

2021.7.19-方法记录

一、对获取到的基因表达数据进行生物学理解/分析流程

(以下主要为用于微阵列和RNA-Seq数据的可视化/生物学解释流程与方法)

1.热图和聚类/Heatmaps and clustering

可视化基因表达数据的常用方法：将其显示为热图。

在与聚类方法结合的热图中，基因可被按照它们的表达模式进行分组

红色——表达上调的基因；蓝色——表达下调的基因；黑色——表达不变的基因。

热图还可以与聚类方法相结合，根据基因表达模式的相似性将基因和/或样本进行分组。

这可以用于识别常规基因，或与特定条件(如疾病或环境条件)相关的生物标记。

2.基因集富集分析与途径分析/Gene set enrichment analysis and pathway analysis

解释基因表达数据的常用方法是基于差异表达基因的功能注释的基因集富集分析，这有助于发现差异表达的基因是否与某种生物学过程或分子功能有关。

用于基因组富集和途径分析的流行工具包括：

- [DAVID](#) (free online tool)
- [GSEA](#) (free)
- [Ingenuity](#) (licence required)
- [Reactome](#) (free)
- WikiPathways
- KEGG (Gene Ontology)

3.网络分析/Network analysis

网络分析是路径分析的补充，可以用来显示不同路径的关键成分是如何相互作用的。这对于识别影响多种生物过程和途径的调控事件是有用的。

二、基因分型、表观遗传及 DNA/RNA 蛋白相互作用方法/Genotyping, epigenetic and DNA/RNA-protein interaction methods

基因分型/Genotyping

DNA 序列的变异 (Variations in the DNA sequence) , 例如拷贝数变异 (copy number variations) 、插入 (insertions) 、缺失 (deletions) 和 SNPs 被称为样本的基因型, 可以用以下方法检测:

1.比较基因组杂交/Comparative genome hybridisation

样本和一个良好特征的参考样本都与同一个微阵列杂交, 并且两个样本之间的相对强度用于推断拷贝数变化

2.全基因组测序/Whole genome sequencing

测试样本的基因组测序和计算比较, 以确定许多类型的遗传变异

3.外显子组测序/Exome sequencing

在溶液中或固定在微阵列上使用序列特定的探针纯化 DNA 编码区。然后对这些区域进行放大、测序和比较, 以确定编码区域的变异

表观遗传修饰/Epigenetic modifications

表观遗传修饰是染色质中可遗传的化学或物理变化。有两种类型的表观遗传修饰: DNA 甲基化和组蛋白修饰。

DNA甲基化/DNA methylation

DNA 甲基化可以使用基于染色质免疫沉淀或亚硫酸氢盐的方法来测量。

组蛋白修饰/Histone modifications

组蛋白的组蛋白修饰有多种类型, 包括组蛋白中特定氨基酸的乙酰化、甲基化、泛素化、瓜氨酸化和磷酸化, 通常朝向蛋白的 c 末端('tail')。通过改变组蛋白与 DNA 的结合方式, 这些修饰可以正向和负向地调节基因表达。

组蛋白修饰可以通过各种技术进行检测, 包括质谱法和基因组学方法, 如 ChIP-chip 和 ChIP-seq。基因组学方法将特定的、修饰过的组蛋白及其相关 DNA 的染色质免疫沉淀与 DNA 分子的微阵列(芯片)或 NGS (seq)结合起来, 以确定与这些修饰相关的基因组区域

蛋白质相互作用/DNA/RNA-protein interactions

DNA 与蛋白质的相互作用包括 DNA 与转录因子或其他调节蛋白质之间的相互作用。RNA 与蛋白质之间的相互作用包括 RNA 与核糖体之间的相互作用，以及其他 RNA 结合蛋白质。

1.核酸-蛋白质复合体的免疫沉淀法/immunoprecipitation of the nucleic acid-protein complex

用基因芯片或 NGS 技术对 DNA/RNA 进行纯化和分析，以确定基因组或转录组中在特定时间点与特定蛋白质相互作用的区域/purification and analysis of the DNA/RNA by microarray or NGS to identify the regions of the genome or transcriptome that are interacting with a specific protein at a specific point in time

蛋白质分析被称为 ChIP-chip 或 ChIP-seq，需要在免疫沉淀法之前用甲醛交联 DNA 蛋白质。对于 RNA 蛋白质分析，交联步骤是可选的。RNA 免疫沉淀(RIP)-芯片/序列分析可以在使用或不使用以甲醛为基础的交联免疫沉淀法前进行。由于没有交联，这种方法不适合于检测瞬时 RNA 与蛋白质的相互作用，而且容易出现高背景值。交联和免疫沉淀法(CLIP)-chip/seq 分析使用 RNA 和蛋白质的紫外交联在免疫沉淀法之前，从而克服了基于 RIP 的方法的许多局限性