

人工智能算法在癌症相关微 RNA 研究中的应用进展

鲁洪宇¹, 张佳¹, 曹一鑫², 吴书铭¹, 魏渊¹, 殷润婷¹

1. 江苏大学药学院, 江苏 镇江 212013

2. 江苏大学附属医院肿瘤内科, 江苏 镇江 212013

[摘要] 微RNA(miRNA)是一类通过不完全碱基互补配对实现后转录调控作用的小分子非编码RNA,其往往在癌症患者的病灶和外周血中表达失调。近年来,基于人工智能算法如机器学习和深度学习的模型逐渐应用于miRNA生物信息学研究。与传统的生物信息学工具比较,基于人工智能算法的miRNA靶点预测工具准确度更高,并实现了miRNA亚细胞定位和亚细胞重分布的预测,进一步深化了科研人员对miRNA的认识。此外,人工智能算法在临床模型构建的应用也显著提升了miRNA生物标志物的挖掘效率。本文总结了近年来人工智能算法在miRNA靶点预测、亚细胞定位和生物标志物挖掘的应用,并探讨了机器学习和深度学习对癌症相关miRNA研究的潜在价值。



[关键词] 微RNA; 机器学习; 深度学习; 靶点预测; 亚细胞分布; 临床预测模型; 综述

[中图分类号] TP181;R966 **[文献标志码]** A

Advances in applications of artificial intelligence algorithms for cancer-related miRNA research

LU Hongyu¹, ZHANG Jia¹, CAO Yixin², WU Shuming¹, WEI Yuan¹, YIN Runting¹

(1. School of Pharmacy, Jiangsu University, Zhenjiang 212013, Jiangsu Province, China;

2. Department of Medical Oncology, Affiliated Hospital of Jiangsu University, Zhenjiang

212013, Jiangsu Province, China)

Corresponding authors: YIN Runting, E-mail: yinrunting@126.com, <https://orcid.org/0000-0009-0008-8652>

收稿日期(Received):2023-10-30 接受日期(Accepted):2024-01-30 网络预发表日期(Online):2024-04-15

基金项目(Funding):国家重点研发计划(2022TFE0205100);中央本级重大增减支项目(2060302-2004-09);镇江市重点研发计划(社会发展)(SH2020036);江苏大学高级人才科研启动基金(21JDG022);江苏省研究生科研与实践创新计划(KYCX23_3752);江苏大学大学生创新创业训练计划(202310299479X)

第一作者(First author):鲁洪宇,硕士研究生,主要从事分子药理学、肿瘤学研究;E-mail:mpbolhy@126.com;<https://orcid.org/0009-0008-8652-6287>

通信作者(Corresponding author):殷润婷,副教授,主要从事分子药理学、肿瘤学研究;E-mail:yinrunting@126.com;<https://orcid.org/0000-0002-4344-6825> 魏渊,教授,主要从事药物代谢和毒理学研究;E-mail:ywei@ujs.edu.cn;<https://orcid.org/0000-0002-4025-6567>

0002-4344-6825; WEI Yuan, E-mail: ywei@ujs.edu.cn, <https://orcid.org/0000-0002-4025-6567>

[**Abstract**] MiRNAs are a class of small non-coding RNAs, which regulate gene expression post-transcriptionally by partial complementary base pairing. Aberrant miRNA expressions have been reported in tumor tissues and peripheral blood of cancer patients. In recent years, artificial intelligence algorithms such as machine learning and deep learning have been widely used in bioinformatic research. Compared to traditional bioinformatic tools, miRNA target prediction tools based on artificial intelligence algorithms have higher accuracy, and can successfully predict subcellular localization and redistribution of miRNAs to deepen our understanding. Additionally, the construction of clinical models based on artificial intelligence algorithms could significantly improve the mining efficiency of miRNA used as biomarkers. In this article, we summarize recent development of bioinformatic miRNA tools based on artificial intelligence algorithms, focusing on the potential of machine learning and deep learning in cancer-related miRNA research.

[**Key words**] MicroRNA; Machine learning; Deep learning; Target prediction; Subcellular distribution; Clinical prediction model; Review

[J Zhejiang Univ (Med Sci), 2024, 53(2): 231-243.]

[**缩略语**] 微RNA(microRNA, miRNA, miR); 信使RNA(messenger RNA, mRNA); 3'-非翻译区(3'-untranslated regions, 3'-UTR); 支持向量机(support vector machine, SVM); 人工神经网络(artificial neural network, ANN); 卷积神经网络(convolutional neural network, CNN); 循环神经网络(recurrent neural network, RNN); 图神经网络(graph neural network, GNN); 图卷积网络(graph convolutional network, GCN); 图注意网络(graph attention network, GAT); 接受者操作特征曲线下面积(area under receiver operating characteristic curve, AUROC); 精确性-召回率曲线下面积(area under the precision-recall curve, AUPRC); 候选靶标结合位点(candidate target site, CTS); 交联免疫共沉淀(crosslinking immunoprecipitation, CLIP); 长链非编码RNA(long non coding RNA, lncRNA); 环状RNA(circle RNA, circRNA); 马修斯相关系数(matthews correlation coefficient, MCC); 癌症基因组图谱(The Cancer Genome Atlas Program, TCGA); 核内不均一核糖核蛋白(heterogeneous nuclear ribonucleo-protein, hnRNP); 成纤维细胞生长因子(fibroblast growth factor, FGF)

miRNA 是一类长度为 18~22 nt 的非编码 RNA。在经典 miRNA 生成途径中, miRNA 先被转录为初级 miRNA, 在细胞核内经微处理器剪切为前体 miRNA, 经 exportin-5 进入细胞质, 前体 miRNA 经过 DICER 和 TRBP 处理生成成熟 miRNA。成熟 miRNA 在细胞质内与 AGO 家族蛋白形成 RNA 诱导沉默复合物, 抑制 mRNA 的翻译活性^[1]。miRNA 通常以 5' 端 2~8 nt 高度保守的种

子序列结合于 3'-UTR 实现对 mRNA 的后转录调控。哺乳动物体内约 50% 的编码基因受到 miRNA 调控^[2]。miRNA 广泛参与如肿瘤^[3]、神经退行性疾病^[4]、排斥反应^[5]等疾病的发展, 具有作为药物靶点^[6]和生物标志物^[4]的潜力。随着测序技术的开发应用和计算机科学的进步, 各类 miRNA 的生物信息学工具取得了长足发展。为了探究 miRNA 的分子机制, 研究人员开发了大量

miRNA 靶点预测工具,如 TargetScan^[7]、miRDB^[8]、PicTar^[9]等。当前 miRNA 生物信息学研究主要聚焦于 miRNA 靶点预测、miRNA 时空分布、基于 miRNA 表达数据的临床预测模型和 miRNA 相关生物分子网络。然而,尽管 miRNA 相关生物信息学工具已经取得了长足进展,现有的 miRNA 生物信息学工具仍有待进一步的发展,如传统 miRNA 靶点预测工具存在准确度不足的问题,miRNA 亚细胞定位预测工具存在记录较少和无法预测 miRNA 重分布的缺陷。基于人工智能算法和新一代测序技术的生物信息学工具开发将有助于缓解这些问题。

随着计算机科学的发展,“人工智能”的概念越发普及。广义的人工智能算法包括机器学习、强化学习和深度学习。机器学习旨在根据数据集间的关联开发预测模型^[10],常用于生物信息学的机器学习算法包括线性回归、逻辑回归、SVM、随机森林和深度神经网络^[11],其中深度神经网络近年来已逐渐成为独立于机器学习模型的热门研究领域。深度学习是一类仍在发展中的机器学习算法^[10]。生物信息学中常用的深度学习算法分为监督学习、非监督学习和半监督学习。常见的监督学习模型包括 ANN、CNN 和 RNN,常用于回归和分类工作。ANN 难以捕捉序列等数据中的顺序信息,且处理多维数据时易出现可训练参数过多的问题,因此更适合处理表格等向量数据;CNN 常用于处理包括图像信息在内的多维数据;RNN 可以处理文本、序列等一维数据^[12-13]。非监督学习模型常用于聚类 and 异常事件检测,常用模型为生成对抗网络和自动编码器^[14]等。半监督学习是一类近年来新兴的深度学习算法,常用于处理基于节点和边的图数据^[15],常用的模型为 GNN、GCN 和 GAT 等。图 1 总结了监督学习、非监督学习和半监督学习的工作流程。深度学习模型的常用评价指标为准确度、灵敏度、特异度、AUROC 和 AUPRC 等^[16],这些参数是基于模型预测结果真阳性、真阴性、假阳性和假阴性四个值的计算结果。其中,AUROC 反映了分类器模型区分样本的总体能力,是最常用的衡量监督学习模型性能的参数之一。相

