

测序与芯片高通量数据挖掘与分析学习班

生物信息学 (bioinformatics) 定义: 是在大分子方面的概念型的生物学, 并且使用了信息学的技术, 这包括了从应用数学、计算机科学以及统计学等学科衍生而来各种方法, 并以此在大尺度上来理解和组织与生物大分子相关的信息。

培训目的:

很多老师都做过测序或者芯片实验, 虽然测序公司已经给出了一部分的数据分析, 比如过滤、质控、比对、定量和差异等, 但是这些分析属于基础傻瓜式的标准分析, 并不能完整的讲一个故事! 因此, 需要我们自己掌握高通量数据挖掘的能力, 从海量信息中获得自己想要的**关键基因!**

培训预期: 通过 2 天的培训, 使学员掌握生信文献分析思路和 8 个实用操作知识点 (具体见最后), 可以独立完成一篇基于公共数据库的高通量数据挖掘分析。

生物信息学魅力:

帮助预测疾病相关的潜在基因, 以及该基因潜在的作用靶点、上游调控转录因子等, 从而指导实验方向、缩小试验范围、简化试验流程。

为基金申请提供支持, 通过强大的信息数据的收集整理, 减少投入增强研究目的性; 且通过整合技术优势, 指导提高临床诊断水平。

丰富的免费的数据资源和生物软件工具: 数据库: NCBI, EBI, DDBJ, PDB, SWISS-PROT 等; 软件工具: Cytoscape, BLAST, R 语言, String, David 等

案例:

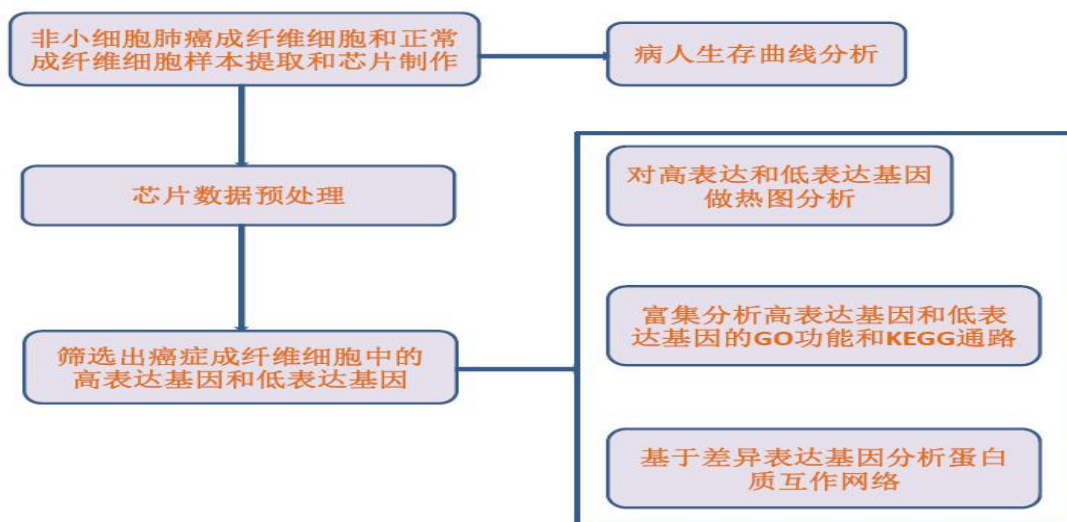
PNAS

Prognostic gene-expression signature of carcinoma-associated fibroblasts in non-small cell lung cancer

Roya Navab^{a,1}, Dan Strumpf^{a,1}, Bizhan Bandarchi^{a,1}, Chang-Qi Zhu^{a,1}, Melania Pintilie^a, Varune Rohan Ramnarine^a, Emin Ibrahimov^a, Nikolina Radulovich^a, Lisa Leung^a, Malgorzata Barczyk^{a,b}, Devang Panchal^a, Christine To^a, James J. Yun^a, Sandy Der^a, Frances A. Shepherd^{a,c}, Igor Jurisica^{a,d,e}, and Ming-Sound Tsao^{a,e,f,2}

