

论著·基础研究

CRISPR/Cas9介导的同源重组在非编码区疾病相关位点功能研究中的应用

徐 宁, 周 甜, 侯国俊, 沈 南, 唐元家

上海交通大学医学院附属仁济医院风湿病科, 上海市风湿病学研究所, 上海 200127

[摘要] **目的**·以全基因组关联研究报道的系统性红斑狼疮相关单核苷酸多态性 (single nucleotide polymorphism, SNP) 位点 rs1978421 为例, 利用规律成簇间隔短回文重复序列/相关蛋白 9 [clustered regularly interspaced short palindromic repeats (CRISPR) / CRISPR associated protein 9, CRISPR/Cas9] 系统建立非编码区疾病相关的遗传突变位点功能的研究策略。**方法**·针对 SNP rs1978421 位点设计单链导向 RNA (single guide RNA, sgRNA), 利用 T7 核酸内切酶 I (T7 endonuclease I, T7E I) 实验在 HEK293T 细胞中验证 sgRNA 对靶点的切割效率。以电穿孔转染的方式将融合 Cas9-绿色荧光蛋白 (green fluorescent protein, GFP) 和 sgRNA 的表达质粒及同源重组的模板转染至 U-937 细胞内, 通过流式分选得到单个的 GFP 阳性 U-937 细胞, 培养 14 d 后进行基因型鉴定。筛选得到同源重组的细胞克隆后, 通过实时荧光定量 PCR (quantitative real-time PCR, qPCR) 对 rs1978421 上下游 2 Mbp 范围内的基因及 miR-146a 进行定量, 分析 rs1978421 不同等位基因对 miR-146a 及其周围基因的表达影响。**结果**·T7E I 实验表明, 约有 24% 的 HEK293T 细胞中的 rs1978421 位点发生切割。在 U-937 细胞内进行的同源重组实验中, 共计得到 141 个细胞株; 其中 6 个同源重组成功, 成功率为 4.3%。qPCR 结果显示, 在野生型 TT 与同源重组成功 CC 的 U-937 细胞中, miR-146a 及所检测基因的表达情况均无显著变化 (均 $P > 0.05$)。**结论**·SNP rs1978421 对其周围 2 Mb 范围内的基因及 miR-146a 并无调控作用, 但该研究证明了 CRISPR 介导的同源重组技术用于非编码区 SNP 位点功能研究的可行性。

[关键词] 规律成簇间隔短回文重复序列; 同源重组; 全基因组关联研究; 单核苷酸多态性

[DOI] 10.3969/j.issn.1674-8115.2021.01.002 **[中图分类号]** Q3-3 **[文献标志码]** A

Application of CRISPR/Cas9-mediated homologous recombination to the functional study of disease-associated genetic variants located in non-coding region

XU Ning, ZHOU Tian, HOU Guo-jun, SHEN Nan, TANG Yuan-jia

Department of Rheumatology, Renji Hospital, Shanghai Jiao Tong University School of Medicine; Shanghai Institute of Rheumatology, Shanghai 200127, China

[Abstract] **Objective**·To take the systemic lupus erythematosus (SLE)-associated single nucleotide polymorphism (SNP) rs1978421 reported in the genome-wide association study (GWAS) as an example, and establish the research strategy for the functional study of disease-associated genetic variants located in the non-coding region through clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (CRISPR/Cas9) system. **Methods**·Single guide RNA (sgRNA) was designed for SNP rs1978421. T7 endonuclease I (T7E I) experiment was used to verify the cleavage efficiency of sgRNA in HEK293T cells. The homologous recombinant template and expression plasmid of Cas9-green fluorescent protein (GFP)/sgRNA were transfected into U-937 cells by electroporation transfection method. GFP positive U-937 single cells were sorted by flow cytometry, and the cells were cultured for the following 14 days. Then the genotype identification was carried out. After the homologous recombinant cell clones were screened, quantitative real-time PCR (qPCR) was used to quantify the upstream and downstream genes of rs1978421 within 2 Mb and miR-146a, and the effects of different alleles of rs1978421 on the expression of miR-146a and its surrounding genes were analyzed. **Results**·T7E I assay showed that about 24% of the total rs1978421 sites in HEK293T were cleaved. In homologous recombination experiment of the U-937 cells, a total of 141 cell lines were obtained and only six were successful, while the ratio of homologous recombination was 4.3%. qPCR showed that there was no significant difference in the expression of miR-146a and the detected genes between wild type TT and homologous recombinant CC in the U-937 cells (all $P > 0.05$). **Conclusion**·SNP rs1978421 has no regulatory effect on the surrounding genes within 2 Mb and miR-146a. However, this study proves the feasibility of CRISPR-mediated homologous recombination technology for the functional study of SNP sites in the non-coding regions.

[Key words] clustered regularly interspaced short palindromic repeats (CRISPR); homologous recombination; genome-wide association study (GWAS); single nucleotide polymorphism (SNP)

[基金项目] 国家自然科学基金 (81871287, 31630021, 31930037); 上海交通大学医学院高水平地方高校创新团队 (SSMU-ZDCX20180100)。

[作者简介] 徐 宁 (1993—), 男, 硕士生; 电子信箱: sjtuxuning@foxmail.com。

[通信作者] 唐元家, 电子信箱: yjtang@sibs.ac.cn。

[Funding Information] National Natural Science Foundation of China (81871287, 31630021, 31930037); Innovative Research Team of High-Level Local Universities in Shanghai (SSMU-ZDCX20180100).

