

文章编号:1671-4229(2024)03-0008-07

# 基于 miRNA-gene 调控网络的 miRNA 重要性评估

顾丽丽<sup>1</sup>, 郑泽昆<sup>2a</sup>, 方紫婵<sup>2a</sup>, 许鹏<sup>2b\*</sup>

(1. 广州康大职业技术学院 信息工程系, 广东 广州 511363; 2. 广州大学 a. 计算机科学与网络工程学院,  
b. 计算科技研究院, 广东 广州 510006)

**摘要:** MicroRNAs (miRNAs) 是一类非编码短序列 RNAs 分子,它能够与靶基因特异性结合,进而影响靶基因对应的蛋白质产量。单个 miRNA 可能调节多个基因的表达,同时,单个基因的表达也可能被多个 miRNAs 所调控。已有研究表明,miRNA 可以作为潜在的肿瘤标志物用于癌症的早期诊断和治疗。评价不同 miRNAs 的重要性对于探究 miRNA 在各种疾病中的作用具有重要的意义。因此,文章基于 miRNA 与基因之间的调控网络,提出了一种 miRNA 重要性评价指标。通过对该指标对 miRNAs 进行打分排序,排名靠前的 miRNAs 与疾病相关可能性越大,GO 富集分析结果表明,排名靠前的 miRNAs 具有重要的生物学功能。基于排名靠前的 100 个 miRNAs 表达特征,训练 SVM 分类模型对各种癌症进行分类,实验证明其平均精度达到 90% 以上。总之,文章提出的 miRNA 重要性指标对于疾病相关 miRNAs 的识别以及应用具有重要的指导意义。

**关键词:** miRNA; 基因调控网络; 基因; 重要性评估

**中图分类号:** TP333;Q811.4 **文献标志码:** A

## Evaluation of the importance of miRNAs based on miRNA-gene regulatory networks

GU Li-li<sup>1</sup>, ZHENG Ze-kun<sup>2a</sup>, FANG Zi-chan<sup>2a</sup>, XU Peng<sup>2b\*</sup>

(1. Department of Information Engineering, Guangzhou Kangda Vocational Technical College, Guangzhou 511363, China;  
2. a. School of Computer Science and Cyber Engineering, b. Institute of Computational Science and Technology, Guangzhou University, Guangzhou 510006, China)

**Abstract:** MicroRNAs (miRNAs) are a class of non-coding short sequence RNA molecules that can specifically bind to target genes, thereby influencing the production of proteins corresponding to the target genes. A single miRNA may regulate the expression of multiple genes, and conversely, the expression of a single gene may be regulated by multiple miRNAs. Existing research has shown that miRNAs can serve as potential tumor markers for early diagnosis and treatment of cancer. Evaluating the importance of different miRNAs is crucial for exploring the role of miRNAs in various diseases. Therefore, based on the regulatory network between miRNAs and genes, this study proposes an evaluation index for miRNA importance. By scoring and ranking miRNAs based on this index, miRNAs ranked higher are more likely to be associated with diseases. GO enrichment analysis results indicate that these top-ranked miRNAs have significant biological functions. Based on the expression characteristics of the top 100 ranked miRNAs, an SVM classification model is trained to classify various

收稿日期: 2024-01-31; 修回日期: 2024-02-20

基金项目: 国家自然科学基金资助项目(62002079); 广东省自然科学基金资助项目(2023A1515011401); 广州市科技局市校联合基金资助项目(SL2022A03J00935)

作者简介: 顾丽丽(1985—),女,博士. E-mail:gulili\_kd@163.com

\*通信作者. E-mail:gdxupeng@gzhu.edu.cn

引文格式: 顾丽丽, 郑泽昆, 方紫婵, 等. 基于 miRNA-gene 调控网络的 miRNA 重要性评估[J]. 广州大学学报(自然科学版), 2024, 23(3): 8-14.

cancers, with an average accuracy of over 90%. The model also performs well in cancer classification problems. In conclusion, the proposed miRNA importance index in this study has important guiding significance for the identification of disease-related miRNAs and the application of miRNAs.

