

文章编号: 2095-2163(2020)08-0135-05

中图分类号: TP312

文献标志码: A

长链非编码 RNA 与疾病关联关系的预测方法研究

阳金豆¹, 袁静², 许敏¹

(1 贵州大学 医学院, 贵阳 550025; 2 贵州省人民医院 肾内科, 贵阳 550002)

摘要: 随着高通量测序技术与计算机技术的高速发展,越来越多的研究表明 lncRNA 与多种人类疾病密切相关,两者间的关联成为了新的研究热点。本文简要概述了一些可供预测和验证使用的公共数据库,以及近几年来开发的预测 lncRNA-疾病关联的算法与工具,分析了这些算法与工具共同存在的问题,并讨论了未来的发展方向。

关键词: lncRNA-疾病关联; 数据库; 算法

Research on the prediction method of the association between long-chain non-coding RNA and disease

YANG Jindou¹, YUAN Jing², XU Min¹

(1 College of Medicine, Guizhou University, Guiyang 550025, China;

2 Department of Nephrology, Guizhou Provincial People's Hospital, Guizhou 550002, China)

[Abstract] With the rapid development of high-throughput sequencing technology and computer technology, more and more studies have shown that lncRNA is closely related to various human diseases, and the relationship between the two has become a new research hotspot. This article briefly outlines some public databases that can be used for prediction and verification, as well as algorithms and tools for predicting lncRNA-disease associations developed in recent years, analyzes the common problems of these algorithms and tools, and then discusses the future development direction.

[Key words] lncRNA-disease association; database; algorithm

0 引言

非编码 RNA (Non-coding RNA, ncRNA), 即不参与编码蛋白质的 RNA。近年来,信息技术与高通量测序技术飞速发展,研究人员发现那些曾经被认为是转录噪声的 ncRNA 在许多生物学进程中发挥着相当大的作用^[1]。其中,因具有重要调控功能且几乎参与到了各种生物学过程、与各种疾病的发生发展关联紧密的长链非编码 RNA (long noncoding RNAs, lncRNA) 成为了近几年的研究热点^[2]。lncRNA, 即长度大于 200 核苷酸的非编码 RNA^[3]。与编码 RNA 相比,人类基因中 lncRNA 数量更多。现阶段,只有少数的 lncRNA 功能及其与疾病间关系较为明确,因而值得进行深入研究^[4]。

相较于传统的生物实验加临床实验的方法,开发有效的算法能够节省大量的时间和精力,以达到快速准确预测关联的目的。因此,近年来研究人员已开发出了不少算法模型和工具以进行 lncRNA 与疾病的关联预测。

本文对近几年提出的一些探寻疾病关联的 lncRNA 的算法、工具进行了综述,并对当前出现的问题和未来的研究方向进行了讨论。

1 相关数据库

随着信息技术与高通量测序技术的飞速发展进步,生物学数据的数量呈现爆炸式增长,因此研究人员建立了专门的数据库以实现数据的存储和管理。与此同时,还建立起了一些 lncRNA 对人类疾病调控数据进行收集整理的数据库,这为验证算法的预测结果提供了较为直接的证据。本节将对一些常用的 lncRNA 数据库、疾病相关数据库及 lncRNA 调控人类疾病关系数据库三类公共数据库进行介绍。

1.1 与 lncRNA 相关的数据库

(1) NONCODE。NONCODE^[5] 是专门针对非编码 RNA (其中不包括 tRNAs 和 rRNAs) 所建立的。包含有 17 个物种(人类、小鼠、拟南芥等)的非编码 RNA 相关数据。NONCODE 数据库可以对 ncRNA 进行注释、评估编码能力、提供位置信息、表达信息

基金项目: 贵州省卫生健康委科学技术基金 (gzwjky2019-1-107)。

作者简介: 阳金豆(1994-),女,硕士研究生,主要研究方向:医学智能信息处理、生物信息学;袁静(1975-),女,博士,教授,硕士生导师,主要研究方向:慢性肾脏病、风湿免疫疾病、生物信息学等;许敏(1995-),女,硕士研究生,主要研究方向:数字图像处理、医学智能信息处理。