[Corresponding Author] TANG Yuan-jia, E-mail: yjtang@sibs.ac.cn.

全基因组关联研究 (genome-wide association study, GWAS) 报道了许多疾病相关的单核苷酸多态性 (single nucleotide polymorphism, SNP) 位点^[1-2], 且进一步研究^[3-4]发现这些位点大多位于具有调控活性的非编码区。由于技术手段的限制, 对上述位点功能性的验证和机制的研究仅通过大样本相关性分析或双荧光素酶报告系统等较为简单又粗糙的方式进行, 尚无法提供能够直接验证 SNP 功能的证据^[5-6]。然而, 规律成簇间隔短回文重复序列/相关蛋白 9 [clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated protein 9, CRISPR/Cas9] 基因组编辑系统的出现改变了这一现状^[7]。该系统由负责靶向基因组的单链导向 RNA (single guide RNA, sgRNA) 与负责对基因组切割的 Cas9 组成。在 sgRNA 的作用下, Cas9 蛋白可诱导靶点 DNA 序列处形成双链断裂 (double-stranded break, DSB), 随后经非同源末端连接 (non-homologous end joining, NHEJ) 或同源重组 (homologous recombination, HR) 修复途径分别产生位点敲除/插入或位点重组的细胞株, 而这些细胞株则可为疾病易感位点的功能研究提供最直接的证据^[8]。

系统性红斑狼疮 (systemic lupus erythematosus, SLE) 是一种复杂的自身免疫病。GWAS 报道显示多个 SNP 位点均与 SLE 的发生与发展密切相关, 但大多数功能未明。因此, 利用 CRISPR/Cas9 基因组编辑系统对这些位点的功能性进行验证可直观地了解其与疾病的关系。研究^[9-11]发现, I 型干扰素信号通路的异常活化是 SLE 的重要特征之一, 且 miR-146a 是干扰素信号通路强有力的负调控分子, 在 SLE 患者的单核细胞中异常低表达。GWAS 发现 SNP rs2431697 是 SLE 的易感位点之一^[12], 表达数量性状位点 (expression quantitative trait locus, eQTL) 分析发现在欧洲人群中 rs2431697 等位基因的变化会影响 miR-146a 的表达^[13]。故而推测, rs2431697 或其连锁位点可能存在对 miR-146a 的调控作用^[14]。借助 HaploReg 数据库 (<https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php>) 对 rs2431697 连锁位点进行分析后发现, SNP rs1978421 与之存在强连锁效应 ($r^2=0.94$)。而后, 本课题组分别对 rs2431697 与 rs1978421 SNP 处的组蛋白修饰信号进行分析 (数据源自 ENCODE 数据库), 结果显示在这 2 个位点处均检测到常见于转录调节活性区域的组蛋白 H3 第 4 位赖氨酸单甲基化 (H3K4me1) 和组蛋白 H3 第 27 位赖氨酸乙酰化 (H3K27ac) 修饰信号, 故而认为这 2 个位点都有可能位于具有转录调节活性的 DNA 元件上, 具备参与 miR-146a 表达调控的能力。因

此, 本研究以 rs1978421 为研究对象, 采用 CRISPR/Cas9 介导的同源重组技术在单核细胞系 U-937 中研究该位点是否参与 miR-146a 的表达调控, 并以此来构建研究非编码区疾病相关 SNP 功能的有效方案。

1 材料与方法

1.1 实验材料

1.1.1 细胞株 HEK293T 细胞和 U-937 细胞均购自中国科学院典型培养物保藏委员会细胞库。经鉴定, 该 2 种细胞无支原体污染。

1.1.2 主要仪器与试剂 DMEM 培养基、RPMI-1640 培养基、胎牛血清 (fetal bovine serum, FBS) 以及 Opti-MEM 无血清培养基 (Gibco, 美国), 电转染系统 NeonTM Transfection System、LipofectamineTM 2000 转染试剂以及 TRIzol 试剂 (Invitrogen, 美国), 高速冷冻离心机 (Eppendorf, 德国), TaqManTM miRNA 反转录试剂盒、miR-146a 与对照 RNU48 定量探针、TaqManTM miRNA 定量试剂盒 (Thermo, 美国), PrimeScriptTM RT 试剂盒与 TB Green[®] Premix Ex TaqTM 试剂盒 (TaKaRa, 日本), ViiATM 7 荧光定量 PCR 系统 (Applied Biosystems, 美国), 质粒 pSpCas9(BB)-2A-GFP (#PX458; Addgene, 美国), 细胞株基因型鉴定试剂盒 TransDirect[®] Animal Tissue PCR Kit (北京全式金生物技术有限公司), I-5TM 2×High-Fidelity Master Mix (北京擎科生物科技有限公司), FACSAria II 流式分选仪 (BD 公司, 美国), 血液/细胞/组织基因组 DNA 提取试剂盒 [天根生化科技 (北京) 有限公司]。

1.2 实验方法

1.2.1 细胞培养 U-937 细胞采用含 10%FBS 的 RPMI-1640 培养基进行培养, HEK293T 细胞采用含 10%FBS 的 DMEM 培养基进行培养。培养条件为 37 °C、5%CO₂。本实验所用细胞均处于指数生长期, 状态良好。

