

# PCL2在乳腺癌中的表达特征及促癌机制研究

廖媛琪<sup>1</sup>, 宋丹蕊<sup>1</sup>, 马娟<sup>2</sup>, 刘仁<sup>1</sup>, 曹相玫<sup>2\*</sup>, 张森<sup>1\*</sup>

<sup>1</sup>宁夏回族自治区人民医院, 宁夏医科大学第三临床医学院, 病理科, 宁夏 银川

<sup>2</sup>宁夏医科大学基础医学院病理学系, 宁夏 银川

收稿日期: 2026年5月25日; 录用日期: 2026年6月18日; 发布日期: 2026年6月29日

## 摘要

乳腺癌(BRCA)是全球最常见的恶性肿瘤之一, 对女性健康构成重大威胁。研究表明, BRCA患者体内多梳蛋白2 (PCL2)表达显著上调。然而, 影响BRCA的PCL2遗传特征及其潜在机制仍不明确。因此, 本研究利用TIMER和TCGA数据库探讨了PCL2在BRCA中的表达水平及其临床意义。结果显示, PCL2过表达与敲低可分别引起329个和489个差异表达基因, 主要富集于细胞黏附、信号传导及肿瘤相关通路。PPI网络提示, PCL2可能通过调控CCL5、NOTCH3、LEF1、CDH1、CXCL8等基因参与细胞恶性表型相关过程。此外, 乳腺BRCA细胞系的PCL2敲低实验显示其表达水平存在高低差异, 表明PCL2可能是癌症进展的潜在靶点。

## 关键词

PCL2, 乳腺癌, 转录组测序, 上皮间质转化, 细胞增殖

# Expression Characteristics and Oncogenic Mechanism of PCL2 in Breast Cancer

Aiqi Liao<sup>1</sup>, Danrui Song<sup>1</sup>, Juan Ma<sup>2</sup>, Ren Liu<sup>1</sup>, Xiangmei Cao<sup>2\*</sup>, Miao Zhang<sup>1\*</sup>

<sup>1</sup>Department of Pathology, The Third Clinical Medical College of Ningxia Medical University, People's Hospital of Ningxia Hui Autonomous Region, Yinchuan Ningxia

<sup>2</sup>Department of Pathology, School of Basic Medicine, Ningxia Medical University, Yinchuan Ningxia

Received: May 25, 2026; accepted: June 18, 2026; published: June 29, 2026

## Abstract

Breast cancer (BRCA) is one of the most common malignant tumors worldwide and poses a major threat to women's health. Studies have shown that the expression of Polycomb-like protein 2 (PCL2)

\*通讯作者。

文章引用: 廖媛琪, 宋丹蕊, 马娟, 刘仁, 曹相玫, 张森. PCL2在乳腺癌中的表达特征及促癌机制研究[J]. 临床医学进展, 2026, 16(6): 2206-2215. DOI: 10.12677/acm.2026.1662442

is significantly upregulated in patients with BRCA. However, the genetic characteristics of PCL2 affecting BRCA and its underlying mechanism remain unclear. Therefore, this study investigated the expression level of PCL2 and its clinical significance in BRCA using the TIMER and TCGA databases. The results showed that PCL2 overexpression and knockdown induced 329 and 489 differentially expressed genes, respectively, which were mainly enriched in cell adhesion, signal transduction, and tumor-related pathways. The PPI network suggested that PCL2 might be involved in cell malignant phenotype-related processes by regulating genes such as CCL5, NOTCH3, LEF1, CDH1, and CXCL8. In addition, PCL2 knockdown experiments in breast BRCA cell lines revealed varying expression levels, indicating that PCL2 could serve as a potential target for cancer progression.

## Keywords

PCL2, Breast Cancer, RNA-Seq, Epithelial-Mesenchymal Transition, Cell Proliferation

Copyright © 2026 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. 引言

乳腺癌(BRCA)现已成为全球最常见的恶性肿瘤,其发病率已超过肺癌。根据世界卫生组织下属国际癌症研究机构的最新数据,乳腺癌在全球女性癌症相关死亡病例数中位居首位[1]。BRCA 的治疗方案包括手术、化疗、放疗、内分泌治疗和免疫疗法。然而,诸多问题依然存在:接受不充分治疗或预后不良的患者复发风险极高。因此,识别影响 BRCA 发展的潜在靶基因至关重要。

多梳组蛋白是表观遗传转录抑制因子,在发育过程中维持着决定细胞命运的关键记忆[2]。这些蛋白可形成复合物,其中多梳抑制复合物 1 和 2 (PRC1/PRC2)最为常见[3]。多梳蛋白 2 (PCL2)——亦称金属调节转录因子 2——属于负责催化 PCL 失活的蛋白家族成员。研究表明,PCL2 在胚胎干细胞(ESCs)中能将 PRC2 招募至靶基因位点[4]。ESCs 中 PCL2 的缺失会导致特定多能性相关因子中 H3K27me3 水平降低,最终引发基因组环境紊乱[5]。由于胚胎干细胞与肿瘤细胞存在诸多相似性,PCL2 的失调可能促进癌症干细胞形成,进而导致肿瘤异质性发展[6]。尽管对 PCL2 在 BRCA 中的功能已有深入认识,但通过慢病毒技术在 BRCA 中探究 PCL2 过表达的转录组效应仍鲜有研究。因此,PCL2 作为 BRCA 治疗靶点的潜力仍有待阐明。需采用更先进的全转录组测序技术并整合大规模数据以进一步揭示该机制。此外,开展基因本体论(Gene Ontology)和京都基因与基因组百科全书(KEGG)分析,将有助于揭示这些差异表达基因的生物学功能及其潜在信号通路。因此,本研究旨在利用 TIMER 和 TCGA 数据库,探究 PCL2 在 BRCA 中的表达水平及其临床意义。该研究对 PCL2 靶向 BRCA 疗法的分子机制具有重要的指导意义。