**Key words:** miRNA; gene regulatory network; gene; importance assessment

MiRNAs 是一类长度约为 22 个核苷酸的非编码单链 RNA 分子,它们通过碱基互补配对原则与靶标 mRNA 的 3' 端非编码区域特异性结合,从而调节转录后基因表达<sup>[1-3]</sup>。一般认为,miRNA 通过与靶 mRNA 的特异性结合对靶基因的表达起到抑制的效果<sup>[4]</sup>。近年来,陆续有研究报道表明,miRNA 也能够促进靶基因的表达<sup>[5-6]</sup>。通过文献检索,课题组前期基于实验验证得到的正调控 miRNA-gene 关系对整理汇总并开发了一个数据库 PmiRtarbase<sup>[7]</sup>。自从 1993 年,miRNA 被首次报道以来,大量的证据表明 miRNA 参与包括细胞衰老、凋亡、分化及增殖等多种生物学过程<sup>[8-11]</sup>。此外,越来越多的研究发现,miRNA 在包括癌症、心血管疾病、神经系统疾病等疾病在内的发生和发展中发挥着重要的调控作用<sup>[12-14]</sup>。miRNAs 可以作为癌症抑制因子和癌基因发挥作用,因此,它们被称为“肿瘤因子”<sup>[15]</sup>。

从大量 miRNAs 中挖掘出关键和重要的 miRNAs,对于 miRNAs 的应用具有重要的意义,如可以用来辅助识别疾病相关的 miRNAs<sup>[16-18]</sup>。在编码基因领域,已有大量的方法被提出用于计算基因/蛋白质的重要性。Li 等<sup>[19]</sup>提出了基于蛋白质网络预测关键蛋白质的方法。Guo 等<sup>[20]</sup>提出了仅需要基因核酸序列的方法,预测人类关键的基因。Peng 等<sup>[21]</sup>基于蛋白质互作网络,提出分子迭代演化的方法预测关键蛋白。以上这些方法对于识别关键基因/蛋白具有重要的贡献作用,遗憾的是这些方法很难涉及聚焦到关键 miRNA 预测问题上,因为 miRNA 之间的互作网络并不存在。2019 年,Cui 等<sup>[22]</sup>定义了疾病谱宽度概念,其假设如果某个 miRNA 相关联的疾病数量越多,那么这个 miRNA 的疾病谱宽度值越大,即重要性越大。

鉴于单个 miRNA 已被证明可以调节多个基因的表达,同时,单个基因的表达也可以被多个 miRNA 调控,基因与基因也存在着复杂的调控关系。本课题以此为切入点,通过对 miRNA-gene 调控网络进行分析,提出了一种 miRNA 重要性的评价标准。

## 1 数据和方法

### 1.1 实验数据集

本文实验数据集包括:miRNA-gene 基因调控网络和癌症 miRNA 表达谱数据。

#### (1) miRNA-gene 调控网络

本文在搭建 miRNA-gene 的调控网络时使用了 miRTarBase 数据库<sup>[23]</sup>。miRTarBase 是一个收集实验验证的 miRNA 靶标数据库,该数据库最新数据显示已经积累了 447 万多条关于人类 miRNA-gene 相互作用关系,其中包含 3 012 个 miRNAs、17 387 个 genes。

#### (2) 基因调控网络

京都基因与基因组百科全书(Kyoto Encyclopedia of Genes and Genomes, KEGG)数据库,长期以来一直是研究生物信号通路的权威数据库<sup>[24]</sup>。该数据库包含人类在内的多种物种的基因调控互作信息。本研究通过 R 工具包“KEGGgraph”<sup>[25]</sup>,获取到 KEGG 信号通路中的基因之间作用关系,作用关系主要包含两类:激活和抑制,该网络实质上是一个有向图,节点是基因,边表示基因之间的作用关系。

