

吕晨恺



研究方向：生物信息学

移动电话： (+86) 18916838506

电子邮件： lvchenkai0812@gmail.com

教育经历

| | | | | |
|--------|--------|--------|----|-----------------|
| 华东师范大学 | 生命科学学院 | 生物医学工程 | 硕士 | 2016.09-2019.06 |
| 华中农业大学 | 信息学院 | 生物信息学 | 学士 | 2012.09-2016.06 |

科研项目

使用深度学习方法整合高危神经母细胞瘤多组学数据进行分型 2018.04-2018.09

- 本研究使用自编码器整合了高危神经母细胞瘤的基因表达和拷贝数变异数据,通过无监督分类器找到两个存在显著预后差异的高危亚型。并对在独立验证集中验证了分型系统的鲁棒性。
- 比较SVM, logistic回归, XGBoost, naïve Bayes等机器学习和深度学习的模型,并不断调整优化模型以达到最好效果
- 撰写英文论文并以共同第一作者完成发表

基因组结构变异的检测以及致病性分析 2017.09-2018.04

- 构建BioSV算法,通过整合全基因组或全外测序数据的分裂读段,不一致读段和读段数目等特征来进行结构变异断点检测,并利用贝叶斯模型进行分型。进一步,通过对点突变数据的系统性整合,利用无向图最大团模型进行特征构建,使用XGBoost法对已经致病性的数据进行训练,构建了准确率召回率均大于0.9的结构变异致病性预测模型。主要完成了特征清洗和提取,以及模型构建步骤。

肝细胞癌的RNA-Seq测序数据分析解读 2016.10-2017.07

- 独立完成测序数据预处理,并分别分析了在癌症中特异表达的miRNA和mRNA,通过已有数据库构建对应调控关系对。并预测新的miRNA靶点,以此构建ceRNA网络,来对未知功能的ncRNA进行注释。并结合血液和肝组织的测序数据来寻找潜在的能用来做液体活检的循环miRNA。

论文成果

- Zhang, L., Lv, C., Jin, Y., Cheng, G., Fu, Y., Yuan, D., Tao, Yi., et al. (2018) Deep learningbased multi-omics data integration reveals two prognostic subtypes in high-risk neuroblastoma. Frontiers in Genetics.(PMID:30405689)
- (Under Revision) Zhang, L., Ding, W., Lv, C., Liu, W., Shi, Y., Hou, C., Christopher E. Mason, et al. (2018) BioSV: an accurate structural variation caller and genotyper with high resolution. Nucleic acids research

掌握技能

- 熟练掌握R, Python ,Shell, Awk, Perl, Matlab等编程语言
- 熟悉各种二代测序数据分析流程,掌握并能熟练运用多种机器学习或深度学习模型以及各种统计学方法
- 熟悉基于PHP或RShiny的网站构建,以及基于SQL的数据库操作
- 熟悉服务器管理以及平台维护工作

实习经历

罗氏全球药品开发中心 数据分析实习生 2018.09-2018.12

- 构建机器学习模型辅助临床诊断病例选择
- 文本挖掘临床试验报告信息,分析常见的不良事件
- 构建webserver辅助临床实验数据进行分析 and 可视化

上海交复科研 服务器管理和平台构建 2018.07-2018.08

- 构建ChIP-Seq, ATAC-Seq, RNA-Seq等测序分析流程的平台,并完成自动化

兴趣特长

- 数学建模 (曾获“华中杯”数学建模大赛一等奖; “高教社杯”数模竞赛湖北省二等奖等)
- 跆拳道 (黑带一段)
- 太极拳