## 2. 材料与方法

### 2.1. 材料

本研究采用 BRCAMDA-MB-231 细胞系。所使用的检测试剂盒包括:Qubit 双链 DNA 检测试剂盒(Life Technologies Q328520)、miRNA 分离试剂盒(Ambion-1561)、Agency AMPure XP 试剂盒(Beckman Coulter A63881)以及 TruSeq 链特异性 mRNA 大样本制备试剂盒(Illumina RS-122-2RNA)。液体样本采集设备来自液态试剂盒研究所(Qantas) 50067;新西兰理工学院;澳大利亚理工学院。凝胶成像系统包括:Tanon 2500 凝胶成像系统、Sky Frozen 离心机(型号 ST16R, Thermo)、台式离心机 Eppendorf Centrifuge

5418R、PCR 仪 Bio-rad My Cycler、紫外分光光度计 NanoDrop 2000 (Thermo)、Magnetogen 磁力支架-96 孔 Bioanalyzer Agilent 2100、Quantifier 定量仪 Invitrogen Qubit 2.0, 以及多种型号的尖端和离心管 Axygen。

## 2.2. 方法

### 2.2.1. RNA 分离与文库构建

总 RNA 提取采用 TRIzol 试剂, 严格按照制造商说明书进行。RNA 纯度通过 NanoDrop 2000 分光光度计(Symer Fischer Science, 美国)进行评估和定量; RNA 完整性则使用 Agilent 2100 Bioanalyzer (Agilent Technologies, 美国加州圣克拉拉)进行检测。随后, 按照制造商说明, 采用 TruSeq Stranded mRNA LT 试剂盒(Illumina, 美国加州圣地亚哥)构建了 RNA 文库。

### 2.2.2. 数据预处理与质量控制

通过高通量测序获得的原始读段均为 FASTQ 格式序列。此外, 我们对原始读段进行了质量过滤, 以获取适用于后续分析的高质量读段。首先使用 Trimomatic [7]软件进行质量控制, 去除连接序列并过滤低质量碱基和 N 碱基, 从而获得高质量的清洁读段。随后使用 hisat2 [8]软件(采用默认参数)将这些清洁读段与物种参考基因组进行比对, 通过基因组比对对样本进行评估。

### 2.2.3. 基因水平定量与差异基因筛选

将清洁读段获得的结果与参考基因组进行比对后, 存储为二进制文件或 BAM 文件。随后使用 Cufflinks [9]软件对基因 FPKM [10]表达水平进行定量分析。为评估基因表达差异, 我们通过 htseqcount 软件统计各样本中的基因读段数量。数据标准化采用 DESeq (2012) R 软件包[11]中的 estimate Size Factor 函数完成, 随后使用 nbinomTest 函数计算 P 值和倍数变化值以识别差异。筛选出 P 值 < 0.05 且差异倍数 > 2 的差异表达基因(DEGs)进行进一步分析。对这些 DEGs 进行 GO (<https://www.geneontology.org/>)和 KEGG (<https://www.genome.jp/kegg/>)通路富集分析, 旨在确定受 DEGs 主要影响的生物学功能或通路。GO 分析包含三个维度: 生物过程(BP)、细胞组分(CC)和分子功能(MF); KEGG 分析用于探索潜在通路。在功能富集分析中, 所有 DEGs 均映射至 GO 数据库条目, 并以 P < 0.05 为阈值检验 DEGs 与 GO 生物功能的显著富集关系; 同时将所有 DEGs 映射至 KEGG 数据库, 以 P < 0.05 阈值鉴定出显著富集的 KEGG 通路。

### 2.2.4. 蛋白质 - 蛋白质相互作用(PPI)网络

所选的异源基因被用于构建能够相互建立网络关系的蛋白质。这些相互作用包括直接(物理层面)和间接(功能层面)关联, 整合了计算机预测结果、已知数据以及其他数据库交互信息。

### 2.2.5. 板式克隆测定法

采用计数板对生长中的 BRCA 细胞进行计数。随后, 将稳定的 hPCL2 细胞与载体细胞分别接种至两个 35 毫米细胞培养皿中, 每皿含 1000 个细胞。细胞在含 10%胎牛血清(FBS)的完全培养基中培养 12 天以进行克隆化。培养结束后, 用磷酸盐缓冲盐水洗涤细胞两次, 再用 4%多聚甲醛固定 15 分钟。弃去固定液后, 将细胞置于结晶紫溶液中染色 30 分钟。染色后的细胞进行克隆化处理并自然风干。菌落数量可通过肉眼或显微镜直接计数; 含 > 50 个细胞的球状结构需在显微镜下计数。