1.2.2 U-937 野生型 rs1978421 位点基因型鉴定 为获得靶细胞 rs1978421 位点等位基因同源重组的细胞系, 首先需对 U-937 细胞 rs1978421 位点处的碱基类型进行鉴定 (该基因型为野生型)。取约 2×10⁵ 个 U-937 细胞, 离心后弃上清液, 使用血液/细胞/组织基因组 DNA 提取试剂盒提取基因组 DNA。取 100 ng 基因组 DNA 进行 rs1978421 位点的 PCR 扩增, 扩增反应采用 I-5TM 2×High-Fidelity Master Mix 高保真酶反应液进行, 反应体系 (50 μL) 如下: 模板 DNA 100 ng、引物 (10 μmol/L) 各 2.5 μL、2×

高保真酶反应液 25 μL。将各成分混合均匀后短时离心，按照说明书中的条件进行PCR反应。引物序列见表1，产物长度 500 bp。当反应结束后，通过测序公司进行取样并测序。

表 1 rs1978421 位点鉴定引物

Tab 1 Identification primer of rs1978421 locus

Primer	Sequence (5'→3')
rs1978421-F	TACTTTTGGCATCCTTGCT
rs1978421-R	GACCAACCTCGGCAACATAG

1.2.3 sgRNA/Cas9 表达载体的构建 选择 pSpCas9 (BB) -2A-GFP 为 sgRNA/Cas9 的表达载体。将其进行酶切并电泳、回收后，与 sgRNA 片段连接。利用 IDT 在线 sgRNA 设计工具 (https://sg.idtdna.com/site/order/designtool/index/CRISPR_CUSTOM) 对 sgRNA 进行设计，原则如下：①SNP 位点在 sgRNA 序列中。②sgRNA 靶向评分 (on-target score) 较高、脱靶评分 (off-target score) 较低。基于该原则，我们选取了靶向评分为 71 分、脱靶评分为 0 分的 sgRNA，并以单链的形式合成 (序列见表 2)。将合成的 sgRNA 正链与负链在体外复性为双链后，与上述酶切载体于 4℃连接过夜，而后行转化实验，次日挑取单菌落进行测序。对测序成功的菌株抽提 px458-rs1978421 质粒备用。

表 2 sgRNA 功能序列及单链引物

Tab 2 sgRNA functional sequence and single strand primers

Name	Sequence (5'→3')
sgRNA functional sequence	TTTCATTCATCACCAAGTC
sgRNA-F	caccgTCTCATTCCATCACCAAGTC
sgRNA-R	aaacGACTTGGTGATGGAATGAGAc
U6 promoter sequencing primer	TGGACTATCATATGCTTACC

1.2.4 质粒转染 于转染前 1 d 向 24 孔板加入 HEK293T 细胞，每孔 2×10^5 个，过夜培养后即可行转染实验。向每孔添加 Lipofectamine™ 2000 转染试剂 2 μL，其余操作均参照说明书进行。当转染结束后，继续培养 48 h 即可行后续实验。将 U-937 细胞离心去培养基后，用磷酸盐缓冲液洗涤 2 次，计数后进行分装，即每 1.5 mL 离心管中加入 1.5×10^6 个细胞。离心后弃上清液并用电转缓冲液重悬，加入 10 μg px458-rs1978421 质粒 (1 μg/μL) 以及 30 μg 单链 DNA 重组模板 (序列见表 3)，电转条件为 1 400 mV、10 ms、3 pulse。在电转后的第 12 h 加入 DNA 连接酶Ⅳ的抑制剂 Scr7 (终浓度为 1 μmol/L)，以抑制 NHEJ 途径对细胞双链 DNA 断裂的修复。

1.2.5 T7 核酸内切酶 I 实验 收集经 px458-rs1978421 质

表 3 同源重组模板序列

Tab 3 Sequence of homologous recombination template

Template	Sequence (5'→3')
rs1978421HR	CCAAACCTGGGCAACAACCCATAAATACATGTAGCCTT-TTTTCTGAGACAGAGTCTCATTCCATCACCAAGTCTGGAGTGCAGTGGTGCATATGGCTCACTGCAGCCTGGACC

粒转染 48 h 后的 HEK293T 细胞，提取基因组 DNA 后经 PCR 反应对包含 SNP 位点的片段进行扩增。取经纯化后的产物 300 ng 进行复性，反应体系为 15 μL，包含 1.5 μL T7 核酸内切酶 I (T7 endonuclease I，T7E I) 缓冲液；反应条件如下：95℃保持 10 min，后以 2℃/s 的速度降至 85℃，再以 0.3℃/s 的速度降至 25℃，以 25℃保持 10 s 后正常降温至 4℃。待反应结束后，向每管复性产物加入 2 μL T7E I，于 37℃孵育 1 h 后用 2% 琼脂糖凝胶电泳检测切割效率。切割效率计算公式如下：

$$\text{切割效率} = \left(1 - \sqrt{1 - \frac{I_{310\text{ bp}} + I_{190\text{ bp}}}{I_{500\text{ bp}} + I_{310\text{ bp}} + I_{190\text{ bp}}}} \right) \times 100\%.$$

其中， $I_{500\text{ bp}}$ 、 $I_{310\text{ bp}}$ 和 $I_{190\text{ bp}}$ 分别为 500 bp、310 bp 和 190 bp 条带的荧光强度值。