收稿日期: 2020-06-07

和潜在功能以及共表达信息等。此外, NONCODE 数据库还对先前的基因芯片数据进行了重注释, 以获得更多的 lncRNA 表达信息。

(2) LncRNome。LncRNome^[6] 是一个收录了超过 18000 条转录本信息的针对人类而构建的数据库。数据库提供有 lncRNA 的分类、染色体位置、二级结构、与其它数据库相关联信息、生物功能描述、与疾病关联的数据。同时提供 lncRNA 与蛋白质相互作用功能预测、SNP 位点等信息。

(3) LncRNadb。LncRNadb^[7] 是针对真核生物所建立的 lncRNA 综合数据库。其中包含有特异的序列结构信息: 转录本数据、基因组位置表达及相关的功能、疾病等。

1.2 与疾病相关数据库

(1) OMIM。OMIM^[8] 的中文名为在线人类孟德尔遗传数据库, 是由美国约翰斯霍普金斯大学医学院 McKusick-Nathans 遗传医学研究所维护的在线人类基因和遗传疾病数据库。OMIM 数据库中包括了现今已知的所有遗传病和 15000 个以上的基因信息, 较侧重疾病表型与其致病基因之间的关联关系。

(2) MalaCards。MalaCards^[9] 综合了 72 个数据库信息的人类疾病数据库。其中包含了疾病总概、疾病间相互关系、所涉通路、基因表达差异等内容。

(3) PUBMED。PubMed^[10] 是由美国国家生物技术信息中心 (NCBI) 开发的基于网路的综合性质生物医学信息检索系统, 提供了与各种综合分子生物学数据库的链接等资源服务。

1.3 lncRNA 调控人类疾病关系数据库

(1) LncRNADisease。LncRNADisease^[11-12] 是一个手动收录各种经实验验证的与 lncRNA 相互作用信息的数据库, 共收录 205959 个 lncRNA-疾病关联。LncRNADisease 通过对疾病的聚类 and 序列的保守性 (SNP 的分布密度) 来预测 lncRNA 与疾病的相关性。与之前的版本相比: ncRNA-疾病关联增强了 40 倍; 提供 lncRNA、mRNA 和 miRNA 之间的转录调控关系; 提供 ncRNA、mRNA 和 miRNA 之间的调控关系; 提供每个 lncRNA 疾病关联的置信度得分。

(2) Lnc2Cancer。Lnc2Cancer^[13-14] 是从文献中收集整理与人类 lncRNA 相关癌症信息的数据库。通过查阅 6500 多篇已发表文献, 目前的 Lnc2Cancer 版本共记录了 4 989 条 1614 个人类 lncRNAs 和 165 个人类癌症亚型之间的关联。此外, 数据库提供了癌症相关 lncRNAs 的实验支持、耐药性、预后, 以及受 microRNA (miRNA)、转录因子 (TF)、变异和甲

基化调控的 lncRNAs。

(3) MNDR。MNDR^[15] 是哺乳动物非编码 RNA-疾病关系在线知识库。该数据库从文献中整理出实验支持的疾病相关的 ncRNA, 同时整合了 LncDisease、miRPD 等数据库资源。通过检索系统, 即可以得到与疾病相关的 ncRNA, 共收集了超过 260000 条 ncRNA-疾病的关联。比之前版本相比较, ncRNA-疾病关联性增强了 220 倍以上; 整合来自 14 种资源的实验和预测证据以及每个 ncRNA-疾病关联性的预测算法; 提供每个 ncRNA-疾病关联的置信度得分; 增加了 6 种哺乳动物的物种覆盖率。

2 算法与工具

相较于传统的生物实验加临床实验的方法, 开发有效的算法和工具能够节省大量的时间和精力, 以达到快速模拟预后与预测关联的目的。本节简要概述了一些对 lncRNA-疾病关联进行预测的计算方法和工具。