^aThe Campbell Family Institute for Cancer Research, Ontario Cancer Institute at Princess Margaret Hospital, University Health Network, Toronto, ON, Canada M5G 2M9; ^bDepartment of Biomedicine, Jonas Lies vei, N-5009 Bergen, Norway; Departments of ^cMedicine, ^dComputer Science, and ^eMedical Biophysics, and ^fLaboratory of Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada M5A 2N4

该文献的分析流程如下图所示:



我们可以看到该文献主要的分析步骤有：差异表达分析、聚类热图构建、GO 和 KEGG 通路分析、蛋白互作网络分析、生存曲线分析等。我们的主题基本上按照类似的文献思路进行讲解，和大家一起分享下一篇 SCI 文章的生信分析步骤。

通过 2 天的学习，我们可以掌握一篇生信 SCI 文献的所有分析模块，并可以独立的结合自己的研究方向进行生信分析和数据挖掘，获得一套属于自己的分析结果。

第一天上午第一节：基因数据库使用介绍

测序标准报告的解读。

NCBI Gene 数据库：最权威的基因综合数据库之一。

NCBI GEO 数据库：数据量最大的高通量公共数据库。

GEO2R 工具：在线做差异分析的 R 工具

bioDBnet：在线基因 ID 转换工具。

Uniprot：最权威的蛋白综合数据库之一。

第一天上午第二节：DAVID 工具——功能富集分析

功能富集的原理和意义

基因 ID 转换

基因功能注释（GO，KEGG pathway）

基因功能富集（GO，KEGG pathway）

基因功能聚类（GO，KEGG pathway）

第一天下午第三节 聚类热图及 pathway map 图

KEGG 数据库：最权威的通路数据库（pathway map 分析）

HEMI 软件：构建聚类热图的本地软件

第一天下午第四节. STRING 工具——基因/蛋白相互作用分析（PPI）

蛋白互作分析的原理和意义；

单个蛋白的蛋白互作分析；

多个蛋白的蛋白互作分析；

STRING 工具的使用步骤：输入、参数设置、输出等。

