

## • 数据挖掘与循证医学 •

## S100A12 在胃癌“炎-癌转化”恶性进程中的免疫调节作用及靶向中药小分子化合物筛选

聂多锐<sup>1</sup>, 康睿哲<sup>2</sup>, 杨梅林<sup>2</sup>, 杨冉<sup>1\*</sup>

1. 南京中医药大学附属医院(江苏省中医院)肿瘤科, 江苏 南京 210029

2. 南京中医药大学第一临床医学院, 江苏 南京 210023

**摘要:** **目的** 基于多组学数据与孟德尔随机化(Mendelian randomization, MR)因果推断方法, 筛选驱动胃黏膜“炎-癌转化”的特征基因, 并探索针对这些基因的潜在中药干预策略。**方法** 整合转录组数据与大规模遗传数据, 通过差异分析、MR分析与蛋白质互作网络构建, 筛选与胃癌风险存在因果关联的特征基因。对筛选出的基因进一步进行生存分析、免疫浸润分析及免疫中介机制解析。以核心特征基因为靶点, 对中药小分子化合物库进行分子对接虚拟筛选, 预测潜在活性成分。**结果** 在“炎-癌转化”进程中鉴定出30个持续上调的差异基因; MR分析从中确定5个与胃癌风险呈正向因果关联的核心驱动基因(S100A8、CXCR1、S100A9、S100A12、ZBED2)。其中, S100A12的高表达与患者不良预后显著相关, 且与中性粒细胞浸润正相关、与B细胞及CD4<sup>+</sup>T细胞浸润负相关。中介分析提示S100A12可能通过负向调控CD25<sup>+</sup>CD4<sup>+</sup>T细胞介导促癌效应。全表型组MR分析显示靶向S100A12具有较好的潜在安全性。虚拟筛选获得多个与S100A12蛋白具有高亲和力的中药小分子化合物(C-箭毒碱、酸浆苦味素D等)。**结论** S100A12是连接慢性炎症与胃癌发生的关键特征基因, 其通过重塑免疫微环境驱动“炎-癌转化”。针对该靶点的中药小分子筛选为干预胃癌前病变提供了先导化合物线索。

**关键词:** 胃癌; 炎癌转化; S100A12; 免疫浸润; 孟德尔随机化; C-箭毒碱; 酸浆苦味素D

中图分类号: Q811.4; R285 文献标志码: A 文章编号: 0253-2670(2026)11-4277-14

DOI: 10.7501/j.issn.0253-2670.2026.11.017

## Immunomodulatory role of S100A12 in malignant progression of inflammation-cancer transformation in gastric cancer and screening of targeted traditional Chinese medicine-derived small molecules

NIE Duorui<sup>1</sup>, KANG Ruizhe<sup>2</sup>, YANG Meilin<sup>2</sup>, YANG Ran<sup>1</sup>

1. Department of Oncology, Affiliated Hospital of Nanjing University of Chinese Medicine (Jiangsu Provincial Hospital of Chinese Medicine), Nanjing 210029, China

2. The First Clinical Medical College, Nanjing University of Chinese Medicine, Nanjing 210023, China

**Abstract: Objective** To screen characteristic genes driving gastric mucosal “inflammation-cancer transformation” based on multi-omics data and Mendelian randomization (MR) causal inference methods, and to explore potential traditional Chinese medicine (TCM) intervention strategies. **Methods** Transcriptomic data and large-scale genetic data were integrated. Differential analysis, MR analysis, and protein-protein interaction network construction were performed to screen characteristic genes causally associated with gastric cancer risk. Subsequently, survival analysis, immune infiltration analysis, and immune mediation mechanism analysis were conducted on the identified genes. Finally, molecular docking-based virtual screening was performed on a TCM small-molecule compound library using core characteristic genes as targets to predict potential active components. **Results** A total of 30 continuously up-regulated differentially expressed genes were identified during the “inflammation-cancer transformation” process. MR analysis identified five core characteristic genes (S100A8, CXCR1, S100A9, S100A12, ZBED2) showing positive causal associations with gastric cancer risk. Among them, high expression of S100A12 was significantly associated with poor prognosis in patients, positively correlated with

收稿日期: 2026-01-04

基金资助: 国家自然科学基金青年基金项目(82505345)

作者简介: 聂多锐, 博士, 住院医师, 研究方向为中西医结合防治恶性肿瘤。E-mail: ndrize@hotmail.com

\*通信作者: 杨冉, 博士, 主治医师, 研究方向为中西医结合防治恶性肿瘤。E-mail: merryang519@126.com

neutrophil infiltration, and negatively correlated with B-cell and CD4<sup>+</sup> T-cell infiltration. Mediation analysis revealed that S100A12 may promote carcinogenic effects by negatively regulating CD25<sup>+</sup>CD4<sup>+</sup> T cells (primarily regulatory T cells). Phenome-wide MR analysis suggested good potential safety for targeting S100A12. Virtual screening identified several TCM small-molecule compounds with high binding affinity to the S100A12 protein (e.g., C-curarine, physalin D). **Conclusion** S100A12 is a key characteristic gene linking chronic inflammation to gastric cancer development, driving “inflammation-cancer transformation” by reshaping the immune microenvironment. The screening of TCM small molecules targeting this gene provides lead compound candidates for intervening in gastric precancerous lesions. The established research framework integrating multi-omics and reverse drug screening offers a new approach for the precise prevention and drug development of gastric cancer.

**Key words:** gastric cancer; inflammation-cancer transformation; S100A12; immune infiltration; Mendelian randomization; C-curarine; physalin D

胃癌是全球范围内重大的公共卫生挑战，其发病率和死亡率分别位居全球癌症的第5位<sup>[1]</sup>。在我国，这一疾病负担尤为突出。根据国家癌症中心的最新统计数据，其发病率位居恶性肿瘤第5位，而死亡率高居第3位<sup>[2]</sup>。更严峻的是，由于缺乏典型的早期症状和有效的筛查普及，临床上大多数患者确诊时已进展至局部晚期或晚期，导致治疗选择有限、疗效不佳，整体预后不良，5年生存率仍有待提高<sup>[3-4]</sup>。这一现状从根本上凸显了将防治关口前移、探索胃癌发生发展的内在规律并寻找源头干预策略的紧迫性。

胃癌的发生发展尤其是肠型胃癌，通常遵循经典的“Correa级联”模式，即从慢性非萎缩性胃炎、萎缩性胃炎、肠上皮化生、低级别上皮内瘤变，逐步进展为高级别上皮内瘤变和最终浸润性胃癌的多阶段序贯过程<sup>[5]</sup>。萎缩性胃炎和肠上皮化生属于癌前状态，胃上皮内瘤变则是癌前病变，它们均与胃癌的风险相关。一项荟萃研究揭示了这一进展序列中各阶段的疾病负担与风险：萎缩性胃炎、肠上皮化生和上皮内瘤变的患病率分别为25.4%、16.2%和2.0%<sup>[6]</sup>。更为关键的是，其疾病进展风险呈阶梯式上升。全球汇总数据显示，每千人年胃癌进展率估计值分别为萎缩性胃炎2.09%、肠化生2.89%，以及不典型增生10.09%<sup>[7]</sup>。鉴于上皮内瘤变是胃癌发生的关键节点且具有较高的进展风险<sup>[8]</sup>，早期诊断与干预癌前病变对预防胃癌至关重要。因此，深入解析此进程中驱动恶性转化的核心分子机制是实现有效预警与干预的关键。

近年来，多组学技术的兴起使得对转录组、蛋白质组等不同维度的生物学数据进行整合分析成为可能，这为发现潜在的生物标志物和治疗靶点提供了交叉验证的有力证据<sup>[9-12]</sup>。然而，这类关联性研究难以区分疾病进展的“驱动者”与“伴随者”，所

发现的众多差异表达基因在胃癌发生中的因果作用仍待严格证实。因此，亟需一种能够从海量数据中有效推断因果关系并甄别关键致病基因的新方法。在此背景下，孟德尔随机化（Mendelian randomization, MR）分析为破解这一难题提供了方法学支撑<sup>[13-14]</sup>。该方法以遗传变异作为工具变量，在很大程度上规避了传统观察性研究中的混杂因素与反向因果关系<sup>[15]</sup>；特别是通过整合表达数量性状位点（expression quantitative trait loci, eQTL）、蛋白质数量性状位点及全基因组关联研究（genome-wide association studies, GWAS）数据，可系统评估基因表达或蛋白质丰度与疾病风险之间的潜在因果关联，为锁定真正的致病驱动因子提供更高等级的证据<sup>[16-17]</sup>。然而，通过因果推断确定关键致病靶点后，如何将其转化为干预策略成为研究的自然延伸。中医药在防治癌前病变及抗肿瘤方面具有多成分、多靶点调节的整体优势<sup>[18]</sup>，但目前大多数研究仍始于临床经验或对单一活性成分的探索，缺乏从明确疾病靶点出发的系统性药物发现路径。

鉴于此，本研究旨在构建一个整合多组学数据与因果推断方法的系统性研究框架。首先利用转录组学数据描绘“炎-癌转化”进程中的基因表达动态；随后，整合eQTL和GWAS数据集，通过MR分析筛选出与胃癌发病存在显著因果关联的关键基因，并通过多维度数据进行验证；最后，以此确立的关键基因为靶标，反向筛选中药小分子化合物，以期发现能够阻断“炎-癌转化”关键环节的潜在干预先导化合物。本研究不仅旨在揭示胃癌发生的新机制，也试图探索一种融合计算生物学与中药药理学的新靶点发现与药物研发模式。

## 1 资料和方法

### 1.1 研究数据集

本研究整合了来自公共数据库的多组学数据。

用于分析“炎-癌转化”基因表达动态的转录组数据源自基因表达综合数据库 (Gene Expression Omnibus, GEO), 系统检索并下载了数据集 GSE55696。该数据集基于微阵列平台, 共包含 77 例胃黏膜样本, 完整覆盖了从慢性炎症、癌前病变到早期癌变的序列, 具体包括慢性胃炎 (chronic gastritis, CG) 19 例、胃癌前病变 (precancerous lesions of gastric cancer, PLGC) 39 例 (低级别上皮内瘤变 19 例、高级别上皮内瘤变 20 例) 及早期胃癌 (early gastric cancer, EGC) 19 例<sup>[19]</sup>。样本来自 46 例男性患者和 31 例女性患者, 年龄 36~78 岁, 平均年龄 (58.87±9.13) 岁。