2.1 基于机器学习的方法

通常, 基于机器学习的方法是使用已知的疾病与 lncRNA 关联来训练学习模型, 然后运用学习后的模型来对新的关联进行预测。这类方法在综合了序列信息、结构和表达数据后才能对新的 lncRNA-疾病关联进行预测整合注释。

Zhao 等^[16] 开发了一种基于朴素贝叶斯分类器方法的计算方法, 对基因组、调控组、转录组的数据进行整合, 共识别出 707 个与癌症相关的潜在 lncRNA。

Hu 等^[17] 提出了一个深度学习框架 NNLD, 用以预测潜在的 lncRNA-disease 关联。NNLD 是第一个使用深度神经网络预测 lncRNA-disease 关联的算法。在 19 166 个 lncRNA 与 529 种疾病之间的 205 959 个相互作用的网络上, 研究者将 NNLD 与另外两种广泛使用的算法进行了比较。结果显示, NNLD 在 lncRNA-disease 关联预测方面优于其它现有算法。另外, 因为利用了小批量随机梯度下降技术, 在规模较大的数据集中 NNLD 应用更容易。

Shi 等^[18] 提出了一个基于图回归的统一框架 (graph regression-based unified framework, GRUF)。此方法可用于之前没有已知关联的疾病与 lncRNAs。研究者将 GRUF 与 3 种最先进的方法进行了比较, 证明了 GRUF 的优越性, 使得工作特性曲线 (AUC) 下的面积提高了 5%~16%。GRUF 还为预测 LDA (lncRNAs and diseases) 提供了一个预测的

置信度评分,同时揭示了评分与 LDAs 中涉及的 RNA 结合蛋白数量之间的显著相关性。GRUF 在新的预测中所产生的前 5 位候选 LDA 中,有 3 位被医学文献和已知生物学事实验证。

Ashis 等^[19]提出了一种作为归纳矩阵完成技术 (IMC) 补救措施的方法 RIMC (一种使用 $\ell_2, 1$ 范数损失函数及基于 $\ell_2, 1$ 范数正规化的鲁棒归纳完备化方法)。研究者将 RIMC 应用于人类 lincRNA (Long intergenic non-coding RNA, 基因间长链非编码 RNA, lincRNA 的一种^[20]) 和 OMIM 疾病表型之间的可用关联数据,以及有关 lincRNA 和疾病的多种辅助信息。该方法为 lincRNA 和通过不同独立数据源获得的疾病的各种类别特征提供了一个集成界面,用于解释两个实体之间的关系。该方法可以处理数据集中的固有噪声和离群值,并且表现出优于基于 ℓ_2 范数的标准 IMC 公式。在对 lincRNAs 进行前 k 位疾病优先排序时,精密度 @ k 和召回率 k 方面表现更好。同时证明,RIMC 在查询新的 lincRNA 方面同样有效,并且对于一组已知的 lincRNA 预测新近已知疾病也同样有效。

2.2 基于网络的方法

相对于数量较少的基于机器学习方法的研究,基于网络的预测方法数量更多。基于网络的预测,通常需要结合基因表达谱等数据,以及多种信息的逻辑关联来构建网络,方法较为复杂。构建完成后,根据 lincRNA 与疾病的关联得分大小对候选的 lincRNA 进行排序来预测致病基因。最常用的是标签传播算法,如随机漫步 (RWR) 和 KATZ,其主要区别在于不同的传播算法所应用的底层网络。

Sun 等人^[21]将 RWR 应用到了 lincRNA 功能相似网络上,提出基于全球网络的计算框架 RWRlncD,实现重启随机游走,以推断潜在的人类 lincRNA 疾病关联。通过留一法交叉验证 RWRlncD 的性能,ROC 曲线下面积为 0.822,证明 RWRlncD 性能良好。研究者通过文献挖掘,手动确认了在前列腺癌和阿尔茨海默氏病的案例研究中具有高度预测性的 lincRNA-疾病关联,这为运用 RWRlncD 方法预测 lincRNA-疾病关联方面的良好性能和潜在价值提供了证据。