第二天上午第五节 cytoscape 网络做图软件（上）

网络图构建

逐个修改节点的属性和修改边的属性

网络展示方式 layout 更美观

网络统计分析

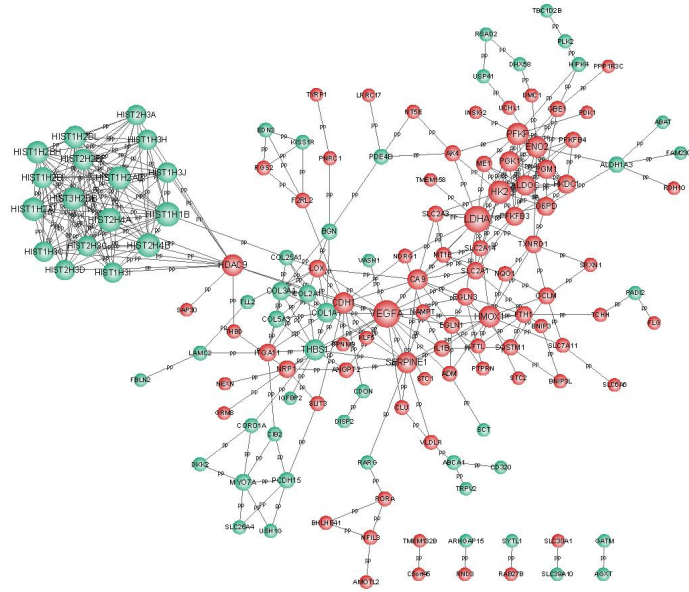
第二天下午第六节 cytoscape 网络做图软件（下）

批量修改节点的属性和修改边的属性

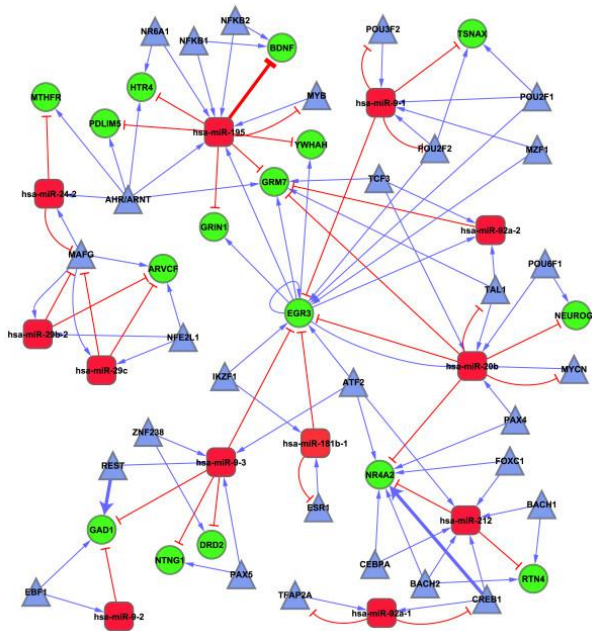
Cytoscape 的常用插件：ClusterONE

Cytoscape 的常用插件：BinGO

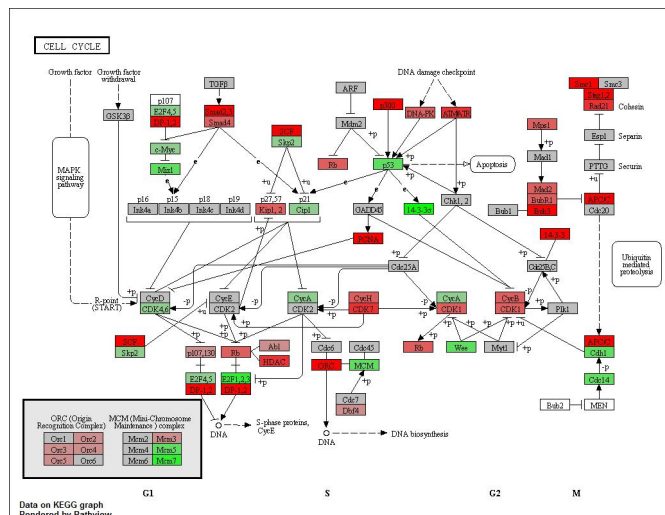
第二天下午第七节 答疑环节



示例图 蛋白互作网络图



示例图 调控网络图



示例图 pathway map 图

讲师宋博士简介：

成果：参与完成了近百篇软件著作权和发明专利的撰写和申请；肺癌、胰腺癌、骨肉瘤、胃癌等数据库的分析和构建；完成个体基因检测流程和无创唐筛流程的开发。

研究方向：有近十年的生信分析经验，擅长方向有转录组测序分析、芯片数据分析、疾病机理研究分析、疾病预后与基因关联分析、项目分析思路设计以及个性化分析等，精通 perl、R 等编程语言。

培训经历：在上海、沈阳、济南、武汉等城市举办过十几场培训班。培训的对象有：医生、学生、科研工作者、生信爱好者等。

培训方向：

《测序与芯片数据分析》、《生物信息学的魅力》、《生信文章实例解读》、《生信与实验的密切关系》、《生信与临床医学的关系》、《生信实用工具培训》、《多组学整合分析流程》、《R 语言培训》等



学习班时间与地点（上海班）

培训时间：2018 年 5 月 26-27 日（2 天）

培训地点：上海好望角大饭店 上海市肇嘉浜路 500 号

收费标准：

注册费：2500 元每位（注册费包含电子版教材、午餐，住宿费自理。）

优惠政策：

1. **提前**确认报名及转账的，可以提前拿到学习材料
 2. 三人组团报名，每人可优惠 100 元
 3. 四人组团报名，每人收费 2300 元，
 4. 五人组团报名缴费，额外带一人免费注册！
- 可以开正规会务发票，纸质邀请函（盖红章）。

学习班报名方式:

报名电话: 17602127353 张老师

邮箱: kf@yanshudata.com, 报名表提交后若 24h 无任何回复, 请及时电联。

报名方式:

1. 在线报名链接: <https://www.wenjuan.in/s/UVfUFb/>
2. 或手机微信扫描下方二维码报名:

