



林 汇亨 **Hui-Heng Lin** 计算机系统生物学 博士

Linkedin: www.linkedin.com/in/res000/ **ORCID:** orcid.org/0000-0003-4060-7336

Github: <https://github.com/devgithwaar?tab=repositories>

Email: lin.hui.heng@connect.um.edu.mo **TEL:** +86 17138862101

专长

计算分析及建模预测 | 系统及计算生物学 | 交叉专业科学研究 | 复杂网络分析

技能

统计分析 / 建模预测 / 机器学习 / 生物信息数据库及分析 / 复杂网络建模分析 / 原始数据清洗及结构化格式化
/ 数据可视化 / 学术英语写作及润色

技术

Linux 操作 & Shell 脚本 / Python / R / SQL / Microsoft Excel Advance Functions / SPSS

学历

2015 ~ 2019: 系统计算生物学 博士

- **澳门大学** (国家重点实验室), 位于中国澳门特别行政区
- 博士毕业论文: Repositioning anti-cancer agents for neurodegenerative diseases via drug-disease network analysis (论文答辩成绩: 优). 论文主题译为: 基于药物-疾病复杂网络链路预测方法的神经退行性疾病的抗癌药重定位研究)
- 核心课程: 高级统计分析 / 图论 / 复杂网络建模分析 / 系统医学 / 网络药理学 / 研究方法导论 / 英语科学写作及演讲

2012 ~ 2014: 医学科学 硕士 (研究方向: 分子肿瘤及免疫学)

- **东京大学**, 位于日本东京
- 核心课程: 肿瘤学 / 免疫学 / 感染及医学微生物学 / 传染病学 / 生物物理学 / 生物化学 / 基因组学 / 生物信息学
- 硕士毕业论文: *in vitro* reconstitution of oncogenic protein CagA-Csk kinase interaction (论文成绩: 优). 论文主题译为: 致癌蛋白 CagA-Csk 激酶相互作用蛋白对的体外生物化学和生物物理特性探索研究)

2007 ~ 2011: 生物技术 理学士 (排名前 10%)

- **华南理工大学** (985、双一流), 位于中国广东省广州市
- 核心课程: 计算机技术 / 线性代数与概率统计 / 高等数学微分积分 / 生物信息学 / 有机无机化学 / 分析化学 / 物理化学 / 微生物学 / 分子与细胞生物学 / 基因组学 / 免疫学 / 遗传学 / 发酵工程
- 毕业论文: 搞分子化合物家族的三维分子结构建模 (研究方向: 化学/生物信息学)

项目经验精选 (详见个人网页 Github、Linkedin 或 ORCID)

项目一、机器学习预测 BRCA1 基因突变的致癌风险.

- (1) 使用正则表达式等脚本工具清洗、整理并结构化、数值化 (生成特征向量) 6k+条 DNA 序列数据.
- (2) 训练朴素贝叶斯、支持向量机、逻辑回归、随即森林等多个机器学习模型, 并与第三方参照模型 (Polyphen 及 SIFT) 进行预测性能的比较.
- (3) 经优化, 最佳模型随机森林的 ROCAUC 从 **80% 提升至 85%** 比第三方参照模型最高高出 **11 %** 。该最佳模型被应用于预测临床人群 (约 60 人) 的 BRCA1 基因突变的致癌风险, 成功辅助临床诊断。
- (4) 项目相关的结果数据被可视化并撰写为英文论文, 发表在 MedRxiv 预印本平台。详见 Lin HH, et al. Predicting ovarian/breast cancer pathogenic risks of BRCA1 gene variants of unknown significance. MedRxiv 2020. <https://doi.org/10.1101/2020.06.04.20120055> **林汇亨为第一作者**
- (5) 该项目荣获韶关市科技局科技计划资助(资助项目号: 200812114531428)。林汇亨为该项目的负责人。
- (6) 核心技术: R 语言编程脚本及包 (Biostrings, Bioconductor, e1071, randomForest, protr, pROC, ROCR, ggplot2, 等)

