

· 医药生物技术 ·

# 应用慢病毒 shRNA 文库结合二代测序筛选白血病细胞系增殖相关 lncRNA

马秋怡, 石得阳, 汪碧忱, 曹牧天, 李浩渊, 袁卫平, 初雅婧

中国医学科学院血液病医院 (中国医学科学院血液学研究所) 北京协和医学院 实验血液学国家重点实验室  
国家血液系统疾病临床医学研究中心 细胞生态海河实验室, 天津 300020

马秋怡, 石得阳, 汪碧忱, 曹牧天, 李浩渊, 袁卫平, 初雅婧. 应用慢病毒 shRNA 文库结合二代测序筛选白血病细胞系增殖相关 lncRNA. 生物工程学报, 2022, 38(9): 3406-3418.

MA QY, SHI DY, WANG BC, CAO MT, LI HY, YUAN WP, CHU YJ. Screening of proliferation related lncRNAs in leukemia cell lines by lentivirus shRNA library combined with second-generation sequencing. Chin J Biotech, 2022, 38(9): 3406-3418.

**摘要:** 长链非编码 RNA (long non-coding RNA, lncRNA) 是包括细胞增殖在内的许多细胞过程的重要调节因子。虽然已有研究表明多种 lncRNA 在造血系统恶性肿瘤的发生发展过程中发挥重要作用, 但是缺少一个更全面和无偏倚的方法同时研究多个 lncRNA 中对白血病细胞系产生功能性影响的 lncRNA。在此, 我们利用短发夹 RNA (short hairpin RNA, shRNA) 文库结合高通量测序的方法, 筛选对白血病细胞系增殖有影响的 lncRNA, 确定了 74 个候选 lncRNAs。从中选取 lncRNA *C20orf204-203* 作为验证研究对象, 发现 *C20orf204-203* 在 K562 和 THP-1 细胞系中均定位于胞质, 敲降 *C20orf204-203* 的 K562 和 THP-1 细胞系增殖能力降低, 早期凋亡细胞增加, *BAD* 基因在 mRNA 水平上表达量增加, TP53、BCL2 蛋白表达量下降, 在 THP-1 细胞系中 Caspase 3 蛋白表达量减少, 激活型 Caspase 3 蛋白表达量上升, 但是二者变化在两种细胞系中不一致。结果表明, 在白血病细胞系中敲降 lncRNA *C20orf204-203* 会使细胞增殖能力降低。但其在不同细胞系作用途径和机制可能存在差异。这一研究表明了利用 shRNA 文库结合高通量测序大规模研究 lncRNA 在白血病细胞系中发挥作用的可行性。

**关键词:** 长链非编码 RNA *C20orf204-203*; 短发夹 RNA 文库; 高通量测序; 白血病细胞系

Received: March 16, 2022; Accepted: May 10, 2022; Published online: May 13, 2022

Supported by: National Natural Science Foundation of China (82170117, 82170135, 81970149)

Corresponding author: CHU Yajing. E-mail: chuyajing@ihcams.ac.cn

基金项目: 国家自然科学基金 (82170117, 82170135, 81970149)

# Screening of proliferation related lncRNAs in leukemia cell lines by lentivirus shRNA library combined with second-generation sequencing

MA Qiuyi, SHI Deyang, WANG Bichen, CAO Mutian, LI Haoyuan, YUAN Weiping, CHU Yajing

State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Haihe Laboratory of Cell Ecosystem, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Tianjin 300020, China

**Abstract:** Long non-coding RNA (lncRNA) has become an important regulator of many cellular processes, including cell proliferation. Although studies have shown that a variety of lncRNAs play an important role in the occurrence and development of hematopoietic malignancies, a more comprehensive and unbiased method to study the function of lncRNAs in leukemia cell lines is lacking. Here, we used short hairpin RNA (shRNA) library combined with high-throughput sequencing to screen lncRNAs that may affect the proliferation of leukemia cell lines, and identified lncRNA *C20orf204-203* among 74 candidate lncRNAs in this study. Further experiments showed that *C20orf204-203* was localized in the cytoplasm in both K562 and THP-1 cell lines. *C20orf204-203* knockdown decreased the proliferation of K562 and THP-1 cell lines accompanied with the increased proportion of early apoptotic cells. We observed the increased mRNA level of *BAD* gene while decreased protein level of TP53 and BCL2. The expression of Caspase 3 decreased and Caspase 3-cleaved protein increased in THP-1 cell line. However, their changes were inconsistent in the two cell lines. Our experimental results showed that knockdown of lncRNA *C20orf204-203* in leukemia cell lines affected cell proliferation although the mechanism of action in different cell lines may differ. Importantly, our research demonstrated the feasibility of using shRNA library combined with high-throughput sequencing to study the role of lncRNA in leukemia cell lines on a large scale.

**Keywords:** lncRNA *C20orf204-203*; shRNA library; high-throughput sequencing; leukemia cell lines

随着对真核生物复杂基因组转录过程的全面解析,人们认为在单个细胞中大约一半的基因组可以转录成 RNA<sup>[1]</sup>。产生 RNA 的基因大致可以分为两种类型:蛋白质编码基因和非蛋白质编码基因。而非蛋白编码基因作为一个高度异质性的集合,可分为小分子非编码 RNA (small non-coding RNA) 和长链非编码 RNA (long non-coding RNA, lncRNA)<sup>[2]</sup>,其中 lncRNA 被定义为长度超过 200 个核苷酸、缺少开放阅

读框且不能翻译成蛋白质的转录产物。与小分子非编码 RNA 相比, lncRNA 曾被认为是不精确转录产生的噪音<sup>[3]</sup>。但是由于其广泛的作用和未知的功能,近年来对 lncRNA 的研究越来越深入。在基因组上,转录 lncRNA 的基因位点既可以位于蛋白编码基因间,也可以与蛋白编码基因的内含子或外显子存在交叉,甚至完全覆盖。功能上,一些 lncRNA 转录本通过其本身募集调节因子到其位点和/或其本身的功

能来调节邻近基因的表达；一些则从转录位点扩散，并以反式作用影响位于不同染色体上的基因；还有一些 lncRNA 的转录和/或剪接过程具有不依赖于 lncRNA 序列本身的调节功能。与特异定位于细胞质的 mRNA 不同，lncRNA 可以位于细胞的不同的核区（例如染色质、核质、亚核结构域）以及细胞质区<sup>[1,4-6]</sup>，且与 mRNA 相比，大多数 lncRNA 的表达水平较低，表达方式具有组织和细胞特异性<sup>[4,7-8]</sup>。近年来，越来越多的研究表明 lncRNA 在血液系统恶性肿瘤病理生理学上发挥重要作用。