对于复杂的基因调控网络,本研究将其抽象为类似于数据结构中的图结构。广度优先遍历(BFS)是一种遍历图存储结构的算法,其核心思想是从一个起始节点开始,访问其所有邻近节点,然后再按照相同的方式访问这些邻近节点的邻近节点。在计算时它首先考虑距离,以便将最接近的节点放入递归队列中,它可以被用于图的搜索,通过遍历所有节点来找到一个目标节点。这种算法的优点在于能够快速找到离起始点最近的节点,在寻找最短路径等问题中非常有用。本文主要采用广度优先遍历策略寻找 miRNA 下游控制基因。

#### (3) 癌症 miRNA 表达谱数据

为了验证关键 miRNAs 的癌症分类性能,本文从 TCGA 数据库网站(<https://tcga-data.nci.nih.gov/tcga>)下载了 miRNA 表达谱数据,并选取样本量大于 200 的癌症数据作为后续分析,删除癌症

数量较少数据集之后,最终共有 19 种癌症被选中。此外,将 19 种癌症的癌旁正常组织样本合并

起来,统称为正常 Normal 类。每种癌症的样本数量信息如表 1 所示。

表 1 不同疾病样本的数量

Table 1 The number of samples of different diseases

序号	癌症名称	癌症样本	癌旁正常样本
1	SKCM (Skin Cutaneous Melanoma)	478	2
2	THCA (Thyroid carcinoma)	515	65
3	SARC (Sarcoma)	265	6
4	PRAD (Prostate adenocarcinoma)	499	67
5	HNSC (Head and Neck squamous cell carcinoma)	530	74
6	COAD (Colon adenocarcinoma)	463	85
7	BRCA (Breast invasive carcinoma)	1 241	23
8	BLCA (Bladder Urothelial Carcinoma)	436	60
9	KIRP (Kidney renal papillary cell carcinoma)	352	407
10	KIRC (Kidney renal clear cell carcinoma)	944	0
11	LAML (Acute Myeloid Leukemia)	200	12
12	OV (Ovarian serous cystadenocarcinoma)	615	120
13	LUAD (Lung adenocarcinoma)	585	59
14	LIHC (Liver hepatocellular carcinoma)	379	47
15	UCEC (Uterine Corpus Endometrial Carcinoma)	548	6
16	GBM (Glioblastoma multiforme)	608	0
17	LGG (Brain Lower Grade Glioma)	530	103
18	STAD (Stomach adenocarcinoma)	478	120
19	LUSC (Lung squamous cell carcinoma)	506	23
20	Normal	0	1 410

## 1.2 miRNA 重要性指标计算

当前已有方法根据相关疾病数量来评价 miRNA 的重要性,本文主要是从系统生物学的角度出发对 miRNA 重要性进行计算。当某个 miRNA 的表达发生变化时,其下游基因的表达将会受到一定程度上的影响。本研究将 miRNA 对下游基因的调控影响用来评价 miRNA 的重要程度。当一个 miRNA 发生变化而影响的基因越多时,这种 miRNA 就越重要。但是,随着遍历层数的增加,miRNA 对基因的影响势必会逐渐减小,最远的一层基因受到 miRNA 的影响远不如 miRNA 直接调控的基因。经过多次观察后,提出一种评估 miRNA 重要性的方法:

步骤 1:统计目标 miRNA 与基因调控关系网络中每一阶(层)邻居基因的个数;

步骤 2:将各层的基因个数除以层数得到目标 miRNA 对该层基因影响的权重;

步骤 3:将 miRNA 对下游各层邻居基因的影响进行累加求和,得到目标 miRNA 的重要性, $n$  为

最深的遍历层数, $num(i)$  表示第  $i$  层的基因总数,则 miRNA 的重要性( $V$ )可表示为

$$V = \sum_{i=1}^n (num(i)/i), \quad (1)$$

得到所有 miRNA 的重要性后,利用公式(2),将 miRNA 的重要性标准化为一个 0 ~ 1 之间的 miRNA 重要性指数。

$$I = \frac{V - V_{\min}}{V_{\max} - V_{\min}}. \quad (2)$$

## 2 结果与讨论

### 2.1 疾病相关 miRNAs 的验证

通过本文提出的 miRNA 重要性指标,计算每一个 miRNA 的重要性,并按照该指标对 miRNA 进行排序。该指标表明了 miRNA 对下游基因的靶向调节能力,指标值越大意味着该 miRNA 越重要。如果某个 miRNA 重要,那么它应该会与很多疾病相关或在其中扮演着重要作用。基于此假设,通过人类疾病相关 miRNAs 数据库 HMDD<sup>[26]</sup>