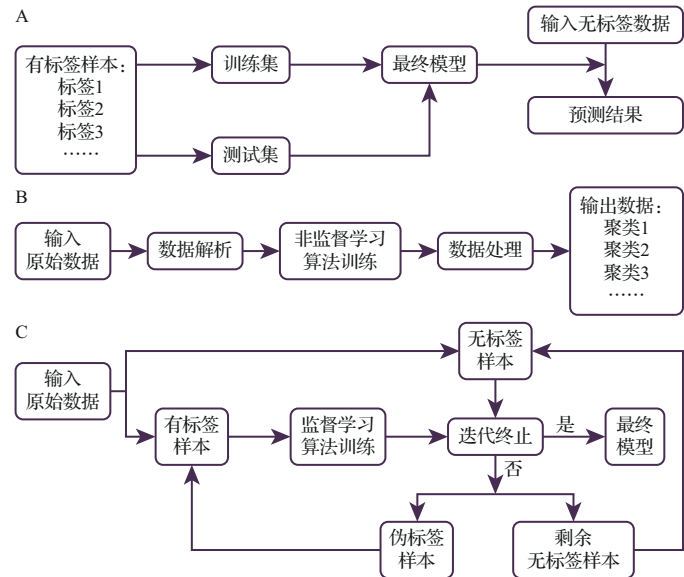
较于传统的 miRNA 生物信息学工具,基于机器学习和深度学习的生物信息学工具能够通过提取原始数据中的抽象特征对数据进行预测和分类,使预测结果更接近数据的实际分布,从而获得更精确的预测结果。

本文主要总结机器学习和深度学习模型在 miRNA 靶点识别、miRNA 亚细胞分布、miRNA 生物标志物挖掘等领域中的应用,并探讨新一代测序技术在 miRNA 生物信息学研究中的应用前景,以期为基础机器学习和深度学习的 miRNA 生物信息学工具进一步开发和应用提供参考。

1 机器学习和深度学习模型用于微 RNA 靶点识别和预测

miRNA 通过高度保守的种子序列识别靶点,科研人员基于这一分子机制开发了多种工具预测 miRNA 的 CTS。传统的靶点预测算法基于 miRNA:CTS 的互补配对、miRNA:CTS 的吉布斯自由能、序列保守性这三条原则^[17],这一类预测算法也称为基于配对原则的预测算法。常见的机器学习和深度学习 miRNA 靶点预测模型基于对训练数据集中的 miRNA 阳性靶点和阴性靶点进行学习,获取隐藏在靶点参数中的抽象特征,从而实现对 miRNA 靶点进行预测^[18-19]。

近年来,机器学习模型在 miRNA 靶点预测领域体现出较传统预测算法更高的精确性。miRDB



A: 监督学习; B: 非监督学习; C: 半监督学习。

图1 深度学习流程示意图

Figure 1 Workflow of deep learning

的最新更新中提出了名为 MirTarget 的靶点预测算法,基于 CLIP 测序和过表达对应 miRNA 的测序数据,采用 SVM 模型预测 miRNA 靶点^[8]。基于核函数构建的 miRNA 靶点预测模型 mintRULS,通过提取 miRNA:CTS 的自由能、miRNA 靶点可及性、miRNA 靶点序列的 AU 含量和 miRNA 靶点的简单重复序列等参数,通过最小二乘法预测 miRNA 靶点。与 miRDB、TargetScan、MBSTAR、RPmirDIP 和 STarMir 等现有 miRNA 预测工具比较,mintRULS 有着更优的精确性和特异性^[18]。尽管机器学习体现出了预测 miRNA 靶点的优越性,但仍需要人为选取靶点特征参数,易引入主观因素导致的误差^[20]。这一局限性限制了机器学习模型的性能。

深度学习模型既可采用人为选取的靶点参数,也可直接提取序列信息,后者由深度神经网络分析原始数据下的规律,可以在一定程度上避免主观误差^[12]。采用深度学习模型的 miRNA 靶点预测算法流程如图 2 所示。miRTDL 是一种采用了 CNN 的 miRNA 靶点预测模型,选取了 9 个种子配对特征、3 个进化保守性特征和 8 个靶点可及性特征为学习样本。研究人员以 1606 个经实验验证的 miRNA 靶点作为测试数据集,发现 miRTDL 具有极高的准确度、灵敏度和特异度^[19]。然而,miRTDL 的学习样本仍需要人为提取特征,其精确性依旧有限。

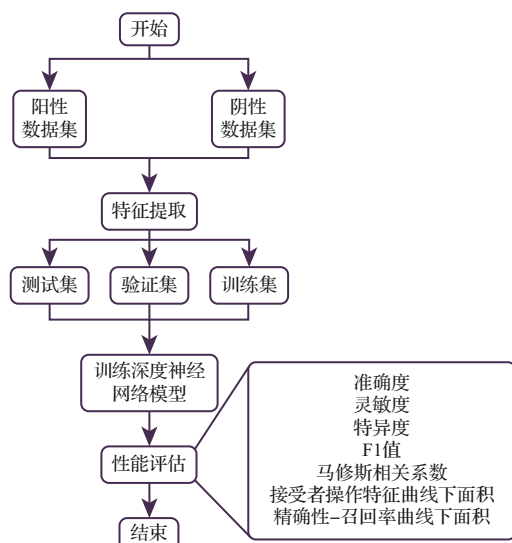


图2 基于深度神经网络的微RNA靶点预测模型构建流程图

Figure 2 Workflow of miRNA target prediction model based on deep neural network

编码序列信息的深度学习模型普遍具备更高的精确性。基于序列的深度学习 miRNA 靶点预测模型需要将核苷酸序列从文本编码为数字向量。常用的编码手段为独热编码和 k-mer 法。独热编码的原理为将序列或碱基匹配信息转化为数字向量输入神经网络^[7, 12-13]。k-mer 法的原理是将 RNA 序列打断成不同长度的序列,并将其转化为二进制向量^[21-22]。miRAW 通过 ANN 提取了来自 CLIP 测序、CLASH 和 iPAR-CLIP 数据集近 20 000 条经验证的 miRNA 靶点序列特征^[12]。miRAW 的编码策略是通过独热编码将碱基和空白配对转化为 0 和 1 构成的向量。然而,由于 CLASH 等数据集中的靶点主要位于 3'-UTR, miRAW 和 DeepMirTar^[23]等算法未将位于 3'-UTR 以外的靶点纳入学习样本,导致预测结果不够全面。miTAR 采用 RNN 和 CNN 读取 miRNA 及其靶点序列以提取 miRNA 靶点特征^[13]。miTAR 采用嵌入层将碱基编码为五维向量,缓解了独热编码导致的资源占用问题。TargetScan 开发者在对 miRNA 靶点结合模式的探索中,以 miRNA 序列中的 1~10 nt 序列、CTS 序列、miRNA:CTS 的自由能、miRNA 靶点下调倍数为学习样本,采用了 CNN 预测 miRNA 的靶点亲和力和靶点抑制效率^[7]。不同于传统的独热编码,此模型将 miRNA 的 1~10 nt 序列、CTS 序列和碱基的 16 种组合构建为 10×12 矩阵输入 CNN。这使统计模型能够获取 miRNA:CTS 中全部碱基的位置信息,为处理多维数据提供了新颖的思路。

lncRNA、circRNA 等非编码 RNA 与 miRNA 的交互作用近年来受到广泛研究^[24-25]。基于深度学习的 miRNA 靶点预测模型在预测 miRNA:lncRNA 和 miRNA:circRNA 交互中也取得了进展。ncRNAInter 基于 GNN,以 lncRNASNP2 中 miRNA:lncRNA 交互的密码子特征、开放阅读框特征、碱基含量占比、转录本序列特征、理化性质和二级结构 6 类共 191 种特征参数为学习样本,构建了精确度高达 93.1% 的预测模型^[26]。值得注意的是,lncMirNet 与 ncRNAInter 的学习样本来源相同,但采用 CNN 的 lncMirNet 在精确度、灵敏度、特异度、F1 值、MCC 和 AUROC 几个指标均低于 ncRNAInter^[21]。KGDCMI 采用 k-mer 编码,通过图嵌入提取 miRNA:circRNA 交互的特征,经过主成分分析进一步提取特征向量,导入深度神经