在 MR 分析中, 用于构建“暴露”的工具变量数据来自 eQTLGen 联盟发布的大规模顺式表达数量性状位点 (cis-eQTL) 汇总数据。该数据是对 37 个队列共 31 684 份样本 (以血液组织为主) 的荟萃分析结果, 提供了 16 699 个基因的 cis-eQTL 信息 [位于基因上下游 1 Mb 内, 次要等位基因频率 (minor allele frequency, MAF) >0.01], 其代表遗传学上基因表达水平<sup>[20]</sup>。作为 MR 分析“结局”的胃癌遗传关联数据, 则取自 FinnGen 项目 (R12) 的 GWAS 汇总统计数据<sup>[21]</sup>。该数据集共计 381 045 个样本, 其中包含 2 296 例胃癌病例与 378 749 例对照。所有使用的公共数据均已在其原始研究中获得伦理批准。

### 1.2 差异表达基因的鉴定

基于 Sangerbox 平台 (<http://www.sangerbox.com/tool>) 的 limma R 包筛选差异表达基因, 为控制多重假设检验导致的假阳性, 使用 Benjamini-Hochberg 方法对  $P$  值进行校正。差异表达基因的筛选条件: 校正后  $P$  值 ( $P_{adj}$ ) <0.05 且  $|\log_2FC| > 1$  [FC 表示差异倍数 (fold change)]。其中, 将  $\log_2FC < -1$  的基因定义为显著下调基因,  $\log_2FC > 1$  的基因定义为显著上调基因。分别对 EGC 组 vs PLGC 组和 PLGC 组 vs CG 组进行差异分析, 取 2 组上调基因的交集 (同理取下调基因的交集), 视为“炎-癌转化”相关差异基因。研究结果采用 ggplot 包进行可视化。

### 1.3 构建蛋白质-蛋白质相互作用 (protein-protein interaction, PPI) 网络和功能富集

为了深入理解“炎-癌转化”相关差异基因的功能, 将上述基因输入 STRING 数据库 (版本 12.0, <http://cn.string-db.org/>), 设置物种为“Homo sapiens”, 将最低相互作用置信度设置为 0.4, 隐藏游离的蛋白节点, 其余参数保留默认设置<sup>[22]</sup>。随后,

将获得的互作关系数据导入 Cytoscape 软件 (版本 3.8.2) 进行网络可视化、拓扑属性计算, 并使用 cytoHubba 插件筛选排名前 10 的基因, 将其视为 PPI 核心基因。

功能富集分析用于系统阐释“炎-癌转化”相关差异基因集合的生物学意义。采用 R 语言软件包 clusterProfiler (版本 4.12.6) 对“炎-癌转化”相关差异基因进行基因本体论 (gene ontology, GO) 和京都基因与基因组百科全书 (Kyoto encyclopedia of genes and genomes, KEGG) 通路富集分析。GO 分析涵盖生物过程 (biological process, BP)、分子功能 (molecular function, MF) 和细胞组分 (cellular component, CC) 3 个层面。富集结果的显著性以  $P_{adj} < 0.05$  被认为具有统计学意义。

### 1.4 工具变量筛选及因果效应评估

为保证 MR 分析结果的可靠性, 遵循以下标准筛选工具变量。首先, 选择与基因表达水平在全血 eQTL 数据中达到全基因组显著关联 ( $P < 5 \times 10^{-8}$ ) 的单核苷酸多态性 (single nucleotide polymorphism, SNPs); 其次, 为消除连锁不平衡的影响, 使用 PLINK 软件进行聚类处理 (参数设置为  $R^2 < 0.1$ 、 $kb = 10\ 000$ ), 仅保留每个基因座中关联性最强的 SNPs 作为代表; 最后, 为排除弱工具变量偏倚, 计算每个 SNPs 的  $F$  统计量<sup>[23]</sup>,  $F = R^2(1 - R^2) \times (N - 2)$ , 并剔除  $F$  值 <10 的 SNPs。其中  $r^2$  指该 SNP 对暴露因素 (即基因表达水平) 的解释方差,  $N$  是指 eQTLGen 数据中对应基因的有效样本量。

采用双样本 MR 分析估计基因表达对胃癌风险的因果效应。对于每个基因, 根据有效工具变量数量采用不同策略 [若仅有 1 个有效 SNP, 则以其 Wald 比率作为因果效应估计; 若包含多个有效 SNPs, 则通过逆方差加权法 (inverse variance weighting, IVW) 合并各 SNPs 的 Wald 比率] 得到总体效应估计<sup>[24]</sup>。为评估结果稳健性, 对多 SNP 基因进一步采用加权中位数法 (weighted median, WM) 和 MR-Egger 回归法进行敏感性分析<sup>[24]</sup>。通过 MR-Egger 截距检验评估水平多效性, 通过 Cochran's  $Q$  检验评估工具变量间的异质性<sup>[25]</sup>。上述分析主要基于 R 软件 (版本 4.2.0) 中的 TwoSampleMR 包 (版本 0.5.6) 完成。阳性因果关系的判定需同时满足以下条件: (1) 主分析方法检验  $P < 0.05$ ; (2) WM 法与 MR-Egger 法估计的效应方向与 IVW 法一致; (3) MR-Egger 截距检验  $P >$

0.05, 提示无显著定向多效性; (4) 留一法敏感性分析未导致因果效应估计发生方向性逆转。

通过上述分析筛选出的基因被视为与胃癌风险存在潜在因果关系的因果关联基因, 将其与“炎-癌转化”相关差异表达基因取交集得到候选关键基因, 用于后续的分析。

### 1.5 候选关键基因的生存分析和免疫浸润分析

从 Kaplan-Meier Plotter 在线数据库 ([https://kmplot.com/analysis/index.php?p=service&cancer=pancancer\\_rnaseq](https://kmplot.com/analysis/index.php?p=service&cancer=pancancer_rnaseq)) 获取胃癌患者转录组与随访数据。利用该数据库内置算法确定每个核心基因表达水平的最佳截断值, 并据此将患者队列划分为高表达组与低表达组。采用 Kaplan-Meier 法绘制生存曲线, 并使用 Log-rank 检验比较两组患者在总生存期的差异, 以评估该基因表达水平对预后的影响<sup>[26]</sup>。利用 TIMER 数据库的算法, 量化胃癌样本中各类免疫细胞 (如 B 细胞、CD4<sup>+</sup> T 细胞、CD8<sup>+</sup> T 细胞、中性粒细胞、巨噬细胞及树突状细胞等) 的浸润丰度。随后, 计算每个核心基因的表达水平与各免疫细胞浸润水平之间的 Spearman 相关系数, 以系统性揭示其与肿瘤免疫微环境特征的相关性<sup>[27]</sup>。

### 1.6 S100A12 在胃癌进展中的机制

为了探究免疫细胞是否介导 S100A12 与胃癌易感性之间的因果关系, 整合了 731 个免疫细胞相关表型的遗传数据<sup>[28]</sup>, 并采用两步法 MR 探索免疫细胞在基因表达与胃癌发病中的潜在中介作用<sup>[29-30]</sup>。分析遵循“先筛选后中介”的逻辑: 首先, 以免疫细胞为暴露, 胃癌为结局, 筛选出与胃癌风险显著相关的免疫特征 ( $P < 0.05$ ); 随后, 以 S100A12 为暴露, 上述筛选出的免疫特征为结局, 进一步筛选受 S100A12 调控的免疫特征 ( $P < 0.05$ )。

### 1.7 全表型组 MR 分析

基于阿斯利康 PheWAS 门户网站 (<https://azphewas.com/>) 和 PheWeb 数据库 (<https://pheweb.org/>) 进行 PheWAS 分析, 评估潜在治疗靶点的多效性效应以及可能的不良反应<sup>[31]</sup>。原始研究中使用了约 15 500 个二元型和 1 500 个连续型表型的数据。外显子测序子组中的个体来自英国生物银行。原始研究中包含了该综合技术的详细描述<sup>[32]</sup>。这项综合的 PheWAS 分析为理解遗传基础的复杂性特征以及评估药物靶点的安全性和有效性提供了依据。

