的正常粘膜样本。而另一数据库则来自于陈等人的研究成果,在他们的研究中,通过转染胃腺癌细胞,导致细胞毒素相关基因 A 磷酸化水平升高,并发生细胞形态改变,这提示来自于幽门螺旋杆菌表达细胞毒素相关基因 A 发挥作用[13]。而后对这些细胞进行了测序,并获得了公开的数据库[13]。Chen 等人的研究在细胞水平上指出了细胞毒素相关基因 A 刺激胃腺癌细胞的潜在基因数据库,而我们的研究是为了进一步在临床样本中探索这些潜在生物学靶点。我们希望能够强调这一点,并绘制了流程图(图 1)。

2. 材料方法

2.1. 临床胃腺癌样本相关基因数据集和微阵列数据

GSE29998 数据集存在于 GPL6947 平台 99 例全部样本来自 Tissue Solutions Ltd. (<http://www.tissue-solutions.com/>)。使用 Illumina mRNA 表达芯片生成基因表达谱。

2.2. 临床胃腺癌样本中相关差异基因的分析

使用 limma 包在数据集中筛选疾病组和对照组间的差异表达基因,得到相应的 $-\log_{10}$ P. Value 值和 logFC 值,差异阈值设定为: $p.val < 0.05 \& |log_2FC| > 1$ 。在疾病组和对照组样本(DN vs Control)共获得 540 个显著差异表达基因,其中上调基因数量为 254 个,下调基因为 316 个,结果中分别展示了上调和下调基因中表达量前 10 的基因。使用 ggplot2 包展示基因差异情况。使用“ComplexHeatmap”包绘制热图,结果中展示了差异表达基因中的表达量绝对值前 20 的基因。

2.3. 细胞毒素相关基因 A 磷酸化水平升高的胃腺癌细胞与对照组的差异基因分析

根据 Chen 等人研究所提供的原始数据,差异阈值设定为: $p.val < 0.05 \& |log_2FC| > 2$ 。在疾病组和对照组样本共获得 1062 个显著差异表达基因,其中上调基因数量为 594 个,下调基因数量为 468 个,使用“ComplexHeatmap”包绘制热图(进行了 log 转换),结果中展示了差异表达基因中的表达量绝对值前 20 的基因。

2.4. 绘制韦恩图

将 DEGs1 (570 个基因)与 DEGs2 (910 个基因, id 转换后有重复值)取交集,交集基因记作候选基因,使用 R 包“VennDiagram”绘制 Venn 图对结果进行可视化处理。

2.5. 临床胃腺癌样本中幽门螺旋杆菌细胞毒素相关基因 A 关联的差异表达基因的富集分析

我们对上述得到的差异表达基因基于 GO、KEGG 通路的富集分析,寻找基因集合内大量基因共同的功能及相关通路。使用统计学方法累计超几何分布分析一组基因在某个功能节点上是否出现(over-

presentation)。京都基因与基因组百科全书(KEGG)是了解高级功能和生物系统(如细胞、生物和生态系统),从分子水平信息,尤其是大型分子数据集生成的基因组测序和其他高通量实验技术的实用程序数据库资源,由日本京都大学生物信息学中心的 Kanehisa 实验室于 1995 年建立,是国际最常用的生物信息数据库之一,以“理解生物系统的高级功能和实用程序资源库”著称。基因本体论(gene ontology, GO)系统包括三个部分:生物学过程(biological process, BP)、分子功能(molecular functions, MF)、细胞组分(cellular components, CC)。我们使用 R 包“clusterProfiler”进行 GO 和 KEGG 功能富集分析,寻找差异表达基因集合内大量基因共同的功能及相关通路。

2.6. 临床胃腺癌样本中幽门螺旋杆菌细胞毒素相关基因 A 关联的差异表达基因的蛋白质-蛋白质相互作用及相关性分析

蛋白质相互作用网络是以蛋白质作为节点,参与同一代谢途径、生物学过程、结构复合体、功能关联或蛋白质间的物理接触作为边的网络。目前来讲,蛋白质互作网络是被研究最充分的生物分子网络之一。蛋白质是组成生物体并行生物功能的重要生物大分子。蛋白质通过相互作用构成网络来参与生物信号传递、基因表达调节、能量和物质代谢及细胞周期调控等生命过程的各个环节。因此,蛋白质互作网络对于理解细胞网络结构及功能,以及疾病发生发展的基础至关重要。为了探究差异基因之间是否存在互作关系,我们利用 STRING (<https://string-db.org>)网站置信度为 0.4 (confidence = 0.4),预测交集基因的互做关系。