Liu 等人^[22]提出了一种基于样本相关的方法,将其应用于前列腺癌中 lincRNA 和 PCGs 的表达谱,结合蛋白质相互作用相关数据,构建 lincRNA-PCG 二分网络,采用随机游走的方法从此二分网络中提取候选肿瘤相关 lincRNA。研究者从 ncRNADisease 数

据库和 MNDR 数据库中获得 14 个前列腺癌相关 lincRNA,发现其中 6 个存在于所构建的网络中。该研究为今后的 lincRNA-疾病关联研究带来新的视角。

Ganegoda 等^[23]提出了一种新算法—KRWRH,通过 lincRNA 的表型信息和组织表达细节来计算疾病的高斯交互作用谱核,然后采用重启的随机游走法进行最终预测。研究将 KRWRH 与 LRLSDAA、TslncRNA、NRWRH、RWRH 这 4 种现有方法进行比较,留一法交叉验证、ROC 曲线、均值富集的结果均表明,KRWRH 能够更有效地预测未知的疾病与 lincRNA 的关联。

Zhou 等^[24]提出了一种新的基于秩的方法 RWRHLD。通过将与 miRNA 相关的 lincRNA - lincRNA 串扰网络、疾病-疾病相似性网络和已知的 lincRNA-疾病关联网络三个网络集成到一个异构网络后,运用重启随机游走,对候选的 lincRNA 疾病关联进行排序,进而预测潜在疾病 lincRNA 关联关系。基于已通过实验证实的 lincRNA-疾病关联,使用留一法进行交叉验证, AUC 值为 0.871,远高于仅基于 lincRNA 网络、超几何测试和随机情况下的 RWR。此外,在卵巢癌和前列腺癌的病例研究中预测的一些新的 lincRNA 疾病关联也已通过文献分析得到证实。

马毅等^[25]使用 HeteSim 法计算疾病基因之间的相关性,预测致病基因。基于路径约束,HeteSim 法将元路径两端节点随机游走到中间节点相遇的概率作为疾病与 lincRNA 之间关联关系的得分,发掘潜在的疾病-lincRNA 关联关系。为提升算法的性能,研究者将编码基因与疾病的关联关系与 lincRNA-疾病网络相整合,得到了基因-疾病异质信息网络。之后进行留一交叉验证, AUC 值为 0.7835。通过整合集成编码基因-疾病关联数据增加网络中边的数量,使得结构更加紧密,潜在 lincRNA 可以从其他基因-疾病关联中获得更多信息传播,进而能够更好地进行预测,算法性能得到提高。研究者将 HeteSim 与文献[26-28]中算法进行比较,结果均优。

文献[27]中考虑到以往算法的局限性,结合已知的 lincRNA-疾病关联、lincRNA 表达谱、lincRNA 功能相似度、疾病语义相似度及高斯交互作用谱核相似性,建立了测量模型 KATZ lincRNA - disease Association prediction (KATZLDA),以发现潜在的 lincRNA-disease 关联。KATZLDA 在局部和全局留一法交叉验证和 5 倍交叉验证中分别获得 AUC 值为:0.717 5、0.788 6、0.771 9。而后研究者通过对结

肠癌、胃癌和肾癌案例进行了分析,生物学实验已证实了前10位预测中的60%。

Xie等^[29]基于功能相似的lncRNAs往往与表型相似的疾病相关联,提出了一种新的算法NCPHLDA(lncRNA-disease associations network consistency projection),用于预测潜在的lncRNA-disease关联。该方法集成了lncRNA余弦相似网络、疾病余弦相似网络和已知的lncRNA-疾病关联网络,无需参数及负样本,在没有任何已知相关疾病的情况下也可预测lncRNA。分别对其进行leave-one-out交叉验证和5倍交叉验证,得到AUC值为0.9273和 0.9179 ± 0.0043 。而后研究者对3种疾病(乳腺癌、宫颈癌和肝癌)的案例研究表明,该算法也具有可靠的预测性能。