造血干细胞的内环境稳态和血液细胞的形成依赖于生长因子、转录因子等不同途径间的相互作用，它们调控造血干细胞自我更新、静止、增殖和分化间的平衡<sup>[9]</sup>。这种关键作用的破坏可以导致肿瘤的发生，比如急性髓系白血病 (acute myelocytic leukemia, AML)、急性淋系白血病 (acute lymphocytic leukemia, ALL)、慢性髓系白血病 (chronic myelogenous leukemia, CML) 等。目前已知数十种 lncRNA 可调节细胞多能性、谱系定向和凋亡等多种过程<sup>[10-12]</sup>，如 lncRNA *HOTAIRM1* 在白血病细胞中作为肿瘤抑制因子能够调控髓系成熟、细胞周期和自噬<sup>[13-15]</sup>；lncRNA *ANRIL* 通过细胞中的 AdipoR1/AMPK/SIRT1 葡萄糖代谢通路促进 AML 细胞的增殖<sup>[16]</sup>；lncRNA *NR-104098* 在体外阻碍 AML 细胞的生长并促进 AML 细胞的分化<sup>[17]</sup>；lncRNA 通过分泌 microRNA-608 调节 *DDA1* 基因的表达，从而促进 AML 细胞的增殖和细胞周期<sup>[18]</sup>。许多 lncRNAs 已被证明通过不同途径调节 AML 细胞的生长。然而，迄今为止，已发现的具有功能的 lncRNA 不超过所有 lncRNA 的 3%<sup>[19]</sup>。因此，需要进一步地研究来寻找可能调节 AML 细胞生长的 lncRNAs。lncRNAs 除了在白血病中发生作用外，最近的研究还表明

lncRNAs 可作为白血病患者诊断、预后和治疗反应的生物标志物<sup>[20]</sup>，如：*LINC01255* 在 AML 患者的骨髓细胞、MDS-AML 患者的 CD34<sup>+</sup> 细胞和 AML 细胞系中高度表达，且 *LINC01255* 的高表达与 AML 患者的低生存率相关<sup>[21]</sup>。

在传统研究中，探究某一基因的作用主要是通过 RNA 干扰 (RNA interference, RNAi) 的方式靶向研究单个候选基因，但是短干扰 RNA (small interfering RNA, siRNA) 在细胞内寿命短，且难以将 siRNA 有效传递给原代细胞；而短发夹 RNA (short hairpin RNA, shRNA) 可以由宿主持续合成，作用时间相对更持久。近年来，shRNA 文库的出现和高通量测序技术的发展使得对多个基因缺失的功能研究成为可能，并且有助于研究与各种生物功能相关的新基因<sup>[22]</sup>。

尽管相关研究已经取得一些进展，但是利用高通量测序结合 shRNA 文库筛选的方法研究 AML 疾病中 lncRNA 的功能作用仍处于起步阶段。在本研究中，我们从 AML/MDS 患者核型异常的造血干祖细胞中筛选出与健康人群相比表达量显著升高的 lncRNA，通过设计包含靶向 74 个 lncRNA 序列和对照序列的 456 条特异 shRNA 文库，按照标准进行严格筛选及实验研究，发现其中 lncRNA *C20orf204-203* 对于白血病细胞系的增殖和凋亡有显著影响。本研究证实利用 shRNA 文库的敲降作用筛选白血病细胞系中的功能性 lncRNA 的可行性，为研究 lncRNA 在白血病的发生和发展过程中的作用提供了新思路。

## 1 材料与方法

### 1.1 细胞培养

293T 细胞系：高糖-DMEM 培养基培养，含 10% 的胎牛血清；K562、NOMO-1、Molm-13 和 THP-1 细胞系：含 10% 的胎牛血清的 RPMI-1640 培养基培养。

## 1.2 慢病毒文库包装

在 10 cm 细胞培养皿中培养工具细胞 293T 至细胞密度达到 80%–90%，将目的质粒文库与包装质粒 psPAX2、pMD2.G 按照质量比为 7:5:3 ( $\mu\text{g}$ ) 的比例加入到 293T 细胞上清中，使其共转染至 293T 中进行慢病毒的包装，6–7 h 后换成新鲜培养基，48 h 和 72 h 后分别收集细胞上清，此时病毒文库存在于上清中，然后利用超速离心机对病毒文库进行 10 倍浓缩。

## 1.3 转基因 K562 细胞系的构建

将浓缩后的病毒在感染复数 (multiplicity of infection, MOI) 为 0.5 的情况下感染 K562 细胞系，使得待感染细胞数量为病毒数量的 2 倍。感染体系中加入聚凝胺 ( $1 \mu\text{g}/\text{mL}$ ) 提高慢病毒感染细胞的效率。感染后 72 h，待病毒已经整合至细胞基因组中，加入嘌呤霉素 ( $3 \mu\text{g}/\text{mL}$ ) 培养 2 d 筛选出被病毒感染的细胞。对于不同细胞系，需要设定嘌呤霉素的浓度梯度 ( $1\text{--}5 \mu\text{g}/\text{mL}$ )，选择 72 h 内可以破坏 90% 以上细胞的浓度。

## 1.4 shRNA 文库筛选分析方法

嘌呤霉素筛选之后将细胞等分成 2 份，一部分细胞冻存于  $-80 \text{ }^\circ\text{C}$  用作基线测序 (baseline sample)，一部分细胞继续培养 5 d (Day 5 sample)，5 d 后提取这两部分细胞的 DNA，扩增标签序列，通过高通量测序测定每一个标签序列的读长，并且将各自标签序列测得的读长标准化到各自标签序列读长总数上。“Day 5 sample”和“baseline sample”中标签标准化读长的比值代表各个 shRNA 频率的分布。因为 shRNA 的频率分布是随机的，所以“Day 5 sample”中对照组对应的 shRNA 频率相对于“baseline sample”中对照组对应的 shRNA 频率的变化作为阈值，若“Day 5 sample”中实验组中某一 lncRNA 对应的 shRNA 有 3 条或 3 条以上的频率相对于“baseline

sample”中对应的 shRNA 频率的变化比阈值低，则认为在 K562 细胞系中敲降该 lncRNA 后使细胞增殖能力降低。

## 1.5 细胞增殖实验

收集对数生长期细胞，计数，用完全培养基重新悬浮细胞，调整细胞浓度至合适浓度，接种于 96 孔板，每孔加  $100 \mu\text{L}$  细胞悬液。细胞在  $37 \text{ }^\circ\text{C}$ 、5%  $\text{CO}_2$  培养箱中培养 1–6 d。每天取一个培养板，每孔中加入  $10 \mu\text{L}$  CCK-8 置于  $37 \text{ }^\circ\text{C}$  培养箱中孵育 4 h。轻轻振荡后在 Synergy H4 Microplate Reader 上测定 450 nm 波长处的吸光度。绘制生长曲线。

## 1.6 DNA 提取

使用血液/细胞/组织基因组 DNA 提取试剂盒 (TIANGEN, DP304-03) 提取细胞 DNA，Qubit 核酸定量分析试剂盒测定微量 DNA 浓度。

## 1.7 核质 RNA 分离实验

使用 Cytoplasmic & Nuclear RNA 纯化试剂盒 (NORGEN, 21000) 分离细胞核和细胞质的 RNA，用 qRT-PCR 进行分析。

## 1.8 qRT-PCR 检测

使用 RNeasy Plus Total RNA 提取试剂盒 (Qiagen, 74104) 从细胞中分离 RNA，通过逆转录试剂盒 (TaKaRa, RR047A) 将  $100\text{--}1000 \text{ ng}$  RNA 逆转录为 cDNA，采用 SYBR Green 定量 RT-PCR 试剂盒 (Roche, 57313500) 进行定量 RT-PCR。