网络进行预测^[22]。上述研究表明,在采用相同学习样本时,不同的算法模型、不同的数据预处理策略都将影响模型的最终性能。

因为模型参数过多和训练样本规模不足的原因,深度学习模型常存在过拟合问题。为了避免过拟合,有的深度学习模型会采用 Dropout 算法,即在迭代过程中随机丢弃一部分神经节点。miRAW^[12]和 miTAR^[13]均采用了 20% 的丢弃率。此外,K 折交叉验证也是常用策略^[21]。总之,miRNA 靶点预测模型的优化主要依赖于数据集的来源、数据预处理方式、数据集的划分以及统计模型的优化。

值得注意的是,上述靶点预测模型大多采用监督学习模型。监督学习模型需要标记学习样本,在 miRNA 靶点预测模型中表现为标记阳性靶点和阴性靶点。miRNA 靶点预测模型的学习样本多来源于 CLIP 测序、CLASH 等阳性样本,阴性样本的匮乏导致正负样本分布不均匀,限制了监督学习模型的功能。随机生成模拟序列作为阴性样本常导致假阳性问题。非监督学习模型常用于聚类,更适合正负样本不均匀的数据集。DeepMirTar 采用了来自 miRMark 和 CLASH 的 miRNA 靶点,基于堆叠去噪自动编码器,以种子配对、自由能、序列组成、靶点位置、保守性、序列编码信息七个类别共 750 个特征参数为学习样本^[23],其真阳性率达到 0.9235 (AUROC: 0.9793),较真阳性率仅为 0.8701 的二维 CNN 模型 (AUROC: 0.9410) 性能更强。然而,DeepMirTar 的学习样本仍依赖人为提取特征,采用原始序列的非监督学习模型亟待开发。除了非监督学习,采用半监督学习的 ncRNAInter 在准确度和灵敏度上也呈现了较监督学习模型更强的靶点预测性能^[21, 26]。表 1、2 比较了几种 miRNA 靶点预测模型的特征及其参数。

随着大数据等技术的发展,单细胞测序已成为分子医学领域的研究热点。由于测序深度不足,单细胞 miRNA 测序仍处于探索阶段^[29-30],而基于单细胞测序数据预测单细胞

miRNA 丰度成为备选方案。Nielsen 等^[31]基于 TCGA 中 3'-UTR 富集程度和基因表达丰度的关联构建了基于种子配对的 miRNA 活性预测算法 miReact,并根据肝组织单细胞测序的 miR-122 靶基因丰度预测了细胞特异性的 miR-122 活性。考虑到 miRNA 功能和分子机制的复杂性,这一预测算法有明显的局限性。此外,由于单细胞测序常存在特征值缺失和数据稀疏的问题,优化单细胞 miRNA 预测模型有赖于更高精度的单细胞测序技术开发。

2 机器学习模型用于解析亚细胞定位

miRNA 呈时空异质性分布,除了在细胞内发挥功能,miRNA 也可从 RNA 诱导沉默复合物解离,与 hnRNP 家族蛋白结合并被分选入外泌体^[32]。近年来,miRNA 在细胞核^[33]、线粒体^[34]以及细胞外^[35]的功能越来越得到重视,其可为肿瘤、神经退行性疾病等难治性疾病新型治疗方案的开发提供思路^[34, 36-38]。以 miRNA 为靶点的疗法已在临床试验中取得了进展。用锁核酸药物 LNA-i-miR-221 治疗 17 例结肠癌患者,结果显示其具有一定安全性和治疗效果^[39];小分子药物 ABX464 能够选择性上调免疫细胞 miR-124 表达水平,目前处于结肠炎治疗的 III 期临床试验阶段^[40]。在累及血管的川崎病病灶中,内皮细胞微囊中的 miRNA 特异性表达^[41],提示外泌体 miRNA 有作为治疗靶点的潜力。外泌体等细胞外囊泡

表 1 微RNA靶点预测模型一览
Table 1 Summary of miRNA target prediction approaches

模型名称	算 法	靶点类型	适用种属	参考文献
TargetScan 7.0	线性回归	mRNA	现代智人、小鼠、果蝇、线虫、斑马鱼	[27]
miranda	SVM	mRNA	现代智人、小鼠、大鼠、果蝇、线虫	[28]
miRDB	SVM	mRNA	现代智人、小鼠、大鼠、家犬、鸡	[8]
mintRULS	核函数	mRNA	现代智人、小鼠	[18]
miRTDL	CNN	mRNA	现代智人	[19]
miRAW	ANN	mRNA	现代智人	[12]
DeepMirTar	SDA	mRNA	现代智人	[23]
miTAR	CNN、RNN	mRNA	现代智人	[13]
ncRNAInter	GNN	lncRNA	现代智人、植物、病毒	[26]
lncMirNet	CNN	lncRNA	现代智人	[21]
KGDCMI	GNN	circRNA	现代智人	[22]

SVM: 支持向量机; CNN: 卷积神经网络; ANN: 人工神经网络; SDA: 堆叠去噪自动编码器; RNN: 循环神经网络; GNN: 图神经网络; mRNA: 信使 RNA; lncRNA: 长链非编码 RNA; circRNA: 环状 RNA。

表2 基于深度学习的微RNA靶点预测模型数据集信息及特征参数

Table 2 Datasets and performance of miRNA target prediction models based on deep learning

模型名称	学习样本来源	训练集规模	测试集规模	准确度	灵敏度	特异度
miRTDL	TargetScanS、TarBase、miRBase	19 000	2915	0.8998	0.8843	0.9644
miRAW	CLASH、iPAR-CLIP	14 000	3500	0.9200	0.9200	0.9400
DeepMirTar	mirMark、CLASH	4408	1469	0.9348	0.9235	0.9479
miTAR	DeepMirTar、miRAW	17 740	5544	0.9627	0.9591	0.9663
ncRNAInter	lncRNASNP2	24 840	2760	0.9309	0.9272	0.9346
lncMirNet	lncRNASNP2	24 617	6154	0.8534	0.9158	0.7910
KGDCMI	circbank、circR2Cancer	6933	990	0.8265	0.8019	0.8510
模型名称	精确率	F1值	MCC	AUROC	AUPRC	参考文献
miRTDL	—	—	—	—	—	[19]
miRAW	—	0.9200	—	0.9600	—	[12]
DeepMirTar	—	0.9348	0.8699	0.9793	—	[23]
miTAR	0.9664	0.9627	—	—	—	[13]
ncRNAInter	0.9342	0.9307	0.8619	0.9715	0.9741	[26]
lncMirNet	—	0.8620	0.7124	0.9381	—	[21]
KGDCMI	0.8435	—	0.6538	0.8930	0.9767	[22]

—:无相关数据.表中用于评价深度学习模型的指标均基于真阳性(TP)、真阴性(TN)、假阳性(FP)和假阴性(FN)计算,计算公式:准确度=(TP+TN)/(TP+TN+FP+FN),灵敏度=TP/(TP+FN),特异度=TN/(TN+FP),精确率=TP/(TP+FP),F1值=2TP/(2TP+FP+FN),MCC=(TP×TN-FP×FN)/sqrt[(TP+FP)×(TP+FN)×(TN+FP)×(TN+FN)].MCC:马修斯相关系数;AUROC:接受者操作特征曲线下面积;AUPRC:精确性-召回率曲线下面积.

在抗癌疗法开发^[42]和标志物挖掘^[4]等领域也展现出潜力。目前,采用外泌体治疗病毒感染^[43]、伤口愈合^[44]等治疗方案已经处于临床试验阶段。miRNA在细胞内外分布的影响因素也引起了科研人员的广泛兴趣。除了探究miRNA与其靶点的交互模式,解析miRNA的亚细胞分布也能够深入理解miRNA功能及其分子机制,并为开发新型药物靶点提供指导。

近年来的研究表明,miRNA的细胞内外分布和外泌体分选存在序列特异性^[32],提示miRNA的亚细胞分布可能同样存在序列特异性。miRLoc基于miRNA的热力学、结构特性和双核苷酸的主成分得分,采用SVM构建了miRNA的亚细胞分布预测算法,预测miRNA在循环血、外泌体、细胞外囊泡、细胞质、微粒体、线粒体和细胞核的分布^[45]。MirLocPredictor采用k-mer法编码miRBase中miRNA的序列,以RNALocate数据库中的miRNA位置信息为标签,采用RNN和CNN预测miRNA的亚细胞分布^[46]。然而,由于miRNA的序列长度较短,k-mer法编码miRNA序列误差较大,而这在mRNA中鲜有出现。此外,研究发现调控线粒体功能的mRNA定位于线粒体^[37, 47],调控线粒体mRNA的miRNA同样定位于线粒