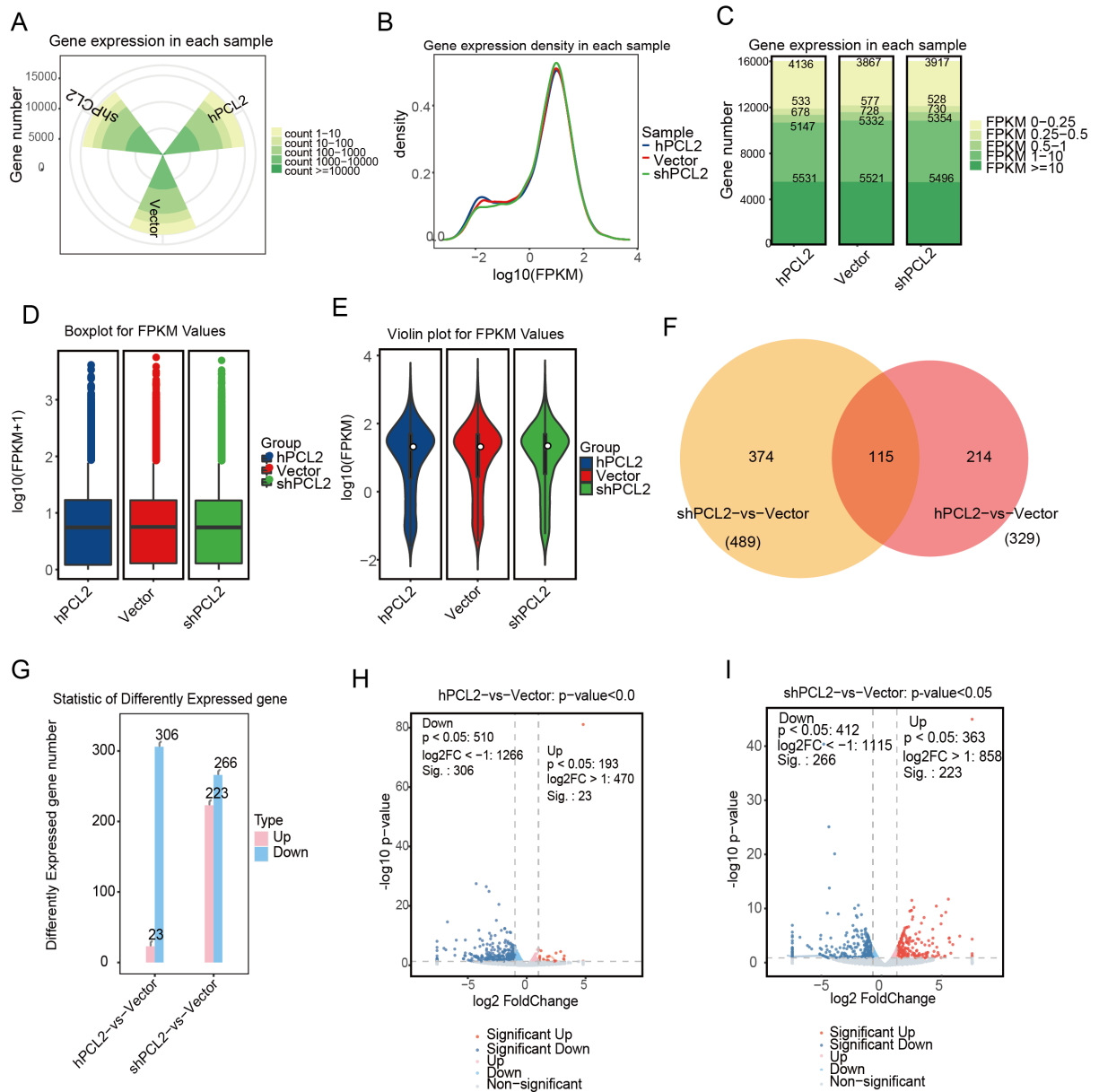
### 2.2.6. 统计分析

统计学差异采用 SPSS 22.0 软件进行分析, P < 0.05 被视为具有统计学显著性。定量数据以均值  $\pm$  标准差表示。在使用单因素或双因素方差分析(ANOVA)比较两组数据时, 数据均符合正态分布及方差齐性检验要求, 统计学显著性差异定义为 P < 0.05。测序数据分析采用 DESeq 软件, 筛选标准为 P < 0.05; 统

计图谱绘制使用 GraphPad Prism 8 软件完成。

### 3. 结果

#### 3.1. RNA-Seq 基因的表达及差异分析



**Figure 1.** Transcriptome sequencing results of gene expression and differential analysis  
**图 1.** 基因表达与差异分析的转录组测序结果

为进一步探究 PCL2 对 BRCA 的影响, 研究人员对表达 PCL2 (hPCL2)、PCL2 敲低 (shPCL2) 及空载体慢病毒转染 (vector) 的 MDA-MB-231 细胞系稳定细胞样本进行了 RNA 测序。各样本的基因计数结果如图 1(A) 所示。通过密度图、区域分布图、箱线图和小提琴图评估了基因表达水平: 密度分布曲线展示了各样本中的 FPKM 值, 反映了蛋白质编码基因的表达模式 (图 1(B)); FPKM 表达的区域分布被划分为多个区间, 便于统

