

Illumina边合成边测序原理

陈明杰 202410

Illumina测序仪

Platform

ABI SOLiD (20,157)

BGISEQ (31,297)

Capillary (47,460)

Complete Genomics (6,096)

Helicos (2,532)

Illumina (10,946,251)

Ion Torrent (211,634)

LS454 (110,307)

Oxford Nanopore (302,604)

PacBio SMRT (611,404)

SRA数据库human
不同平台检索结果



NovaSeq



HiSeq X Ten



NextSeq 550AR
(国械标准: 20173400330)



HiSeq 2500

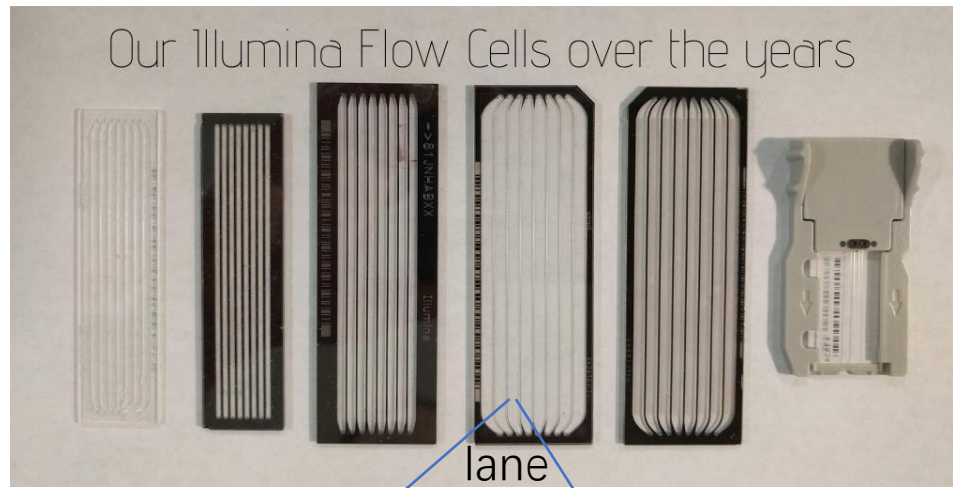


HiSeq 4000



MiSeq