## 1.9 细胞周期的检测

K562-shC20-scramble、K562-shC20-B、K562-shC20-C、THP-1-shC20-scramble、THP-1-shC20-B 和 THP-1-shC20-C 组分别取  $1 \times 10^6$  的细胞培养于 6 孔板中，培养 12 h 后用 FITC-Ki67 (BD Bioscience 公司) 和 Hoechst 33342 (Sigma 公司) 抗体标记细胞，流式细胞仪检测细胞周期比例。

### 1.10 细胞凋亡的检测

K562-shC20-scramble、K562-shC20-B、K562-shC20-C、THP-1-shC20-scramble、THP-1-shC20-B 和 THP-1-shC20-C 组分别取  $1 \times 10^6$  的细胞培养于 6 孔板中, 培养 12 h 后用 FITC-Annexin V 和 7-AAD 抗体 (BD Bioscience 公司) 标记凋亡细胞, 流式细胞仪检测凋亡细胞比例。

### 1.11 免疫印迹分析

K562-shC20-scramble、K562-shC20-B、K562-shC20-C、THP-1-shC20-scramble、THP-1-shC20-B 和 THP-1-shC20-C 组分别取  $1 \times 10^6$  的细胞, 1 mL PBS 重悬细胞清洗一遍, 离心弃上清, 加入 1 mL RIPA 裂解液和蛋白酶抑制剂 (Invitrogen 公司), 使用超声破碎仪超声。将超声后样本上清转移到新 EP 管。BCA 法测定蛋白浓度后进行免疫印迹实验。经 12% SDS-PAGE 凝胶电泳和转膜后, 依次孵育一抗和二抗, 加显影液并用 ImageQuantLas4000 (GE 公司) 检测。

### 1.12 统计学方法

数据采用 GraphPad Prism 8 版本进行统计学分析, 用非配对 Student's *t* 检验进行统计学差异分析,  $P < 0.05$  为有统计学差异(\*:  $P < 0.05$ , \*\*:  $P < 0.01$ , \*\*\*:  $P < 0.001$ )。

## 2 结果与分析

### 2.1 设计针对 74 个 lncRNAs 的 shRNA 文库

本研究首先排除 AML/MDS 患者造血干祖细胞中异常的染色体核型中包含已知原癌基因和抑癌基因的区域, 注释剩余区段内的 lncRNAs, 分析这些 lncRNAs 在健康人群造血干祖细胞和 AML/MDS 患者造血干祖细胞中的表达量是否有差异, 对患者造血干祖细胞中表达量高于健康人群造血干祖细胞的 lncRNAs 进行下一步研究。最终筛选出 74 个 lncRNAs 参与下一步实验 (图 1A)。

设计针对 74 个 lncRNAs 和对照序列的 shRNA 文库, 将 shRNA 序列整合到 pLVX 慢病毒载体骨架上, 此载体上包含 *mCherry* 荧光报告基因和嘌呤霉素抗性基因, 通过流式细胞分析仪 (fluorescence activating cell sorter, FACS) 测定荧光报告基因呈阳性的细胞所占百分比以确定病毒对细胞的转导效率, 并利用嘌呤霉素筛选出被病毒感染的细胞。在此载体上 U6 启动子控制 shRNA 的表达, 每条包含 shRNA 的载体上均带有一个独特的标签用于高通量测序后的鉴别。文库中针对每条 lncRNA 分别设计了 6 条对应的 shRNA, 还包含有 12 条 shRNA 不靶向任何基因序列作为对照组。由于作为对照组的 12 条 shRNA 不靶向任何序列, 所以其测序后的频率分布将作为筛选的阈值 (图 1B)。

### 2.2 利用 shRNA 文库筛选抑制 K562 细胞系增殖的 lncRNA

将 shRNA 质粒文库包装成慢病毒文库感染 K562 细胞系, MOI 值设为 0.5, 以此确保每个细胞只被一个携带 shRNA 的病毒感染。感染后 72 h 收集细胞并用嘌呤霉素筛选 2 d 后, 将细胞等分成两份, 一份冻存于  $-80^\circ\text{C}$  用作基线样本, 另一份不再加嘌呤霉素, 继续培养 5 d, 5 d 后同时提取两份细胞的基因组 DNA 进行高通量测序。为了确保文库的完整性, 我们对质粒文库和基线样本测序后的读长数量进行了相关性分析, 结果显示两个样本的读长数量相关性良好, 说明基线样本对质粒文库覆盖度良好 (图 1C)。对基线样本和培养 5 d 后的细胞样本的测序结果进行分析, 确定每一个标签的频率分布。若一个 lncRNA 所对应的 6 条 shRNA 中, 有 3 条或 3 条以上 shRNA 的读长数量低于对照的读长数量, 例如 *C20orf204-203* (图 1D), 则认为该 lncRNA 被敲降后会抑制 K562 细胞系的增殖。按照此筛选标准, 最终筛选出 6 个 lncRNA,

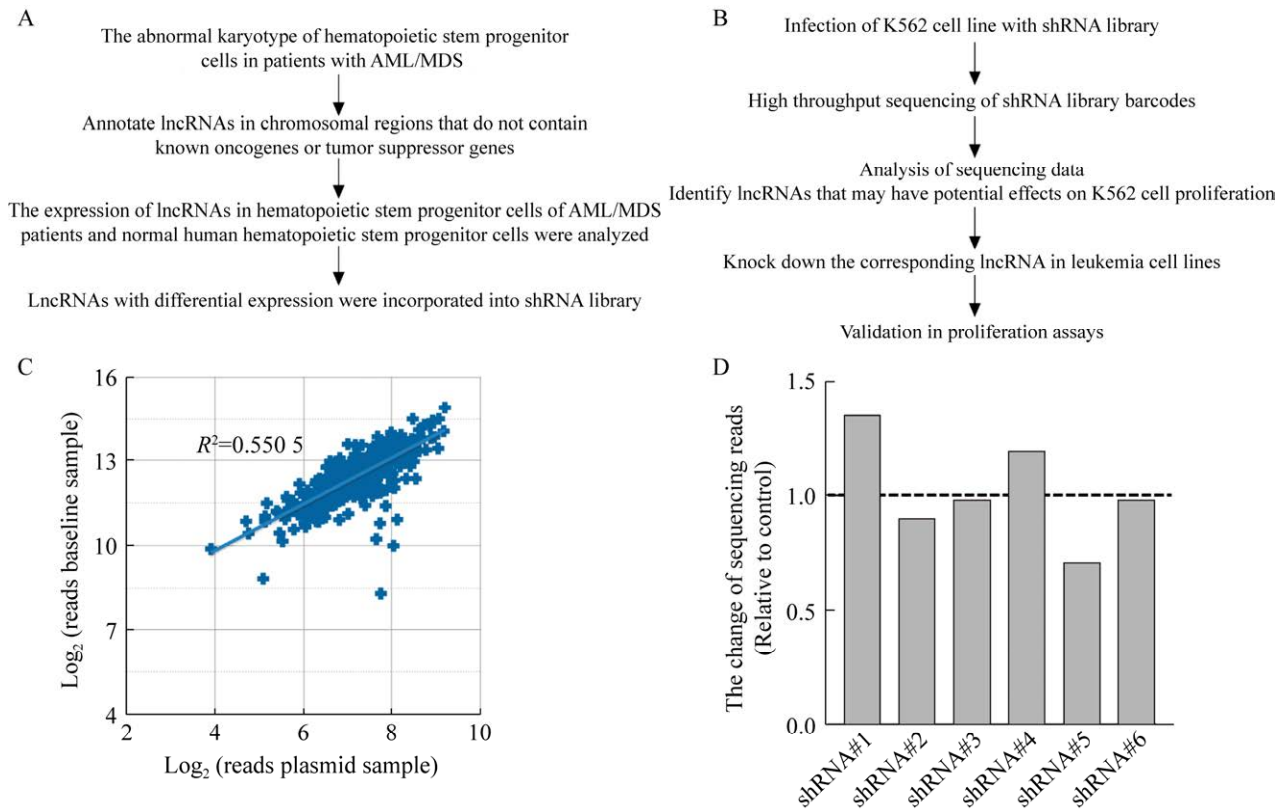


图 1 候选 lncRNAs 的选择策略和影响增殖的 lncRNAs 的筛选流程

Figure 1 Selection strategy of candidate lncRNAs (A) and screening process of lncRNAs affecting proliferation (B). (C) Library infection quality control: scatter plot of the log 2-transformed DNA sequencing read values from baseline sample versus the sequencing reads of the input pooled plasmid library. (D) Sequencing reads of shRNA barcodes against lncRNA *C20orf204-203*. Data are relative to barcode reads of shRNAs of control.

依次为 *PAN3-204*、*C20orf204-203*、*RPL21-209*、*NELFCD-202*、*MKLN1-211* 和 *SUGTIP3-201*。

### 2.3 *C20orf204-203* 在 K562 和 THP-1 细胞中定位于胞质

在筛选出的 6 个 lncRNAs 中, 我们选择 lncRNA *C20orf204-203* 进行下一步研究。*C20orf204* 基因位于 20 号染色体上, ENSEMBL 数据库 Human (GRCh38.p13) 显示其有 5 个转录本, 其中转录本 *C20orf204-203* 为长度大于 200 个核苷酸的非编码 RNA。qRT-PCR 实验表明, 在 K562、THP-1、NOMO-1 和 Molm-13 这 4 个白血病细胞系中, THP-1 细胞系中 lncRNA

*C20orf204-203* 的表达量相对较高 (图 2A), 同时核质分离试验验证在 K562 和 THP-1 细胞系中, 以定位于细胞核 (以字母“N”表示) 的 *U6* 基因和定位于胞质 (以字母“C”表示) 的  $\beta$ -Actin 为阳性对照, lncRNA *C20orf204-203* 相对定位在胞质 (图 2B-2C)。

### 2.4 *C20orf204-203* 敲降的 K562 细胞和 THP-1 细胞增殖能力降低

从 shRNA 文库中选取 3 条靶向 *C20orf204-203* 的 shRNA, 分别命名为 shC20-A、shC20-B 和 shC20-C, 将 3 条序列分别整合到 pLVX 慢病毒载体骨架上, 包装成慢病毒, 构建稳转细胞系:

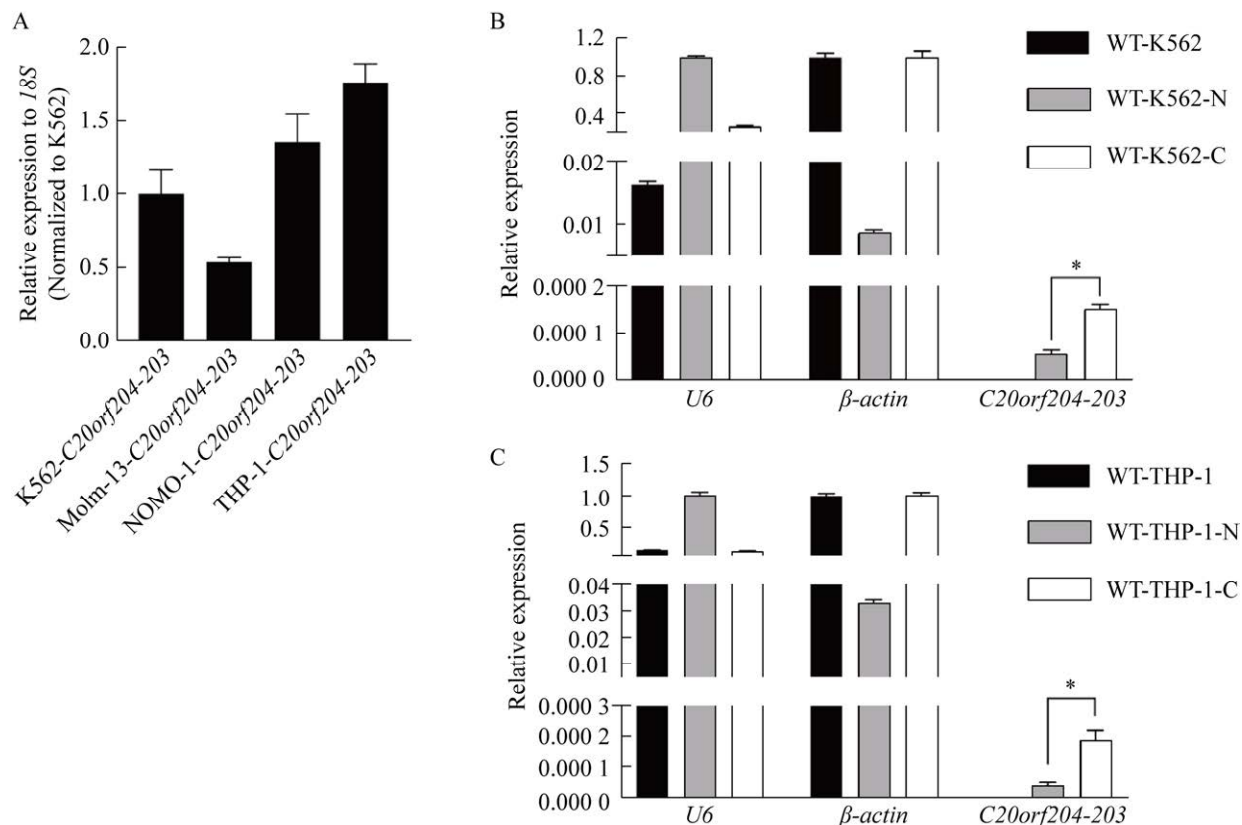


图 2 *C20orf204-203* 在 K562 和 THP-1 细胞中定位于胞质

Figure 2 *C20orf204-203* was localized in the cytoplasm of K562 and THP-1 cells. (A) Expression of *C20orf204-203* in leukemia cell lines. (B–C) Subcellular distribution of *C20orf204-203* in K562 and THP-1 subfractions, *U6* expression level in nucleic fraction and  $\beta$ -actin expression level in cytoplasmic fraction were measured and used as positive controls. \*:  $P < 0.05$ .