1.2.6 RNA 抽提与 qPCR 检测 使用 TRIzol 法提取 RNA，以三氯甲烷萃取 TRIzol 中的水相，并以异丙醇沉淀其中的 RNA，用 75% 乙醇洗涤沉淀 2 次后室温晾干。待 RNA 充分干燥后，用 20 μL 焦碳酸二乙酯 (diethyl pyrocarbonate, DEPC) 处理水进行溶解并测定浓度，将其浓度统一调至 50 ng/μL 后利用 TaqMan™ miRNA 反转录试剂盒进行 miRNA qPCR 模板的制备，并将制备好的模板稀释 3 倍后再行定量。利用 PrimeScript™ RT 反转录试剂盒将提取的总 RNA 中的 mRNA 反转录为 cDNA，稀释 6 倍后作为 qPCR 的模板。利用 TB Green® Premix Ex Taq™ 试剂盒对稀释后的 cDNA 模板进行 qPCR 检测。结果以相对表达量 $2^{-\Delta\Delta C_t}$ 展示。qPCR 引物序列见表 4。

1.2.7 细胞分选 U-937 细胞转染 48 h 后即可进行流式分选。以未转染的 U-937 细胞作为阴性对照来调节电压及相关参数；将活细胞与单个细胞依次设门圈出，通过调节异硫氰酸荧光素 (fluorescein isothiocyanate, FITC) 通道的电压将阴性对照的细胞集中在横坐标 10^3 左侧，即可将转染过的 U-937 细胞上样；分选 FITC 荧光通道横坐标 10^3 右侧的 FITC 阳性的 U-937 细胞群至 96 孔细胞培养板，每孔 1 个细胞。分选结束后于 37℃、5%CO₂ 培养箱中培养约 14 d 后即可进行基因型鉴定实验。

1.2.8 细胞基因型鉴定 在分选后的第 14 日，96 孔板底部已有肉眼可见的细胞群落。使用移液器将细胞群落均匀

表4 qPCR引物序列

Tab 4 qPCR primers

Primer	Sequence (5'→3')
<i>GAPDH</i> -F	GGAGCGAGATCCCTCCAAAT
<i>GAPDH</i> -R	GGCTGTTGTCATACTTCTCATGG
<i>PTTG1</i> -F	ACCCGTGTGGTTGCTAAGG
<i>PTTG1</i> -R	ACGTGGTGTGAACTTGAGAT
<i>SLU7</i> -F	AAGAACAGCGAAAATTGGGCA
<i>SLU7</i> -R	CCCTCTGTACCATTCTCCAGA
<i>TTC1</i> -F	GAGCGGACAAGGTTGAGAACA
<i>TTC1</i> -R	CTCCTCCTTTAGTCTAGTGCTCT
<i>PWWP2A</i> -F	CTTGTCGTGTCGTTCCGCTT
<i>PWWP2A</i> -R	ACCATTGCTTCACACTTGACTT
<i>UBLCP1</i> -F	CTCGCAGAGTGAAAGAGTACAAA
<i>UBLCP1</i> -R	GCACAAGACCTGTGGTCAAATA

Note: *GAPDH*—*Homo sapiens* glyceraldehyde-3-phosphate dehydrogenase; *PTTG1*—*Homo sapiens* pituitary tumor-transforming 1; *SLU7*—*Homo sapiens* SLU7 homolog splicing factor; *TTC1*—*Homo sapiens* tetratricopeptide repeat domain 1; *PWWP2A*—*Homo sapiens* PWWP domain containing 2A; *UBLCP1*—*Homo sapiens* ubiquitin-like domain containing CTD phosphatase 1.

地吹打成细胞悬液,吸取5~10 μ L至96孔PCR反应板中,用来制备鉴定基因型的PCR模板,剩余细胞悬液转至24孔板中继续培养。模板制备采用TransDirect® Animal Tissue PCR试剂盒,反应体系为50 μ L,包括4~8 μ L模板。其余实验步骤按照试剂盒说明书进行。

1.3 统计学分析

应用GraphPad Prism 7.0软件对qPCR数据进行分析。通过非配对 t 检验比较野生型与重组型细胞中的对应基因表达的差异。 $P<0.05$ 表示差异具有统计学意义。

2 结果

2.1 rs1978421位点sgRNA设计及效率鉴定

利用CRISPR/Cas9基因组编辑系统研究非编码区SNP位点的功能,具体流程如图1A。为设计针对rs1978421位点进行同源重组的sgRNA,我们首先对U-937细胞rs1978421位点处的碱基类型进行鉴定;结果(图1B)发现在U-937和HEK293T细胞内,rs1978421位点的等位基因均为胸腺嘧啶(T)。之后,我们以含有T