提问: 报告中的测序仪是什么型号



flowcell
8 lane

22CLJHLT4:3

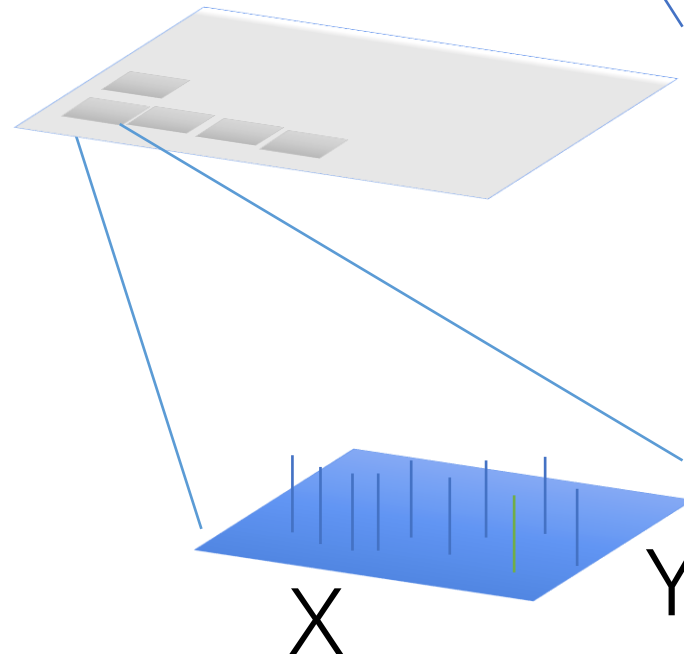
lane
3

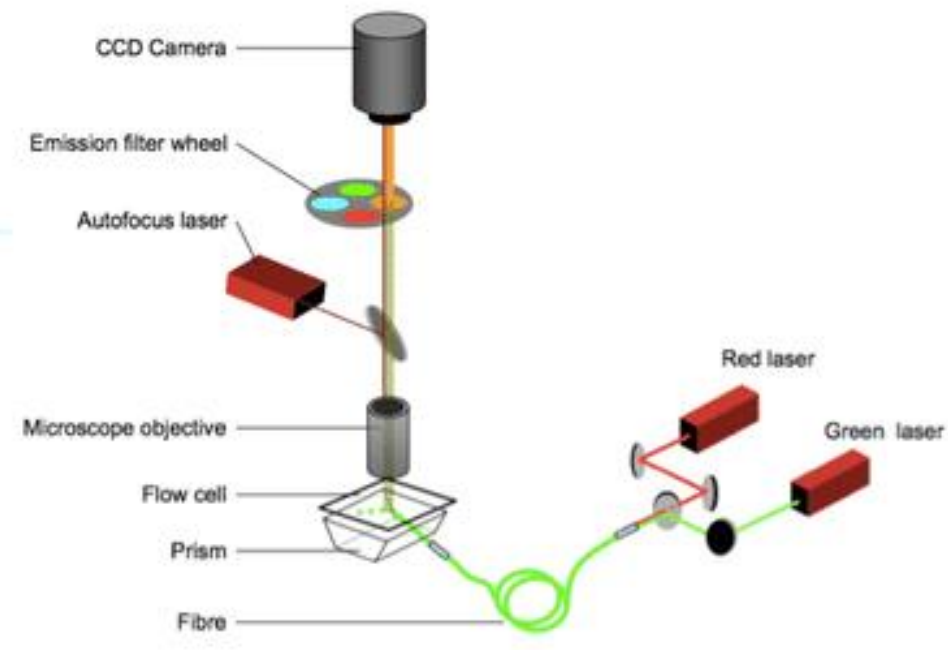
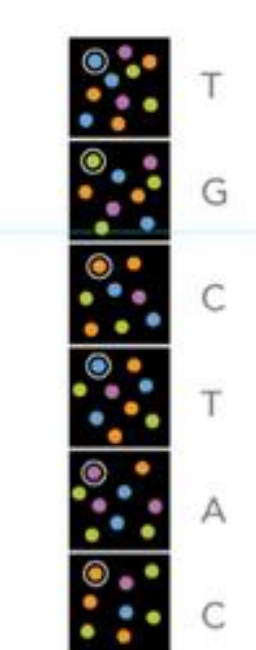
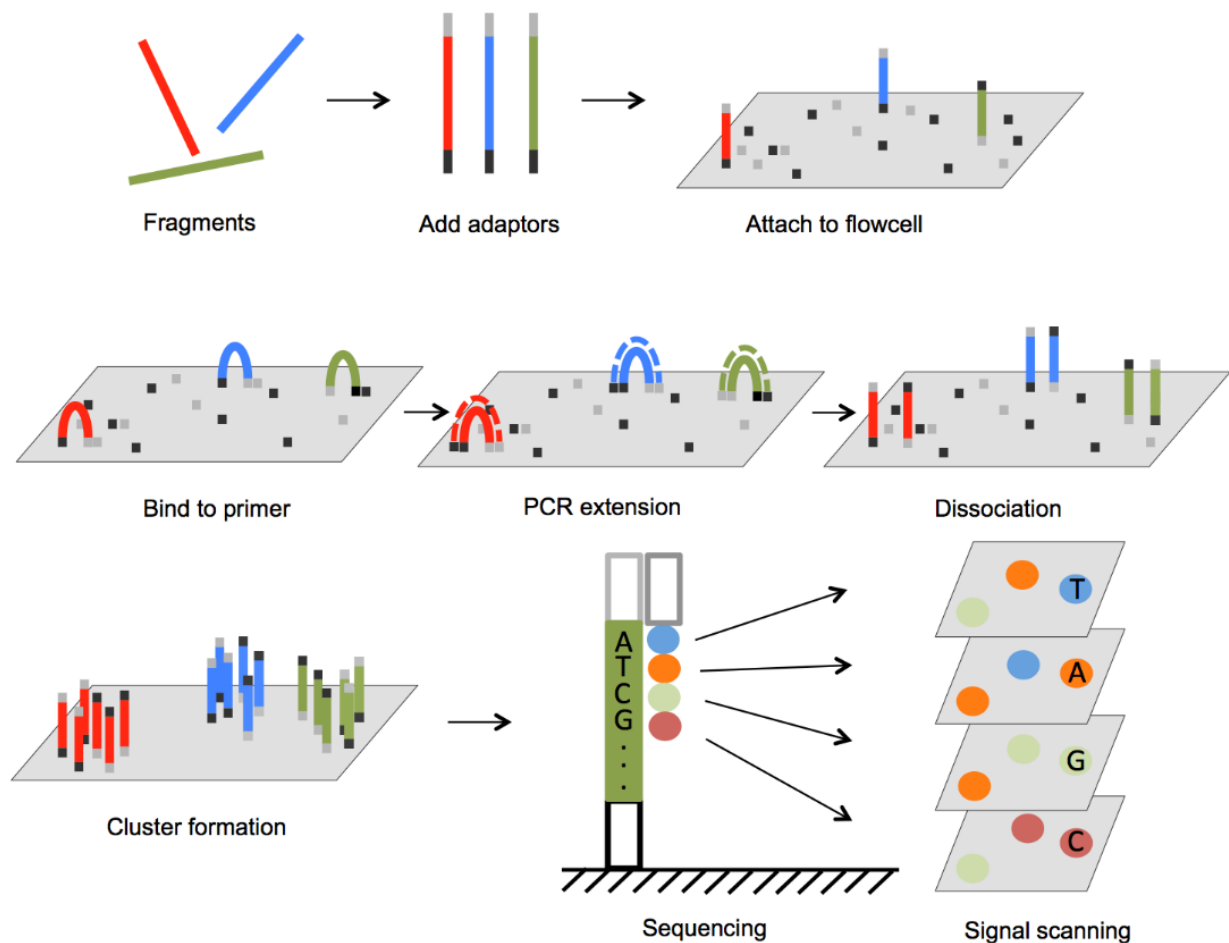
Tiles