K562-shC20-A、K562-shC20-B、K562-shC20-C、THP-1-shC20-B 和 THP-1-shC20-C, 在 K562 细胞系中敲降 *C20orf204-203*, 发现 shC20-B 和 shC20-C 针对该基因有敲降作用, 而 shC20-A 对该基因的敲降作用并不明显 (图 3A–3B)。CCK-8 实验证实: 相比于对照组, K562-shC20-B、K562-shC20-C 和 THP-1-shC20-C 细胞增殖相对减慢 (图 3C–3D)。以上结果说明在 K562 细胞系和 THP-1 细胞系中敲降 lncRNA *C20orf204-203* 可以使细胞增殖能力降低。

## 2.5 敲降 *C20orf204-203* 的 K562 和 THP-1 细胞中凋亡相关基因表达量发生变化

为了进一步探究 K562 和 THP-1 细胞中敲

降 lncRNA *C20orf204-203* 后对细胞的影响, 用 Ki67 和 Hoechst 33342 标记 K562-shC20-B、K562-shC20-C、THP-1-shC20-B 和 THP-1-shC20-C, 通过流式细胞仪检测细胞周期发现, 与对照组 K562-shScramble 和 THP-1-shScramble 相比, 实验组细胞周期没有明显差异 (图 4A–4B); 但是通过 Annexin V 和 7-AAD 标记细胞, 流式细胞仪检测显示, 实验组早期凋亡细胞比例增加 (图 4C–4D)。qRT-PCR 实验显示 K562-shC20-B 和 K562-shC20-C 细胞中基因 *TP53* 和抗凋亡基因 *BCL2* 表达量变化不明显, 但是促凋亡基因 *BAD* 表达水平上升; THP-1-shC20-B 和 THP-1-shC20-C 细胞中 *TP53* 和 *BCL2* 基因表达量下降,

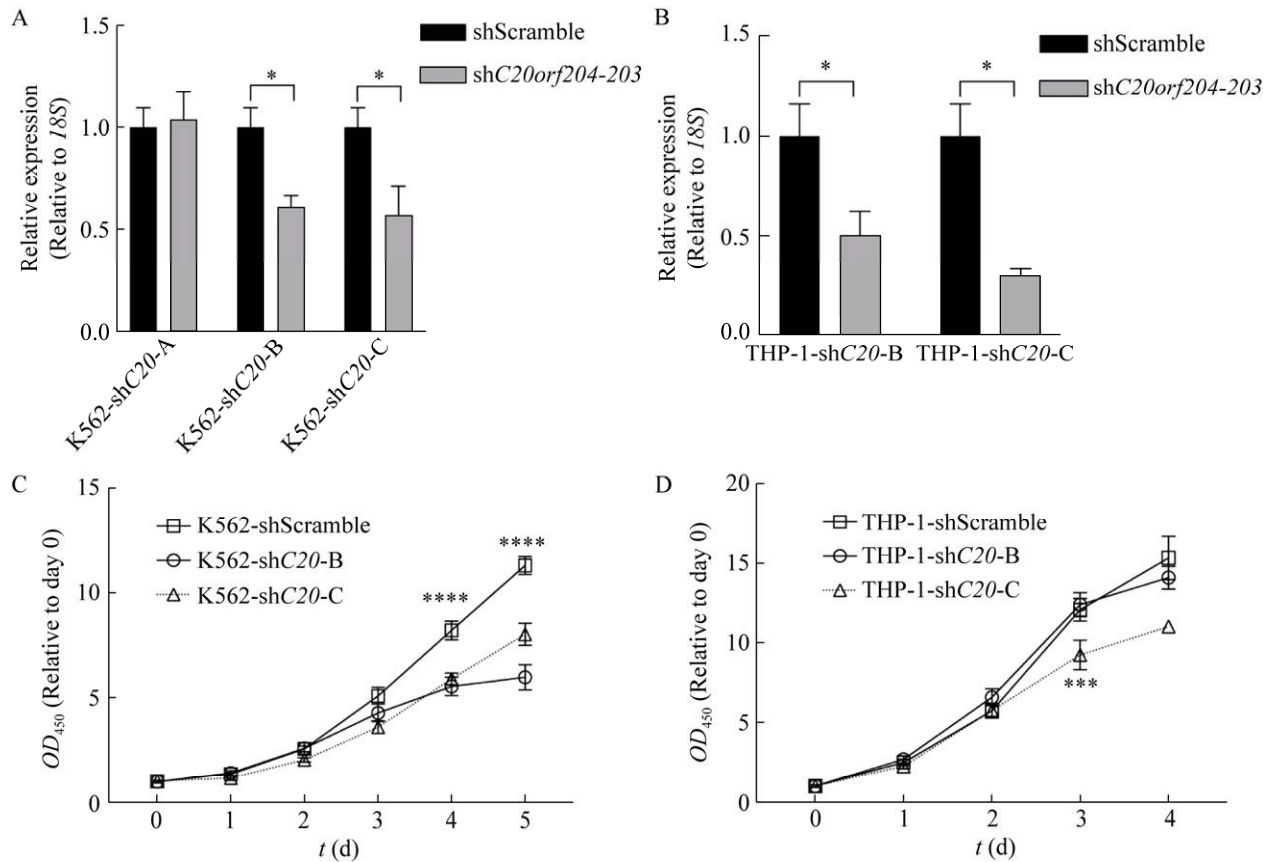


图3 *C20orf204-203* 敲降的 K562 和 THP-1 细胞增殖能力降低

Figure 3 Knockdown of *C20orf204-203* decreased K562 and THP-1 proliferation. (A–B) Relative expression of *C20orf204-203* in K562-shC20-B, K562-shC20-C, THP-1-shC20-B and THP-1-shC20-C. Expression level was measured with qPCR. (C–D) Cell growth curve. \*:  $P < 0.05$ ; \*\*\*:  $P < 0.001$ ; \*\*\*\*:  $P < 0.0001$ .

THP-1-shC20-B 中 *BAD* 基因表达水平上升 (图 4E)。蛋白水平检测发现在 THP-1-shC20-B 和 THP-1-shC20-C 细胞中, *BCL2*、*TP53* 蛋白表达量下降, *Caspase 3* 蛋白表达量下降, 激活型 *Caspase 3* 蛋白表达量上升, 但是在 K562-shC20-B 和 K562-shC20-C 细胞中变化不明显 (图 4F)。