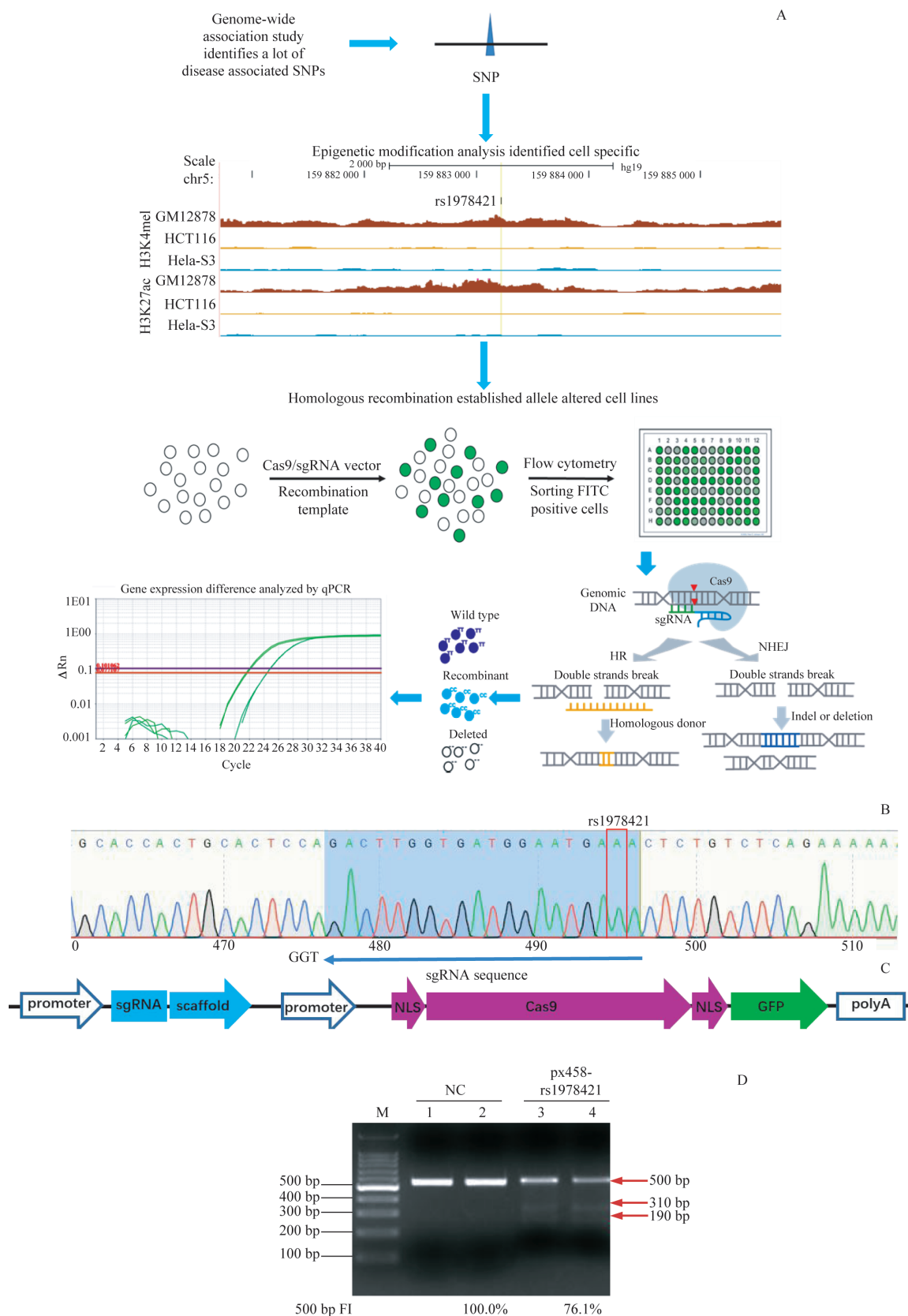
等位基因的DNA片段进行sgRNA设计及载体px458-rs1978421的构建(图1C)。成功获得载体后,我们在HEK293T细胞中对sgRNA切割效率进行检测。T7E I核酸内切酶切割实验结果(图1D)显示,目的sgRNA在HEK293T中rs1978421位点的切割效率约为24%,表明该sgRNA能够对rs1978421位点进行有效切割,可以应用于同源重组实验。

2.2 同源重组细胞株的构建与鉴定

鉴于rs1978421可能与miR-146a的表达相关,而miR-146a在HEK293T中低表达,在U-937中高表达,因此我们选择在U-937细胞中进行同源重组实验。将载体以及同源重组模板通过电穿孔方法转入U-937细胞48 h后,通过流式分选仪将单个绿色荧光蛋白(green fluorescent protein, GFP)强阳性的细胞分选至96孔板中,确保每孔只含有1个细胞。在进行流式分选时发现,活细胞中GFP阳性的细胞数占总数的2.29%(图2A)。经分选共得到960个只含有1个GFP阳性U-937细胞的培养孔。经2周培养后,有141个细胞增殖至肉眼可见的细胞群落;将rs1978421位点附近的基因组DNA经PCR扩增并测序发现,有86个细胞克隆基因为杂合子,占总数的61.0%;有25个细胞基因组未发生改变,占总数的17.7%;有9个细胞在rs1978421位点旁侧检测到不明来源DNA序列的插入,占总数6.4%;有15个细胞在rs1978421旁侧发生不同数目碱基删除,占总数的10.6%;另有6个细胞rs1978421位点的碱基成功由胸腺嘧啶(T)转变为胞嘧啶(C),重组成功细胞数占总数的4.3%(图2B、C和表5)。