项目二、使用复杂网络模型的链路预测方法预测抗癌药物对神经退行性疾病 (如阿尔兹海默病, 俗称老年痴呆, 和帕金森病) 的药物再定位

- (1) 把药物的分子结构转化为分子描述符向量 (化合物指纹图谱和原子对描述符), 并使用 Tanimoto 系数计算 **40,000 对药物的结构相似度**. 使用余弦相似度系数计算 **1024 对神经退行性疾病的表征相似度**, 并可视化两个相似度系数矩阵, 绘制为热图.
- (2) 经数据标准化, 矩阵被整合为一个邻接矩阵。可视化该矩阵生成**药物-疾病二元网络 (图)**。该网络包含 232 个节点和 1007 条边。对网络进行**拓扑学分析**了解该网络的基本属性如模块性等。

- (3) 对上述构建的网络模型应用 8 种网络链路预测算法(均为非监督机器学习法, 包括基于模块聚类的算法、基于共同邻居计算的多种、基于路径计算的多种算法、及基于特征值计算的多种算法等), 且编写自定义脚本以留一法 (Leave One Out) 和计算排列分值评估每种链路预测算法的预测性能。绘制箱图/小提琴图可视化每种算法的排序分值分布。经比较发现 Rooted PageRank 为预测性能最佳的链路预测算法。
- (4) 选用 Rooted PageRank 预测网络中的 **6367 对抗癌药物-神经退行性疾病组合是否存在潜在连边**。于文献数据库中, 检索预测所得的高分值的药物-疾病连边的相关文献。在 7 对预测所得的高分值抗癌药物-神经退行性疾病组合中发现 3 对抗癌药物-神经退行性疾病组合有文献报道支持。
- (5) 该项目结果数据被可视化并撰写为英文学术论文, 预备投递至 SCI 学术期刊出版。
- (6) 本项目荣获 **Ernst Mach 学术交流项目资助**。资助编号: ICM-2018-10230。赞助方为欧亚太平洋基金会 (EurAsia pacific network foundation) 及奥地利 OeAD 组织。
- (7) 核心技术: Python 编程脚本及模块 (igraph, networkX, numpy, pandas, RDkit, matplotlib, linkprediction), R 语言脚本及包 (igraph, SANDS, LinkPrediction, ChemmineR, dSimier, pheatmap, disgen2r, Bioconductor, ggplot2), Gephi, 等。

项目三、对医疗保险控费措施调查问卷数据的统计分析 (贝叶斯概率计算及网络)

- (1) 对医疗保险控费措施所造成的患者诊疗差异进行问卷设计, 并使用 Cronbach alpha test 验证问卷的可靠性及因子分析 Kaiser-Meyer-Olkin and Bartlett's test 验证问卷的有效性。发放问卷至 110 为临床医生并进行访谈, 收集整理回答数据, 对某些非数值数据转化为数值并进行统计分析
- (2) 进行 2860 次大规模卡方检验以检测医保控费措施及诊疗受影响事件之间的关联性。基于具有显著性 ($P\text{-value} < 0.01$) 的卡方检验结果数据构建包含 25 个节点及 43 条边的贝叶斯(概率)网络 (Bayesian network)。对网络进行多变量分析及贝叶斯概率计算以推断医保控费措施和诊疗受影响的事件之间的因果关系
- (3) 该项目结果数据被可视化并撰写为英文学术论文, 且发表在国际 SCI 期刊《BMC Healthcare Service Research》上。详见 “Yan J, **Lin HH***, (本人作为共同第一作者) et al., China's new policy for healthcare cost-control based on global budget. *BMC Healthcare Service Research*, 2019, 19:84. <https://doi.org/10.1186/s12913-019-3921-8> (*本人负责该项目的研究方法设计、数据分析及整理、统计分析及建模英文手稿的撰写, 及投稿的修订、审核)
- (4) 核心技术: R (igraph, ggplot2, pheatmap, etc), Gephi, SPSS, Microsoft Excel, etc。
- (5) 本研究受到以下各期刊的编辑关注 Journal of quality in healthcare & economics、Journal of health economics and outcomes research、Journal of health policy and outcomes research。