体,提示基于miRNA靶点构建miRNA亚细胞分布预测模型具有可行性。miRLoc采用重启随机游走算法和消息传递算法,整合了miRNA-mRNA交互网络、miRNA功能相似性网络和mRNA的亚细胞定位信息以预测miRNA的亚细胞分布^[48]。这提示生物分子网络预测亚细胞分布的潜力,GNN、GCN等半监督学习算法常用于构建生物分子网络。DAmiRLocGNet利用自动编码器提取miRNA序列特征,采用GCN构建了基于miRNA功能相似性网络和miRNA-疾病网络的异质性网络,这两者结合得到的亚细胞分布预测模型具有最强的预测性能(AUROC:0.8049)^[49]。表3比较了几种常见miRNA亚细胞分布预测算法的特点及其参数。值得注意的是,这些算法均采用了K折交叉验证以避免深度学习模型的过拟合。

然而,这些算法未收录miRNA在所有细胞器的分布记录,如调控溶酶体的miRNA^[50]。这主要因为数据来源匮乏,miRNA亚细胞定位数据库RNALocate中相关记录极少^[51]。这在RNALocate v2.0的更新中并未得到改善^[52]。值得注意的是,这些模型不能预测外部刺激导致的miRNA重分布^[53]。结合miRNA-疾病网络的深度学习模型能够在一定程度上解决这个问题,并为挖掘生物标

志物和后转录调控靶点提供思路。此外,近年来大量研究报道了miRNA与hnRNP家族等RNA结合蛋白的交互作用^[54]。RNA结合蛋白在miRNA的外泌体分选^[54]、线粒体转运^[55]等生物过程中的作用越发受到关注,RNA结合蛋白免疫沉淀高通量测序和细胞器小RNA测序将有助于开发新的miRNA亚细胞分布预测算法。

近年来,空间转录组受到广泛关注。空间转录组主要分为基于测序的空间转录组和基于影像的空间转录组^[56]。当前商业化的空间转录组主要基于对切片的原位杂交或高通量测序,活细胞空间转录组能更直观地研究RNA的细胞内外再分布。2014年,Lovatt等^[57]首次标记了活组织中单个细胞的mRNA,开启了活细胞空间转录组的研究和开发工作。在miRNA检测方面,Wei等^[58]在单细胞尺度实现了对乳腺癌中9个miRNA的活细胞成像。这一成果不仅在miRNA测序领域结合了空间转录组和单细胞测序,还能通过细胞器靶向技术直接观测miRNA亚细胞重分布。可以预见,空间转录组技术的应用和发展将一定程度缓解miRNA的亚细胞分布数据稀缺和研究不足的问题。

3 机器学习模型用于挖掘微RNA生物标志物

血液、血清、尿液、胸腔积液以及外泌体

miRNA可作为一类极具潜力的非侵入性生物标志物^[4,59-62]。探究循环miRNA、外泌体miRNA和其他细胞外miRNA有助于深入理解相关疾病,并为开发新型miRNA分布预测算法提供支持^[49]。

基于体液miRNA表达谱和临床数据的机器学习模型适用于挖掘生物标志物。Moisoiu等^[63]基于尿液miRNA测序,采用逻辑回归、朴素贝叶斯和随机森林构建诊断模型,发现尿液中miR-34a-5p、miR-205-3p、miR-210-3p高表达与膀胱癌发生相关。除此之外,检测尿液中的miR-615-3p和miR-185-5p可区分管腔型与基底型膀胱癌,miR-615-3p在基底型膀胱癌中高表达,miR-185-5p在管腔型膀胱癌中高表达。Irlam-Jones等^[64]在一项三期临床试验中发现,在接受放疗和缺氧修饰剂治疗后,miR-210高表达膀胱癌患者的5年生存期显著改善。Sathipati等^[65]采用SVM和双目标组合遗传算法,基于TCGA的miRNA表达谱和临床数据发现,let-7f、miR-1237、miR-98、miR-933和miR-889的丰度与卵巢癌患者的生存期显著相关。其中,卵巢癌患者血浆let-7f表达水平较健康受试者低,且血浆let-7f丰度与上皮性卵巢癌预后相关^[66]。Morokoff等^[67]采用随机森林算法,在47例2级胶质瘤患者、44例4级胶质瘤患者和17名健康受试者的外周血miRNA测序中发现,血清miR-320e、mir-223和miR-21与肿瘤体积相关,为胶质瘤患者

表3 微RNA亚细胞分布预测模型一览
Table 3 Summary of miRNA subcellular location prediction models

模型名称	算法	亚细胞分布			数据库			数据集划分和验证方法
MiRNALoc	SVM	轴突、循环系统、细胞质、外泌体、细胞核、胞外微囊、微囊、线粒体			RNAlocate v1.0、miRBase			五折交叉验证
MirLocPredictor	CNN、RNN	细胞质、外泌体、细胞核、微囊、线粒体			RNAlocate v1.0			十折交叉验证
MiRLoc	RWR、mPAT	细胞质、外泌体、细胞核、核仁、胞外微囊、微囊、线粒体			RNAlocate v2.0、miRBase、HMDD v3.0、MISIM v2.0			十折交叉验证
DAmiRLocGNet	GCN、自动编码器	细胞质、外泌体、细胞核、胞外微囊、微囊、线粒体			RNAlocate v2.0、miRBase、HMDD v3.2			十折交叉验证
模型名称	灵敏度	特异度	精确率	F1 值	MCC	AUROC	AUPRC	参考文献
MiRNALoc	0.6949	0.7619	—	0.7238	0.6588	0.7197	0.7562	[45]
MirLocPredictor	0.6784	—	0.6878	0.6178	—	0.6098	0.4990	[46]
MiRLoc	—	—	—	—	—	0.7270	0.6280	[48]
DAmiRLocGNet	—	—	—	—	—	0.8049	0.7281	[49]

—:无相关数据。表中用于评价深度学习模型的指标均基于真阳性(TP)、真阴性(TN)、假阳性(FP)和假阴性(FN)计算,计算公式:灵敏度=TP/(TP+FN),特异度=TN/(TN+FP),精确率=TP/(TP+FP),F1值=2TP/(2TP+FP+FN),MCC=(TP×TN−FP×FN)/sqrt[(TP+FP)×(TP+FN)×(TN+FP)×(TN+FN)]。SVM:支持向量机;CNN:卷积神经网络;RNN:循环神经网络;RWR:重启随机游走算法;mPAT:消息传递算法;GCN:图卷积网络;MCC:马修斯相关系数;AUROC:接受者操作特征曲线下面积;AUPRC:精确性-召回率曲线下面积。

提供了一组预后标志物。Wong 等^[68]基于 ANN、随机森林算法、梯度增强分类器和逻辑回归算法,在中期原发性肝癌患者、结直肠癌肝转移癌患者和健康志愿者的血浆 miRNA 表达谱中发现,与健康志愿者比较,中期原发性肝癌患者血浆 miR-221-3p、miR-223-3p、miR-26a-5p 和 miR-30c-5p 下调,而结直肠癌肝转移癌患者血浆 miR-365a-3p 和 miR-423-3p 上调,从而实现原发性肝癌和结直肠癌肝转移的区分。Kim 等^[69]采用复合协变量模型、对角线性判别模型、最近质心模型和 SVM,在 108 个胃癌标本的微阵列数据中发现 miR-628-5p、miR-1587、miR-3175、miR-3620-5p、miR-4459、miR-4505、miR-4507、miR-4720-5p、miR-4742-5p、miR-6779-5p 丰度上调或 miR-106b-3p、miR-125a 丰度下调可预测胃癌淋巴转移。Huang 等^[70]采用复合协变量、对角线性判别和 SVM 发现,血清高 miR-21、低 miR-22 和低 miR-29c 可作为胃癌的诊断指标。Savareh 等^[71]组合粒子群优化、ANN 和邻域成分分析算法发现,胰腺癌患者外周血高 miR-1469、高 miR-663a 和低 miR-532-5p 丰度与预后不良显著相关。值得注意的是,由于算法和学习样本的差异,同类肿瘤的候选生物标志物常存在差异,基于机器学习的诊断模型精度有待进一步提升。