计各区间内的表达基因数量(图 1(C)); 箱线图呈现了各样本基因的 FPKM 值分布(图 1(D)); 小提琴图则显示各样本基因的 FPKM 值, 表明样本间基因表达丰度普遍较高(图 1(B)~(E))。采用 Wynn 图和条形图比较了不同组别间差异表达基因(DEGs)的统计学差异(图 1(F)、图 1(G))。通过对 DEGs 分布进行火山图注释分析(图 1(H)、图 1(I)), 发现与空载体组相比, hPCL2 组共鉴定出 329 个 DEGs, 其中 23 个基因上调、306 个基因下调; shPCL2 细胞中共鉴定出 489 个 DEGs, 上调和下调基因数分别为 223 个和 266 个(图 1(H)、图 1(I))。

### 3.2. GO 2 级水平的差异表达基因分析

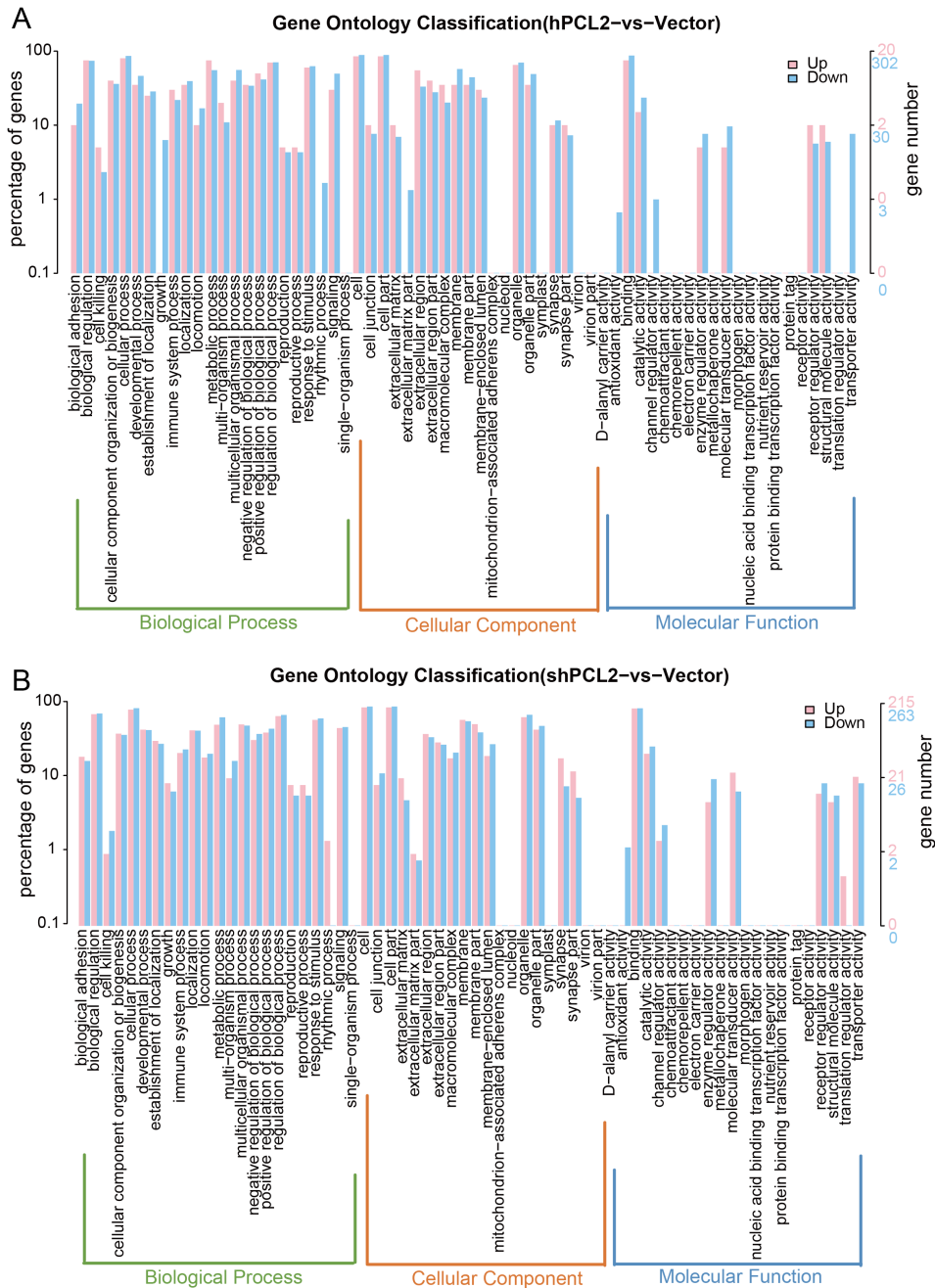


Figure 2. Comparison of differentially expressed gene distribution at the second level of GO  
图 2. GO 第 2 级差异基因分布的比较