tiles是测序荧光扫描的最小单位
1101

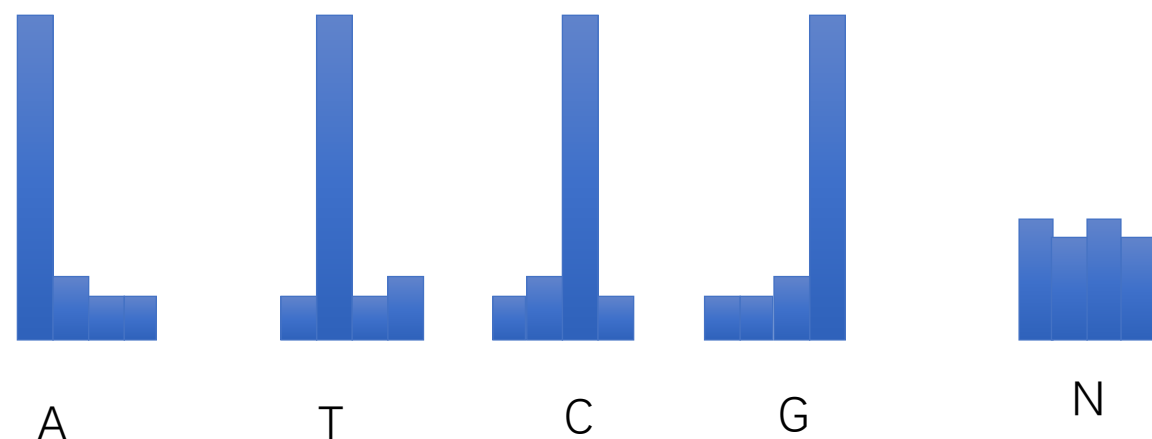
clusters

2400:1042



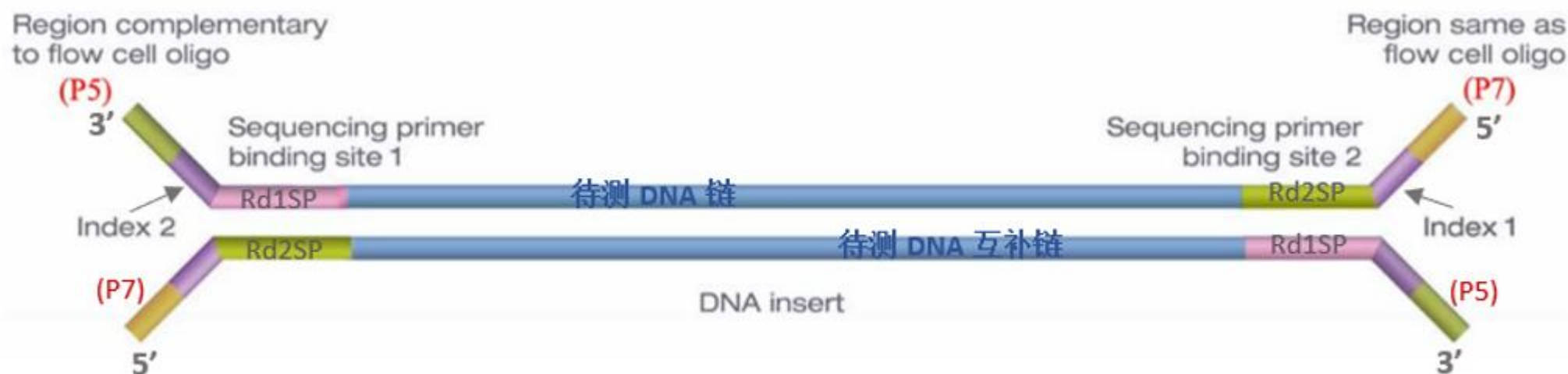


测序	芯片
“150”张图	1张图
单碱基	探针集
3D	2D



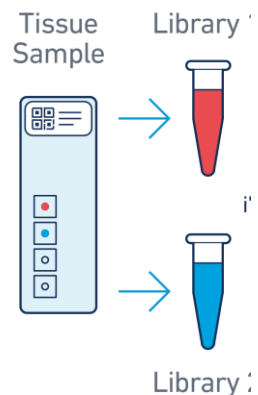
视频

- <https://www.bilibili.com/video/BV1W44y1373N/>



- P5和P7序列：允许文库结合且能在flow cell上生成簇的序列
- Rd1 SP和Rd2 SP：启动测序的测序引物结合位点
- Index 1和Index 2：用于样本区分的index序列，可允许单次测序或单个Flow Cell通道中混合多个样本