此外,基于肿瘤组织 miRNA 表达谱和临床数据的机器学习模型也能提高候选药物靶点的挖掘效率。Liu 等^[72]采用 SVM 在 TCGA 的 miRNA 表达谱和临床数据中发现,miR-383、miR-615 和 miR-877 可作为头颈部鳞状细胞癌的诊断和预后标志物,其中 miR-383^[73]和 miR-877^[74]已证明能够调控头颈部鳞状细胞癌的进展。Ye 等^[75]采用 SVM 和主成分分析在 TCGA 数据库和 GSE16025 数据集中发现,miR-143、miR-100、miR-101-1、miR-101-2、miR-182、miR-183、miR-205、miR-21、miR-30a、miR-30d 的表达可区分肺鳞状细胞癌与正常对照样本。在这些候选标志物中,miR-21 靶向 LZFTL1 诱导肺鳞状细胞癌的耐药性^[76];miR-30d 能够促进肺鳞状细胞癌增殖^[77]。此外,在一项随机对照试验中,Zhu 等^[78]发现血清低 miR-21 丰度可作为肺癌脑转移患者的预后标志物。Zhang 等^[79]采用随机森林算法,在肺癌患者的外周血 miRNA 表达谱中发现,miR-92a、miR-140-5p、miR-331-3p、miR-223、miR-374a 表现出作

为预后标志物的潜力。其中,miR-92a 可靶向 Sprouty4 促进肺癌的上皮间质转化^[80],也可诱导肺癌的耐药性^[81];miR-140-5p 可靶向 FGF9 抑制肺癌进展^[82],也可靶向 WEE1 抑制肺癌的耐药性^[83];miR-331-3p^[84]、miR-223^[85]和 miR-374a^[86]均能够抑制肺小细胞肺癌的迁移。在一项针对烟民的临床研究中发现,与无肺癌烟民比较,发生肺癌的烟民支气管上皮中 miR-223 表达较低^[87]。Zhao 等^[88]通过多变量逻辑回归,在 77 例肺癌患者的肿瘤组织微阵列中发现,miR-210、miR-214 和 miR-15a 的表达数据可预测肺癌患者的脑转移。提示机器学习模型能够对细胞表型和肿瘤类型进一步分类^[89],并针对肿瘤异质性问题开发新型靶点。

4 结 语

近年来,多组学联合分析成为生物学的热门领域。miRNA 测序与代谢组学、mRNA 测序、lncRNA 测序、circRNA 测序和 DNA 甲基化测序等高通量数据的结合将帮助科研人员更加系统地了解生物过程。Olgun 等^[90]开发了基于极梯度提升算法的 miRSCAPE,以 TCGA、基因型-组织表达和癌细胞系百科全书数据库中的 miRNA 表达谱和 mRNA 表达谱为训练数据,成功基于单细胞测序数据预测了单细胞 miRNA 丰度,并在乳腺癌和胰腺癌的单细胞测序数据中预测了发生上皮间质转化细胞的 miRNA 表达谱。Kutlay 等^[91]整合了 TCGA 黑色素瘤的转录组数据、miRNA 表达谱和 DNA 甲基化数据,采用多种机器学习模型精确预测了黑色素瘤转移的相关通路。Moisoiu 等^[63]构建的膀胱癌诊断模型结合了尿液 miRNA 表达数据和尿液的表面增强拉曼光谱,实现了对膀胱癌分子亚型的精确预测。上述研究表明,机器学习和深度学习将为科研人员提供更加高效的生物信息学工具。

miRNA 常在不同肿瘤中共同表达,但功能不尽相同,如 miR-381 可促进肾癌获得耐药性^[92]并抑制乳腺癌的上皮间质转化^[93];miR-155 可促进肺癌进展^[94]并提升三阴性乳腺癌干性^[95];miR-125b-5p 可促进胶质母细胞瘤侵袭^[96]和结直肠癌迁移侵袭^[97];miR-378 可提升白血病细胞干性^[98]并促进胰腺癌进展^[99]。miRNA-疾病网络^[100]从统计学角度为这一现象提供了解释,被聚类的疾病往往存

在共同表达的 miRNA。网络医学的发展为研究人员系统了解发病机制提供了机会^[101],也为挖掘广谱抗肿瘤靶点和生物标志物提供了依据。生物分子网络的构建基于生物分子间的关系,除了分子交互作用^[102],发病过程中出现的相似表达水平变化同样可以作为构建生物分子网络的数据来源。MuCoMiD 采用 GCN 提取 miRNA 相关网络、编码基因相关网络和疾病相关网络的特征,构建了 miRNA-编码基因-疾病相关网络,首次提出了预测基因与疾病的相关性的多任务学习方法^[103]。采用 GNN 的模型 NGCICM 整合了 miRNA: circRNA 交互、miRNA-疾病网络和 circRNA-疾病网络以探究疾病中的 circRNA-miRNA 轴^[104]。除了生物分子-疾病交互,同一信号通路和生物过程中常有不同的 miRNA 参与^[105-106]并调控不同疾病的进展^[107-108],提示 miRNA-生物过程-疾病网络挖掘潜在药物靶点的潜力,也为非编码 RNA 的功能注释提供了思路。此外,近年来中药活性单体的药理作用研究也取得了长足进展^[109-112],中药单体对同一疾病的后转录调控呈多靶点特征^[113-114]。中药网络药理学^[115-117]的广泛应用也提示构建中药活性单体-miRNA-mRNA-疾病网络的可能。总之,基于深度学习算法的 miRNA 靶点预测模型、亚细胞分布预测模型与 miRNA 相关网络等生物分子网络结合^[118],将会更加系统地为 miRNA 分子机制、miRNA 生物功能、后转录调控药物开发、miRNA 生物标志物挖掘等研究工作提供依据。

志谢 研究受到国家重点研发计划(2022TFE0205100)、中央本级重大增减支项目(2060302-2004-09)、镇江市重点研发计划(社会发展)(SH2020036)、江苏大学高级人才科研启动基金(21JDG022)、江苏省研究生科研与实践创新计划(KYCX23_3752)、江苏大学大学生创新创业训练计划(202310299479X)支持。江苏大学药学院王星艳、白荣裕、赵瑒、朱俊同学参与资料查询和收集工作

Acknowledgements This work was supported by the National Key R&D Program of China (2022TFE0205100), Central-level Major Increases and Decreases in Expenditure Project (2060302-2004-09), Social Development Project of Zhenjiang City (SH2020036), Scientific Research Foundation for the Senior Personnel of Jiangsu University (21JDG022), Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX23_3752), College Student Innovation Program of Jiangsu University (202310299479X). WANG

Xingyan, BAI Rongyu, ZHAO Chang and ZHU Jun from School of pharmacy, Jiangsu University participated in data searching and collection

利益冲突 所有作者均声明不存在利益冲突

Conflict of Interests The authors declare that there is no conflict of interests

©The author(s) 2024. This is an open access article under the CC BY-NC-ND 4.0 License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

参考文献(References)

- [1] GALLACH S, CALABUIG-FARIÑAS S, JANTUS-LEWINTRE E, et al. MicroRNAs: promising new antiangiogenic targets in cancer[J]. *Biomed Res Int*, 2014, 2014: 878450.
- [2] KROL J, LOEDIGE I, FILIPOWICZ W. The widespread regulation of microRNA biogenesis, function and decay [J]. *Nat Rev Genet*, 2010, 11(9): 597-610.
- [3] LIN X, PING J, WEN Y, et al. The mechanism of ferroptosis and applications in tumor treatment[J]. *Front Pharmacol*, 2020, 11: 1061.
- [4] WEI H, XU Y, XU W, et al. Serum exosomal miR-223 serves as a potential diagnostic and prognostic biomarker for dementia[J]. *Neuroscience*, 2018, 379: 167-176.
- [5] WANG H, FAN H, TAO J, et al. MicroRNA-21 silencing prolongs islet allograft survival by inhibiting Th17 cells[J]. *Int Immunopharmacol*, 2019, 66: 274-281.
- [6] RUI M, QU Y, GAO T, et al. Simultaneous delivery of anti-miR21 with doxorubicin prodrug by mimetic lipoprotein nanoparticles for synergistic effect against drug resistance in cancer cells[J]. *Int J Nanomedicine*, 2017, 12: 217-237.
- [7] MCGEARY S E, LIN K S, SHI C Y, et al. The biochemical basis of microRNA targeting efficacy[J]. *Science*, 2019, 366(6472): eaav1741.
- [8] CHEN Y, WANG X. MiRDB: an online database for prediction of functional microrna targets[J]. *Nucleic Acids Res*, 2020, 48(D1): D127-D131.
- [9] LALL S, GRÜN D, KREK A, et al. A genome-wide map of conserved microRNA targets in *C. elegans*[J]. *Curr Biol*, 2006, 16(5): 460-471.
- [10] CAMACHO D M, COLLINS K M, POWERS R K, et al. Next-generation machine learning for biological networks[J]. *Cell*, 2018, 173(7): 1581-1592.
- [11] SCHÄFER M, CIAUDO C. Prediction of the miRNA interactome—established methods and upcoming perspectives[J]. *Comput Struct Biotechnol J*, 2020, 18: 548-557.
- [12] GU T, ZHAO X, BARBAZUK W B, et al. MiTAR: a hybrid deep learning-based approach for predicting miRNA targets[J]. *BMC Bioinformatics*, 2021, 22

- (1): 96.
- [13] SAPOVAL N, AGHAZADEH A, NUTE M G, et al. Current progress and open challenges for applying deep learning across the biosciences[J]. **Nat Commun**, 2022, 13(1): 1728.
 - [14] ZHOU Z H. A brief introduction to weakly supervised learning[J]. **Natl Sci Rev**, 2018, 5(1): 44-53.
 - [15] JIANG T, GRADUS J L, ROSELLINI A J. Supervised machine learning: a brief primer[J]. **Behav Ther**, 2020, 51(5): 675-687.
 - [16] MIN H, YOON S. Got target? Computational methods for microRNA target prediction and their extension[J]. **Exp Mol Med**, 2010, 42(4): 233-244.
 - [17] SHAKYAWAR S, SOUTHEKAL S, GUDA C. MintRULS: prediction of miRNA-mRNA target site interactions using regularized least square method[J]. **Genes (Basel)**, 2022, 13(9): 1528.
 - [18] CHENG S, GUO M, WANG C, et al. MiRTDL: a deep learning approach for miRNA target prediction[J]. **IEEE/ACM Trans Comput Biol Bioinform**, 2016, 13(6): 1161-1169.
 - [19] MIN S, LEE B, YOON S. Targetnet: functional microRNA target prediction with deep neural networks[J]. **Bioinformatics**, 2022, 38(3): 671-677.
 - [20] PLA A, ZHONG X, RAYNER S. MiRAW: a deep learning-based approach to predict microRNA targets by analyzing whole microRNA transcripts[J/OL]. **PLoS Comput Biol**, 2018, 14(7): e1006185.
 - [21] YANG S, WANG Y, LIN Y, et al. LncMirNet: predicting lncRNA-miRNA interaction based on deep learning of ribonucleic acid sequences[J]. **Molecules**, 2020, 25(19): 4372.
 - [22] WANG X F, YU C Q, LI L P, et al. KGDCMI: a new approach for predicting circRNA-miRNA interactions from multi-source information extraction and deep learning[J]. **Front Genet**, 2022, 13: 958096.
 - [23] WEN M, CONG P, ZHANG Z, et al. DeepMirTar: a deep-learning approach for predicting human miRNA targets[J]. **Bioinformatics**, 2018, 34(22): 3781-3787.
 - [24] YANG L, LI T. LncRNA TUG1 regulates ApoM to promote atherosclerosis progression through mir-92a/FXR1 axis[J]. **J Cell Mol Med**, 2020, 24(15): 8836-8848.
 - [25] LIANG Y, CEN J, HUANG Y, et al. CircNTNG1 inhibits renal cell carcinoma progression via HOXA5-mediated epigenetic silencing of Slug[J]. **Mol Cancer**, 2022, 21(1): 224.
 - [26] ZHANG H, WANG Y, PAN Z, et al. ncRNAInter: a novel strategy based on graph neural network to discover interactions between lncRNA and miRNA[J]. **Brief Bioinform**, 2022, 23(6): bbac411.
 - [27] AGARWAL V, BELL G W, NAM J W, et al. Predicting effective microRNA target sites in mammalian mRNAs [J/OL]. **Elife**, 2015, 4: e05005.
 - [28] JOHN B, ENRIGHT A J, ARAVIN A, et al. Human microRNA targets[J/OL]. **PLoS Biol**, 2004, 2(11): e363.
 - [29] WU C C, HSIEH K T, YEH S Y, et al. Simultaneous detection of miRNA and mRNA at the single-cell level in plant tissues[J]. **Plant Biotechnol J**, 2023, 21(1): 136-149.
 - [30] ALBERTI C, MANZENREITHER R A, SOWEMIMO I, et al. Cell-type specific sequencing of microRNAs from complex animal tissues[J]. **Nat Methods**, 2018, 15(4): 283-289.
 - [31] NIELSEN M M, PEDERSEN J S. MiRNA activity inferred from single cell mRNA expression[J]. **Sci Rep**, 2021, 11(1): 9170.
 - [32] GARCIA-MARTIN R, WANG G, BRANDÃO B B, et al. MicroRNA sequence codes for small extracellular vesicle release and cellular retention[J]. **Nature**, 2022, 601(7893): 446-451.
 - [33] ROBERTS T C. The microRNA biology of the mammalian nucleus[J/OL]. **Mol Ther Nucleic Acids**, 2014, 3(8): e188.
 - [34] YE Y, FANG Y, XU W, et al. 3, 3'-Diindolylmethane induces anti-human gastric cancer cells by the miR-30e-ATG5 modulating autophagy[J]. **Biochem Pharmacol**, 2016, 115: 77-84.
 - [35] WEI H, XU Y, CHEN Q, et al. Mesenchymal stem cell-derived exosomal miR-223 regulates neuronal cell apoptosis[J]. **Cell Death Dis**, 2020, 11(4): 290.
 - [36] WANG Q, DU X, ZHOU B, et al. Mitochondrial dysfunction is responsible for fatty acid synthase inhibition-induced apoptosis in breast cancer cells by Pdpmn[J]. **Biomed Pharmacother**, 2017, 96: 396-403.
 - [37] DU X, ZHANG P, FU H, et al. Smart mitochondrial-targeted cancer therapy: subcellular distribution, selective TrxR2 inhibition accompany with declined antioxidant capacity[J]. **Int J Pharm**, 2019, 555: 346-355.
 - [38] CHEN D, ZHANG X Y, SUN J, et al. Asiatic acid protects dopaminergic neurons from neuroinflammation by suppressing mitochondrial ros production[J]. **Biomol Ther (Seoul)**, 2019, 27(5): 442-449.
 - [39] TASSONE P, DI MARTINO M T, ARBITRIO M, et al. Safety and activity of the first-in-class locked nucleic acid (LNA) miR-221 selective inhibitor in refractory advanced cancer patients: a first-in-human, phase 1, open-label, dose-escalation study[J]. **J Hematol Oncol**, 2023, 16(1): 68.
 - [40] VERMEIRE S, SANDS B E, TILG H, et al. ABX464 (obefazimod) for moderate-to-severe, active ulcerative colitis: a phase 2b, double-blind, randomised, placebo-controlled induction trial and 48 week, open-label extension[J]. **Lancet Gastroenterol Hepatol**, 2022, 7(11): 1024-1035.
 - [41] NAKAOKA H, HIRONO K, YAMAMOTO S, et al. MicroRNA-145-5p and microRNA-320a encapsulated in endothelial microparticles contribute to the pro-