为深入解析差异表达基因(DEGs)的功能,我们对 hPCL2 和 shPCL2 组的 DEGs 进行了 GO 富集分析。GO 通常根据功能分级分为三个层级:层级 1、层级 2 和层级 3。层级 1 包含三个条目:BP (生物过程)、CC (细胞过程)和 MF (分子功能);层级 2 包含 64 个具体条目,如生物黏附、细胞及结合等;层级 3 则涵盖数万个常规富集条目。层级 1 至层级 3 的功能范围从更具体到更通用逐步递进。GO 层级 2 的分布结果显示, hPCL2 和 shPCL2 细胞中的 DEGs 在 BP、CC 和 MF 三个层级均呈现富集现象。其中, hPCL2 组在层级 2 的多个类别中表现出显著的 DEGs 富集现象,尤其是生物调控、细胞过程和信号传导在 BP 层级集中度最高;最丰富的 CC 条目涉及细胞组分和细胞外基质;而 MF 层级中最主要的条目为分子转导活性和细胞外信号受体活性。在 shPCL2 组中, DEGs 主要集中在 BP 层级,特别是生物调控、刺激响应和信号传导领域,最丰富的条目为细胞外区域部分组分和跨膜蛋白;MF 层级中最突出的条目则是结构分子活性和转运蛋白活性(图 2(A)、图 2(B))。

### 3.3. KEGG 中差异表达基因(DEGs)的富集分析

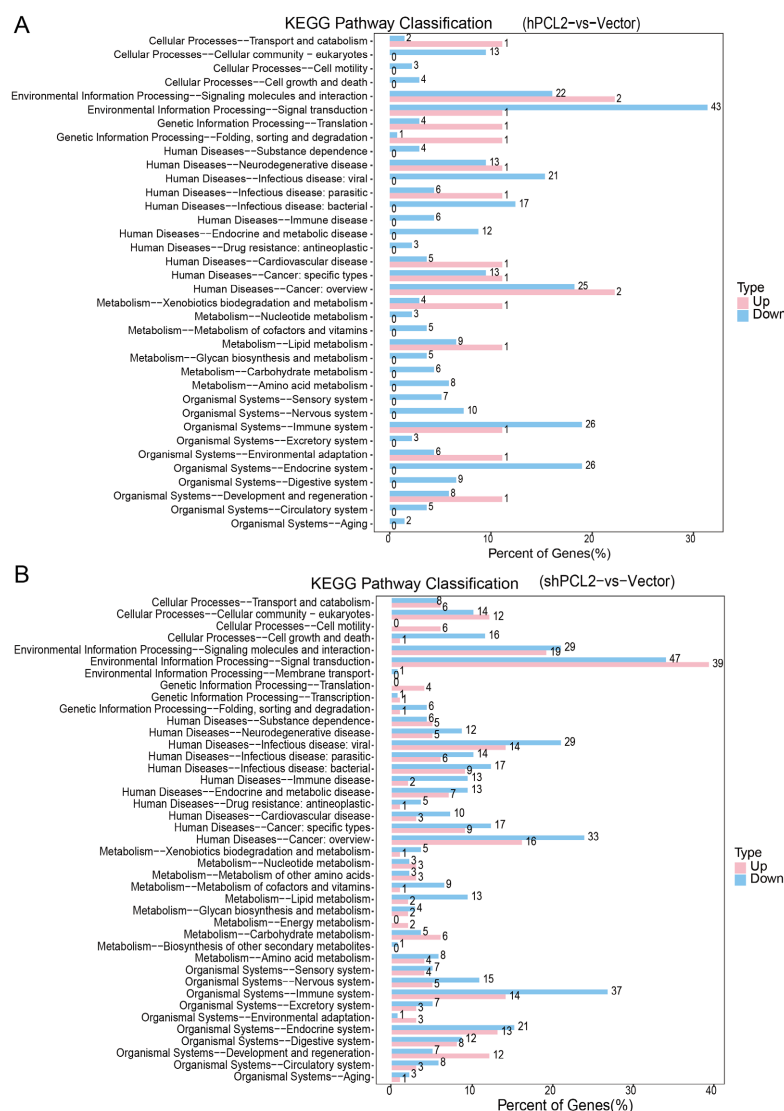
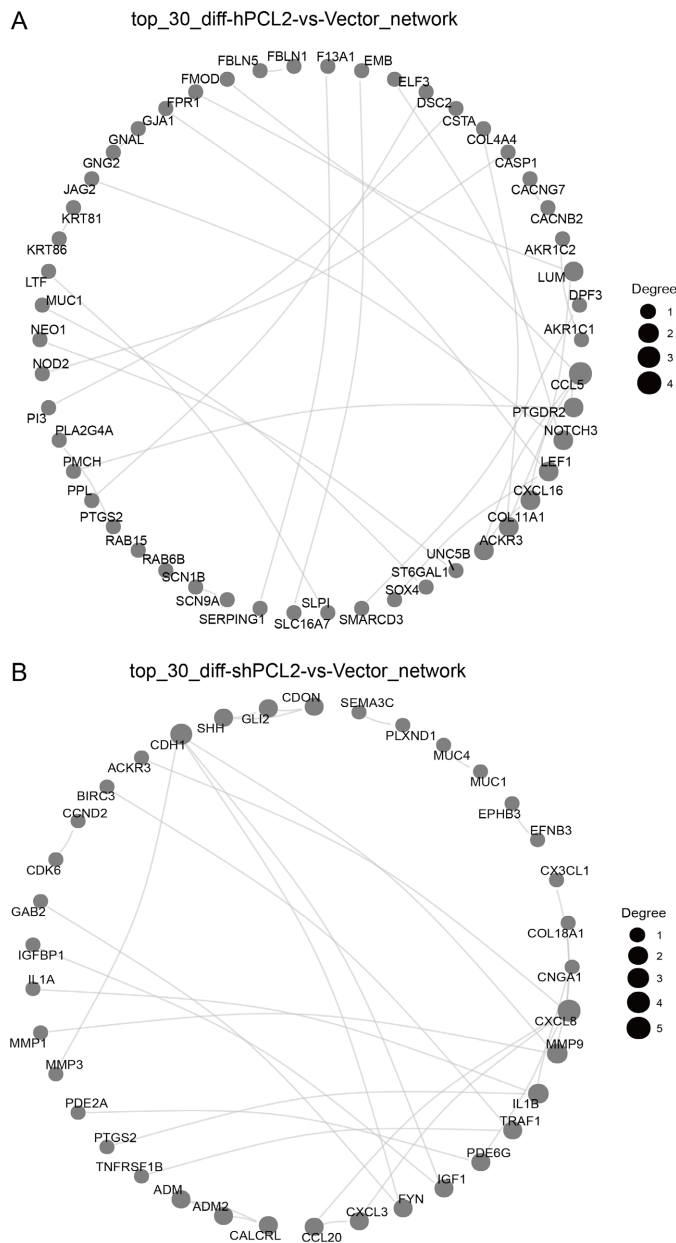