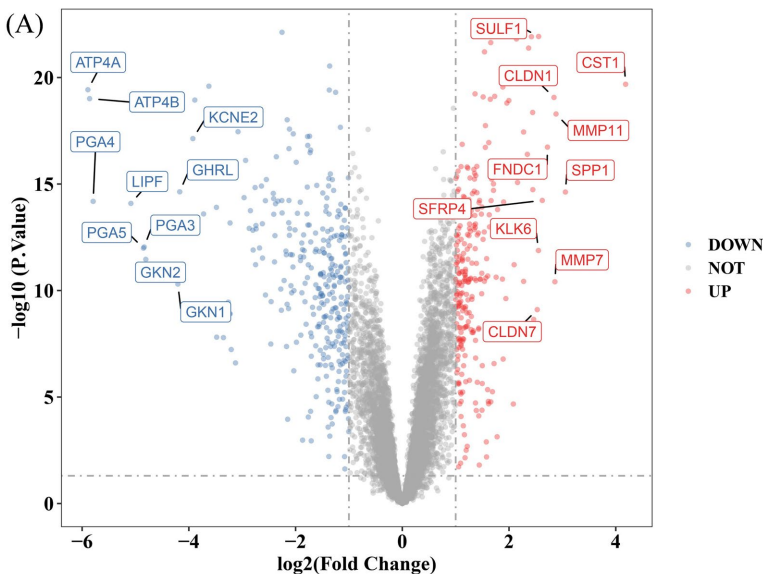
2.7. 统计分析

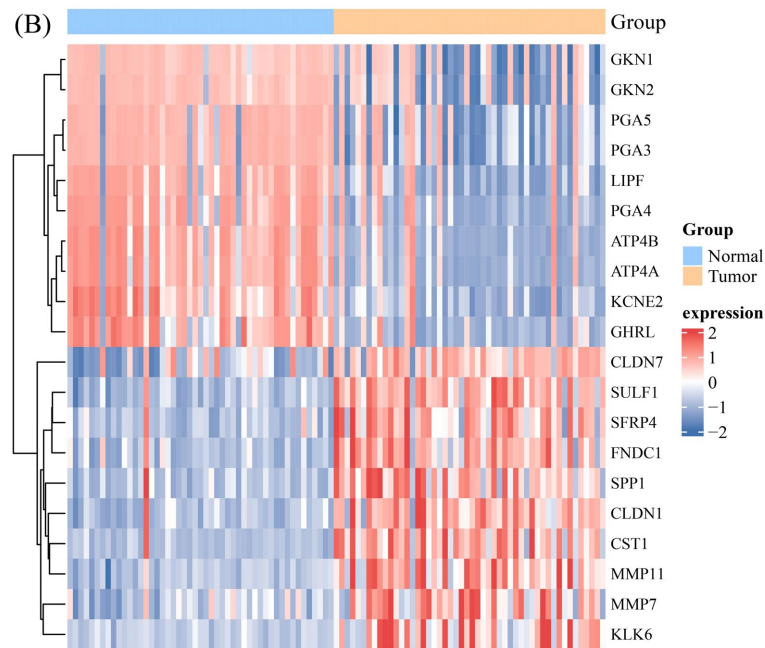
使用 R 软件(4.0.0 版)对生物信息学数据进行统计分析。Student t 检验用于评估临床样本中的基因表达水平。P 值 < 0.05 表示差异具有统计学意义。

3. 结果

3.1. 临床胃腺癌样本中相关差异基因的回顾性分析

在疾病组和对照组样本(DN vs Control)共获得 540 个显著差异表达基因,其中上调基因数量为 254 个,下调基因为 316 个,火山图和热图中分别展示了上调和下调基因中表达量前 10 的基因(图 2(A)和图 2(B))。