2.3 rs1978421不同等位基因对周围基因表达影响

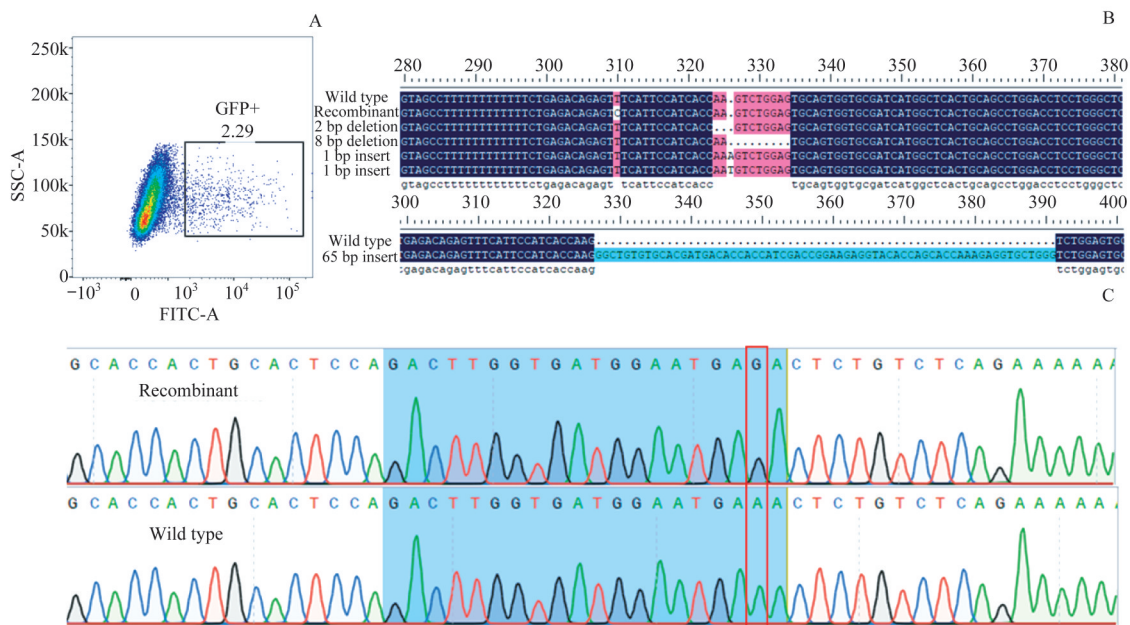
rs1978421位于5号染色体*PTTG1*基因下游约27.4 $\times 10^3$ bp以及miR-146a的初级转录本miR3142HG上游约12.1 $\times 10^3$ bp处,其邻近2 Mbp区域内的基因包括*SLU7*、*TTC1*、*PWWP2A*等(图3A)。为了检测rs1978421位点的调节基因,我们首先采用qPCR法检测miR-146a在不同基因型细胞中的表达水平;结果(图3B)显示,与TT基因型相比,在CC型、SNP位点邻近序列切除8 bp以及SNP位点邻近序列插入65 bp的细胞克隆中,miR-146a表达无差异。同时,对rs1978421邻近的高表达基因检测表明,不同基因型对这些基因的表达也无显著影响(图3C)。基于以上数据,我们认为rs1978421位点可能不是一个具有功能性的调节位点。



Note: A. Schematic diagram of the research strategy for SNP located in non-coding region. B. Wild type U-937 SNP rs1978421 genotype identification (reverse primer-mediated sequencing). C. pSpCas9(BB)-2A-GFP plasmid element instructions (NLS—nuclear location signal). D. T7E1 experiments for sgRNA cutting efficiency test in HEK293T cells (M—marker; FI—fluorescence intensity).

图1 非编码区SNP位点研究策略及rs1978421位点sgRNA设计及切割效率检测

Fig 1 Research strategy for SNP located in non-coding region and rs1978421 sgRNA design and cutting efficiency test



Note: A. FACS analyzed the GFP positive U-937 cell ratio. B. Genotypes of part of homozygous cell clones. C. Sequencing data of rs1978421 in successfully recombined clones and wild type clones.

图2 流式分选与细胞基因型鉴定

Fig 2 FACS sorting and cell clones genotype identification

表5 SNP rs1978421 同源重组实验所得基因型细胞株汇总

Tab 5 Summary of genotypes of cell lines obtained in SNP rs1978421 homologous recombination

Genotype	Count/n	Ratio/%
Heterozygote	86	61.0
Wild type	25	17.7
Successfully recombined (T-C)	6	4.3
Insert (1 bp)	5	3.5
Insert (65 bp)	4	2.8
Deletion (2 bp)	9	6.4
Deletion (8 bp)	6	4.3

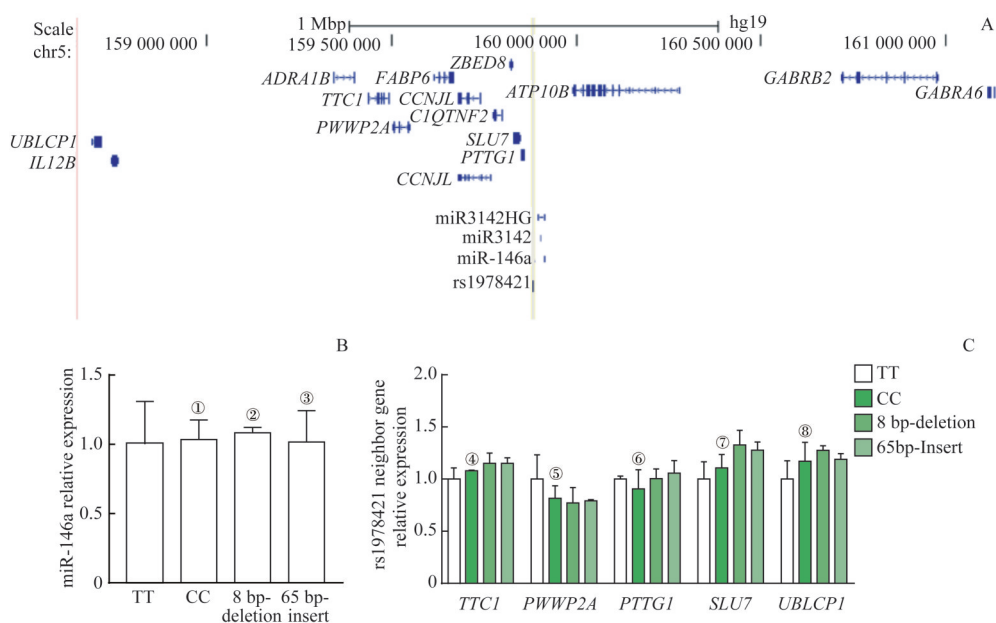
3 讨论

CRISPR/Cas9 基因组编辑系统是一种可靠的非编码区疾病相关性 SNP 功能研究手段。得益于该研究方法，许多疾病相关的功能性位点参与基因表达调控的方式被报道：如 SNP rs1421085 等位基因的改变可加剧肥胖^[15]，SNP rs11672691 等位基因的变化与前列腺癌细胞的增殖相关^[16]。在本研究中，利用 CRISPR/Cas9 基因组编辑系统介导的同源重组成功得到了 rs1978421 等位基因重组的 U-937 细胞克隆，通过荧光定量 PCR 检测 rs1978421 周围基因的表达变化并以此明确了 rs1978421 位点不是功能性的疾病易感位点。经过对 rs1978421 和 rs2431697 位点遗