验证本研究计算得到重要 miRNAs。

筛选出 miRNA 重要性最高的前 100 个,然后根据 HMDD 数据库统计与 miRNA 重要性最高的前 100 个 miRNA 关联的疾病数量。本次验证设置了 5 个梯度,分别为排名前 20、前 40、前 60、前 80 和前 100 的 miRNAs,结果如图 1 所示。

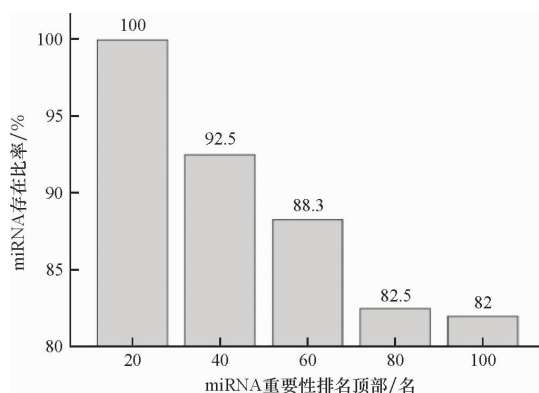


图 1 miRNA 在 HMDD 数据库中的存在情况

Fig. 1 Presence of miRNA in HMDD database

由图 1 可见,前 20 名的 miRNAs 在 HMDD 数据库中的存在率为 100%,这成功地验证了通过 miRNA 重要性指标筛选出的基因的确能对人类的

疾病产生影响。此外,随着排名的递增,与疾病相关的 miRNAs 比率在降低,这也从侧面说明本研究指标的准确性,排名越靠前越重要。尽管随着 miRNA 排名的增加,存在率有一定程度的减弱,但是前 100 名 miRNAs 中仍有 82% 的 miRNAs 被证实与一些疾病有关。

## 2.2 重要 miRNAs 的生物学功能分析

根据 miRNAs 的重要性得分排序,前 20 名的 miRNAs 如表 2 所示。通过检索,展示了这些 miRNAs 的重要生物学功能,如 hsa-miR-335-5p miRNA 在脂肪来源的干细胞中显著失调;hsa-miR-26b-5p miRNA 具有调控 EZH2 表达的重要因子,有望成为治疗干预的靶点。此外,选取了 miRNA 重要性排名前 20 的 miRNAs 所调控的基因进行 GO 富集分析,探究这些基因会富集在哪些生物学功能上。图 2 所示为排名前 20 miRNAs 靶基因富集的信号通路,可见,这些基因显著富集在一些重要的生物学信号通路上,如血管生成(angiogenesis)、细胞分裂(cell division)、机体衰老(aging)、免疫系统(immune system),这些功能与人类的健康以及癌症的产生息息相关。