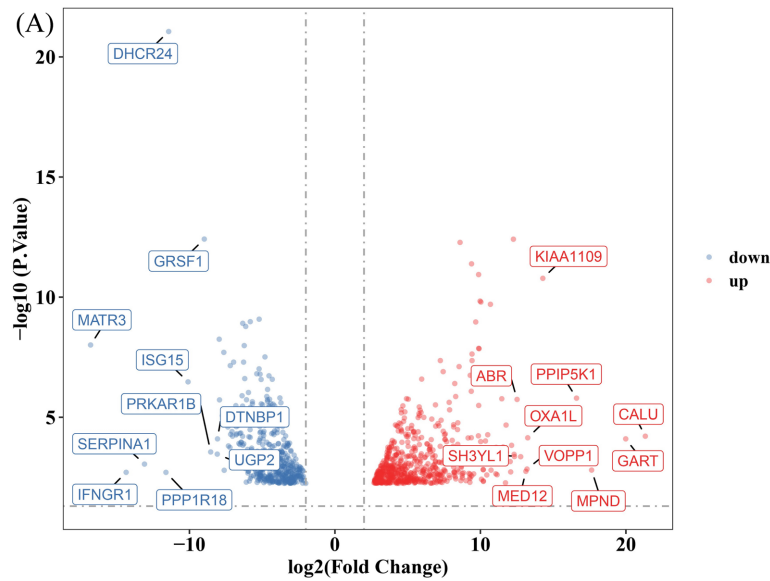


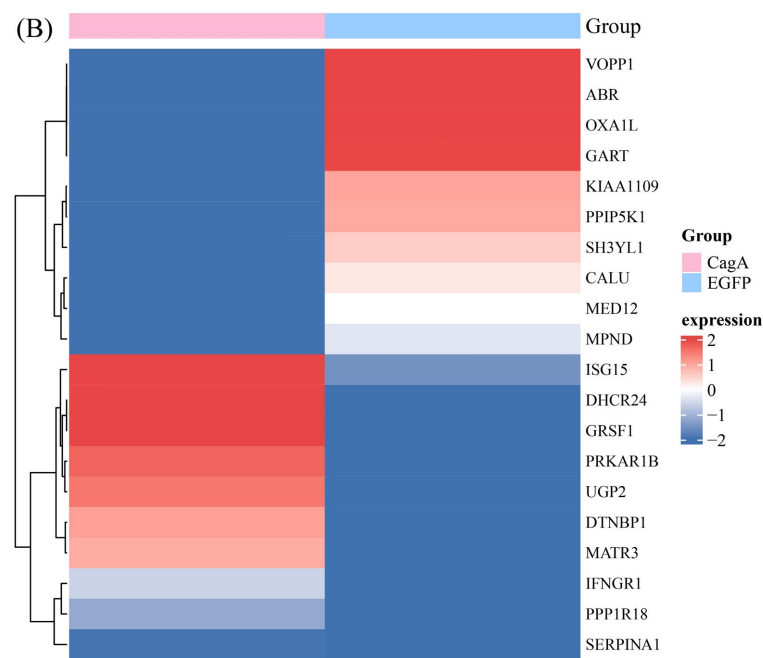
注：图 2 为 GSE29998 差异性分析，疾病组和对照样本 DEG 的鉴定。图 A 差异表达基因的火山图，此时横轴为 $\log_2FC$ ，纵轴为 $-\log_{10}(P.val)$ ，每一个红点代表一个上调基因，每一个蓝点代表一个下调基因；图 B 差异表达基因的热图，红色和蓝色分别表示具有高、低表达水平的差异表达基因，左侧树状图表示对来自不同基因的聚类分析结果。

Figure 2. Analysis of related differentially expressed genes in clinical gastric adenocarcinoma samples
图 2. 临床胃腺癌样本中相关差异基因的分析

3.2. 细胞毒素相关基因 A 磷酸化水平升高的胃腺癌细胞与对照组的差异基因分析

在疾病组和对照组样本共获得 1062 个显著差异表达基因，其中上调基因数量为 594 个，下调基因数量为 468 个，同时绘制了火山图和热图，结果中展示了差异表达基因中的表达量绝对值前 20 的基因(图 3(A)和图 3(B))。