Li等^[30]使用OMIM数据库中的表型与基因的关系信息,连接基因网络和表型网络,构建了一个异构网络,然后将重启随机游走扩展到这个异构网络,对基因和表型同时进行排序。使用留一法交叉验证来评估,性能比其它算法有所提升。研究者使用该算法揭示出一些仅靠基因网络或表型网络无法发现的隐藏疾病关联,共确定了18个隐藏的疾病关联,大部分得到了文献证据的支持。

Zhang等^[31]开发了一个新的基于全球网络的框架LncRDNetFlow,以区分与疾病相关的lncRNA的优先级。基于lncRNA相似性、疾病相似性以及它们之间的关联性等多种生物信息,LncRDNetFlow利用流传播算法将多个网络集成到一个网络来预测lncRNA疾病关联性。交叉验证AUC值为0.841,性能明显优于现有技术水平。通过文献筛选出卵巢癌、神经胶质瘤和宫颈癌相关的lncRNA,并通过生物学研究证实了排名靠前的许多预测的出的lncRNA。

Wang等^[32]开发了一种基于序列的生物信息学方法LncDisease,以lncRNA和miRNAs之间的串扰为基础,预测lncRNA与疾病的关联。研究者运用LncDisease预测了与乳腺癌和高血压相关的lncRNA,同时进行生物实验,qRT-PCR结果显示12种预测的lncRNA中有11种(91.7%)得到验证。同时,研究者还评估了受血管紧张素II(Ang II)调节的人血管平滑肌细胞(VSMCs)中与高血压相关的lncRNAs。qRT-PCR结果显示,4个预测的lncRNAs中有3个(75.0%)得到验证。此外,研究者还对lncRNAGAS5相关的6种疾病进行预测,并通过文献挖掘验证了其中的4种(66.7%)。

Yu等^[33]开发了一个通过双随机漫步来预测新

的lncRNA疾病关联(BRWLDA)的双向关系网络模型。双向关系网络由从各种生物学数据得出的lncRNA-lncRNA功能相似性网络和从疾病之间的本体结构得出的疾病-疾病相似性网络两个子网络,以及这两个子网络之间的互连关系lncRNA-疾病关联所构成。与其它计算模型相比,BRWLDA在lncRNA和疾病导向的LOOCV中获得了最高的AUC值0.7952和0.7940。在递归掩蔽的实验中,BRWLDA再次获得了最高的AUC值,分别为0.9888、0.9548、0.9275。研究者对乳腺癌、结肠癌、膀胱癌、肝癌、肺癌和胃癌进行了案例研究,BRWLDA获得了更多的相关lncRNA。此外,手工文献挖掘也证实,BRWLDA预测出的前20个合理的lncRNA确实与乳腺癌、结肠癌和肺癌有关。

除了提出上述预测算法工具外,研究者还对预测算法进行优化,以提高预测准确性。

Duc-Hau Le等^[34]提出了一种新的方法ORIENT(亲邻权重强化),通过适当加强与已知疾病基因相互作用的权重,来提高RWR的性能,帮助研究人员确定临界点。通过对数百种疾病进行仿真,研究者发现ORIENT比传统的RWR算法具有更好的性能。当只有所涉及疾病基因的最邻近基因的相互作用的权重增加时算法效果最好。

王波等^[35]提出了一种基于动态衰变群的优化算法,将其应用于与疾病关联的lncRNA预测中,有效提高了预测的准确性。

3 结束语

随着高通量测序技术与计算机技术的高速发展,越来越多的研究表明lncRNA与多种人类疾病密切相关,两者间的关联成为了新的研究热点。研究两者间关联有助于加深对疾病致病机理的理解,为后续的疾病诊疗和防治打下基础。与传统的生物实验和临床实验相比较,有效的算法和工具能够节省大量的时间和精力且达到预测关联的目的。