- gression of vasculitis in acute Kawasaki disease[J]. **Sci Rep**, 2018, 8(1): 1016.
- [42] LI T, ZHOU X, WANG J, et al. Adipose-derived mesenchymal stem cells and extracellular vesicles confer antitumor activity in preclinical treatment of breast cancer[J]. **Pharmacol Res**, 2020, 157: 104843.
- [43] CHU M, WANG H, BIAN L, et al. Nebulization therapy with umbilical cord mesenchymal stem cell-derived exosomes for COVID-19 pneumonia[J]. **Stem Cell Rev Rep**, 2022, 18(6): 2152-2163.
- [44] JOHNSON J, LAW S, SHOJAEE M, et al. First-in-human clinical trial of allogeneic, platelet-derived extracellular vesicles as a potential therapeutic for delayed wound healing[J/OL]. **J Extracell Vesicles**, 2023, 12(7): e12332.
- [45] MEHER P K, SATPATHY S, RAO A R. MiRNALoc: predicting miRNA subcellular localizations based on principal component scores of physico-chemical properties and pseudo compositions of di-nucleotides[J]. **Sci Rep**, 2020, 10(1): 14557.
- [46] ASIM M N, MALIK M I, ZEHE C, et al. MirLoc-Predictor: a ConvNet-based multi-label microRNA subcellular localization predictor by incorporating k-Mer positional information[J]. **Genes (Basel)**, 2020, 11(12): 1475.
- [47] JOHN A, KUBOSUMI A, REDDY P H. Mitochondrial microRNAs in aging and neurodegenerative diseases [J]. **Cells**, 2020, 9(6): 1345.
- [48] XU M, CHEN Y, XU Z, et al. MiRLoc: predicting miRNA subcellular localization by incorporating miRNA-mRNA interactions and mRNA subcellular localization [J]. **Brief Bioinform**, 2022, 23(2): bbac044.
- [49] BAI T, YAN K, LIU B. DAmiRLocGNet: miRNA subcellular localization prediction by combining miRNA-disease associations and graph convolutional networks [J]. **Brief Bioinform**, 2023, 24(4): bbad212.
- [50] LONG J, HE Q, YIN Y, et al. The effect of miRNA and autophagy on colorectal cancer[J/OL]. **Cell Prolif**, 2020, 53(10): e12900.
- [51] ZHANG T, TAN P, WANG L, et al. RNALocate: a resource for RNA subcellular localizations[J]. **Nucleic Acids Res**, 2017, 45(D1): D135-D138.
- [52] CUI T, DOU Y, TAN P, et al. RNALocate v2.0: an updated resource for RNA subcellular localization with increased coverage and annotation[J]. **Nucleic Acids Res**, 2022, 50(D1): D333-D339.
- [53] TURUNEN T A, ROBERTS T C, LAITINEN P, et al. Changes in nuclear and cytoplasmic microRNA distribution in response to hypoxic stress[J]. **Sci Rep**, 2019, 9(1): 10332.
- [54] O'GRADY T, NJOCK M S, LION M, et al. Sorting and packaging of RNA into extracellular vesicles shape intracellular transcript levels[J]. **BMC Biol**, 2022, 20(1): 72.
- [55] GUO Q Q, GAO J, WANG X W, et al. RNA-binding protein MSI2 binds to miR-301a-3p and facilitates its distribution in mitochondria of endothelial cells[J]. **Front Mol Biosci**, 2020, 7: 609828.
- [56] TIAN L, CHEN F, MACOSKO E Z. The expanding vistas of spatial transcriptomics[J]. **Nat Biotechnol**, 2023, 41(6): 773-782.
- [57] LOVATT D, RUBLE B K, LEE J, et al. Transcriptome *in vivo* analysis (TIVA) of spatially defined single cells in live tissue[J]. **Nat Methods**, 2014, 11(2): 190-196.
- [58] WEI W, DAI W, YANG F, et al. Spatially resolved, error-robust multiplexed microRNA profiling in single living cells[J/OL]. **Angew Chem Int Ed Engl**, 2022, 61(20): e202116909.
- [59] WU J, LI G, YAO Y, et al. MicroRNA-421 is a new potential diagnosis biomarker with higher sensitivity and specificity than carcinoembryonic antigen and cancer antigen 125 in gastric cancer[J]. **Biomarkers**, 2015, 20(1): 58-63.
- [60] AFTAB M, POOJARY S S, SESHAN V, et al. Urine miRNA signature as a potential non-invasive diagnostic and prognostic biomarker in cervical cancer[J]. **Sci Rep**, 2021, 11(1): 10323.
- [61] ZHU L R, YUAN R X, XIA X B, et al. Assessment of a panel of miRNAs in serum and pleural fluid for the differential diagnosis of malignant and benign pleural effusion[J]. **Cancer Biomark**, 2022, 33(1): 71-82.
- [62] ZHU Y, WANG S, XI X, et al. Integrative analysis of long extracellular RNAs reveals a detection panel of noncoding RNAs for liver cancer[J]. **Theranostics**, 2021, 11(1): 181-193.
- [63] MOISOIU T, DRAGOMIR M P, IANCU S D, et al. Combined miRNA and SERS urine liquid biopsy for the point-of-care diagnosis and molecular stratification of bladder cancer[J]. **Mol Med**, 2022, 28(1): 39.
- [64] IRLAM-JONES J J, EUSTACE A, DENLEY H, et al. Expression of miR-210 in relation to other measures of hypoxia and prediction of benefit from hypoxia modification in patients with bladder cancer[J]. **Br J Cancer**, 2016, 115(5): 571-578.
- [65] SATHIPATI S Y, HO S Y. Identification of the miRNA signature associated with survival in patients with ovarian cancer[J]. **Aging (Albany NY)**, 2021, 13(9): 12660-12690.
- [66] ZHENG H, ZHANG L, ZHAO Y, et al. Plasma miRNAs as diagnostic and prognostic biomarkers for ovarian cancer[J/OL]. **PLoS One**, 2013, 8(11): e77853.
- [67] MOROKOFF A, JONES J, NGUYEN H, et al. Serum microRNA is a biomarker for post-operative monitoring in glioma[J]. **J Neurooncol**, 2020, 149(3): 391-400.
- [68] WONG V C, WONG M I, LAM C T, et al. Hallmark microRNA signature in liquid biopsy identifies hepatocellular carcinoma and differentiates it from liver metastasis[J]. **J Cancer**, 2021, 12(15): 4585-4594.
- [69] KIM S, BAE W J, AHN J M, et al. MicroRNA signatures associated with lymph node metastasis in intramucosal

- gastric cancer[J]. **Mod Pathol**, 2021, 34(3): 672-683.
- [70] HUANG Y, ZHU J, LI W, et al. Serum microRNA panel excavated by machine learning as a potential biomarker for the detection of gastric cancer[J]. **Oncol Rep**, 2018, 39(3): 1338-1346.
- [71] SAVAREH B A, AGHDAIE H A, BEHMANESH A, et al. A machine learning approach identified a diagnostic model for pancreatic cancer through using circulating microRNA signatures[J]. **Pancreatology**, 2020, 20(6): 1195-1204.
- [72] LIU C, YU Z, HUANG S, et al. Combined identification of three miRNAs in serum as effective diagnostic biomarkers for HNSCC[J]. **EBioMedicine**, 2019, 50: 135-143.
- [73] GAO C, LU W, LOU W, et al. Long noncoding RNA HOXC13-AS positively affects cell proliferation and invasion in nasopharyngeal carcinoma via modulating miR-383-3p/HMGA2 axis[J]. **J Cell Physiol**, 2019, 234(8): 12809-12820.
- [74] YU W, HE X, ZHANG C, et al. Circular RNA circSLC7A11 contributes to progression and stemness of laryngeal squamous cell carcinoma via sponging miR-877-5p from LASP1[J/OL]. **Heliyon**, 2023, 9(7): e18290.
- [75] YE Z, SUN B, XIAO Z. Machine learning identifies 10 feature miRNAs for lung squamous cell carcinoma [J]. **Gene**, 2020, 749: 144669.
- [76] LIANG Z Y, ZHANG Z M, SUN G R, et al. LncRNA ASBEL and lncRNA Erbb4-IR reduce chemoresistance against gemcitabine and cisplatin in stage IV lung squamous cell carcinoma via the microRNA-21/LZTFL1 axis[J]. **Am J Cancer Res**, 2023, 13(6): 2732-2750.
- [77] SHI J, LV X, ZENG L, et al. CircPVT1 promotes proliferation of lung squamous cell carcinoma by binding to miR-30d/e[J]. **J Exp Clin Cancer Res**, 2021, 40(1): 193.
- [78] ZHU Z, LI Q, XU M, et al. Effect of whole-brain and intensity-modulated radiotherapy on serum levels of miR-21 and prognosis for lung cancer metastatic to the brain[J/OL]. **Med Sci Monit**, 2020, 26: e924640.
- [79] ZHANG Y H, JIN M, LI J, et al. Identifying circulating miRNA biomarkers for early diagnosis and monitoring of lung cancer[J]. **Biochim Biophys Acta Mol Basis Dis**, 2020, 1866(10): 165847.
- [80] ZHANG X, WANG X, CHAI B, et al. Downregulated miR-18a and miR-92a synergistically suppress non-small cell lung cancer via targeting Sprouty 4[J]. **Bio-engineered**, 2022, 13(4): 11281-11295.
- [81] LARRUE R, FELLAH S, BOUKROUT N, et al. MiR-92a-3p regulates cisplatin-induced cancer cell death [J]. **Cell Death Dis**, 2023, 14(9): 603.
- [82] LI X, LV F, LI F, et al. Long non-coding RNA H19 facilitates small cell lung cancer tumorigenesis through miR-140-5p/FGF9 axis[J]. **Onco Targets Ther**, 2020, 13: 3525-3534.
- [83] FU J, CAI H, WU Y, et al. Elevation of FGD5-AS1 contributes to cell progression by improving cisplatin resistance against non-small cell lung cancer cells through regulating miR-140-5p/WEE1 axis[J]. **Gene**, 2020, 755: 144886.
- [84] TIAN Q Q, XIA J, ZHANG X, et al. MiR-331-3p inhibits tumor cell proliferation, metastasis, invasion by targeting MLLT10 in non-small cell lung cancer[J]. **Cancer Manag Res**, 2020, 12: 5749-5758.
- [85] ZHU S, KONG X, SONG M, et al. MiR-223-3p attenuates the migration and invasion of NSCLC cells by regulating NLRP3[J]. **Front Oncol**, 2022, 12: 985962.
- [86] GUO Q, WANG H, XU Y, et al. MiR-374a-5p inhibits non-small cell lung cancer cell proliferation and migration via targeting NCK1[J]. **Exp Ther Med**, 2021, 22(3): 943.
- [87] PAVEL A B, CAMPBELL J D, LIU G, et al. Alterations in bronchial airway miRNA expression for lung cancer detection[J]. **Cancer Prev Res (Phila)**, 2017, 10(11): 651-659.
- [88] ZHAO S, YU J, WANG L. Machine learning based prediction of brain metastasis of patients with IIIa-n2 lung adenocarcinoma by a three-mirna signature[J]. **Transl Oncol**, 2018, 11(1): 157-167.
- [89] KACZMAREK E, PYMAN B, NANAYAKKARA J, et al. Discriminating neoplastic from nonneoplastic tissues using an miRNA-based deep cancer classifier [J]. **Am J Pathol**, 2022, 192(2): 344-352.
- [90] OLGUN G, GOPALAN V, HANNENHALLI S. MiRSCAPE-inferring miRNA expression from scRNA-seq data[J]. **iScience**, 2022, 25(9): 104962.
- [91] KUTLAY A, AYDIN SON Y. Integrative predictive modeling of metastasis in melanoma cancer based on MicroRNA, mRNA, and DNA methylation data[J]. **Front Mol Biosci**, 2021, 8: 637355.
- [92] CHEN B, DUAN L, YIN G, et al. MiR-381, a novel intrinsic WEE1 inhibitor, sensitizes renal cancer cells to 5-FU by up-regulation of Cdc2 activities in 786-O [J]. **J Chemother**, 2013, 25(4): 229-238.
- [93] YU Y Z, MU Q, REN Q, et al. MiR-381-3p suppresses breast cancer progression by inhibition of epithelial-mesenchymal transition[J]. **World J Surg Oncol**, 2021, 19(1): 230.
- [94] LI F, HU Q, PANG Z, et al. LncRNA MAGI2-AS3 upregulates cytokine signaling 1 by sponging miR-155 in non-small cell lung cancer[J]. **Cancer Biother Radiopharm**, 2020, 35(1): 72-76.
- [95] YANG L W, WU X J, LIANG Y, et al. MiR-155 increases stemness and decitabine resistance in triple-negative breast cancer cells by inhibiting TSPAN5[J]. **Mol Carcinog**, 2020, 59(4): 447-461.
- [96] SHI L, WAN Y, SUN G, et al. MiR-125b inhibitor may enhance the invasion-prevention activity of temozolomide in glioblastoma stem cells by targeting PIAS3[J].