工作 / 研究经历

2019 ~ 至今: 汕头大学附属医院附属粤北人民医院 助理研究员 (博士后。于中国广东省韶关市)

1. **科研:** 使用机器学习、统计分析、复杂网络分析等方法独立开展生物信息学、计算系统生物医学等多个科研项目并撰写论文投稿至 SCI 学术期刊
2. **科研经费:** 独立构思并撰写多份研究计划书用于申请多项科研基金。荣获粤北人民医院博士后启动科研经费及韶关市科技计划项目资助 (资助编号: 200812114531428; 本人作为项目负责人)
3. **奖项/荣誉:** 获广东省博士后人才计划 (原扬帆计划) 及韶关市博士后人才计划
4. **指导 (协助):** 协助指导 3 位临床型硕士研究生的科研项目及论文撰写。如实验流程设计、统计分析方法、及数据挖掘、基因表达分析方法。
5. **科研合作:** 作为主要负责人带领团队成员合作申请多项科研基金。同时主导多项合作研究, 如探索复方穿心莲对癌症治疗的网络药理学机制及临床遗传、基因突变数据的检测及生物信息学分析
6. **科研咨询:** 为医院多个科室、多个部门、多位临床医师、技师、药师提供专业科研咨询服务。如 meta 分析、生物信息学分析、统计建模分析、网络药理学及中医药研究方面的知识及方法。

2015 ~ 2019: 澳门大学 研究助理 (于中国澳门特别行政区)

1. **科研:** 使用统计分析、复杂网络分析、机器学习、生物信息学分析独立进行多项生命科学及交叉学科研究课题

2. **学术研究论文出版:** 于 SCI 国际学术期刊高产发表多篇原创研究型论文。其中 2 篇第一作者论文, 2 篇共同第一作者论文, 1 篇第二作者论文。同时发表 4 篇原创研究型论文作为专著章节 (详见 ORCID: <https://orcid.org/0000-0003-4060-7336> 或下方论文列表部分)
3. **奖项及荣誉:** (1) 获 2018~2019 年度 Ernst Mach 科研学术交流访问**资助**: 资助本人至奥地利维也纳学术机构 (维也纳医科大学及维也纳生命科学与农业资源大学) 进行为期 9 个月的学术访问及科研合作。(赞助方: 欧亚太平洋基金会及奥地利 OeAD 组织。资助编号: ICM-2018-10230); (2) 获 2017~2018 年度香港理工大学合作研究**资助** (资助编号: G-SB81)
4. **合作/指导经验:** 与研究所内部及外部合作伙伴开展 3 项合作研究并成功产出发表论文。且协助指导多名硕士研究生的科研项目, 和学术论文撰写。
5. **奖项:** 澳门大学中华医药研究院**英语学术**论坛演讲奖项 (2016)
6. **社团服务:** 澳门大学吕志和书院导师及书院纪律委员 (2016 ~ 2017 年)
7. **教学经历** (全英语授课): Research methodologies. Section of Introduction to research in medical administration (2017 秋学期, 课程代码 CMED710)

2011 ~ 2015: 东京大学 研究助理 (于日本东京)

1. 独立开展分子生物学、生物化学、分子肿瘤学实验。(湿实验技能包括但不限于: 基因克隆及定点诱导突变、动物细胞及微生物细胞培养、DNA 测序、蛋白表达与纯化(液相色谱仪操作)、分子体外互相作用测定(分子膜共振技术、Biacore)等)
2. 辅助其他同事进行分子肿瘤学实验研究
3. 实验室管理: 实验室安全性检查、生物化学试剂溶液的制配和库存管理、剧毒危险试剂的管理、实验仪器的维护保养、生物样本库的管理和维护