本文对近几年提出的一些可供使用的公共数据库、开发的预测lncRNA-疾病关联的算法与工具进行简要概述。这些算法与工具虽有一定的成果,但也存在一些不足,后续需要进行更深入的研究改进。首先是数据量的问题,上述算法都是基于数据库构建,但有的疾病在数据库中的病例数量较少,无法满足机器学习或是网络构建的需求。然后是整合数据的问题,疾病与lncRNA间以及不同数据库间的数据整合存在困难,需要针对这一问题提出适当的解决方法或整合标准。最后是相似性评估标准的问

题,不同数据间相似性如何评估处理也需要提出相应的解决方法与标准。总之,预测 lncRNA-疾病关联研究未来还有很长的路要走。

参考文献

- [1] LOURO R, SMIRNOVA A S, VERJOVSKI-ALMEIDAS. Long intronic noncoding RNA transcription; Expression noise or expression choice? [J]. *Genomics*, 2009, 93(4): 291-298.
- [2] CHEN X, YAN C C, LUO C, et al. Constructing lncRNA functional similarity network based on lncRNA - disease associations and disease semantic similarity [J]. *entific reports*, 2015, 5:11338.
- [3] Orly Wapinski, Howard Y. Chang. Long noncoding RNAs and human disease. [J] *Trends in Cell Biology*. 2011, 21(6).
- [4] HARRIES, LORNA W. Long non-coding RNAs and human disease[J]. *Biochemical Society Transactions*, 2012, 40(4):902-906.
- [5] YI Z, HUI L, SHUANGSANG F, et al. NONCODE 2016: an informative and valuable data source of long non-coding RNAs [J]. *Nucleic Acids Research(D1)*:D1.
- [6] DEEKSHA B, KOUSTAV P, SOURAV G, et al. lncRNome: a comprehensive knowledgebase of human long noncoding RNAs [J]. *Database*.
- [7] CHENG Q X, THOMSON D W, MAAG JESPER L V, et al. lncRNADB v2.0: expanding the reference database for functional long noncoding RNAs[J]. *Nucleic Acids Research(D1)*:D1.
- [8] AMBERGER J S, BOCCHINI C A, SCHIETTECATTE F, et al. OMIM.org: Online Mendelian Inheritance in Man (OMIM 库), an online catalog of human genes and genetic disorders. [J]. *Nucleic Acids Research*, 2015, 43(Database issue):789-98.
- [9] Rappaport, Noa, Twik, Michal, Nativ, Noam, et al. MalaCards: A Comprehensive Automatically - Mined Database of Human Diseases[M]// *Current Protocols in Bioinformatics*. John Wiley & Sons, Inc. 2014.
- [10] 吴蓉, 羨秋盛, 刘一洋. PubMed 数据库检索系统及相关检索方法[J]. *实用医药杂志*, 2010, 27(3):276-277.
- [11] BAO Z, YANG Z, HUANG Z, et al. lncRNADisease 2.0: an updated database of long non-coding RNA-associated diseases [J]. *Nucleic Acids Research*, 2018.
- [12] CHEN G, WANG Z, WANG D, et al. lncRNADisease: A database for long - non - coding RNA-associated diseases [J]. *Nucleic Acids Research*, 2013, 41(Database issue): 983-986.
- [13] SHANGWEI N, JIZHOU Z, PENG W, et al. lnc2Cancer: a manually curated database of experimentally supported lncRNAs associated with various human cancers[J]. *Nucleic Acids Research (D1)*:D1.
- [14] GAO Y, WANG P, WANG Y, et al. lnc2Cancer v2.0: updated database of experimentally supported long non-coding RNAs in human cancers[J]. *Nucleic acids research*, 2019, 47(D1):D1028.
- [15] TIANYU C, LIN Z, YAN H, et al. MNDR v2.0: an updated resource of ncRNA-disease associations in mammals[J]. *Nucleic Acids Research(D1)*:D1.
- [16] ZHAO T, XU J, LIU L, et al. Identification of cancer-related lncRNAs through integrating genome, regulome and transcriptome features[J]. *Mol Biosyst*, 2015, 11(1):126-136.
- [17] HU J, GAO Y, LI J, et al. 2019. Deep Learning Enables Accurate Prediction of Interplay Between lncRNA and Disease. [J] *Front Genet*, 2019, 10: 937.