表 2 排名前 20 的 miRNAs 功能描述

Table 2 Description of the top 20 miRNAs functions

排名	miRNA 名称	功能描述
1	hsa-miR-335-5p	一种在脂肪来源的干细胞(ASCs)中失调的微小 RNA
2	hsa-miR-26b-5p	调控 EZH2 表达的重要因子,有望成为治疗干预的靶点。
3	hsa-miR-16-5p	通过靶向调节 BCL2 的表达,影响细胞周期和凋亡
4	hsa-miR-124-3p	是 HPV DNA 检测中经常甲基化的位点
5	hsa-miR-92a-3p	miR17-92 簇已被发现在 medaka 和人类黑色素瘤模型中上调
6	hsa-miR-17-5p	具有调控 STAT3 基因表达,从而影响各种细胞过程
7	hsa-miR-93-5p	通过调控 PKMYT1 基因表达,从而影响细胞增值、迁移及侵袭
8	hsa-let-7b-5p	会导致脉络膜氧化应激和衰老
9	hsa-miR-106b-5p	调控 ESR1 表达影响肝细胞癌患者生存预后
10	hsa-miR-20a-5p	可作为癌基因或者抑癌基因参与肿瘤的发生发展过程
11	hsa-miR-155-5p	其在造血、炎性反应、免疫和细胞系分化中具有调控作用
12	hsa-miR-1-3p	与调控血管生成、细胞凋亡和内皮功能有关
13	hsa-miR-34a-5p	可靶向调控巨噬细胞活化功能
14	hsa-miR-20b-5p	可显著增加体外牵张损伤神经元神经突起数量及总长度
15	hsa-miR-519d-3p	能够促进 CRC 细胞的增殖、迁移和侵袭
16	hsa-miR-218-5p	可通过调节 HOXA5 促进非小细胞肺癌细胞增殖和侵袭
17	hsa-miR-15b-5p	可调控 TNFR2 的表达,进而调节代谢相关信号通路
18	hsa-miR-615-3p	通过抑制 IGF2 抑制肿瘤生长和转移
19	hsa-miR-192-5p	能诱导骨髓瘤细胞的凋亡
20	hsa-miR-335-5p	可促进白细胞介素 1 $\beta$ 作用的软骨细胞增殖和抑制细胞凋亡

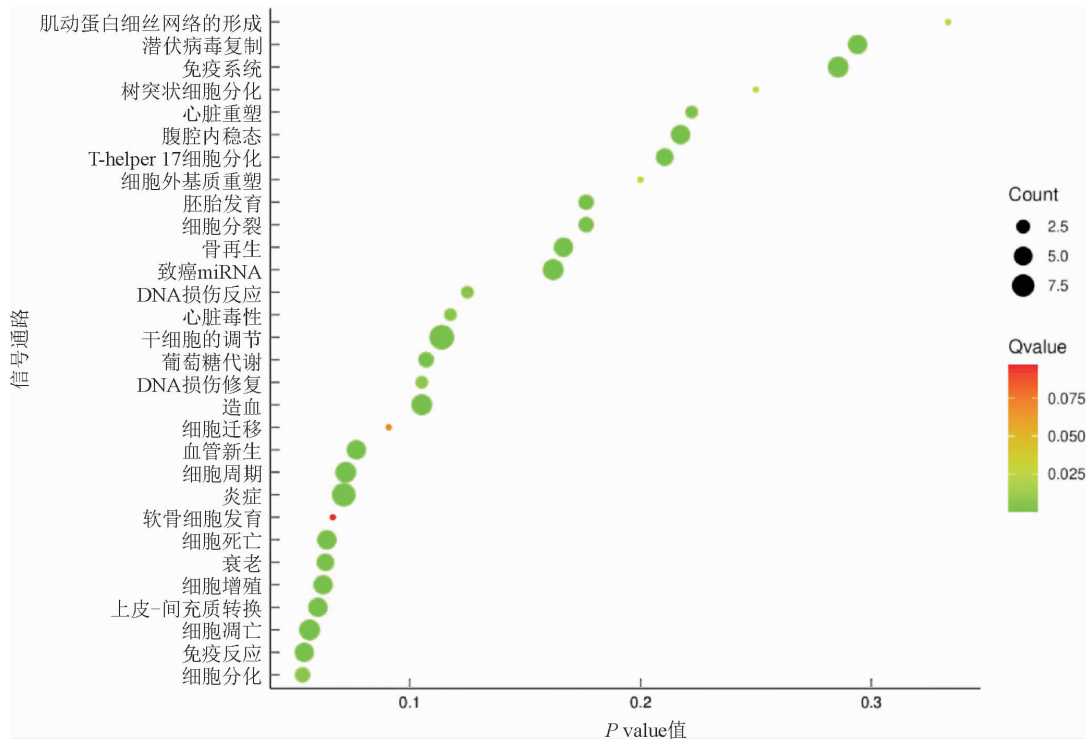


图 2 排名前 20 的 miRNAs 调控的基因富集分析结果

Fig. 2 The results of the enrichment analysis of the top 20 miRNAs regulated genes

2.3 基于聚类分析的验证

本研究运用了 19 种癌症以及正常人的 miRNA 表达谱样本, 总共 20 种类别。以筛选出的 miRNA 重要性前 100 的 miRNAs 作为参考指标, 利用 t-SNE 聚类图(t-SNE 聚类图能够代表 t 分布随机邻居嵌入, 使高维数据点投影到低维空间的

特点) 观察前 100 的 miRNA 是否能够区分出 19 种疾病样本以及正常人的样本。t-SNE 聚类图如图 3 所示, 可见, 通过本文方法识别的前 100miRNAs 可以非常明显地将 20 种数据样本分开, 说明通过 miRNA 重要性指标得到的重要 miRNAs 具有作为癌症分类标识物的潜力。

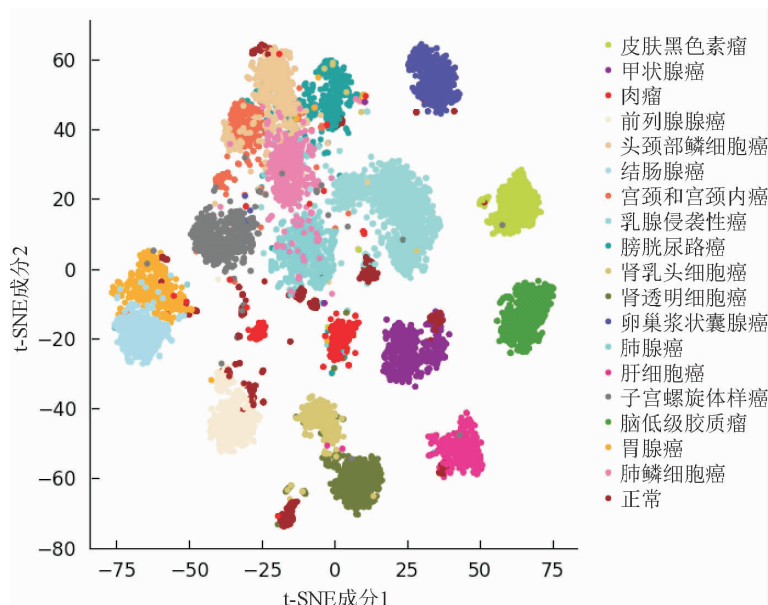


图 3 多种疾病的 t-SNE 聚类分析结果

Fig. 3 t-SNE cluster analysis of multiple diseases

## 2.4 基于 SVM 算法的验证

支持向量机(Support Vector Machine, SVM)是一种监督学习算法,可用于分类和回归分析,它可以处理高维数据,具有较高的准确性和泛化能力。以本文筛选得到的前 100 miRNAs 为特征,训练

SVM 模型,检验重要 miRNAs 在癌症样本上的分类能力。在 SVM 分类器训练过程中,学习率设置为  $5 * 10^{-6}$ ,正则化参数  $C = 5$ ,  $\gamma = 0.05$ 。见表 3。