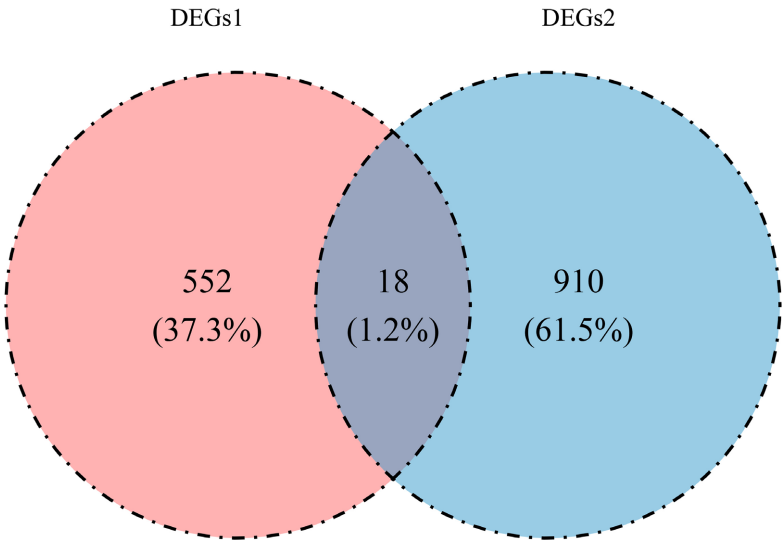


注：图 3 为 Cheng 等人研究结论中数据库的差异性分析。图 A 差异表达基因的火山图，此时横轴为 $\log_2FC$ ，纵轴为 $-\log_{10}(P.val)$ ，每一个红点代表一个上调基因，每一个蓝点代表一个下调基因，图 B 差异表达基因的热图，红色和蓝色分别表示具有高、低表达水平的差异表达基因，左侧树状图表示对来自不同基因的聚类分析结果。

Figure 3. Analysis of differential genes between gastric adenocarcinoma cells with elevated phosphorylation level of cytotoxin-related gene A and the control group

图 3. 细胞毒素相关基因 A 磷酸化水平升高的胃腺癌细胞与对照组的差异基因分析

3.3. 绘制韦恩图



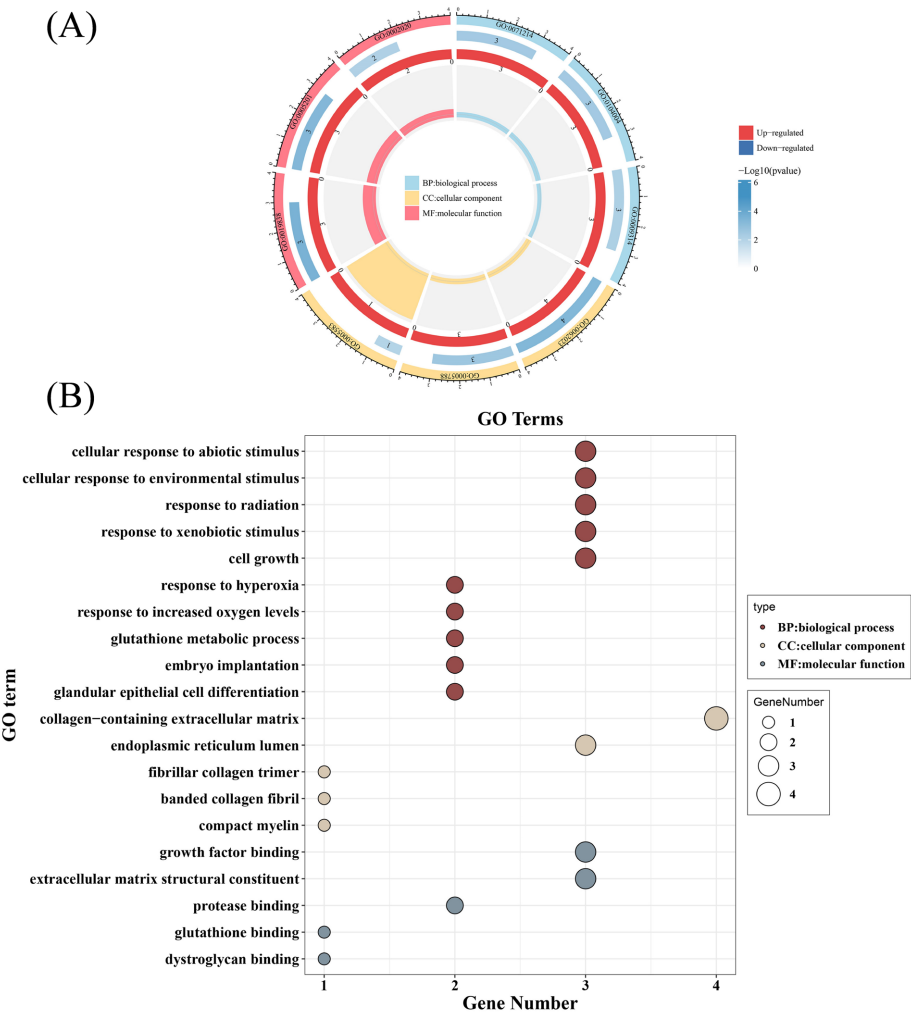
注：图 4 将 DEGs1 (570 个基因)与 DEGs2 (910 个基因，id 转换后有重复值)取交集，交集基因记作候选基因。

