

第二次作业

崔晏菲 自硕 21 2021210976

一、这三个指标都试图对测序深度和基因长度进行标准化。

RPKM:

1. 计算样本中的总 reads 数，并将该数字除以 1,000,000，得到了“每百万”缩放比例。
2. 将读取计数除以“每百万”缩放比例。这会规范化测序深度，使 reads 数达到百万分之一（RPM）
3. 将 RPM 值除以基因长度（以千碱基为单位）。就得到了 RPKM。

FPKM 和 RPKM 定义相同，惟一的区别是 FPKM 适用于双端测序文库，而 RPKM 适用于单端测序文库。FPKM 将会配对比对到一个 fragment 上的两个 reads 计算一次，之后和 RPKM 的计算过程一样。

TPM:

1. 将 reads 数计数除以每个基因的长度（以千碱基为单位）。这样就可以得到每千个基数（RPK）的 reads 数。
2. 计算样本中所有 RPK 值，然后将其除以 1,000,000。得到了“每百万”缩放比例。
3. 将 RPK 值除以“每百万”比例因子。就得到 TPM。

每个样本的 TPM 的总和是相同的，因此 TPM 数值能体现出比对中某个基因的 reads 的比例，因此 TPM 可以直接进行样本间的比较。

二、

精准医学的定义:精准医学是应用现代遗传技术、分子影像技术、生物信息技术，结合患者生活环境和临床数据，实现精准的疾病分类及诊断，制定具有个性化的疾病预防和治疗方案。

中国实施精准医学计划的战略意义总共有 4 点：提高疾病诊治水平，惠及民生与国民健康;推动医学科技前沿发展，增强国际竞争力;发展医药生物技术，促进医疗体制改革;形成经济新增长点，带动大健康产业发展。

其指导思想是：贯彻创新驱动发展战略，面向我国重大疾病防治和人口健康保障需求，与深化医疗卫生体系改革紧密结合，与发展生物医药和健康服务等新兴产业紧密结合，发挥举国体制优势和市场配置资源决定性作用，通过政府推动、科技支撑和体系建立，提升自主创新能力，形成引领世界的精准医学发展的有效力量和途径。

新一代测序技术（**NGS**）在精准医疗中的地位不可或缺，可谓是精准医疗的“先行者”。最近，在《**Genome Medicine**》杂志上发表的一项研究报告显示，大多数在医学上具有重要意义的基因，都位于人类基因组中比较复杂、难以分析的区域，目前难以获得准确的序列信息。通常，这些医学上重要的基因中有五分之一，用今天的基因测序方法来进行破译，是具有挑战性的，这些基因区域中的信息，可能对于患者的诊断或治疗计划至关重要。

例如儿童肿瘤专家不仅将 **NGS** 应用于日常工作，并进一步借助此工具推动儿童肿瘤精准诊疗的研究——美国大溪城 **Helen Devos** 儿童医院血液肿瘤科的 **Giselle Sholler** 博士利用脉络丛瘤（一种罕见的侵袭性儿童脑癌）临床试验的 **NGS** 数据，来对脑癌患儿进行疗法筛选。

三、

无创产前基因检测是利用新一代高通量测序检测胎儿染色体异常的新一代产前检测技术。通过采集孕妇外周血，对母体外周血血浆中的游离 DNA 片段（包括胎儿游离 DNA）进行测序，结合生物信息分析，计算胎儿患染色体非整倍体的风险。此技术能同时检测 21-三体、18 三体及 13-三体，还可发现其他染色体非整倍体及染色体缺失/重复。

以 21 号染色体为例，解述原理如下，其他染色体以此类推。

- 1) 假设每毫升母体外周血中的染色体有 1000 份基因组当量，母体的染色体所占比例为 900 份，胎儿的染色体所占比例为 100 份。
- 2) 对于怀有正常胎儿的孕妇，胎儿的 100 份中含有 200 条 21 号染色体单体，母亲的 900 份中含有 1800 条 21 号染色体单体，两者相加，共有 21 号染色体单体 2000 条。
- 3) 对于怀有 21-三体胎儿的孕妇，由于胎儿的 21 号染色体多出一条，则胎儿的 100 份基因组当量中含有 300 条 21 号染色体单体，母亲的 900 份基因组当量中含有 1800 条 21 号染色体单体，两者相加，共有 21 号染色体单体 2100 条。
- 4) 根据数理统计的原理和基因测序的结果，可计算出孕妇怀有 21-三体综合征患儿的危险指数。