表 3 各种样本在 SVM 算法下的表现

Table 3 The performance of various samples under SVM algorithm

疾病	精确率	召回率	F1 分数	疾病	精确率	召回率	F1 分数
SKCM	0.974 ± 0.015	0.989 ± 0.009	0.981 ± 0.008	KIRC	0.970 ± 0.015	0.931 ± 0.027	0.949 ± 0.016
THCA	0.976 ± 0.016	0.972 ± 0.014	0.974 ± 0.010	LAML	1 ± 0	1 ± 0	1 ± 0
SARC	0.886 ± 0.041	0.939 ± 0.032	0.911 ± 0.029	OV	0.989 ± 0.013	0.979 ± 0.013	0.984 ± 0.010
PRAD	0.930 ± 0.024	0.992 ± 0.009	0.960 ± 0.013	LUAD	0.912 ± 0.021	0.891 ± 0.027	0.901 ± 0.017
HNSC	0.736 ± 0.041	0.906 ± 0.030	0.811 ± 0.027	LIHC	0.897 ± 0.049	0.991 ± 0.011	0.941 ± 0.029
COAD	0.959 ± 0.021	0.959 ± 0.020	0.959 ± 0.013	UCEC	0.907 ± 0.030	0.899 ± 0.024	0.902 ± 0.019
CESC	0.811 ± 0.062	0.652 ± 0.057	0.720 ± 0.046	LGG	0.991 ± 0.008	0.997 ± 0.005	0.994 ± 0.004
BRCA	0.948 ± 0.015	0.966 ± 0.012	0.957 ± 0.009	STAD	0.717 ± 0.041	0.938 ± 0.023	0.812 ± 0.029
BLCA	0.829 ± 0.032	0.771 ± 0.046	0.798 ± 0.031	LUSC	0.830 ± 0.032	0.848 ± 0.033	0.838 ± 0.023
KIRP	0.946 ± 0.031	0.923 ± 0.027	0.934 ± 0.022	Normal	0.863 ± 0.035	0.771 ± 0.044	0.814 ± 0.031

本研究使用的是表 1 中所提及的 19 种疾病样本以及正常样本。使用 5 折交叉验证,通过将数据分成 5 等份,在 4 份上进行模型训练,剩下的一份进行测试,可以捕获模型在测试数据上的分类性能;同时,通过使用不同的折叠划分重复交叉验证过程 5 次,可以获得对模型性能更客观的估计,有助于避免数据中过拟合或随机性的影响。表 3 展示了排名前 100 个 miRNAs 在癌症样本上的分类性能,从结果上可见,通过观察样本这 100 个 miRNAs 的表达,对于大部分的癌症来说都能够拥有一个很高的分类效果。特别地,对于 LAML 样本,分类器在精确率、召回率和 F1-分数指标上均达到 100% 的分类效果。此外,在 OV 和 LGG 等疾病分类中,分类精度均能达到 98% 以上。以上这些结果表明,通过重要性排序得到的 miRNAs 对于疾病的分类具有重要的意义。倘若再结合其他诸如 miRNAs 的差异表达信息,相信定能得到更好的疾病分类特征。

## 3 总结

在 miRNA 与基因的调控关系网络中,miRNA 直接调控的基因个数极大程度地影响了 miRNA 的重要性,也就是说,miRNA 直接调控的基因个数足够多,则该 miRNA 的重要性会更高,在对人类疾病的预测和监控中会更为重要。随着调控层数的增加,miRNA 对基因的调控能力会越来越小。基于上述思想,本文提出了 miRNA 重要性的评价指标。通过一系列实验及分析发现:①基于重要性指标筛选到的重要 miRNAs,排名靠前的 miRNAs 更倾向于与疾病有关联;②通过生物功能富集分析,表明排名靠前的 miRNAs 参与多种重要的生物学功能模块;③通过 SVM 算法利用筛选出的 miRNAs 对 19 种癌症进行分类,得到了优秀的预测结果。本研究方法给重要 miRNAs 的筛选提供了一种思路,有望加速 miRNAs 的应用研究。

## 参考文献:

- [1] Ambros V. The functions of animal microRNAs[J]. Nature, 2004,431(7006):350-355.
- [2] Bartel D P. MicroRNAs: Genomics, biogenesis, mechanism, and function[J]. Cell, 2004, 116(2):281-297.
- [3] 许鹏, 咎乡镇, 路德杨, 等. 疾病相关 miRNAs 识别方法研究进展[J]. 广州大学学报(自然科学版), 2019, 18(1): 54-64.
- [4] Wu Z W, Lu H F, Sheng J F, et al. Inductive microRNA-21 impairs anti-mycobacterial responses by targeting IL-12 and