### 3 讨论

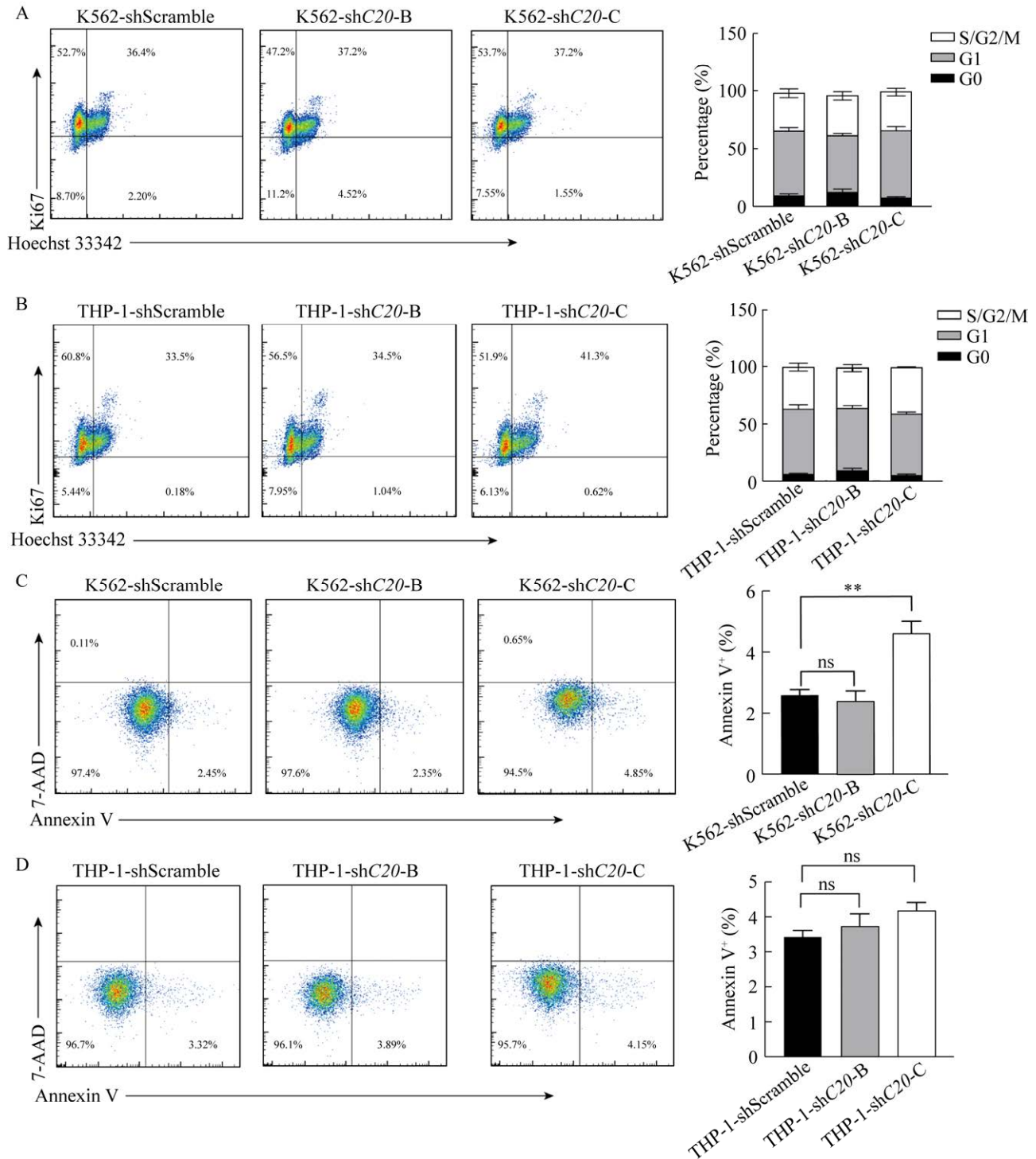
MDS 是一种无效造血性疾病, 其特征是造血细胞前体分化缺陷和异常克隆扩增<sup>[23-25]</sup>。大

约 30% 患有这种骨髓衰竭综合征的 MDS 患者会发展为 AML<sup>[26]</sup>。而 AML 是由于造血细胞前体基因改变导致克隆性髓系干细胞产生过度的造血干细胞的疾病。近年人们发现 lncRNAs 是多种细胞机制的新型调节因子。事实上, 由于 lncRNAs 数量多、作用模式广泛, 它们很可能几乎在任何细胞途径中起着关键作用<sup>[27]</sup>。已有文献系统地总结了 lncRNAs 在正常造血和恶性造血中扮演的角色<sup>[28]</sup>, 并且发现在血液细胞中, lncRNA 对不同分化谱系和免疫反应等生理环境具有高度特异性<sup>[29-30]</sup>。



目前利用 shRNA 逆转录病毒和慢病毒载体文库, RNAi 筛查已成为从基因上剖析正常和恶性 HSC 功能的有力工具<sup>[31]</sup>, 并且 RNAi 筛查成功地在体外<sup>[32]</sup>和体内<sup>[33]</sup>鉴定了恶性造血中的

几个重要驱动基因, 尤其是在 AML 疾病中。但是, AML 的病因尚未完全了解, 而且由于 AML 患者, 尤其是复发和难治性 AML 患者的复杂性和异质性, 改善 AML 的治疗仍然是一个巨大的