### 1.8 基于分子对接的中药小分子虚拟筛选

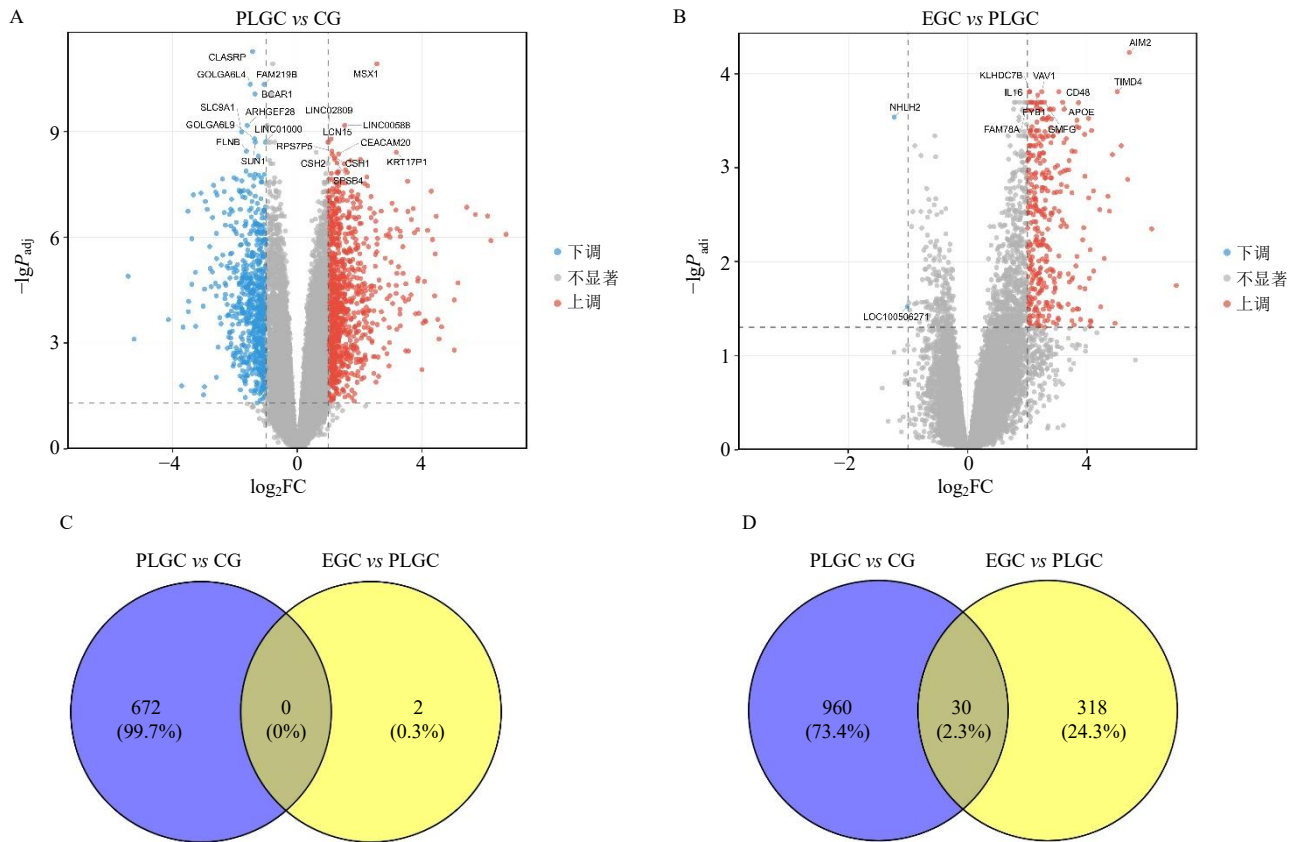
采用基于结构的虚拟筛选策略寻找靶向

S100A12 蛋白的潜在中药小分子化合物。首先, 从 PDB 数据库获取人源 S100A12 蛋白的晶体结构 (PDB ID: 2M9G), 并利用 AutoDock Tools 软件进行预处理, 包括添加极性氢原子、修复缺失侧链、补充结晶水分子并执行能量最小化, 以获得稳定的受体构象。随后, 基于 TCMSP 2.3 数据库 (<https://www.tcmsp-e.com/tcmsp.php>) 构建中药活性成分库, 并依托 QuickVina 2 软件开展大规模分子对接。对接过程中, 以 S100A12 的活性口袋为中心定义搜索空间, 中心坐标设定为 (-10.863 7, -7.152 1, -8.711 4), 网格尺寸设置为 (14.5 Å, 15.5 Å, 13.5 Å, 1 Å=0.1 nm), 其余参数保持默认, 以结合亲和力作为筛选指标。最后, 利用 PyMOL 2.5.0 和 Discovery Studio Visualizer 2019 软件对高分复合物进行三维构象及二维相互作用图谱绘制, 以揭示关键的分子间作用力。

## 2 结果

### 2.1 差异表达基因的鉴定

对 GSE55696 数据集进行差异表达基因分析, 首先比较 PLGC 组与 CG 组, 结果共鉴定出 1 662 个差异表达基因, 其中 990 个基因表达上调, 672 个基因表达下调; EGC 组与 PLGC 组的比较中, 发现了 350 个差异表达基因, 包括 348 个上调基因和 2 个下调基因, 两组比较的火山图分别见图 1-A、B。进一步分析显示, 在上述 2 个连续发展阶段的差异基因中未发现共同下调的基因 (图 1-C), 但鉴定出共同上调基因 30 个 (图 1-D)。这 30 个基因的表达在疾病进展中均呈持续升高趋势, 因此被视为可能参与“CG-PLGC-EGC”这一恶性进展过程的关键差异表达基因。具体如下: G 蛋白信号调节因子 1 (regulator of G protein signaling 1, RGS1)、免疫球蛋白 G Fc 受体 1b (Fc fragment of Immunoglobulin G receptor 1b, FCGR1B)、巨噬细胞清道夫受体 1 (macrophage scavenger receptor 1, MSR1)、慢性血栓栓塞性肺动脉高压相关长链非编码 RNA 1 (chronic thromboembolic pulmonary hypertension-associated long non-coding RNA1, CTEPHA1)、唾液酸结合免疫球蛋白样凝集素 10 (sialic acid binding Ig like lectin 10, SIGLEC10)、核糖核酸酶 A 家族成员 2 (ribonuclease A family member 2, RNASE2)、免疫球蛋白超家族成员 6 (immunoglobulin superfamily member 6, IGSF6)、CC 趋化因子配体 3 (C-C motif chemokine ligand 3, CCL3)、黏附 G



A-胃癌前病变组与慢性胃炎组差异表达基因的火山图；B-早期胃癌组与胃癌前病变组差异表达基因的火山图；C-PLGC vs CG 下调基因的韦恩图；D-EGC vs PLGC 上调基因的韦恩图。

A-volcano plot of differentially expressed genes between precancerous lesions of gastric cancer (PLGC) group and chronic gastritis (CG) group; B-volcano plot of differentially expressed genes between early gastric cancer (EGC) group and PLGC; C-Venn diagram of downregulated genes in PLGC vs CG; D-Venn diagram of upregulated genes in EGC vs PLGC.

图1 “慢性胃炎-胃癌前病变-早期胃癌”的差异基因筛选

Fig. 1 Screening of differentially expressed genes for “chronic gastritis-precancerous lesions of gastric cancer-early gastric cancer”

蛋白偶联受体 E3 (adhesion G protein-coupled receptor E3, ADGRE3)、去整合素金属蛋白酶样蜕膜素 1(a disintegrin and metalloproteinase like decysin 1, ADAMDEC1)、白介素 2 受体  $\alpha$  亚基(interleukin 2 receptor subunit alpha, IL2RA)、S100 钙结合蛋白 A9 (S100 calcium binding protein A9, S100A9)、S100A8、S100A12、纤维胶凝蛋白 1 (ficolin 1, FCN1)、前列腺素 D2 合酶 (prostaglandin D2 synthase, PTGDS)、IgG Fc 受体 IIIa (Fc gamma receptor IIIa, FCGR3A)、抑瘤素 M (oncostatin M, OSM)、卷曲螺旋结构域蛋白 178(coiled-coil domain containing 178, CCDC178)、同源异型盒 A7 (homeobox A7, HOXA7)、甲酰肽受体 1 (formyl peptide receptor 1, FPR1)、干扰素诱导跨膜蛋白 1 (interferon induced transmembrane protein 1, IFITM1)、水通道蛋白 9(aquaporin 9, AQP9)、CXC

趋化因子受体 1 (C-X-C motif chemokine receptor 1, CXCR1)、锌指 BED 结构域蛋白 2 (zinc finger BED type containing 2, ZBED2)、C 型凝集素结构域家族 4 成员 E (C-type lectin domain family 4 member E, CLEC4E)、髓细胞核分化抗原 (myeloid cell nuclear differentiation antigen, MND4)、趋化素样因子样囊泡运输与膜连接相关 MAL 结构域跨膜结构域蛋白 2 (chemokine-like factor like MAL and related proteins for vesicle trafficking and membrane link transmembrane domain containing 2, CMTM2)、基质金属蛋白酶 9(matrix metalloproteinase 9, MMP9)、软骨肉瘤相关基因家族成员 2 (chondrosarcoma-associated gene family member 2, CSAG2)。

## 2.2 PPI 分析和功能富集分析

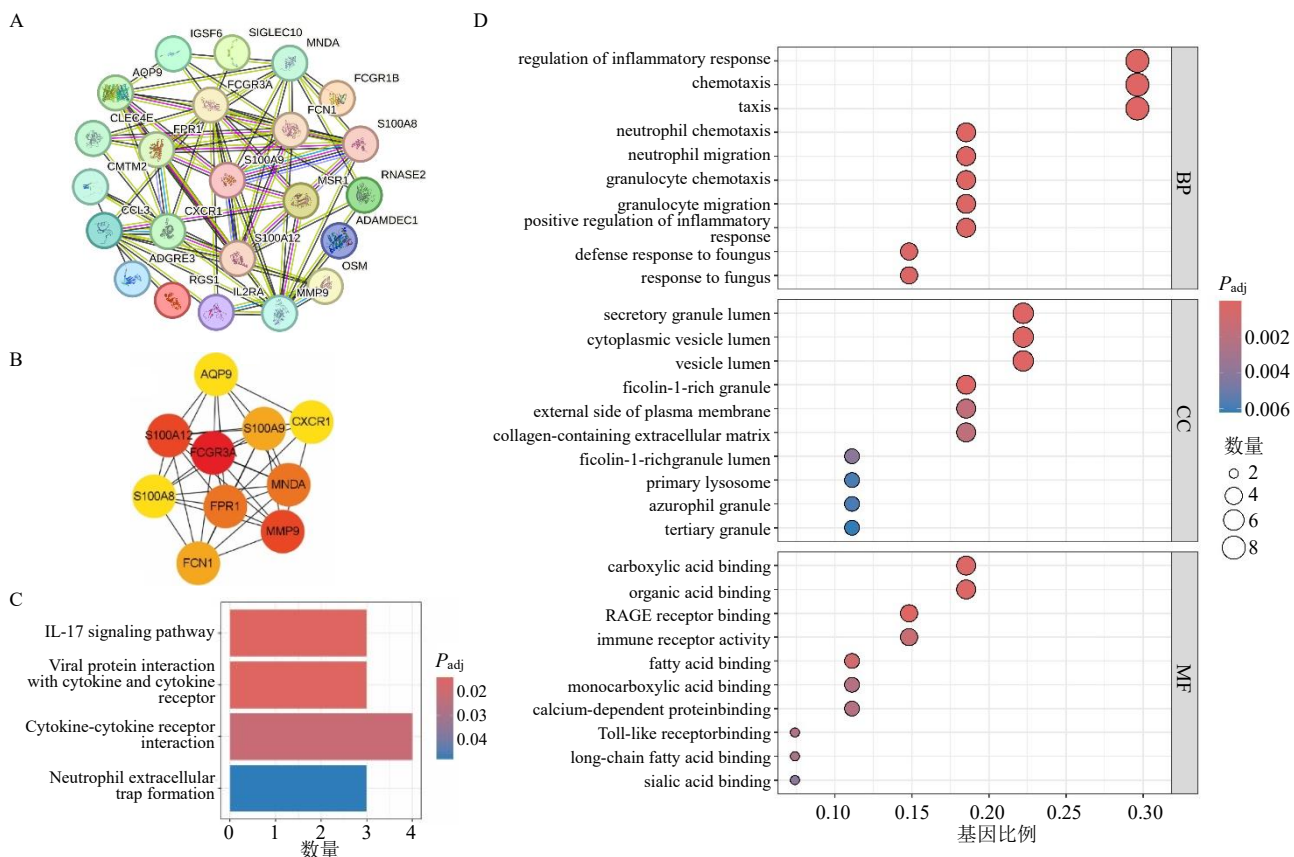
将上述 30 个“炎-癌转化”相关差异基因导入 String 数据库,进行 PPI 网络分析,隐藏游离节点得