Figure 3. KEGG enrichment analysis of differentially expressed genes  
图 3. 差异基因 KEG 富集分析

为深入探究差异表达基因(DEGs)中显著富集的信号通路,我们采用KEGG软件解析了PCL2在BRCA中的功能作用(图3)。KEGG通常根据功能层级将通路划分为三个层级:层级1、层级2和层级3。层级1包含六个类别:代谢、遗传信息处理、环境信息处理、细胞过程、有机系统及人类疾病;层级2涵盖44个类别:细胞生长与死亡、物质运输及发育;层级3则包含数百条用于常规富集分析的通路,其功能范围从层级1到层级3逐步扩展。分析结果显示,在hPCL2组和shPCL2组分别包含的329个和489个基因中,与信号传导及分子相互作用相关的通路在hPCL2组中显著富集;而shPCL2组则表现出与人类癌症通路及生物免疫系统相关的通路丰度更高。

### 3.4. PPI网络



**Figure 4.** Top 30 protein-protein interactions (PPI)  
**图4.** 前30种蛋白质-蛋白质相互作用(PPI)

我们进一步探究了差异表达基因(DEGs)之间的相互作用模式。蛋白质相互作用(PPI)分析显示各组内存在显著的关联性(图4)。在 hPCL2 组中, 关键基因包括趋化因子(C-C 基序)配体 5 (CCL5)、神经源性位点 Notch 同源蛋白-3 (NOTCH3)以及淋巴样增强子结合因子 1 (LEF1); 相反, shPCL2 组则以上皮细胞钙黏蛋白(CDH1)、白细胞介素 8 (CXCL8)及 TNF 受体相关因子 1 (TRAF1)相关的基因为主。这些基因与肿瘤发展密切相关, 尤其是 CDH1——作为上皮-间质转化(EMT)相关因子, 在肿瘤迁移和侵袭过程中发挥关键作用。

### 3.5. PCL2 对 BRCA 影响的细胞学分析

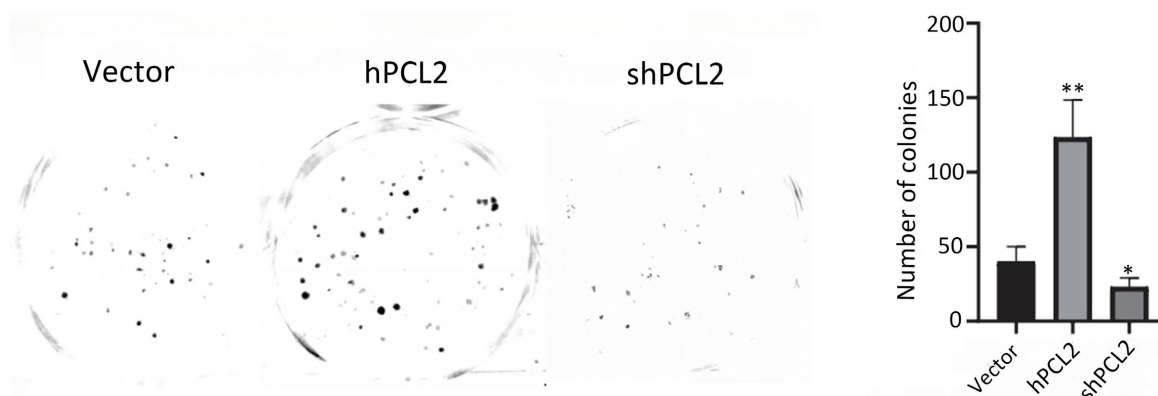


Figure 5. Cytological analysis of the role of PCL2 in BRCA

图 5. PCL2 在 BRCA 中作用的细胞学分析

为验证 PCL2 对 BRCA 的影响, 研究人员通过细胞学实验确认其增殖效应。实验组包括载体组、hPCL2 组以及载体组与 shPCL2 组。平板克隆实验显示(图 5), hPCL2 组中增殖细胞数量增加, 而 shPCL2 组则呈现减少趋势。