Index-区分文库, 凑样/拆样



@LH00348:184:22CLJHLT4:3:1101:2400:1042 1:N:0:GCGGTATT+CCAATCAC

@LH00348:184:22CLJHLT4:3:1101:2400:1042 1:N:0:ACGGGTAC+GGATGCCT

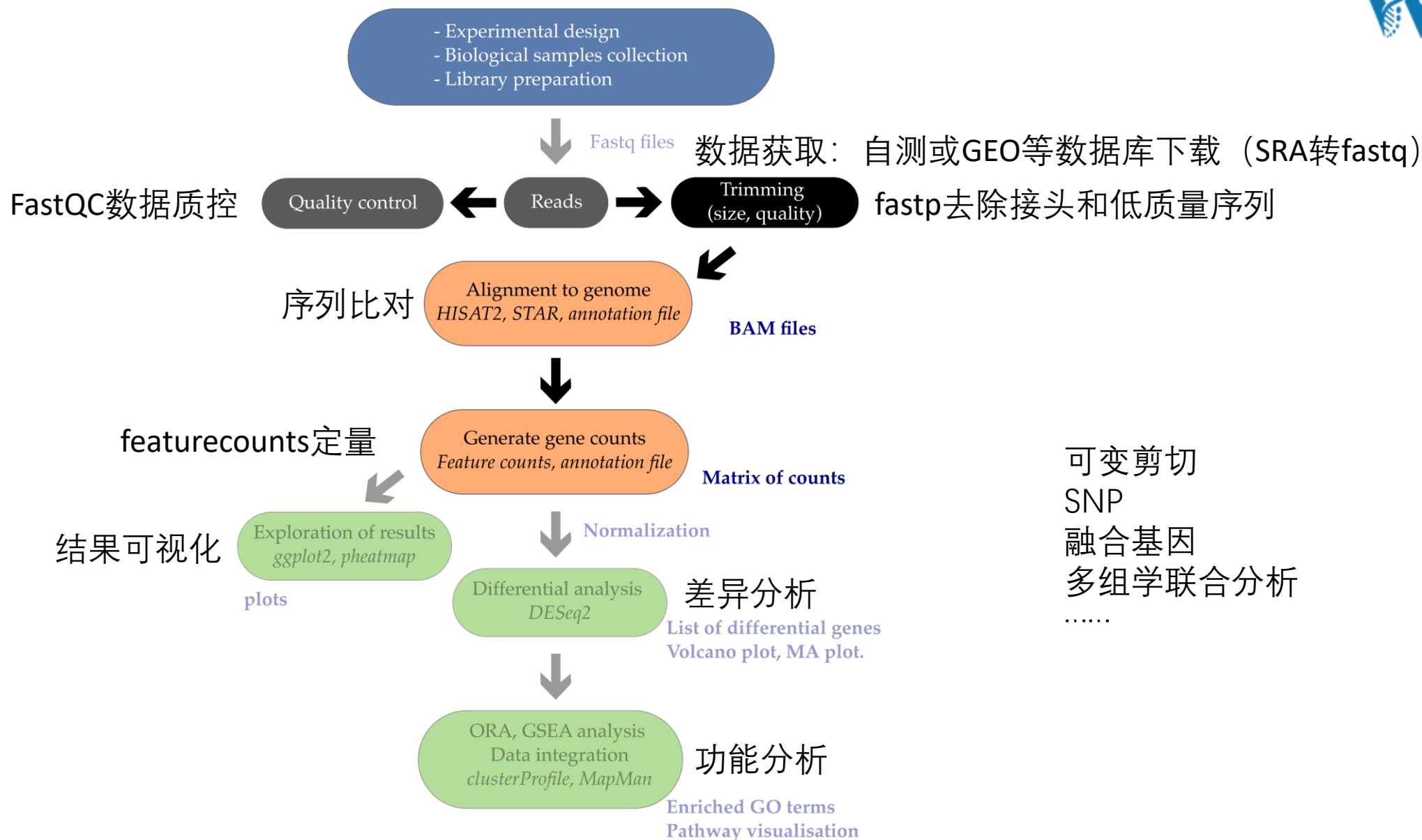
一条lane 产出120G, 测20个mRNA-seq样品

张三: s1,s2,s3, t1,t2,t3

李四: m1,m2,m3, n1,n2,n3

王五: i1,i2,i3,i4, j1,j2,j3,j4

注意: index冲突
碱基平衡, 不平衡文库



准备文件

- 原始数据 (3 vs 3) fastq.gz
- 基因组 (fa) hisat2 (STAR) 建索引
- 注释 (gtf)
- fasta、注释文件预处理