到 23 个蛋白的互作网络, 如图 2-A 所示。使用 CytoHub 按照度 (degree) 值排名前 10 进行核心基因的筛选, 核心基因为 FCGR3A、FCN1、AQP9、MNDA、FPR1、S100A12、MMP9、S100A9、S100A8、CXCR1, 如图 2-B 所示。对上述关键基因进行 GO 和 KEGG 富集分析, KEGG 富集分析结果显示共有 4 条显著的通路 ( $P < 0.05$ ), 分别为白细胞介素-17 (interleukin-17, IL-17) 信号通路、病毒蛋白与细胞因子及细胞因子受体的相互作用、细胞因子-细胞因子受体相互作用、中性粒细胞胞外陷阱形成, 见图 2-C。GO 富集分析中 BP 共有 73 条、CC 有 14 条、MF 有 12 条满足  $P < 0.05$ , 对每个项目前 10 条进行可视化, 气泡图见图 2-D, BP 分析主要涉及炎症反应的调节、中性粒细胞趋化作用、中性粒细胞迁移、粒细胞趋化作用、对真菌的防御反应、粒细胞迁移、炎症反应的正向调节等; CC 方面主要涉及分泌颗粒腔、细胞质小泡腔囊泡腔、含胶原蛋白的细胞外基质、富含纤连蛋白-1 的颗粒腔、原生溶酶体、嗜碱性颗粒、

次级颗粒富集居多; MF 主要涉及晚期糖基化终产物受体 (receptor for advanced glycosylation end products, RAGE) 结合、羧酸结合、有机酸结合、脂肪酸结合、免疫受体活性、Toll 样受体结合、单羧酸结合、钙依赖性蛋白结合、长链脂肪酸结合、唾液酸结合、模式识别受体活性、细胞因子活性等。

### 2.3 因果关联基因的筛选

对上述 30 个“炎-癌转化”相关差异基因与 eQTL 中基因取交集, 共得到 22 个交集基因, 按照工具变量筛选标准, 使用 MR 对上述基因进行筛选, 22 个基因相关结果如图 3 所示, 其中 CXCR1、ZBED2、S100A12、S100A8、S100A9 与胃癌的发病呈正相关 [比值比 (odds ratio, OR)  $> 1$ ], 而 AQP9、IGSF6 与胃癌的发病呈负相关 ( $OR < 1$ )。与“炎-癌转化”相关差异基因核对, 得到 5 个差异表达一致且为上调的基因 (S100A8、CXCR1、S100A9、S100A12、ZBED2), 并视为胃癌的核心驱动基因, 散点图如图 4 所示。并且 S100A8、S100A9、S100A12



A-关键差异基因 PPI 网络; B-基于 CytoHub 筛选的关键核心基因; C-关键差异基因的 KEGG 富集分析条形图; D-关键差异基因的 GO 富集分析气泡图。  
A-PPI network of key differentially expressed genes; B-core key genes screened by CytoHub; C-bar chart of KEGG enrichment analysis of key differentially expressed genes; D-bubble chart of GO enrichment analysis of key differentially expressed genes.

图 2 “炎-癌转化” 相关差异基因的 PPI 网络与功能富集分析

Fig. 2 PPI network and functional enrichment analysis of differentially expressed genes related to “inflammation-cancer transformation”

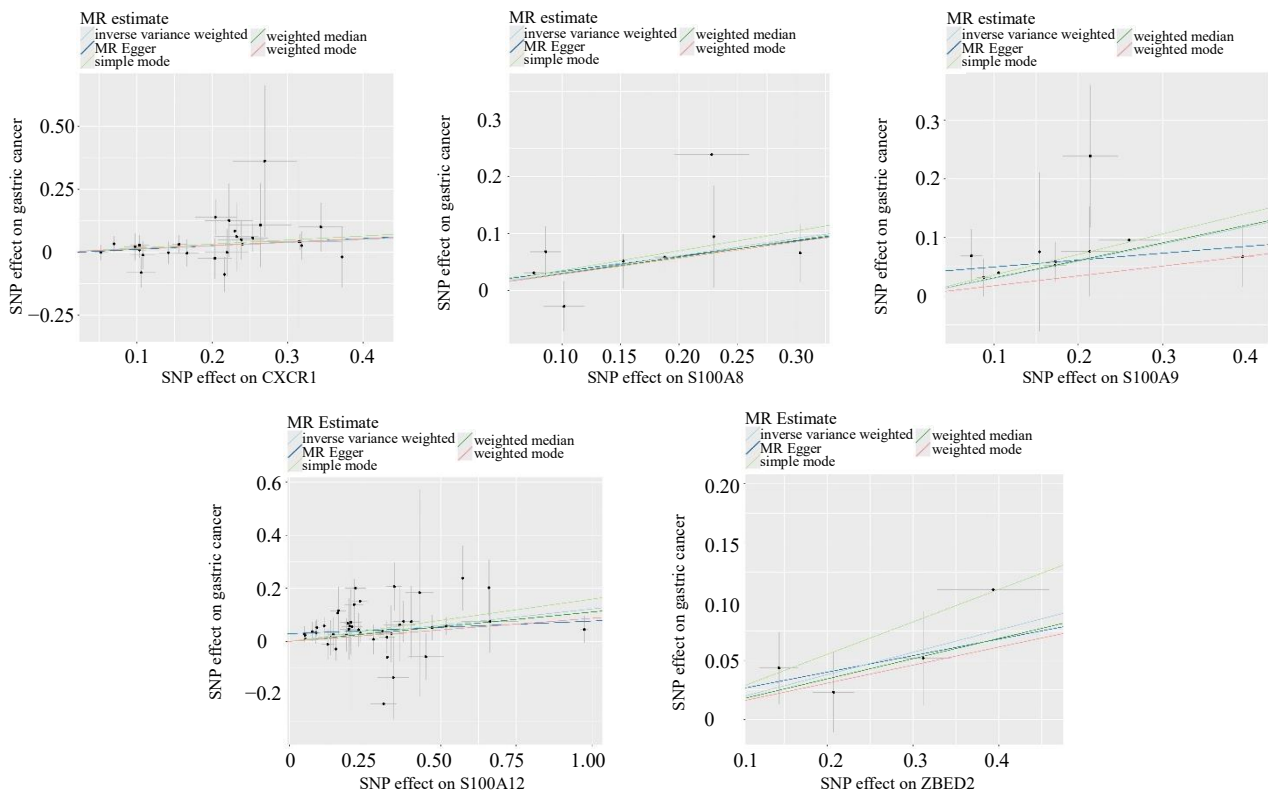
ID	nSNP	OR(95% CI)	Fstatistics	P value
S100A12	43	1.13 (1.07 - 1.20)	344.927	2.15E-05
S100A9	10	1.34 (1.13 - 1.59)	178.396	0.001
S100A8	8	1.36 (1.11 - 1.65)	168.159	0.002
AQP9	8	0.64 (0.47 - 0.87)	91.770	0.005
ZBED2	4	1.61 (1.06 - 2.43)	62.526	0.024
CXCR1	27	1.13 (1.01 - 1.27)	139.077	0.026
IGSF6	11	0.83 (0.70 - 0.99)	111.858	0.038
IFITM1	24	1.15 (1.00 - 1.33)	120.798	0.055
MSR1	31	0.91 (0.83 - 1.01)	182.654	0.075
ADAMDEC1	8	0.89 (0.73 - 1.10)	128.013	0.290
IL2RA	6	0.81 (0.53 - 1.24)	99.110	0.333
MNDA	25	1.06 (0.93 - 1.21)	169.259	0.384
CCL3	1	0.82 (0.45 - 1.48)	45.857	0.506
CLEC4E	22	0.98 (0.90 - 1.05)	333.538	0.527
MMP9	11	0.94 (0.76 - 1.15)	103.894	0.536
SIGLEC10	11	0.97 (0.83 - 1.13)	204.494	0.671
CMTM2	1	0.90 (0.50 - 1.59)	145.492	0.705
FCGR3A	2	0.96 (0.68 - 1.37)	77.812	0.831
PTGDS	12	0.98 (0.84 - 1.14)	113.610	0.832
RNASE2	41	1.00 (0.93 - 1.09)	275.352	0.931
FPR1	49	1.00 (0.94 - 1.06)	262.125	0.986
FCN1	43	1.00 (0.93 - 1.08)	249.692	0.998

OR>1 提示该基因可能增加疾病风险, OR<1 提示可能降低风险, OR=1 (虚线) 为无关联参考线。

Odds ratio (OR) greater than 1 indicates that the gene may increase disease risk, an OR less than 1 suggests it may reduce the risk, and an OR equal to 1 (dotted line) serves as the reference line for no correlation.

图 3 22 个候选基因与胃癌发病风险的森林图

Fig. 3 Forest plot of 22 candidate genes associated with gastric cancer risk



横轴为 SNP 对暴露基因的效应值, 纵轴为 SNP 对结局表型的效应值; 不同颜色线条代表不同 MR 方法的回归拟合结果, 黑色点为单个 SNP 效应估计, 误差线为 95%置信区间。

Horizontal axis represents the effect size of SNPs on exposure genes, and vertical axis represents the effect size of SNPs on outcome phenotypes; lines in different colors indicate the regression fitting results of different MR methods, black dots stand for the effect estimates of individual SNPs, and error bars represent the 95% confidence intervals.

图 4 胃癌核心驱动基因的散点图

Fig. 4 Scatter plot of core driver genes of gastric cancer

在 PPI 分析和 MR 分析中均被鉴定为关键基因。

### 2.4 核心驱动基因的生存分析和免疫浸润

由于胃癌是“炎-癌转化”的最终阶段，首先评估了 5 个核心驱动基因对胃癌患者总生存期的影响。生存分析结果显示，在 *S100A8*、*CXCR1*、*S100A9*、*S100A12* 和 *ZBED2* 个基因中，仅有 *S100A12* 的表达水平与胃癌总生存期存在显著相关性，其高表达预示着更差的预后，风险比为 1.65 (95% 置信区间: 1.16~2.35)，对数秩检验  $P=0.0045$ 。而 *S100A8*、*CXCR1*、*S100A9* 及 *ZBED2* 的表达水平与总生存期均无显著关联 ( $P>0.05$ )，所有生存分析结果如图 5-A 所示。

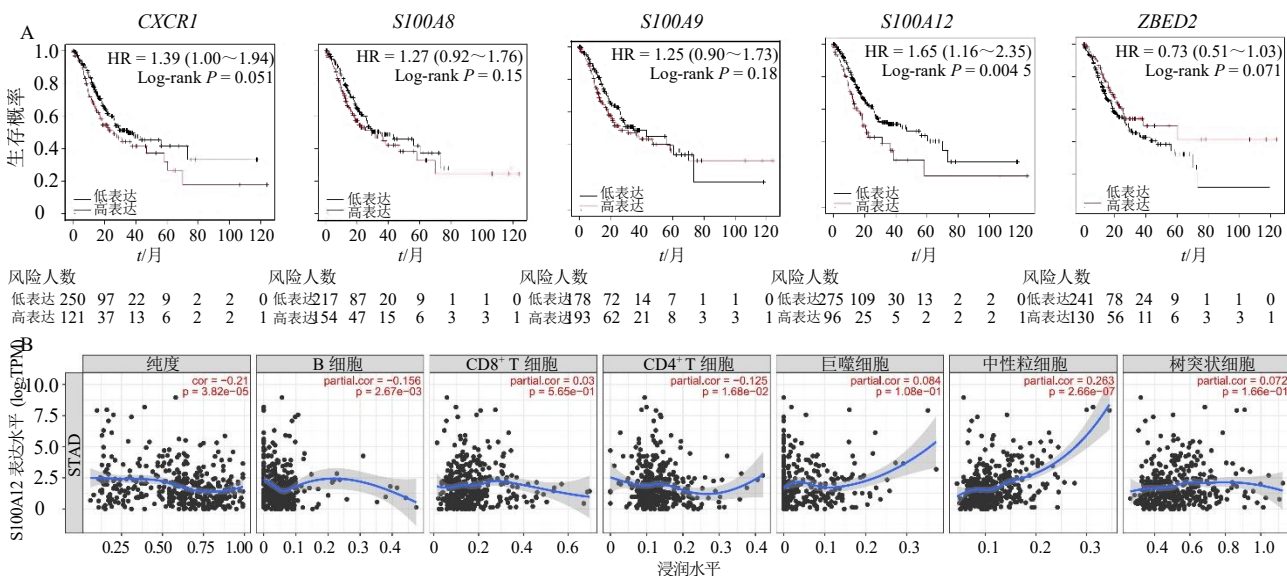
鉴于 *S100A12* 显著的预后价值，进一步探究了其表达水平与肿瘤免疫微环境中特定免疫细胞浸润程度的相关性。相关性分析发现，*S100A12* 的高表达与 B 细胞 ( $cor=-0.156$ ,  $P=2.67\times 10^{-3}$ ) 及  $CD4^+$ T 细胞 ( $cor=-0.125$ ,  $P=0.0168$ ) 的浸润水平呈负相关，而与中性粒细胞的浸润水平呈显著正相关 ( $cor=0.263$ ,  $P=2.66\times 10^{-7}$ )，与  $CD8^+$ T 细胞、巨噬细胞及树突状细胞的浸润无明显相关性 ( $P>0.05$ )，见图 5-B。

### 2.5 免疫细胞介导 *S100A12* 促进胃癌发生机制研究

鉴于 *S100A12* 在调节免疫调控中的重要作用，进一步采用 MR 进行了中介分析。结果发现 11 种与胃癌风险增加呈正相关的免疫细胞特征，包括

Treg 细胞 ( $n=2$ )、B 细胞 ( $n=7$ )、单核细胞 ( $n=1$ ) 及经典树突状细胞 ( $n=1$ )。相反，有 23 种特征与胃癌风险呈负相关，涉及成熟期 T 细胞 ( $n=1$ )、TBNK 细胞 (T 细胞、B 细胞和自然杀伤细胞 3 类免疫细胞的统称,  $n=6$ )、单核细胞 ( $n=2$ )、Treg 细胞 ( $n=4$ )、B 细胞 ( $n=5$ ) 和 cDC 细胞 ( $n=5$ )，见表 1。随后分析了 *S100A12* 对上述的免疫细胞特征的调控，结果 *S100A12* 可以调控 4 种免疫细胞特征，包括  $CD25^+CD4^+$ T 细胞占  $CD4^+$ T 百分比、 $CD4$  阳性  $CD8$  弱表达 T 细胞绝对计数 ( $CD4^+CD8^{dim}$  T cell absolute count)、 $CD4$  阳性  $CD8$  弱表达 T 细胞占淋巴细胞百分比 ( $CD4^+CD8^{dim}$  T cell %lymphocyte)、 $CD4$  阳性  $CD8$  弱表达 T 细胞占白细胞百分比 ( $CD4^+CD8^{dim}$  T cell %leukocyte)。

如图 6 所示，*S100A12* 对  $CD25^+CD4^+$ T 细胞 (即  $CD25^+CD4^+$ T 细胞% $CD4^+$ T 细胞) 的因果效应为负向 ( $OR<1$ )，提示 *S100A12* 表达升高会显著抑制该群细胞的浸润水平。而中介分析的第 1 步已证实， $CD25^+CD4^+$ T 细胞与胃癌发病风险呈显著负相关 ( $OR<1$ ，表 1)，即该群细胞浸润水平越高，胃癌风险越低。2 个步骤的效应方向高度一致，共同提示 *S100A12* 可能通过负向调控  $CD25^+CD4^+$ T 细胞来促进胃癌的发生。



A-5 个核心驱动基因的 Kaplan-Meier 生存曲线；B-S100A12 表达水平与胃癌组织中多种免疫细胞浸润丰度的相关性分析。

A-Kaplan-Meier survival curves of five core driver genes; B-correlation analysis between *S100A12* expression level and infiltration abundance of various immune cells in gastric cancer tissues.

图 5 核心驱动基因的预后分析以及 *S100A12* 的免疫相关性分析

Fig. 5 Prognostic analysis of core driver genes and immune correlation analysis of *S100A12*

表 1 与胃癌发病风险显著相关的免疫细胞特征

Table 1 Immune cell characteristics significantly associated with risk of gastric cancer onset

免疫细胞特征	nSNP	OR (95% CI)	P 值
髓系树突状细胞表面 CCR2 表达 (CCR2 on myeloid dendritic cell)	10	0.864 (0.795~0.940)	0.001
CD28 阳性 CD45RA 阳性 CD8 阳性 T 细胞表面 CD8 表达 (CD8 on CD28 <sup>+</sup> CD45RA <sup>+</sup> CD8 <sup>+</sup> T cell)	41	0.917 (0.872~0.965)	0.001
高表达 CD25 的 CD4 阳性 T 细胞占 CD4 阳性 T 细胞百分比 (CD25 <sup>+</sup> CD4 <sup>+</sup> T cell%CD4 <sup>+</sup> T cell)	3	0.737 (0.601~0.904)	0.003
CD8 阳性 T 细胞表面 CD8 表达 (CD8 on CD8 <sup>+</sup> T cell)	15	0.872 (0.792~0.960)	0.005
CD14 阴性 CD16 阳性单核细胞表面 CD16 表达 (CD16 on CD14 <sup>-</sup> CD16 <sup>+</sup> monocyte)	79	0.957 (0.927~0.987)	0.006
CD39 阳性静息型 CD4 调节性 T 细胞绝对计数(CD39 <sup>+</sup> resting CD4 regulatory T cell absolute count)	79	1.043 (1.011~1.075)	0.008
CD14 阳性 CD16 阴性单核细胞绝对计数 (CD14 <sup>+</sup> CD16 <sup>-</sup> monocyte absolute count)	3	1.420 (1.096~1.841)	0.008
记忆 B 细胞表面 CD27 表达 (CD27 on memory B cell)	24	1.093 (1.022~1.168)	0.010
CD4 阳性 CD8 弱表达 T 细胞绝对计数 (CD4 <sup>+</sup> CD8 dim T cell absolute count)	9	0.864 (0.774~0.965)	0.010
CD4 阳性 CD8 弱表达 T 细胞占淋巴细胞百分比 (CD4 <sup>+</sup> CD8 dim T cell %lymphocyte)	13	0.862 (0.768~0.967)	0.011
CD4 调节性 T 细胞占 CD4 阳性 T 细胞百分比 (CD4 regulatory T cell %CD4 <sup>+</sup> T cell)	4	0.777 (0.640~0.945)	0.011
CD4 阳性 CD8 弱表达 T 细胞占白细胞百分比 (CD4 <sup>+</sup> CD8 dim T cell %leukocyte)	12	0.855 (0.757~0.966)	0.012
IgD 阴性 CD38 弱表达 B 细胞表面 CD27 表达 (CD27 on IgD <sup>-</sup> CD38 dim B cell)	31	1.072 (1.015~1.132)	0.013
HLA-DR 阳性自然杀伤细胞侧向散射光强度 (SSC-A on HLA DR <sup>+</sup> natural killer)	17	0.907 (0.840~0.980)	0.013
髓系树突状细胞占总树突状细胞百分比 (myeloid dendritic cell %dendritic cell)	16	1.112 (1.022~1.210)	0.014
CD14 阳性 CD16 阳性单核细胞表面 CD16 表达 (CD16 on CD14 <sup>+</sup> CD16 <sup>+</sup> monocyte)	31	0.938 (0.890~0.988)	0.016
CD62L 阳性浆细胞样树突状细胞表面 CCR2 表达 (CCR2 on CD62L <sup>+</sup> plasmacytoid dendritic cell)	7	0.880 (0.792~0.977)	0.017
浆细胞样树突状细胞表面 CCR2 表达 (CCR2 on plasmacytoid dendritic cell)	7	0.880 (0.792~0.977)	0.017
IgD 阳性 CD38 阴性 B 细胞表面 CD25 表达 (CD25 on IgD <sup>+</sup> CD38 <sup>-</sup> B cell)	10	0.874 (0.782~0.977)	0.018
IgD 阳性 CD38 阴性未转换记忆 B 细胞表面 CD27 表达 (CD27 on IgD <sup>+</sup> CD38 <sup>-</sup> unswitched memory B cell)	16	1.074 (1.011~1.140)	0.021
效应记忆型 CD4 阳性 T 细胞占总 T 细胞百分比 (effector memory CD4 <sup>+</sup> T cell %T cell)	11	0.875 (0.781~0.980)	0.021
单核细胞表面 CD62L 表达 (CD62L on monocyte)	9	0.895 (0.815~0.984)	0.022
CD39 阳性静息型 CD4 调节性 T 细胞占 CD4 调节性 T 细胞百分比 (CD39 <sup>+</sup> resting CD4 regulatory T cell %CD4 regulatory T cell)	85	1.034 (1.005~1.064)	0.022
B 细胞表面 CD25 表达 (CD25 on B cell)	13	0.900 (0.822~0.987)	0.025
转换型记忆 B 细胞表面 CD27 表达 (CD27 on switched memory B cell)	33	1.062 (1.007~1.120)	0.027
IgD 阴性 CD38 阴性 B 细胞表面 CD27 表达 (CD27 on IgD <sup>-</sup> CD38 <sup>-</sup> B cell)	24	1.079 (1.007~1.155)	0.030
CD4 调节性 T 细胞占总 T 细胞百分比 (CD4 regulatory T cell %T cell)	3	0.762 (0.596~0.974)	0.030
浆细胞样树突状细胞占总树突状细胞百分比 (plasmacytoid dendritic cell %dendritic cell)	16	0.910 (0.835~0.993)	0.033
IgD 阳性 CD24 阳性 B 细胞表面 CD27 表达 (CD27 on IgD <sup>+</sup> CD24 <sup>+</sup> B cell)	20	1.077 (1.005~1.156)	0.037
初始成熟 B 细胞表面 CD25 表达 (CD25 on naive-mature B cell)	12	0.913 (0.837~0.996)	0.039
CD24 阳性 CD27 阳性 B 细胞表面 CD27 表达 (CD27 on CD24 <sup>+</sup> CD27 <sup>+</sup> B cell)	34	1.058 (1.003~1.117)	0.040
CD4 阳性 T 细胞侧向散射光强度 (SSC-A on CD4 <sup>+</sup> T cell)	9	0.845 (0.718~0.994)	0.042
IgD 阳性 CD24 阴性 B 细胞表面 CD25 表达 (CD25 on IgD <sup>+</sup> CD24 <sup>-</sup> B cell)	13	0.917 (0.842~0.999)	0.046
转换型记忆 B 细胞表面 CD25 表达 (CD25 on switched memory B cell)	10	0.923 (0.852~0.999)	0.048

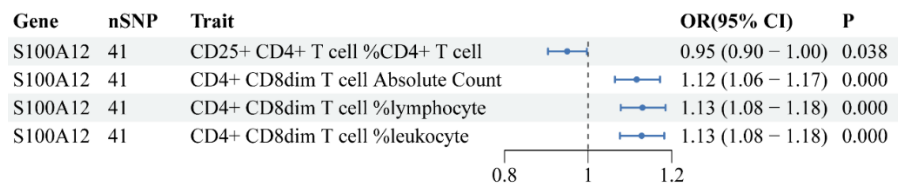


图 6 S100A12 与免疫细胞之间的因果关系

Fig. 6 Causal relationship between S100A12 and immune cells

## 2.6 全表型组关联研究

利用 PheWAS Portal 数据库进行全表型组 MR 分析,以确定靶向 S100A12 可能产生的副作用。结果显示,在全基因组显著性水平 ( $P < 5 \times 10^{-8}$ ) 下, S100A12 与 PheWAS Portal 数据库(图 7)中的其他表型之间均未发现显著相关性。这些结果强化了研究结果的有效性,并提示如果以 S100A12 为靶点进行治疗,发生药物不良反应或意外的水平多效性的风险较低。

## 2.7 靶向中药筛选

基于结构的虚拟筛选结果显示,多种中药活性小分子化合物能够稳定结合于 S100A12 的活性口袋区域。以结合亲和力作为筛选标准,结合能  $< -6$  kcal/mol (1 cal = 4.2 J) 的化合物共有 342 个,结合能排名前 15 的小分子化合物见表 2。进一步筛选发现,结合能  $\leq -7$  kcal/mol 的化合物共有 6 个,分别为 C-箭毒碱、酸浆苦味素 D、原金丝桃素、酸浆苦味素 B、前杯苋甾酮和原伪金丝桃素(图 8)。其中, C-箭毒碱、酸浆苦味素 D 表现出较优的结合能力,其结合能分别为  $-7.8$ 、 $-7.5$  kcal/mol。

三维结构分析结果表明, C-箭毒碱能够稳定嵌入 S100A12 的活性口袋,并与多个关键氨基酸残基形成相互作用。相比之下,酸浆苦味素 D 在结合口袋内形成更加丰富的氢键网络。二维相互作用分析显示, C-箭毒碱分别与 LYS83、GLN72、GLU73 以

及 THR27 等关键残基形成多重氢键作用,其相互作用距离约为  $1.82 \sim 2.74$  Å。这些氢键相互作用有助于增强配体与 S100A12 之间的结合稳定性。结果提示上述候选化合物能够通过氢键及  $\pi$  相关相互作用稳定结合于 S100A12 的活性口袋区域,可能对该蛋白的功能产生潜在调控作用。

## 3 讨论

本研究通过整合多阶段转录组差异分析与 MR 因果推断,系统筛选并验证了驱动胃黏膜“炎-癌转化”进程的关键基因。结果发现该进程中持续上调基因有 30 个,5 个与胃癌发病存在潜在因果关系的驱动基因 (S100A8、CXCR1、S100A9、S100A12、ZBED2)。其中, S100A12 因其与患者不良预后的显著关联及独特的免疫调控特征,被确立为核心枢纽。进一步的机制研究阐明, S100A12 可能通过调控如  $CD25^+CD4^+$  T 细胞等特定免疫细胞,介导其促癌的因果效应,驱动免疫微环境向促癌表型重塑。

S100A12 是主要由中性粒细胞分泌的一种损伤相关分子模式蛋白,作为连接固有免疫与慢性炎症的关键介质<sup>[33]</sup>,在幽门螺杆菌感染的胃黏膜中即被显著激活,且表达水平与组织炎症评分呈正相关<sup>[34]</sup>。本研究进一步证实, S100A12 在“炎-癌转化”序列中表达持续升高,为其作为胃癌早期预警标志物提供了分子基础。S100A12 不仅展现出良好的诊断效能——由 S100A12 等组成的 4 蛋白模型区分胃癌与

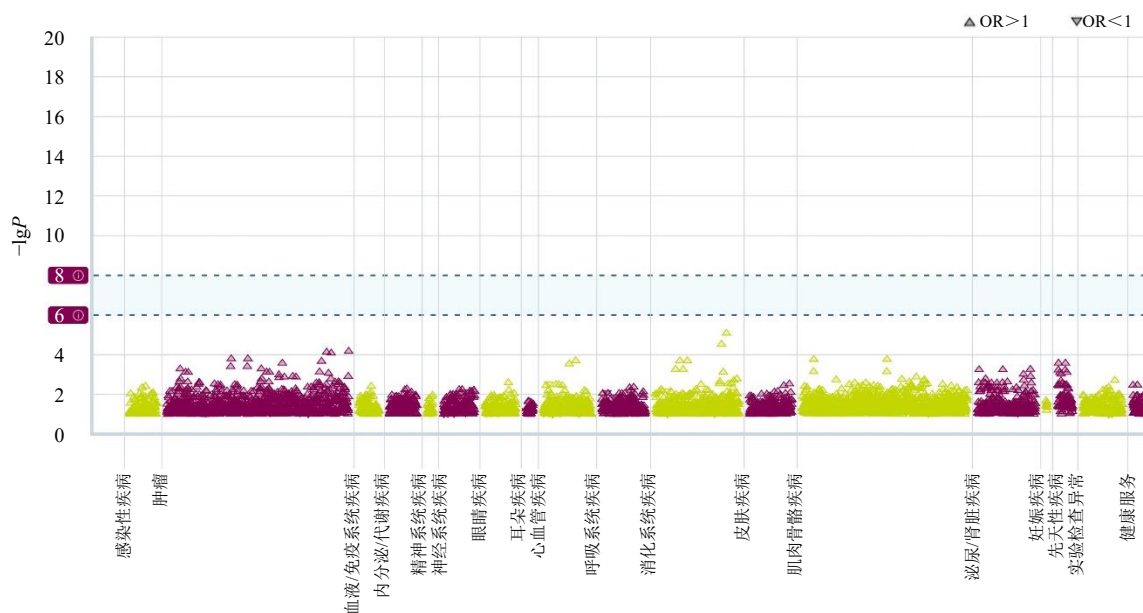


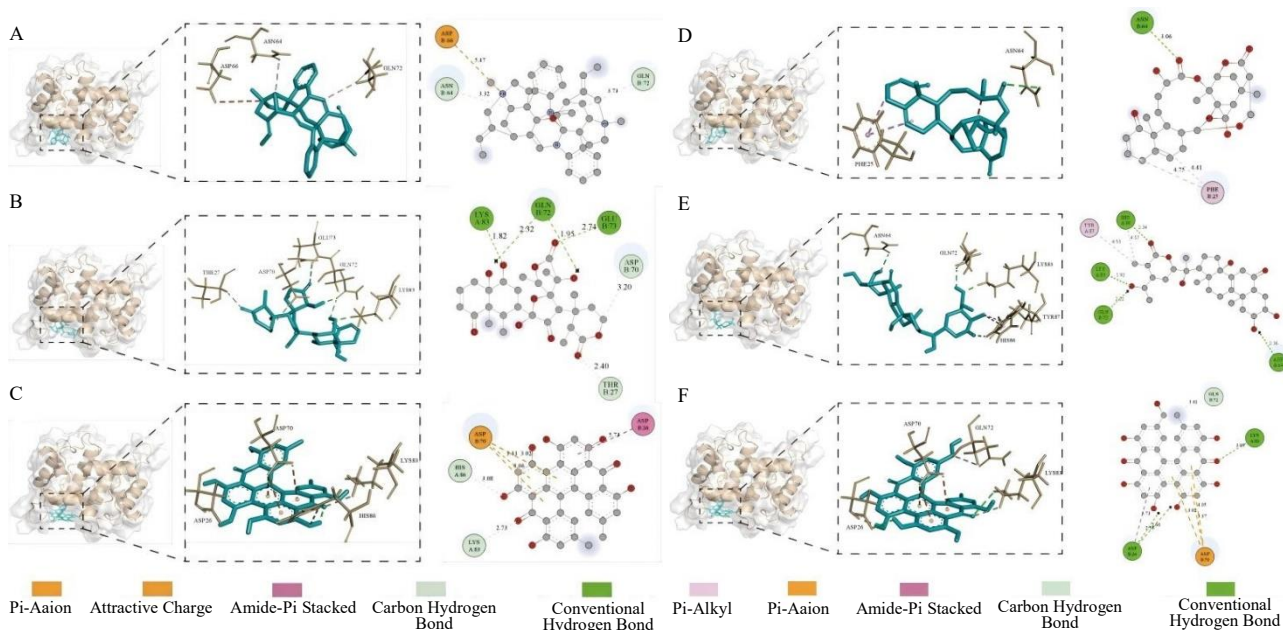
图 7 S100A12 的全表型组 MR 分析结果

Fig. 7 Results of full phenome-wide MR analysis of S100A12

表 2 靶向 S100A12 蛋白的高亲和力中药小分子化合物 (前 15 位)

Table 2 Small molecule compounds with high-affinity in traditional Chinese medicine targeting S100A12 protein (top 15)

MOL_ID	化合物名称	结合能/ (kcal·mol <sup>-1</sup> )	来源中药
MOL002041	C-箭毒碱 (C-curarine)	-7.8	侧柏叶
MOL007236	酸浆苦味素 D (physalin D)	-7.5	锦灯笼
MOL003350	原金丝桃素 (protohypericin)	-7.1	连翘
MOL007238	酸浆苦味素 B (physalin B)	-7.1	锦灯笼
MOL012296	前杯菟甾酮 (precyasterone)	-7.0	川牛膝
MOL003349	原伪金丝桃素 (protopseudohypericin)	-7.0	连翘
MOL002252	掌叶大黄素 B (palmidin B)	-6.9	大黄
MOL003424	准番木瓜碱 (pseudocarpaine)	-6.9	马钱子
MOL002583	3-O-苯甲酰基-20-脱氧巨大戟醇 (3-O-benzoyl-20-deoxyingenol)	-6.9	甘遂
MOL002067	金丝桃素 (hypericin)	-6.9	垂盆草、地耳草、连翘、贯叶金丝桃
MOL011098	9α,13α-二羟基-1-异亚丙基菘蓝素 A (9α,13α-dihydroxyl-isopropylidenylisatisine A)	-6.9	蓼大青叶、青黛
MOL006277	臭椿双内酯 (shinjudilactone)	-6.9	椿皮
MOL000988	没药甾酮 [4,17(20)-cis-pregnadiene-3,16-dione]	-6.9	没药
MOL009078	大叶骨碎补苷 A (davallioside a qt)	-6.9	骨碎补
MOL009398	羽扇豆烷-3-酮 (lupan-3-one)	-6.8	百部



A-S100A12-C-箭毒碱; B-S100A12-酸浆苦味素 D; C-S100A12-原金丝桃素; D-S100A12-酸浆苦味素 B; E-S100A12-前杯菟甾酮; F-S100A12-原伪金丝桃素。

A-S100A12-C-curarine; B-S100A12-physalin D; C-S100A12-protohypericin; D-S100A12-physalin B; E-S100A12-precyasterone; F-S100A12-protopseudohypericin.

图 8 靶向 S100A12 的小分子化合物 (结合能 ≤ -7 kcal·mol<sup>-1</sup>) 的分子对接示意图

Fig. 8 Schematic diagram of molecular docking of small molecule compounds targeting S100A12 (binding energy ≤ -7 kcal·mol<sup>-1</sup>)

非癌病例的受试者工作特征曲线下面积达 94.1%~98.2%<sup>[35]</sup>,且在膀胱癌、结直肠癌等多种肿瘤中亦具诊断潜力<sup>[36-39]</sup>。同时其具有预后价值,高表达与胃癌患者总生存期缩短显著相关,这与 S100A12 促进肿瘤进展的跨癌种机制报道相一致<sup>[40-42]</sup>。上述证据共同确立了 S100A12 从早期炎症到晚期癌症进程中持续的关键分子地位。

S100A12 能够募集肥大细胞、单核细胞等免疫细胞至炎症部位,为其在肿瘤微环境重塑中的核心作用提供了机制基础<sup>[33]</sup>。本研究发现 S100A12 表达与 CD25<sup>+</sup>CD4<sup>+</sup> T 细胞浸润水平呈负相关,中介分析提示其可能通过负向调控这群 T 细胞来驱动胃癌发生,这是本研究发现的核心机制。结合文献,推测 S100A12 对 Treg 的负向调控可能通过以下途径实现:一方面, S100A12 与 RAGE 结合激活核因子- $\kappa$ B (nuclear factor- $\kappa$ B, NF- $\kappa$ B) 通路,诱导趋化因子产生,改变 Treg 的趋化环境;另一方面, S100A12 促进中性粒细胞募集和活化,释放大炎症介质,营造不利于 Treg 存活和扩增的炎性微环境<sup>[33]</sup>。值得注意的是, Treg 在胃癌发生中具有阶段性双重角色:癌前阶段发挥抑制炎症的保护作用,晚期则转为促进免疫逃逸<sup>[43]</sup>。因此, S100A12 的促癌效应可能在于其异常高表达提前瓦解了 Treg 在早期炎症阶段的保护性屏障,导致局部炎症持续加剧、组织损伤累积,最终推动“炎-癌”转化。

值得注意的是, S100A12 的促癌效应根植于一个更广泛的分子网络之中,尤其与同属 S100 钙结合蛋白家族的 S100A8、S100A9 存在密切协同关系。本研究的 PPI 网络显示, S100A8、S100A9 与 S100A12 之间存在直接的蛋白互作,三者均处于网络核心位置;功能富集分析表明这些基因共同富集于炎症反应调节、中性粒细胞趋化、RAGE 受体结合等与“炎-癌转化”密切相关的生物学过程和信号通路。这些结果提示, S100A8/A9 与 S100A12 可能通过以下方式协同驱动胃癌发生:首先, S100A8/A9 常以异源二聚体形式存在,与 S100A12 共同作为损伤相关分子模式蛋白,通过识别 RAGE、Toll 样受体 (Toll-like receptor, TLR) 等模式识别受体激活下游 NF- $\kappa$ B 信号通路,放大炎症反应<sup>[33]</sup>;其次,三者均可由活化的中性粒细胞分泌,形成局部“炎症放大环路”,相互诱导表达、协同促进中性粒细胞募集<sup>[33]</sup>;最后, CXCR1 作为介导中性粒细胞趋化的关键受体,其共同上调进一步加剧了这一过程。

在转化医学层面,本研究的意义具有双重性。首先,全表型组 MR 分析未发现靶向 S100A12 会显著增加其他复杂疾病的风险,提示其作为治疗靶点的潜在安全性。其次,基于此靶点开展的“从基因到药物”反向筛选,成功从 TC MSP 中药化合物库中鉴定出多个与 S100A12 具有高预测结合亲和力的天然小分子(如酸浆苦味素 B、前杯苋甾酮等),其中部分化合物在研究中已被证实具有抗肿瘤活性<sup>[44]</sup>,这为开发干预 PLGC 的潜在先导化合物奠定了基础。

本研究存在一定局限性。第一,核心机制结论“S100A12 通过负向调控 CD25<sup>+</sup>CD4<sup>+</sup> T 细胞介导促癌效应”主要基于 2 步法 MR 中介分析的统计学推断。该方法虽能有效控制混杂因素、提示因果关系,但最终确立该机制仍需来自细胞或动物模型的直接功能实验证据(如基因敲除/过表达后观察 Treg 变化及肿瘤表型)。第二, MR 分析中使用的 *cis*-eQTL 数据源自血液,其前提假设是基因表达的遗传调控机制在血液与胃组织中具有一致性。尽管这是该领域常用方法,但可能不完全适用于 S100A12 这类在特定炎症部位高表达的基因,未来的研究若能整合胃黏膜组织的 eQTL 数据,将提供更精确的评估。第三,差异基因筛选仅基于 GSE55696 单数据集,虽完整覆盖“炎-癌转化”各阶段且后续经遗传学与预后数据多维验证,但仍需独立数据集外部验证。第四,分子对接筛选仅为计算机模拟预测,化合物的实际生物活性、药动学特性及体内安全性亟需实验验证;第五,尽管全表型组 MR 分析提示靶向 S100A12 的潜在脱靶风险较低,但其长期干预的安全性仍需在动物模型中全面评估。

综上所述,本研究通过系统性分析,不仅将 S100A12 确立为连接慢性炎症、免疫微环境失衡与胃癌发生的关键分子枢纽和潜在预后标志物,更描绘了一个以 S100 蛋白家族及 CXCR1 为核心的炎症网络在驱动“炎-癌转化”中的协同作用机制。针对该枢纽及其网络,尤其是利用本研究筛选出的高亲和力中药小分子进行干预,有望为阻断 PLGC 进展提供创新的靶向预防策略。未来的研究需在实验模型中验证 S100A12 对 CD25<sup>+</sup>CD4<sup>+</sup> T 细胞等功能的确切调控机制,并评估所筛选中药先导化合物在体内外模型中的确切疗效。

**利益冲突** 所有作者均声明不存在利益冲突

#### 参考文献

- [1] Bray F, Laversanne M, Sung H, *et al.* Global cancer

- statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. *CA*, 2024,74(3):229-263.
- [2] Han B F, Zheng R S, Zeng H M, *et al.* Cancer incidence and mortality in China, 2022 [J]. *J Natl Cancer Cent*, 2024, 4(1): 47-53.
- [3] Luan L, Yang L, Zhang Y T, *et al.* Highly sensitive multiplexed sensing of miRNAs in a gastric cancer patient's liquid biopsy [J]. *Anal Chem*, 2024, 96(50): 20015-20025.
- [4] Valero-Almingol A, Montori S, Felípez N, *et al.* Targetable pathways for drug repurposing in gastric cancer [J]. *World J Gastroenterol*, 2025, 31(39): 110986.
- [5] Correa P. Human gastric carcinogenesis: A multistep and multifactorial process: First American Cancer Society Award Lecture on Cancer Epidemiology and Prevention [J]. *Cancer Res*, 1992, 52(24): 6735-6740.
- [6] Mülder D T, Hahn A I, Huang R J, *et al.* Prevalence of gastric precursor lesions in countries with differential gastric cancer burden: A systematic review and Meta-analysis [J]. *Clin Gastroenterol Hepatol*, 2024, 22(8): 1605-1617.e46.
- [7] Hahn A I, Mülder D T, Huang R J, *et al.* Global progression rates of precursor lesions for gastric cancer: A systematic review and Meta-analysis [J]. *Clin Gastroenterol Hepatol*, 2025, 23(9): 1514-1524.
- [8] Pimentel-Nunes P, Libânio D, Marcos-Pinto R, *et al.* Management of epithelial precancerous conditions and lesions in the stomach (MAPS II): European Society of Gastrointestinal Endoscopy (ESGE), European *Helicobacter* and Microbiota Study Group (EHMSG), European Society of Pathology (ESP), and Sociedade Portuguesa de Endoscopia Digestiva (SPED) guideline update 2019 [J]. *Endoscopy*, 2019, 51(4): 365-388.
- [9] Yeoh K G, Tan P. Mapping the genomic *Diaspora* of gastric cancer [J]. *Nat Rev Cancer*, 2022, 22(2): 71-84.
- [10] Totoki Y, Saito-Adachi M, Shiraishi Y, *et al.* Multiancestry genomic and transcriptomic analysis of gastric cancer [J]. *Nat Genet*, 2023, 55(4): 581-594.
- [11] Kim J, Park C, Kim K H, *et al.* Single-cell analysis of gastric pre-cancerous and cancer lesions reveals cell lineage diversity and intratumoral heterogeneity [J]. *NPJ Precis Oncol*, 2022, 6(1): 9.
- [12] Kreitmaier P, Katsoula G, Zeggini E. Insights from multi-omics integration in complex disease primary tissues [J]. *Trends Genet*, 2023, 39(1): 46-58.
- [13] Emdin C A, Khera A V, Kathiresan S. Mendelian randomization [J]. *Jama*, 2017, 318(19): 1925.
- [14] 蒋先伟, 王明航, 李慧茹, 等. 基于孟德尔随机化和 GEO 数据库识别支气管哮喘的潜在靶点及干预中药预测 [J]. *中草药*, 2025, 56(3): 919-932.
- [15] Thanassoulis G, O'Donnell C J. Mendelian randomization: Nature's randomized trial in the post-genome era [J]. *Jama*, 2009, 301(22): 2386.
- [16] Yao C, Chen G, Song C, *et al.* Genome-wide mapping of plasma protein QTLs identifies putatively causal genes and pathways for cardiovascular disease [J]. *Nat Commun*, 2018, 9: 3268.
- [17] Zhang N Q, Li Y N, Sundquist J, *et al.* Identifying actionable druggable targets for breast cancer: Mendelian randomization and population-based analyses [J]. *EBioMedicine*, 2023, 98: 104859.
- [18] Li K, Ma X Y, Li Z H, *et al.* A natural peptide from a traditional Chinese medicine has the potential to treat chronic atrophic gastritis by activating gastric stem cells [J]. *Adv Sci*, 2024, 11(20): 2304326.
- [19] Xu X, Feng L, Liu Y, *et al.* Differential gene expression profiling of gastric intraepithelial neoplasia and early-stage adenocarcinoma [J]. *World J Gastroenterol*, 2014, 20(47): 17883-17893.
- [20] Vösa U, Claringbould A, Westra H J, *et al.* Large-scale *cis*- and *trans*-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression [J]. *Nat Genet*, 2021, 53(9): 1300-1310.
- [21] Kurki M I, Karjalainen J, Palta P, *et al.* FinnGen provides genetic insights from a well-phenotyped isolated population [J]. *Nature*, 2023, 613(7944): 508-518.
- [22] Szklarczyk D, Kirsch R, Koutrouli M, *et al.* The STRING database in 2023: Protein-protein association networks and functional enrichment analyses for any sequenced genome of interest [J]. *Nucleic Acids Res*, 2023, 51(D1): D638-D646.
- [23] Burgess S, Thompson S G, Collaboration C C G. Avoiding bias from weak instruments in Mendelian randomization studies [J]. *Int J Epidemiol*, 2011, 40(3): 755-764.
- [24] Bowden J, Davey Smith G, Haycock P C, *et al.* Consistent estimation in Mendelian randomization with some invalid instruments using a weighted Median estimator [J]. *Genet Epidemiol*, 2016, 40(4): 304-314.
- [25] Bowden J, Smith G D, Burgess S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression [J]. *Int J Epidemiol*, 2015, 44(2): 512-525.
- [26] Györfy B. Integrated analysis of public datasets for the discovery and validation of survival-associated genes in solid tumors [J]. *Innovation*, 2024, 5(3): 100625.
- [27] Li B, Severson E, Pignon J C, *et al.* Comprehensive analyses of tumor immunity: Implications for cancer immunotherapy [J]. *Genome Biol*, 2016, 17(1): 174.
- [28] Orrù V, Steri M, Sidore C, *et al.* Complex genetic signatures in immune cells underlie autoimmunity and inform therapy [J]. *Nat Genet*, 2020, 52(10): 1036-1045.
- [29] Relton C L, Davey Smith G. Two-step epigenetic Mendelian randomization: A strategy for establishing the

- causal role of epigenetic processes in pathways to disease [J]. *Int J Epidemiol*, 2012, 41(1): 161-176.
- [30] Burgess S, Daniel R M, Butterworth A S, *et al.* Network Mendelian randomization: Using genetic variants as instrumental variables to investigate mediation in causal pathways [J]. *Int J Epidemiol*, 2015, 44(2): 484-495.
- [31] Gagliano Taliun S A, VandeHaar P, Boughton A P, *et al.* Exploring and visualizing large-scale genetic associations by using PheWeb [J]. *Nat Genetics*, 2020, 52(6): 550-552.
- [32] Wang Q L, Dhindsa R S, Carss K, *et al.* Rare variant contribution to human disease in 281, 104 UK Biobank exomes [J]. *Nature*, 2021, 597(7877): 527-532.
- [33] Xia P P, Ji X D, Yan L, *et al.* Roles of S100A8, S100A9 and S100A12 in infection, inflammation and immunity [J]. *Immunology*, 2024, 171(3): 365-376.
- [34] Leach S T, Mitchell H M, Geczy C L, *et al.* S100 calgranulin proteins S100A8, S100A9 and S100A12 are expressed in the inflamed gastric mucosa of *Helicobacter pylori*-infected children [J]. *Can J Gastroenterol*, 2008, 22(5): 461-464.
- [35] Giai Gianetto Q, Michel V, Douché T, *et al.* Plasma protein biomarkers to detect early gastric preneoplasia and cancer: A prospective study [J]. *Int J Mol Sci*, 2025, 26(20): 10114.
- [36] Karihtala P, Leivonen S K, Puistola U, *et al.* Serum protein profiling reveals an inflammation signature as a predictor of early breast cancer survival [J]. *Breast Cancer Res*, 2024, 26(1): 61.
- [37] Elamin A A, Klunkelfuß S, Kämpfer S, *et al.* A specific blood signature reveals higher levels of S100A12: A potential bladder cancer diagnostic biomarker along with urinary engrailed-2 protein detection [J]. *Front Oncol*, 2020, 9: 1484.
- [38] Thierolf M, Hagmann M L, Pfeffer M, *et al.* Towards a comprehensive proteome of normal and malignant human colon tissue by 2-D-LC-ESI-MS and 2-DE proteomics and identification of S100A12 as potential cancer biomarker [J]. *Proteomics Clin Appl*, 2008, 2(1): 11-22.
- [39] Holub M, Bartáková E, Stráníková A, *et al.* Calprotectin and calgranulin C as biomarkers of pancreatic tumors: Baseline levels and level changes after surgery [J]. *Mediat Inflamm*, 2019, 2019: 6985703.
- [40] Liu Z, Pu X F. Orosomucoid 1 interacts with S100A12 and activates ERK signalling to expedite the advancement of bladder cancer [J]. *Cell Adh Migr*, 2025, 19(1): 1-11.
- [41] Wang X J, Sun Z X, Tian W, *et al.* S100A12 is a promising biomarker in papillary thyroid cancer [J]. *Sci Rep*, 2020, 10(1): 1724.
- [42] Lu C H, Liu J, Yao M Z, *et al.* Downregulation of S100 calcium binding protein A12 inhibits the growth of glioma cells [J]. *BMC Cancer*, 2020, 20(1): 261.
- [43] Negura I, Pavel-Tanasa M, Danciu M. Regulatory T cells in gastric cancer: Key controllers from pathogenesis to therapy [J]. *Cancer Treat Rev*, 2023, 120: 102629.
- [44] Ma Y M, Han W, Li J, *et al.* Physalin B not only inhibits the ubiquitin-proteasome pathway but also induces incomplete autophagic response in human colon cancer cells *in vitro* [J]. *Acta Pharmacol Sin*, 2015, 36(4): 517-527.

[责任编辑 潘明佳]