- [18] SHI J Y, HUANG H, ZHANG Y N, et al. Predicting binary, discrete and continued lncRNA-disease associations via a unified framework based on graph regression[J]. *Bmc Medical Genomics*, 2017, 10(S4):65.
- [19] ASHIS B, MINGON K, DONGCHUL K, et al. Robust Inductive Matrix Completion Strategy to Explore Associations between lncRNAs and Human Disease Phenotypes [J]. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2018: 1-1.
- [20] ULITSKY I, BARTEL D. lincRNAs: Genomics, Evolution, and Mechanisms[J]. *Cell*, 2013, 154(1):26-46.
- [21] SUN J, SHI H, WANG Z, et al. Inferring novel lncRNA-disease associations based on a random walk model of a lncRNA functional similarity network[J]. *Molecular Biosystems*, 2014, 10(8):2074.
- [22] LIU Y, ZHANG R, QIU F, et al. Construction of a lncRNA-PCG bipartite network and identification of cancer-related lncRNAs; a case study in prostate cancer [J]. *Molecular BioSystems*, 2015, 11(2):384-393.
- [23] GANEGODA G U, LI M, WANG W, et al. Heterogeneous Network Model to Infer Human Disease-Long Intergenic Non-Coding RNA Associations [J]. *NanoBioscience, IEEE Transactions on*, 2015, 14(2):175-183.
- [24] ZHOU M, WANG X, LI J, et al. Prioritizing candidate disease-related long non-coding RNAs by walking on the heterogeneous lncRNA and disease network [J]. *Mol Biosyst*, 2015, 11(3):760-769.
- [25] 马毅, 郭杏莉, 孙宇彤, 等. 基于 HeteSim 的疾病关联长非编码 RNA 预测[J]. *计算机研究与发展*, 2019(9).
- [26] XING C, ZHU-HONG Y, GUI-YING Y, et al. IRWRLDA: improved random walk with restart for lncRNA-disease association prediction[J]. *Oncotarget*, 2016, 7(36).
- [27] CHEN Xing. KATZLDA: KATZ measure for the lncRNA-disease association prediction[J]. *Scientific Reports*, 2015, 5(1):16840.
- [28] YANG X, GAO L, GUO X, et al. A Network Based Method for Analysis of lncRNA - Disease Associations and Prediction of lncRNAs Implicated in Diseases [J]. *Plos One*, 2014, 9(1): 87797.
- [29] XIE G, HUANG Z, LIU Z, et al. NCPHLDA: a novel method for human lncRNA - disease association prediction based on network consistency projection[J]. *Molecular Omics*, 2019, 15.
- [30] YONGJIN L, PATRA J C. Genome-wide inferring gene-phenotype relationship by walking on the heterogeneous network [J]. *Bioinformatics(9)*:9.
- [31] ZHANG J, ZHANG Z, CHEN Z, et al. Integrating Multiple Heterogeneous Networks for Novel lncRNA-disease Association Inference[J]. *IEEE/ACM Transactions on Computational Biology & Bioinformatics*, 2017:1-1.
- [32] WANG Junyi, RUI Xia, et al. lncDisease: a sequence based bioinformatics tool for predicting lncRNA - disease associations. [J]. *Nucleic Acids Research*, 2016.
- [33] YU G, FU G, LU C, et al. BRWLDA: bi-random walks for predicting lncRNA-disease associations[J]. *Oncotarget*, 2017, 8(36).
- [34] LE D H, KWON Y K. Neighbor-favoring weight reinforcement to improve random walk - based disease gene prioritization [J]. *Computational biology and chemistry*, 2013, 44(jun.):1-8.
- [35] 王波, 张菁. 基于异步消减群优化的疾病关联 lncRNA 预测方法[J]. *现代信息科技*, 2018.