Figure 4. Venn diagrams of the two databases

图 4. 两数据库的韦恩图

将两个差异数据集取交集，交集基因记作候选基因，共得到 18 个基因为候选基因：Collagen type I alpha 1 (**COL1A1**), tissue inhibitor of metalloproteinase 1 (**TIMP1**), Sushi repeat-containing protein X-linked 2 (**SRPX2**), NEDD4-like E3 ubiquitin protein ligase (**NEDD4L**), Rap1 GTPase-activating protein (**Rap1GAP**), **Superoxide dismutase 2 (SOD2)**, syntaxin-16 (**STX16**), tubulin-binding cofactor B (**TBCB**), Sorbin and SH3 domain-containing protein 2 (**SORBS2**), Plasmolipin (**PLLP**), RAD51-associated protein 1 (**RAD51AP1**), Insulin-like growth factor-binding protein 7 (**IGFBP7**), Homeobox C6 (**HOXC6**), anterior gradient protein 2 (**AGR2**), Interferon-stimulated gene product 15 (**ISG15**), NK6 Homeobox 2 (**NKX6-2**), glutathione S-transferase mu 3 (**GSTM3**), phospholipase A2 group IIA (**PLA2G2A**)。占总比例的 1.2% (图 4)。

3.4. 临床胃腺癌样本中幽门螺旋杆菌细胞毒素相关基因 A 关联的差异表达基因的富集分析



注：图 5 候选基因的 GO 分析。A 图为八卦图，最外圈中蓝色代表生物学过程(BP)，黄色代表细胞成分(CC)，红色代表功能(MF)；第三圈代表的是通路中富集到的基因数量；第二圈代表的是富集到的基因上调和下调基因的占比；B 图为气泡图，横坐标为注释通路的基因个数，纵坐标为通路，颜色是以 $p.adjust$ 的值来决定的。

Figure 5. GO enrichment analysis of differentially expressed genes associated with *Helicobacter pylori* cytotoxin-related gene A in clinical gastric adenocarcinoma samples
图 5. 临床胃腺癌样本中幽门螺旋杆菌细胞毒素相关基因 A 关联的差异表达基因的 GO 富集分析



Figure 6. KEGG enrichment analysis of differentially expressed genes associated with helicobacter pylori cytotoxin-related gene A in clinical gastric adenocarcinoma samples