## 4. 讨论

根据国际癌症研究机构的 GLOBOCAN 数据, 2020 年全球新发 BRCA 病例达 230 万例, 占有新诊断癌症病例的 11.7%。肺癌在全球癌症发病率中位居首位[12]。研究表明, 三阴性 BRCA 是指雌激素受体(ER)、孕激素受体(PR)和人表皮生长因子受体 2 均呈阴性表达的 BRCA。三阴性 BRCA 的患病率极高, 占 BRCA 病例的 10%~20% [13]。与激素受体阳性或人表皮生长因子受体 2 过表达的 BRCA 相比, 三阴性 BRCA 通常发病更早、侵袭性更强且预后较差。PCL 是 PRC2 (Polycomb 抑制复合体 2)复合物的关键亚基, 其主要功能是抑制基因转录, 并作为染色质处于抑制状态时的表达标志物。该复合物包含四个核心组分: EED、EZH2、RbAp46/RbAp48 和 SUZ12。PCL2 在将整个复合物募集至染色质的过程中起关键作用[14]。在 PRC2 中, EZH2 是核心催化亚基。临床肿瘤治疗包括针对 EZH2 的化疗药物, 但国内外关于 PCL2 与肿瘤关系的研究仍较为有限; 此外, 利用数据挖掘和转录组测序技术对 PCL2 的研究尚未深入开展。因此, 本研究旨在通过生物信息学工具和细胞学实验, 评估 PCL2 对 BRCA 的影响。

随后, 我们对 BRCA-MDA-MB-231 细胞系进行了 hPCL2 过表达、shPCL2 过表达以及空载体慢病毒转染处理。进一步对处理后的样本进行 RNA-seq 数据分析, 获得了庞大的数据集。这些全面的数据使我们能够解析 PCL2 作用于 BRCA 的复杂机制。总体而言, 与空载体处理组相比, hPCL2 组共鉴定出 329 个差异表达基因(DEGs), 其中未调控基因 23 个, 下调基因 306 个; 而 shPCL2 组则鉴定出 489 个 DEGs, 包括 223 个上调基因和 266 个下调基因。我们发现部分基因可能与 BRCA 相关, 但仍有相当比例的 DEGs

功能未知,这可能源于该领域相关研究的不足。

对 DEGs 进行了 GO 和 KEGG 分析: GO 富集分析显示, hPCL2 组和 shPCL2 组的 DEGs 在 BP (生物过程)、CC (细胞周期)和 MF (功能模块)中显著富集,其中 BP 和 CC 的富集程度尤为突出;功能富集分析表明,与 hPCL2 相关的 BP 中最主要的条目包括细胞黏附、轴突导向和细胞外基质组织;而 CC 分析则显示胶原细胞外基质是最丰富的条目;MF 中最为显著的是细胞外信号受体活性。与 shPCL2 相关的最丰富的生物过程(BPs)涉及细胞外基质组织、细胞黏附以及细胞因子介导的信号通路。最常见的功能类别(CC)是细胞外空间,而最丰富的分子功能(MF)条目则是细胞外基质(ECM)的结构组分。

KEGG 分析显示,两组中差异表达基因(DEGs)最显著富集的领域均与环境信息处理及信号转导相关。此外,在 hPCL2 组中,最普遍的通路是细胞因子-细胞因子受体相互作用通路。在 shPCL2 组中,最显著的通路是人类癌症通路和生物免疫系统。这一癌症相关通路可能揭示了 PCL2 在 BRCA 中表达的分子机制。在细胞外基质、细胞黏附和轴突导向方面富集的差异基因,有助于全面理解 BRCA 恶性发展与 PCL2 表达之间的关联。

此外,我们利用蛋白质相互作用网络(PPIs)分析了前 30 个差异基因间的相互作用关系。包括 NOTCH3、LEF1、CDH1、CXCL8 等,这些基因集中参与上皮间质转化(EMT)、肿瘤干性维持、炎症反应等恶性表型调控。其中 CDH1 是 EMT 关键标志物,其表达改变直接影响肿瘤迁移与侵袭能力,提示 PCL2 可能通过 EMT 通路促进乳腺癌转移。研究发现:在 hPCL2 组中,CCL5、NOTCH3 和 LEF1 表现出最强的基因关联性;而在 shPCL2 组中,CDH1、CXCL8 和 TRAF1 显示出最显著的关联。这些基因被认为在 PCL2 调控 BRCA 发展过程中发挥重要作用。