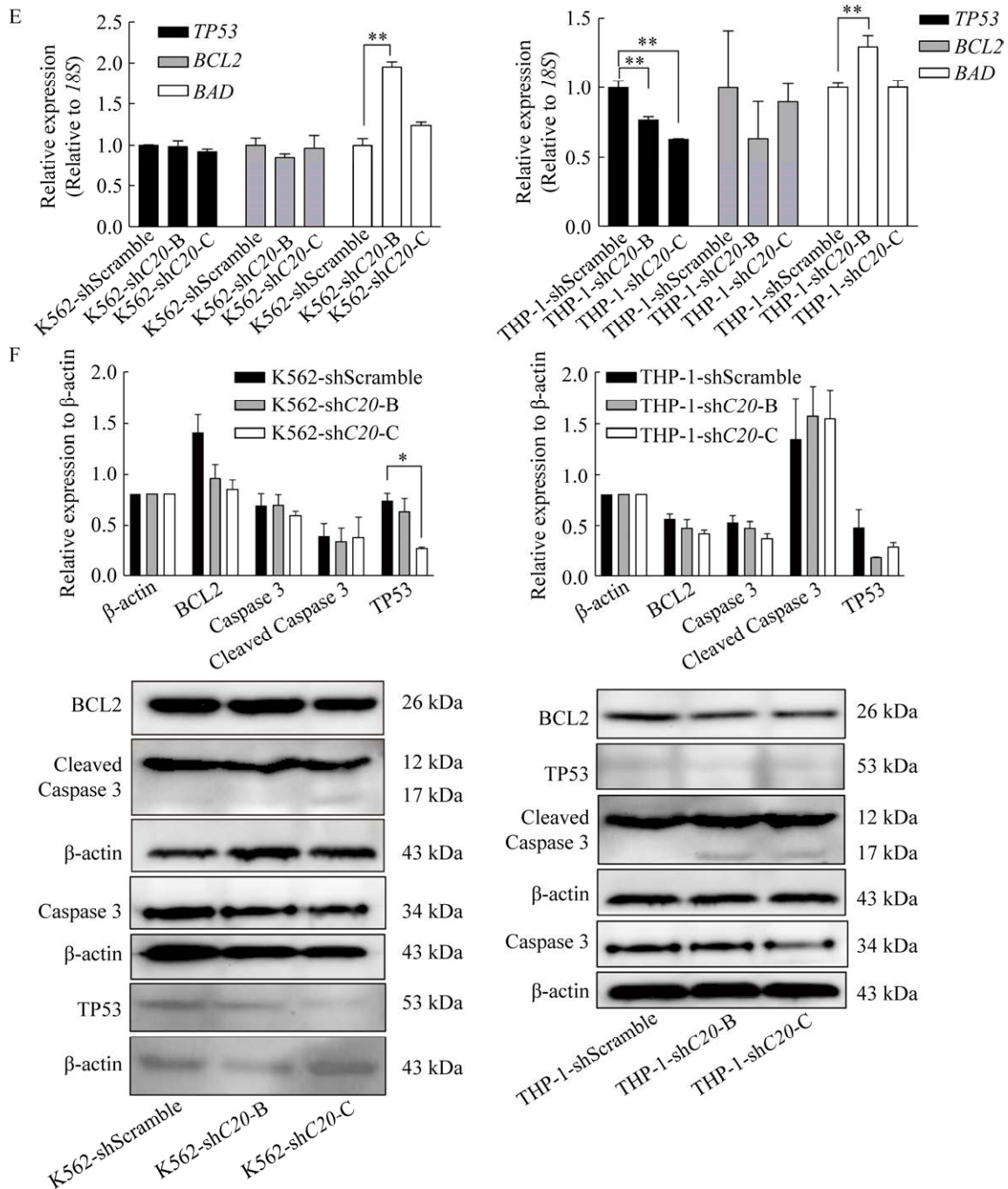


图 4 *C20orf204-203* 敲降的 K562 和 THP-1 凋亡增加, 细胞周期变化不明显

Figure 4 Knockdown of *C20orf204-203* increased K562 and THP-1 apoptosis without obvious changes of cell cycle. (A–B) K562-shC20-B, K562-shC20-C, THP-1-shC20-B and THP-1-shC20-C cell cycle detection by Ki67. (C–D) K562-shC20-B, K562-shC20-C, THP-1-shC20-B and THP-1-shC20-C cell apoptosis analysis by Annexin V. (E) Expression level of *TP53*, *BCL2* and *BAD* in K562-shC20-B, K562-shC20-C, THP-1-shC20-B and THP-1-shC20-C. Expression level was measured with qRT-PCR. (F) Validation of *TP53*, *BCL2*, *Caspase 3* and *Cleaved Caspase 3* protein level in K562-shC20-B, K562-shC20-C, THP-1-shC20-B and THP-1-shC20-C. Expression level was measured using Western blotting. \*\*:  $P < 0.01$ ; ns: not significant.

挑战。K562 细胞系是人红白血病细胞系, 因 9 号和 22 号染色体间的易位导致 *BCR* 和 *ABL* 基因融合, 使胞质中 Bcr-Abl 激酶表达<sup>[34]</sup>; 并且其在蛋白和 mRNA 水平都有 *TP53* 基因表达缺陷; 在一定的体外条件下, K562 细胞系可被多种化学物质诱导向红系分化, 因为其染色体核型异常且多变, 所以是临床和基础领域中筛选新的抗肿瘤药物的常用工具细胞。在本实验中, 我们首先通过分析健康人群的造血干祖细胞和核型异常的 AML/MDS 患者造血干祖细胞中表达量有差异的 lncRNA 确定我们的研究对象, 然后利用 shRNA 慢病毒文库结合高通量测序的方法筛选出对 K562 细胞系增殖有影响的 lncRNAs, 其中 lncRNA *PAN3-204* 在数据库 ENSEMBL Human (GRCh37) 中显示位于 13 号染色体 28 712 643–28 846 450 位置, 而 *circPAN3* 基因位于 13 号染色体 28 830 428–28 841 538 位置, 虽然二者之间的关系目前尚不明确, 但是有研究表明, *circPAN3* 被发现与 AML 细胞的增殖、凋亡和细胞周期有关<sup>[35]</sup>, 并且 *circPAN3* 的上调参与 AML 细胞耐药性的形成<sup>[36]</sup>。同时 CCK8 实验显示, 在 K562 和 THP-1 细胞系中敲降 lncRNA *C20orf204-203* 会导致细胞增殖减慢, 说明我们在 AML 细胞系中利用 shRNA 文库筛选功能性 lncRNA 的方法是可行的。

前期研究表明, 包括信号和激酶途径 (FLT3、KRAS、NRAS、KIT、PTPN11 和 NF1)、表观遗传修饰物 (DNA 甲基化和染色质修饰)、核磷蛋白 (NPM1)、转录因子 (CEBPA、RUNX1 和 GATA2)、肿瘤抑制因子 (TP53)、剪接体复合物 (SRSF2、U2AF1、SF3B1 和 ZRSR2) 和黏着素复合体 (RAD21、STAG1、STAG2、SMC1A 和 SMC3) 等许多基因突变在发病机制、预后和治疗中发挥着重要作用<sup>[37]</sup>。其中 TP53 介导多种细胞途径, 包括 DNA 损伤修复、细胞周期阻滞和调

亡的启动<sup>[38-39]</sup>。TP53 蛋白在肿瘤细胞中积累, 从而导致并促进肿瘤的发生和化疗耐药。在包括造血干细胞在内的大多数细胞中, TP53 在非活性状态下维持在低水平<sup>[40-41]</sup>。在我们的实验中发现敲降 lncRNA *C20orf204-203* 的 K562 和 THP-1 细胞 *TP53*、*BCL2* 基因在 mRNA 水平上表达降低; *BAD* 基因在 mRNA 水平上表达升高。*TP53*、*BCL2* 和 *Caspase 3* 基因在蛋白水平上表达明显降低。以上结果与早期凋亡细胞比例增加结果一致。虽然在我们的实验中发现敲降 lncRNA *C20orf204-203* 会导致细胞增殖能力降低, 但是通过 Ki67 检测细胞周期实验发现敲降 lncRNA *C20orf204-203* 后, K562 和 THP-1 细胞周期变化没有明显差异, 并且靶向同一 lncRNA 的不同 shRNA 在不同细胞系的敲降效率以及对同一 lncRNA 在同种细胞系中的敲降效率存在差异。同时实验结果显示, 在敲降 *C20orf204-203* 的 K562 和 THP-1 细胞系中, 活化的 *Caspase 3* 蛋白水平变化不一致, 可能由于文库筛选是在 K562 细胞系中进行, 所以导致在 K562 和 THP-1 细胞系中分别验证时结果出现差异, 也可能因此导致一些在 AML/MDS 患者中表达量相对高的 lncRNA 在筛选过程中被遗漏。这也表明该文库在其他细胞系中的筛选结果有待验证, 敲降 lncRNA *C20orf204-203* 后对细胞周期的影响以及导致细胞系凋亡增加的作用机制还需要进一步实验探究。

综上所述, 我们的研究显示在 K562 和 THP-1 细胞系中敲降 lncRNA *C20orf204-203* 后使细胞增殖能力降低, 促进了细胞的凋亡, *TP53*、*BCL2* 基因在 mRNA 和蛋白水平表达量下调, *BAD* 在 mRNA 水平表达量升高, 同时在蛋白水平上, *Caspase 3* 表达量下降, 激活型 *Caspase 3* 表达水平上升。本课题研究结果表明, 利用 shRNA 文库靶向 AML/MDS 患者核型异常