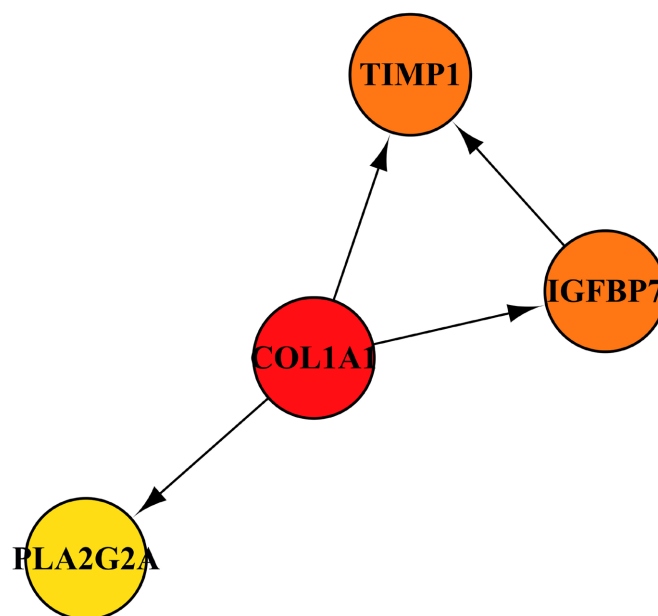
临床医学进展

使用“clusterProfiler”和“GOplot”包对 18 个交集基因进行 GO 富集和 KEGG 富集分析，其中 GO 分析筛选标准为 $p.adjust < 0.05$ ，结果显示共富集了 262 条 GO 通路，其中生物学过程(BP)有 223 条(结果展示了 TOP3，下同)，主要在细胞对非生物刺激的反应(cellular response to abiotic stimulus)、细胞对环境刺激的反应(cellular response to environmental stimulus)、对辐射的反应(response to radiation)等生物学过程显著富集；细胞组分(CC)有 12 条，主要在含胶原蛋白的细胞外基质(collagen-containing extracellular matrix)、内质网腔(endoplasmic reticulum lumen)及纤维胶原三聚体(fibrillar collagen trimer)等细胞成分显著富集；分子功能(MF)有 27 条，主要在生长因子结合(growth factor binding)、细胞外基质结构成分(extracellular matrix structural constituent)及蛋白酶结合(enzyme binding)等分子功能显著富集(图 5(A)和图 5(B))。

KEGG 富集分析筛选标准为 $p.adjust < 0.05$ ，共富集到 8 条 KEGG 通路，主要富集于化学致癌-活性氧(chemical carcinogenesis-reactive oxygen species)、亚麻酸代谢(alpha-Linolenic acid metabolism)、亚油酸代谢(linoleic acid metabolism)、囊泡运输中的 SNARE 相互作用(SNARE interactions in vesicular transport)及醛固酮调节的钠重吸收(aldosterone-regulated sodium reabsorption)等。利用“clusterProfiler”R 包进行可视化(图 6(A)和图 6(B))。

3.5. 临床胃腺癌样本中幽门螺旋杆菌细胞毒素相关基因 A 关联的差异表达基因的蛋白质 - 蛋白质相互作用及相关性分析

蛋白质 - 蛋白质相互作用网络(PPI, protein-protein interaction)分析显示只有少量分候选基因中存在着密切的相互作用(图 7)。而处于枢纽地位的是 COL1A1。因此我们认定 COL1A1 可能是幽门螺旋杆菌通过细胞毒素相关蛋白 A 参与胃腺癌进程的关键靶点。



注：图 7 每一个圆点代表一个差异表达的主动脉夹层中细胞焦亡相关的候选基因，蓝色代表显著下调的候选基因，红色代表显著上调的候选基因，与其他候选基因间的桥对最多的，则被认为是处于核心位置。

Figure 7. Protein-protein interaction and correlation analysis of differentially expressed genes associated with *Helicobacter pylori* cytotoxin-related gene A in clinical gastric adenocarcinoma samples

图 7. 临床胃腺癌样本中幽门螺旋杆菌细胞毒素相关基因 A 关联的差异表达基因的蛋白质 - 蛋白质相互作用及相关性分析

4. 讨论

幽门螺旋杆菌是一种具有螺旋弯曲的形状和鞭的微需氧革兰阴性芽孢杆菌，定植于人胃上皮表面。幽门螺旋杆菌的感染率在世界范围内已超过 50%，被认为是包括胃腺癌在内的消化道疾病的高危因素[14] [15]。值得一提的是幽门螺旋杆菌耐药性是世界性的难题[14]。因此如何阻断幽门螺旋杆菌的毒性是值得探讨的话题。鉴于表达细胞毒素相关基因 A 的幽门螺旋杆菌对于胃粘膜具有强烈的炎性刺激作用，因此寻找细胞毒素相关基因 A 在胃腺癌中发挥作用的靶点意义重大[10] [13]。