- BioDrugs**, 2014, 28(1): 41-54.
- [97] ZHANG X, LI T, HAN Y N, et al. MiR-125b promotes colorectal cancer migration and invasion by dual-targeting CFTR and CGN[J]. **Cancers (Basel)**, 2021, 13(22): 5710.
- [98] MA J, WU D, YI J, et al. MiR-378 promoted cell proliferation and inhibited apoptosis by enhanced stem cell properties in chronic myeloid leukemia K562 cells[J]. **Biomed Pharmacother**, 2019, 112: 108623.
- [99] LIU L, HAN S, XIAO X, et al. Glucocorticoid-induced microRNA-378 signaling mediates the progression of pancreatic cancer by enhancing autophagy[J]. **Cell Death Dis**, 2022, 13(12): 1052.
- [100] LU M, ZHANG Q, DENG M, et al. An analysis of human microrna and disease associations[J/OL]. **PLoS One**, 2008, 3(10): e3420.
- [101] BARABÁSI A L, GULBAHCE N, LOSCALZO J. Network medicine: a network-based approach to human disease[J]. **Nat Rev Genet**, 2011, 12(1): 56-68.
- [102] UTHAYOPAS K, DE SÁ A, ALAVI A, et al. TSMDA: target and symptom-based computational model for miRNA-disease-association prediction[J]. **Mol Ther Nucleic Acids**, 2021, 26: 536-546.
- [103] DONG N, MUCKE S, KHOSLA M. MuCoMiD: a multitask graph convolutional learning framework for miRNA-disease association prediction[J]. **IEEE/ACM Trans Comput Biol Bioinform**, 2022, 19(6): 3081-3092.
- [104] MA Z, KUANG Z, DENG L. NGCIM: a novel deep learning-based method for predicting circRNA-miRNA interactions[J]. **IEEE/ACM Trans Comput Biol Bioinform**, 2023, 20(5): 3080-3092.
- [105] XU W P, LIU J P, FENG J F, et al. MiR-541 potentiates the response of human hepatocellular carcinoma to sorafenib treatment by inhibiting autophagy[J]. **Gut**, 2020, 69(7): 1309-1321.
- [106] AHANGARI F, PRICE N L, MALIK S, et al. MicroRNA-33 deficiency in macrophages enhances autophagy, improves mitochondrial homeostasis, and protects against lung fibrosis[J/OL]. **JCI Insight**, 2023, 8(4): e158100.
- [107] YANG B, BAI Y, YIN C, et al. Activation of autophagic flux and the Nrf2/ARE signaling pathway by hydrogen sulfide protects against acrylonitrile-induced neurotoxicity in primary rat astrocytes[J]. **Arch Toxicol**, 2018, 92(6): 2093-2108.
- [108] DUAN Y, XU H, LUO X, et al. Procyanidins from *nelumbo nucifera* gaertn. Seedpod induce autophagy mediated by reactive oxygen species generation in human hepatoma G2 cells[J]. **Biomed Pharmacother**, 2016, 79: 135-152.
- [109] YANG B, YIN C, ZHOU Y, et al. Curcumin protects against methylmercury-induced cytotoxicity in primary rat astrocytes by activating the Nrf2/ARE pathway independently of PKCδ[J]. **Toxicology**, 2019, 425: 152248.
- [110] LIANG Z, WU R, XIE W, et al. Effects of curcumin on tobacco smoke-induced hepatic MAPK pathway activation and epithelial-mesenchymal transition *in vivo*[J]. **Phytother Res**, 2017, 31(8): 1230-1239.
- [111] GENG J, LIU W, XIONG Y, et al. Andrographolide sulfonate improves Alzheimer-associated phenotypes and mitochondrial dysfunction in APP/PS1 transgenic mice[J]. **Biomed Pharmacother**, 2018, 97: 1032-1039.
- [112] GENG J, LIU W, GAO J, et al. Andrographolide alleviates Parkinsonism in MPTP-PD mice via targeting mitochondrial fission mediated by dynamin-related protein 1[J]. **Br J Pharmacol**, 2019, 176(23): 4574-4591.
- [113] LI B, SHI C, LI B, et al. The effects of curcumin on HCT-116 cells proliferation and apoptosis via the miR-491/PEG10 pathway[J]. **J Cell Biochem**, 2018, 119(4): 3091-3098.
- [114] LIU C, ROKAVEC M, HUANG Z, et al. Curcumin activates a ROS/KEAP1/NRF2/miR-34a/b/c cascade to suppress colorectal cancer metastasis[J]. **Cell Death Differ**, 2023, 30(7): 1771-1785.
- [115] GE Q, CHEN L, TANG M, et al. Analysis of mulberry leaf components in the treatment of diabetes using network pharmacology[J]. **Eur J Pharmacol**, 2018, 833: 50-62.
- [116] WANG N, ZHU F, SHEN M, et al. Network pharmacology-based analysis on bioactive anti-diabetic compounds in *Potentilla discolor bunge*[J]. **J Ethnopharmacol**, 2019, 241: 111905.
- [117] GE Q, CHEN L, YUAN Y, et al. Network pharmacology-based dissection of the anti-diabetic mechanism of *lobelia chinensis*[J]. **Front Pharmacol**, 2020, 11: 347.
- [118] NIU Z, GAO X, XIA Z, et al. Prediction of small molecule drug-miRNA associations based on GNNs and CNNs[J]. **Front Genet**, 2023, 14: 1201934.

[本文编辑 余 方 沈 敏]