传学数据的进一步挖掘和分析，我们推测后者可能是参与 miR-146a 的表达调控的功能性位点。后续我们将进一步对这个位点进行功能性鉴定。

本研究在 SNP 位点的选择时仅以 H3K4me1 与 H3K27ac 这 2 个组蛋白修饰信号作为依据，并未对 SNP 位点处的信息学数据进行深入挖掘。由于同源重组技术工作量大且实验周期长，因此对位点的选择尤为重要。在位点选择时，不仅要分析 SNP 位点所在区域的组蛋白修饰情况，还要从其他角度进行信息挖掘：如参考 ENCODE 数据库中转录因子 Chip-seq 的数据对 SNP 位点所在区域转录因子的结合情况进行分析，或是根据 Haploreg 数据库对 SNP 位点处转录因子结合位点进行预测以及分析 SNP 等位基因的变化是否会影响转录因子结合等信息。由于增强子对转录调控不受基因组上距离的限制，因此在鉴定 SNP 的功能性时，最好的办法是对不同等位基因细胞株进行全转录组测序，通过比较两者存在差异表达的基因来寻找 SNP 潜在的靶基因。

CRISPR/Cas9 基因组编辑系统及其衍生工具系统的出现极大地推动了基因组非编码区调控功能的研究，如 CRISPR 激活（CRISPR activation, CRISPRa）系统可对新的调控元件进行鉴定^[17]，CRISPR 干扰（CRISPR interference, CRISPRi）系统可特异性地抑制基因表达^[18]。由于这些转录调控的系统均依赖于 sgRNA 与基因



Note: A. Neighbor genes within 2 Mbp around rs1978421 [*IL12B*—*Homo sapiens* interleukin 12B (natural killer cell stimulatory factor 2, cytotoxic lymphocyte maturation factor 2, p40); *ADRA1B*—*Homo sapiens* adrenoceptor alpha 1B; *FABP6*—*Homo sapiens* fatty acid binding protein 6, ileal; *CCNJL*—*Homo sapiens* cyclin J-like; *C1QTNF2*—*Homo sapiens* C1q and tumor necrosis factor related protein 2; *miR3142HG*—transcript *miR3142HG*; *miR3142*—*Homo sapiens* microRNA 3142; *ZBED8*—*Homo sapiens* zinc finger BED-type containing 8; *ATP10B*—*Homo sapiens* ATPase, class V, type 10B; *GABRB2*—*Homo sapiens* gamma-aminobutyric acid (GABA) A receptor, beta 2; *GABRA6*—*Homo sapiens* gamma-aminobutyric acid (GABA) A receptor, alpha 6]. B. *miR-146a* expressions in wild type, homologous recombined rs1978421 cell lines and rs1978421 downstream DNA deletion or insertion cell lines (^① $P=0.898$, ^② $P=0.741$, ^③ $P=0.955$, compared with TT genotype). C. rs1978421 neighbor genes expressions in wild type, homologous recombined rs1978421 cell lines and rs1978421 downstream DNA deletion or insertion cell lines (^④ $P=0.408$, ^⑤ $P=0.423$, ^⑥ $P=0.548$, ^⑦ $P=0.543$, ^⑧ $P=0.443$, compared with TT genotype).

图3 不同基因型细胞中rs1978421邻近基因表达分析

Fig 3 rs1978421 neighbor genes expressions analysis in cell clones with different genotypes

组DNA序列的碱基互补配对, 因此我们设想了一种等位碱基特异性的表达调控系统: 已知sgRNA与基因组DNA之间的碱基错配会影响两者之间互补配对效率^[19-20], 那么如果能够证明rs2431697对miR-146a存在转录调控作用, 则可以把该区域作为靶点, 利用CRISPR激活系统上调miR-146a的表达, 以增强对过度活化的干扰素信号通路的负调控作用。sgRNA设计策略如下: 通过分析SLE患者基因组样本确定rs2431697位点在患者中的高频碱基, 以携带高频碱基的DNA片段为模板设计包含SNP位点的sgRNA序列, 在含有rs2431697不同等位碱基的细胞系中验证sgRNA与DNA的结合效率, 筛选与携带高频等

位碱基的基因组DNA结合能力强而与携带低频等位碱基的基因组DNA结合能力弱的sgRNA。而此sgRNA设计策略使得该基因表达激活系统仅在携带SNP风险等位基因的患者体内发挥效应, 对携带正常等位基因的患者则表现为低干预效应或无干预效应, 以减少对其免疫稳态的影响^[21], 从而建立了一种致病等位基因特异性的干预策略。

总之, 环境因素和遗传因素共同参与了诸多疾病的发展。虽然这些疾病发生的机制仍未明确, 但随着疾病易感基因和相关位点的功能机制研究日趋深入, 人们对疾病的理解也将会更加全面。