COL1A1 又称 I 型胶原 $\alpha 1(I)$ ，是胶原蛋白家族的一员，参与上皮-间充质转化，在多种恶性肿瘤中高表达[16]。其本质是一种异源三聚体分子，由两条 $\alpha 1(I)$ 链和一条 $\alpha 2(I)$ 链组成[17]。胶原蛋白是细胞外基质的主要成分，是正常组织功能的重要组成部分；它在维持组织和器官的稳定性和完整性方面起着至关重要的作用[16]。根据 COL1A1 的功能，有可能是因为其改变了细胞外基质的构成，从而为胃腺癌的转移和入侵提供了协助，但其与幽门螺杆菌细胞毒素相关蛋白 A 的具体互作关系尚不得知。有研究指出 COL1A1 的过表达通过 SMAD3 激活了 TGF- β 信号通路，胃腺癌中过量的 COL1A1 可能通过激活 TGF- β 信号通路显著增强了其迁移和侵袭能力，这一结论符合也符合细胞外基质的功能[18]。通过 TGF- β 信号通路，COL1A1 促使腺癌细胞获得间充质表型，这最终增强了胃腺癌的入侵能力[18]。Li 等人的文章指出，在胃腺癌中 COL1A1 的表达量与生存率呈负相关，可能可以作为胃腺癌的潜在生物学靶点，这与我们的研究结论相互支持[16]。此外，COL1A1 高表达水平的胃腺癌细胞通过靶向抑制 miR-129-5p 和 miRNA let-7i 发挥的保护作用获得更强的增殖能力和迁徙能力[16] [19]。目前，COL1A1 在胃腺癌领域的研究相对不足，但这在未来将是一个非常有前景的方向[16]。总之我们的研究首次通过生物信息学分析，在已有研究的基础上提出胃腺癌中的高表达的 COL1A1 很可能来自于表达细胞毒素相关基因 A 的幽门螺旋杆菌这一结论。

本研究存在一些局限性，首先，我们缺乏具体的机制研究来讨论 COL1A1 是如何具体与来自于幽门螺旋杆菌的细胞毒素相关基因 A 相互作用的。其次，由于临床样本相比单纯的胃腺癌细胞更加复杂，所以可能有很多潜在靶向基因被忽略了，这需要通过更大样本量的测序工作来完善。

参考文献

- [1] Park, J.Y., Forman, D., Waskito, L.A., Yamaoka, Y. and Crabtree, J.E. (2018) Epidemiology of *Helicobacter pylori* and CagA-Positive Infections and Global Variations in Gastric Cancer. *Toxins*, **10**, Article No. 163. <https://doi.org/10.3390/toxins10040163>
- [2] Piazzuelo, M.B., Epplin, M. and Correa, P. (2010) Gastric Cancer: An Infectious Disease. *Infectious Disease Clinics of North America*, **24**, 853-869. <https://doi.org/10.1016/j.idc.2010.07.010>
- [3] Joshi, S.S. and Badgwell, B.D. (2021) Current Treatment and Recent Progress in Gastric Cancer. *CA: A Cancer Journal for Clinicians*, **71**, 264-279. <https://doi.org/10.3322/caac.21657>
- [4] Schistosomes (1994) Liver Flukes and *Helicobacter pylori*. IARC Monographs on the Identification of Carcinogenic Hazards to Humans, Vol. 61, 1-241.
- [5] Helicobacter and Cancer Collaborative Group (2001) Gastric Cancer and *Helicobacter pylori*: A Combined Analysis of 12 Case Control Studies Nested within Prospective Cohorts. *Gut*, **49**, 347-353. <https://doi.org/10.1136/gut.49.3.347>
- [6] Kamangar, F., Dawsey, S.M., Blaser, M.J., Perez-Perez, G.I., Pietinen, P., Newschaffer, C.J., et al. (2006) Opposing Risks of Gastric Cardia and Noncardia Gastric Adenocarcinomas Associated with *Helicobacter pylori* Seropositivity. *JNCI: Journal of the National Cancer Institute*, **98**, 1445-1452. <https://doi.org/10.1093/jnci/djj393>
- [7] Uemura, N., Okamoto, S., Yamamoto, S., Matsumura, N., Yamaguchi, S., Yamakido, M., et al. (2001) *Helicobacter pylori* Infection and the Development of Gastric Cancer. *New England Journal of Medicine*, **345**, 784-789. <https://doi.org/10.1056/nejmoa001999>
- [8] El-Omar, E.M., Carrington, M., Chow, W., McColl, K.E.L., Bream, J.H., Young, H.A., et al. (2000) Interleukin-1 Polymorphisms Associated with Increased Risk of Gastric Cancer. *Nature*, **404**, 398-402. <https://doi.org/10.1038/35006081>
- [9] El-Omar, E.M., Carrington, M., Chow, W., McColl, K.E.L., Bream, J.H., Young, H.A., et al. (2001) Correction: The