最后,平板克隆形成实验(图 5)从细胞功能层面证实,PCL2 可显著增强乳腺癌细胞的克隆形成与增殖能力,过表达促进恶性表型,敲低则抑制增殖,进一步确立 PCL2 作为乳腺癌促癌因子的角色。结合 TCGA 数据库中 PCL2 在乳腺癌中高表达且与预后不良相关的结果,本研究充分说明 PCL2 可作为乳腺癌潜在诊断标志物与治疗靶点。

## 5. 结论

本研究通过转录组测序与体外细胞实验证实,PCL2 可显著调控乳腺癌 MDA-MB-231 细胞的全局基因表达谱。PCL2 过表达与敲低分别筛选出 329 个和 489 个差异表达基因,主要富集于细胞黏附、信号传导及肿瘤相关通路。PPI 网络分析提示 PCL2 可能通过调控 CCL5、NOTCH3、LEF1、CDH1、CXCL8 等核心基因,参与上皮间质转化相关的细胞表型调控。体外细胞学实验表明,PCL2 可促进乳腺癌细胞克隆形成与增殖能力。综上,PCL2 在乳腺癌细胞中发挥增殖调控作用,是具有潜在研究价值的分子靶点,其在乳腺癌体内进展与临床预后中的作用仍有待进一步实验验证。

## 基金项目

宁夏自然科学基金项目(2026AAC030221; 2024AAC03248)。

## 参考文献

- [1] Torre, L.A., Bray, F., Siegel, R.L., Ferlay, J., Lortet-Tieulent, J. and Jemal, A. (2015) Global Cancer Statistics, 2012. *CA: A Cancer Journal for Clinicians*, **65**, 87-108. <https://doi.org/10.3322/caac.21262>
- [2] Walker, E., Manias, J.L., Chang, W.Y. and Stanford, W.L. (2011) PCL2 Modulates Gene Regulatory Networks Controlling Self-Renewal and Commitment in Embryonic Stem Cells. *Cell Cycle*, **10**, 45-51. <https://doi.org/10.4161/cc.10.1.14389>
- [3] Simon, J.A. and Kingston, R.E. (2013) Occupying Chromatin: Polycomb Mechanisms for Getting to Genomic Targets, Stopping Transcriptional Traffic, and Staying Put. *Molecular Cell*, **49**, 808-824. <https://doi.org/10.1016/j.molcel.2013.02.013>

- 
- [4] Li, X., Isono, K., Yamada, D., Endo, T.A., Endoh, M., Shinga, J., *et al.* (2011) Mammalian Polycomb-Like Pcl2/Mtf2 Is a Novel Regulatory Component of PRC2 That Can Differentially Modulate Polycomb Activity Both at the *Hox* Gene Cluster and at *Cdkn2a* Genes. *Molecular and Cellular Biology*, **31**, 351-364. <https://doi.org/10.1128/mcb.00259-10>
- [5] Zhang, Z., Jones, A., Sun, C., Li, C., Chang, C., Joo, H., *et al.* (2011) PRC2 Complexes with JARID2, MTF2, and esPRC2p48 in ES Cells to Modulate ES Cell Pluripotency and Somatic Cell Reprograming. *Stem Cells*, **29**, 229-240. <https://doi.org/10.1002/stem.578>
- [6] Li, H. (2018) PCL2, a Novel Tumor Suppressor in Breast Cancer. *Science Bulletin*, **63**, 597-598. <https://doi.org/10.1016/j.scib.2018.04.004>
- [7] Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics*, **30**, 2114-2120. <https://doi.org/10.1093/bioinformatics/btu170>
- [8] Kim, D., Langmead, B. and Salzberg, S.L. (2015) HISAT: A Fast Spliced Aligner with Low Memory Requirements. *Nature Methods*, **12**, 357-360. <https://doi.org/10.1038/nmeth.3317>
- [9] Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., *et al.* (2010) Transcript Assembly and Quantification by RNA-Seq Reveals Unannotated Transcripts and Isoform Switching during Cell Differentiation. *Nature Biotechnology*, **28**, 511-515. <https://doi.org/10.1038/nbt.1621>
- [10] Roberts, A., Trapnell, C., Donaghey, J., Rinn, J.L. and Pachter, L. (2011) Improving RNA-Seq Expression Estimates by Correcting for Fragment Bias. *Genome Biology*, **12**, Article No. R22. <https://doi.org/10.1186/gb-2011-12-3-r22>
- [11] Anders, S. and Huber, W. (2010) Differential Expression Analysis for Sequence Count Data. *Genome Biology*, **11**, article No. R106. <https://doi.org/10.1186/gb-2010-11-10-r106>
- [12] Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., *et al.* (2021) Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, **71**, 209-249. <https://doi.org/10.3322/caac.21660>
- [13] Jiang, Y.Z., *et al.* (2019) Genomic and Transcriptomic Landscape of Triple-Negative Breast Cancers: Subtypes and Treatment Strategies. *Cancer Cell*, **35**, 428-440.e5.
- [14] Margueron, R. and Reinberg, D. (2011) The Polycomb Complex PRC2 and Its Mark in Life. *Nature*, **469**, 343-349. <https://doi.org/10.1038/nature09784>