参 · 考 · 文 · 献

- [1] Zhao J, Ma JY, Deng Y, et al. A missense variant in *NCF1* is associated with susceptibility to multiple autoimmune diseases[J]. *Nat Genet*, 2017, 49(3): 433-437.
- [2] Sun CL, Molineres JE, Looger LL, et al. High-density genotyping of immune-related loci identifies new SLE risk variants in individuals with Asian ancestry[J]. *Nat Genet*, 2016, 48(3): 323-330.
- [3] Maurano MT, Humbert R, Rynes E, et al. Systematic localization of common disease-associated variation in regulatory DNA[J]. *Science*, 2012, 337(6099): 1190-1195.
- [4] Hnisz D, Abraham BJ, Lee TI, et al. Super-enhancers in the control of cell identity and disease[J]. *Cell*, 2013, 155(4): 934-947.
- [5] Shen C, Wang QZ, Shen ZY, et al. Genetic association between the *NLRP3*

- gene and acne vulgaris in a Chinese population[J]. Clin Exp Dermatol, 2019, 44(2): 184-189.
- [6] Wang XJ, Chen X, Yang ZZ, et al. Correlation of TET2 SNP rs2454206 with improved survival in children with acute myeloid leukemia featuring intermediate-risk cytogenetics[J]. Genes Chromosomes Cancer, 2018, 57(8): 379-386.
- [7] Ran FA, Hsu PD, Lin CY, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity[J]. Cell, 2013, 154(6): 1380-1389.
- [8] Gupta RM, Hadaya J, Trehan A, et al. A genetic variant associated with five vascular diseases is a distal regulator of endothelin-1 gene expression[J]. Cell, 2017, 170(3): 522-533. e15.
- [9] Higgs BW, Liu Z, White B, et al. Patients with systemic lupus erythematosus, myositis, rheumatoid arthritis and scleroderma share activation of a common type I interferon pathway[J]. Ann Rheum Dis, 2011, 70(11): 2029-2036.
- [10] Ivashkiv LB, Donlin LT. Regulation of type I interferon responses[J]. Nat Rev Immunol, 2014, 14(1): 36-49.
- [11] Tang YJ, Luo XB, Cui HJ, et al. MicroRNA-146a contributes to abnormal activation of the type I interferon pathway in human lupus by targeting the key signaling proteins[J]. Arthritis Rheum, 2009, 60(4): 1065-1075.
- [12] Bentham J, Morris DL, Cunningham Graham DS, et al. Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus[J]. Nat Genet, 2015, 47(12): 1457-1464.
- [13] Löfgren SE, Frostegård J, Truedsson L, et al. Genetic association of miRNA-146a with systemic lupus erythematosus in Europeans through decreased expression of the gene[J]. Genes Immun, 2012, 13(3): 268-274.
- [14] Hazelett DJ, Conti DV, Han Y, et al. Reducing GWAS complexity[J]. Cell Cycle, 2016, 15(1): 22-24.
- [15] Claussnitzer M, Dankel SN, Kim KH, et al. FTO obesity variant circuitry and adipocyte Browning in humans[J]. N Engl J Med, 2015, 373(10): 895-907.
- [16] Hua JT, Ahmed M, Guo HY, et al. Risk SNP-mediated promoter-enhancer switching drives prostate cancer through lncRNA PCAT19[J]. Cell, 2018, 174(3): 564-575. e18.
- [17] Konermann S, Brigham MD, Trevino AE, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex[J]. Nature, 2015, 517(7536): 583-588.
- [18] Qi LS, Larson MH, Gilbert LA, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression[J]. Cell, 2013, 152(5): 1173-1183.
- [19] Kleinstiver BP, Pattanayak V, Prew MS, et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects[J]. Nature, 2016, 529(7587): 490-495.
- [20] Pattanayak V, Lin S, Guilinger JP, et al. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity[J]. Nat Biotechnol, 2013, 31(9): 839-843.
- [21] Kim D, Lim K, Kim ST, et al. Genome-wide target specificities of CRISPR RNA-guided programmable deaminases[J]. Nat Biotechnol, 2017, 35(5): 475-480.

[收稿日期] 2020-02-17

[本文编辑] 邢宇洋

学术快讯

上海交通大学医学院附属新华医院沈立松团队与附属瑞金医院李鹤成团队发表食管癌联合研究成果

正常的机体免疫功能是抵御肿瘤发生和发展的重要环节，肿瘤微环境中的免疫细胞是机体抗肿瘤的重要组成部分。充分理解肿瘤微环境中免疫细胞的功能状态及调控网络将有效提高肿瘤的免疫治疗效果，从而延长患者的生存期。食管癌是最常见的消化道恶性肿瘤之一，治疗效果较差，肿瘤微环境中免疫细胞的组成和功能仍未得到充分的认识。

2020年12月，上海交通大学医学院附属新华医院检验科沈立松教授团队联合美国 Rutgers 大学 CAO Jian 教授和上海交通大学医学院附属瑞金医院心胸外科李鹤成教授团队在 *Nature Communications* 杂志发表题为“Immune suppressive landscape in the human esophageal squamous cell carcinoma microenvironment”的论文。该研究揭示了食管鳞状细胞癌（esophageal squamous cell carcinoma, ESCC）微环境中免疫细胞的组成和功能状态，全面系统地分析了 ESCC 浸润免疫细胞的特征，揭示了 ESCC 肿瘤微环境中抑制性的免疫状态，为应用和开发全新的 ESCC 免疫疗法奠定了基础。