- Role of Interleukin-1 Polymorphisms in the Pathogenesis of Gastric Cancer. *Nature*, **412**, 99. <https://doi.org/10.1038/35083631>
- [10] Crabtree, J.E., Taylor, J.D., Heatley, R.V., Shallcross, T.M., Rathbone, B.J., Wyatt, J.I., *et al.* (1991) Mucosal IgA Recognition of *Helicobacter pylori* 120 Kda Protein, Peptic Ulceration, and Gastric Pathology. *The Lancet*, **338**, 332-335. [https://doi.org/10.1016/0140-6736\(91\)90477-7](https://doi.org/10.1016/0140-6736(91)90477-7)
 - [11] Ohnishi, N., Yuasa, H., Tanaka, S., Sawa, H., Miura, M., Matsui, A., *et al.* (2008) Transgenic Expression of *Helicobacter pylori* CagA Induces Gastrointestinal and Hematopoietic Neoplasms in Mouse. *Proceedings of the National Academy of Sciences*, **105**, 1003-1008. <https://doi.org/10.1073/pnas.0711183105>
 - [12] Nishikawa, H. and Hatakeyama, M. (2017) Sequence Polymorphism and Intrinsic Structural Disorder as Related to Pathobiological Performance of the *Helicobacter pylori* CagA Oncoprotein. *Toxins*, **9**, Article No. 136. <https://doi.org/10.3390/toxins9040136>
 - [13] Chen, D., Li, C., Zhao, Y., Zhou, J., Wang, Q. and Xie, Y. (2021) Bioinformatics Analysis for the Identification of Differentially Expressed Genes and Related Signaling Pathways in *H. pylori*-CagA Transfected Gastric Cancer Cells. *PeerJ*, **9**, e11203. <https://doi.org/10.7717/peerj.11203>
 - [14] Xu, W., Xu, L. and Xu, C. (2022) Relationship between *Helicobacter pylori* Infection and Gastrointestinal Microecology. *Frontiers in Cellular and Infection Microbiology*, **12**, Article ID: 938608. <https://doi.org/10.3389/fcimb.2022.938608>
 - [15] Zavros, Y. and Merchant, J.L. (2022) The Immune Microenvironment in Gastric Adenocarcinoma. *Nature Reviews Gastroenterology & Hepatology*, **19**, 451-467. <https://doi.org/10.1038/s41575-022-00591-0>
 - [16] Li, X., Sun, X., Kan, C., Chen, B., Qu, N., Hou, N., *et al.* (2022) COL1A1: A Novel Oncogenic Gene and Therapeutic Target in Malignancies. *Pathology—Research and Practice*, **236**, Article ID: 154013. <https://doi.org/10.1016/j.prp.2022.154013>
 - [17] Van Doren, S.R. (2015) Matrix Metalloproteinase Interactions with Collagen and Elastin. *Matrix Biology*, **44**, 224-231. <https://doi.org/10.1016/j.matbio.2015.01.005>
 - [18] Guo, Y., Lu, G., Mao, H., Zhou, S., Tong, X., Wu, J., *et al.* (2020) miR-133b Suppresses Invasion and Migration of Gastric Cancer Cells via the COL1A1/TGF- β Axis. *OncoTargets and Therapy*, **13**, 7985-7995. <https://doi.org/10.2147/ott.s249667>
 - [19] Wang, Q. and Yu, J. (2018) miR-129-5p Suppresses Gastric Cancer Cell Invasion and Proliferation by Inhibiting COL1A1. *Biochemistry and Cell Biology*, **96**, 19-25. <https://doi.org/10.1139/bcb-2016-0254>