语言

普通话、粤语 \ **英语(流利)**: 国际商务交流水平及国际学术研讨水平 (**托福 iBT 99 分** & 大学英语六级) \ **日语**: 日本语能力考试 1 级 (N1) \ **德语** (基础)

出版学术论文列表 (SCI 期刊收录)

1. Lin HH, Xu H, Hu H, Ma Z, Zhou J, Liang Q. predicting ovarian/breast cancer pathogenic risks of BRCA1 gene variants of unknown significance. *MedRxiv* 2020.
2. Lin HH. Repositioning anti-cancer agents for neurodegenerative diseases based on drug-disease analysis. University of Macau Ph.D. graduation dissertation thesis 2019.
3. Wang Z, Lin HH*, Linghu K, Huang RY, Li G, Zuo H, Chan G, Hu Y. Novel compound-target interactions prediction for the herbal formula Hua-Yu-Qiang-Shen-Tong-Bi-Fang, *Chemical and Pharmaceutical Bulletin*. 2019. (*Co-first authorship, **Science citation-Indexed**)
4. Yan J, Lin HH*, Zhao D, Hu Y, Shao R. Deterioration of the health quality: a study on healthcare cost control policy and actions in China. *BMC health service research*. 2019, 19:84. DOI:10.1186/s12913-019-3921-8 (*Co-first authorship, **Science citation-Indexed**).
5. Lin HH, Ouyang D, Hu Y. Intelligent Classifier: a Tool to Impel Drug Technology Transfer from Academia to Industry. *Journal of Pharmaceutical Innovation*. 2018:1-7. (**Science citation-Indexed**)
6. Liu K, Lin HH, Pi R, Mak S, Han Y, Hu Y. Research and development of anti-Alzheimer's disease drug: an update from the perspective of technology flows. *Expert Opinion On Therapeutic Patents*. 2018 (**Science citation-Indexed**).
7. Lin HH, Zhang LL, Yan R, Lu JJ, Hu Y. Network Analysis of Drug-target Interactions: A Study on FDA-approved New Molecular Entities Between 2000 to 2015. *Scientific reports*. 2017 Sep 25;7(1):12230.
8. Huang C, Lin HH*, Wan JB, He C, Hu Y. Research and Development of Hepatitis B Drugs: An Analysis Based on Technology Flows Measured by Patent Citations. *PloS one*. 2016 Oct 11;11(10):e0164328. (*Co-first authorship, **Science citation-Indexed**)

出版专著 (书章节)

1. Huang C, Lin HH*, Wan J, He C, Hu Y. Research and Development of Hepatitis B Drugs: An Analysis Based on Technology Flows Measured by Patent Citations. *Innovation Value Chain of Chinese Pharmaceuticals*, United States-China Intellectual Property Institute, 2017, Chapter 8. (*Co-first authorship)
2. Tian S, Lin HH*, Mak MSH, Han Y, Hu Y. Analysis of Technology Flows of Antidepressants Based on Patent Citation Network. *Innovation Value Chain of Chinese Pharmaceuticals*, United States-China Intellectual Property Institute, 2017, Chapter 9. (*Co-first authorship)
3. Li P, Lin HH, Hu Y. Pharmaceutical Innovation Network: Global Patterns and the Role of China. *Innovation Value Chain of Chinese Pharmaceuticals*, United States-China Intellectual Property Institute, 2017, Chapter 13.
4. Lin HH, P Li, Kong XJ, Cai J, Su SB, Hu Y. An analysis of evolutionary conservation based on genomic orthology and protein-protein interaction networks. *Biotechnology in Hong Kong III*. United States-China Intellectual Property Institute, 2017. Chapter 4.