的造血干祖细胞中表达量与正常造血干祖细胞中表达量有差异的 lncRNAs, 进而筛选出功能性 lncRNAs 是可行的。

## REFERENCES

- [1] Djebali S, Davis CA, Merkel A, et al. Landscape of transcription in human cells. *Nature*, 2012, 489(7414): 101-108.
- [2] Ali T, Grote P. Beyond the RNA-dependent function of lncRNA genes. *eLife*, 2020, 9: e60583.
- [3] Mowel WK, Kotzin JJ, McCright SJ, et al. Control of immune cell homeostasis and function by lncRNAs. *Trends Immunol*, 2018, 39(1): 55-69.
- [4] Derrien T, Johnson R, Bussotti G, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res*, 2012, 22(9): 1775-1789.
- [5] Cabili MN, Dunagin MC, McClanahan PD, et al. Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome Biol*, 2015, 16(1): 20.
- [6] Kopp F, Mendell JT. Functional classification and experimental dissection of long noncoding RNAs. *Cell*, 2018, 172(3): 393-407.
- [7] Guttman M, Garber M, Levin JZ, et al. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat Biotechnol*, 2010, 28(5): 503-510.
- [8] Cabili MN, Trapnell C, Goff L, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev*, 2011, 25(18): 1915-1927.
- [9] Orkin SH, Zon LI. SnapShot: hematopoiesis. *Cell*, 2008, 132(4): 712.
- [10] Sheik Mohamed J, Gaughwin PM, Lim B, et al. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. *RNA*, 2010, 16(2): 324-337.
- [11] Guttman M, Donaghey J, Carey BW, et al. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature*, 2011, 477(7364): 295-300.
- [12] Williams GT, Mourtada-Maarabouni M, Farzaneh F. A critical role for non-coding RNA GAS5 in growth arrest and rapamycin inhibition in human T-lymphocytes. *Biochem Soc Trans*, 2011, 39(2): 482-486.
- [13] Zhang XQ, Lian Z, Padden C, et al. A myelopoiesis-associated regulatory intergenic noncoding RNA transcript within the human HOXA cluster. *Blood*, 2009, 113(11): 2526-2534.
- [14] Zhang XQ, Weissman SM, Newburger PE. Long intergenic non-coding RNA HOTAIRM1 regulates cell cycle progression during myeloid maturation in NB<sub>4</sub> human promyelocytic leukemia cells. *RNA Biol*, 2014, 11(6): 777-787.
- [15] Chen ZH, Wang WT, Huang W, et al. The lncRNA HOTAIRM1 regulates the degradation of PML-RARA oncoprotein and myeloid cell differentiation by enhancing the autophagy pathway. *Cell Death Differ*, 2017, 24(2): 212-224.
- [16] Sun LY, Li XJ, Sun YM, et al. lncRNA ANRIL regulates AML development through modulating the glucose metabolism pathway of AdipoR1/AMPK/SIRT1. *Mol Cancer*, 2018, 17(1): 127.
- [17] Feng YB, Hu S, Li LL, et al. lncRNA NR-104098 inhibits AML proliferation and induces differentiation through repressing EZH2 transcription by interacting with E2F1. *Front Cell Dev Biol*, 2020, 8: 142.
- [18] Zhuang MF, Li LJ, Ma JB. lncRNA HOTTIP promotes proliferation and cell cycle progression of acute myeloid leukemia cells. *Eur Rev Med Pharmacol Sci*, 2019, 23(7): 2908-2915.
- [19] Robinson EK, Covarrubias S, Carpenter S. The how and why of lncRNA function: an innate immune perspective. *Biochim Biophys Acta Gene Regul Mech*, 2020, 1863(4): 194419.
- [20] Cruz-Miranda GM, Hidalgo-Miranda A, Bárcenas-López DA, et al. Long non-coding RNA and acute leukemia. *Int J Mol Sci*, 2019, 20(3): 735.
- [21] Liu Q, Dong J, Li J, et al. LINC01255 combined with BMI1 to regulate human mesenchymal stromal senescence and acute myeloid leukemia cell proliferation through repressing transcription of MCP-1. *Clin Transl Oncol*, 2021, 23(6): 1105-1116.
- [22] Sims D, Mendes-Pereira AM, Frankum J, et al. High-throughput RNA interference screening using pooled shRNA libraries and next generation sequencing. *Genome Biol*, 2011, 12(10): R104.
- [23] Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*, 2016, 127(20): 2391-2405.
- [24] Huang HH, Chen FY, Chou WC, et al. Long

- non-coding RNA HOXB-AS3 promotes myeloid cell proliferation and its higher expression is an adverse prognostic marker in patients with acute myeloid leukemia and myelodysplastic syndrome. *BMC Cancer*, 2019, 19(1): 617.
- [25] Larsson CA, Cote G, Quintás-Cardama A. The changing mutational landscape of acute myeloid leukemia and myelodysplastic syndrome. *Mol Cancer Res*, 2013, 11(8): 815-827.
- [26] Zeidan AM, Wang R, Davidoff AJ, et al. Disease-related costs of care and survival among Medicare-enrolled patients with myelodysplastic syndromes. *Cancer*, 2016, 122(10): 1598-1607.
- [27] Geisler S, Coller J. RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. *Nat Rev Mol Cell Biol*, 2013, 14(11): 699-712.
- [28] Alvarez-Dominguez JR, Lodish HF. Emerging mechanisms of long noncoding RNA function during normal and malignant hematopoiesis. *Blood*, 2017, 130(18): 1965-1975.
- [29] Alvarez-Dominguez JR, Hu WQ, Gromatzky AA, et al. Long noncoding RNAs during normal and malignant hematopoiesis. *Int J Hematol*, 2014, 99(5): 531-541.
- [30] Atianand MK, Fitzgerald KA. Long non-coding RNAs and control of gene expression in the immune system. *Trends Mol Med*, 2014, 20(11): 623-631.
- [31] Karlsson C, Rak J, Larsson J. RNA interference screening to detect targetable molecules in hematopoietic stem cells. *Curr Opin Hematol*, 2014, 21(4): 283-288.
- [32] Zuber J, Shi JW, Wang E, et al. RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature*, 2011, 478(7370): 524-528.
- [33] Miller PG, Al-Shahrour F, Hartwell KA, et al. In Vivo RNAi screening identifies a leukemia-specific dependence on integrin beta 3 signaling. *Cancer Cell*, 2013, 24(1): 45-58.
- [34] Nemkov T, D'Alessandro A, Reisz JA. Metabolic underpinnings of leukemia pathology and treatment. *Cancer Rep (Hoboken)*, 2019, 2(2): e1139.
- [35] Li W, Zhong CQ, Jiao J, et al. Characterization of hsa\_circ\_0004277 as a new biomarker for acute myeloid leukemia *via* circular RNA profile and bioinformatics analysis. *Int J Mol Sci*, 2017, 18(3): 597.
- [36] Shang J, Chen WM, Liu S, et al. CircPAN<sub>3</sub> contributes to drug resistance in acute myeloid leukemia through regulation of autophagy. *Leuk Res*, 2019, 85: 106198.
- [37] DiNardo CD, Cortes JE. Mutations in AML: prognostic and therapeutic implications. *Hematology Am Soc Hematol Educ Program*, 2016, 2016(1): 348-355.
- [38] Kojima K, Konopleva M, Samudio IJ, et al. MDM2 antagonists induce p53-dependent apoptosis in AML: implications for leukemia therapy. *Blood*, 2005, 106(9): 3150-3159.
- [39] Carvajal LA, Neriah DB, Senecal A, et al. Dual inhibition of MDMX and MDM2 as a therapeutic strategy in leukemia. *Sci Transl Med*, 2018, 10(436): eaa03003.
- [40] Pant V, Quintás-Cardama A, Lozano G. The p53 pathway in hematopoiesis: lessons from mouse models, implications for humans. *Blood*, 2012, 120(26): 5118-5127.
- [41] Bykov VJN, Eriksson SE, Bianchi J, et al. Targeting mutant p53 for efficient cancer therapy. *Nat Rev Cancer*, 2018, 18(2): 89-102.

(本文责编 陈宏宇)