- Bcl-2[J]. *Febs Letters*, 2012, 586(16):2459-2467.
- [5] Tsai N P, Lin Y L, Wei L N. MicroRNA mir-346 targets the 5'-untranslated region of receptor-interacting protein 140 (RIP140) mRNA and up-regulates its protein expression[J]. *Biochemical Journal*, 2009, 424:411-418.
- [6] Vasudevan S, Steitz J A. AU-rich-element-mediated upregulation of translation by FXR1 and Argonaute 2[J]. *Cell*, 2007, 128(6):1105-1118.
- [7] Xu P, Li X, Liang Y, et al. PmiRtarbase: A positive miRNA-target regulations database[J]. *Computational Biology and Chemistry*, 2022, 98:107690.
- [8] Bartel D P. MicroRNAs: Target recognition and regulatory functions[J]. *Cell*, 2009, 136(2):215-233.
- [9] Xu P, Wu Q, Lu D, et al. A systematic study of critical miRNAs on cells proliferation and apoptosis by the shortest path[J]. *BMC Bioinformatics*, 2020, 21(1):396-409.
- [10] Xu P, Wu Q, Rao Y, et al. Predicting the influence of microRNAs on drug therapeutic effects by random walking[J]. *IEEE Access*, 2020, 8:117347-117353.
- [11] Xu P, Zhu Y, Sun B, et al. Colorectal cancer characterization and therapeutic target prediction based on microRNA expression profile[J]. *Scientific Reports*, 2016, 6:20616-20626.
- [12] Yu Z, Li Z, Jolicoeur N, et al. Aberrant allele frequencies of the SNPs located in microRNA target sites are potentially associated with human cancers[J]. *Nucleic Acids Research*, 2007, 35(13):4535-4541.
- [13] Liu W, Cui Z, Zan X. Identifying cancer-related microRNAs based on subpathways[J]. *IET Systems Biology*, 2018, 12(6):273-278.
- [14] Lu J, Getz G, Miska E A, et al. MicroRNA expression profiles classify human cancers[J]. *Nature*, 2005, 435(7043):834-838.
- [15] Esquela-Kerscher A, Slack F J. Oncomirs-microRNAs with a role in cancer[J]. *Nature Reviews Cancer*, 2006, 6(4):259-269.
- [16] 许鹏, 谢斌, 鲍振申, 等. 基于自编码器的疾病相关 miRNAs 的预测方法[J]. *广州大学学报(自然科学版)*, 2024, 23(1):12-19.
- [17] Zhou F, Yin M M, Jiao C N, et al. Predicting miRNA-disease associations through deep autoencoder with multiple kernel learning[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2023, 34(9):5570-5579.
- [18] Ha J, Park S. NCMD: Node2vec-based neural collaborative filtering for predicting miRNA-disease association[J]. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2023, 20(2):1257-1268.
- [19] Li M, Li W, Wu F X, et al. Identifying essential proteins based on sub-network partition and prioritization by integrating subcellular localization information[J]. *Journal of Theoretical Biology*, 2018, 447:65-73.
- [20] Guo F B, Dong C, Hua H L, et al. Accurate prediction of human essential genes using only nucleotide composition and association information[J]. *Bioinformatics*, 2017, 33(12):1758-1764.
- [21] Peng W, Wang J, Wang W, et al. Iteration method for predicting essential proteins based on orthology and protein-protein interaction networks[J]. *BMC Systems Biology*, 2012, 6(1):87-103.
- [22] Cui C, Shi B, Shi J, et al. Defining the importance score of human microRNAs and their single nucleotide mutants using random forest regression and sequence data[J]. *Advanced Theory and Simulations*, 2019, 2(9):1900083.
- [23] Chou C H, Shrestha S, Yang C D, et al. miRTarBase update 2018: A resource for experimentally validated microRNA-target interactions[J]. *Nucleic Acids Research*, 2018, 46(D1):D296-D302.
- [24] Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes[J]. *Nucleic Acids Research*, 2000, 28(1):27-30.
- [25] Zhang J D, Wiemann S. KEGGgraph: A graph approach to KEGG pathway in R and bioconductor[J]. *Bioinformatics*, 2009, 25(11):1470-1471.
- [26] Huang Z, Shi J, Gao Y, et al. HMDD v3.0: A database for experimentally supported human microRNA-disease associations[J]. *Nucleic Acids Research*, 2018, 47(D1):D1013-D1017.