

# 碧云天高通量测序服务介绍



碧云天网站



微信公众号

碧云天生物技术/Beyotime Biotechnology  
订购热线：400-168-3301或800-8283301  
技术咨询：info@beyotime.com  
高通量测序服务：ngs@beyotime.com  
网址：<http://www.beyotime.com>

# 高通量测序技术

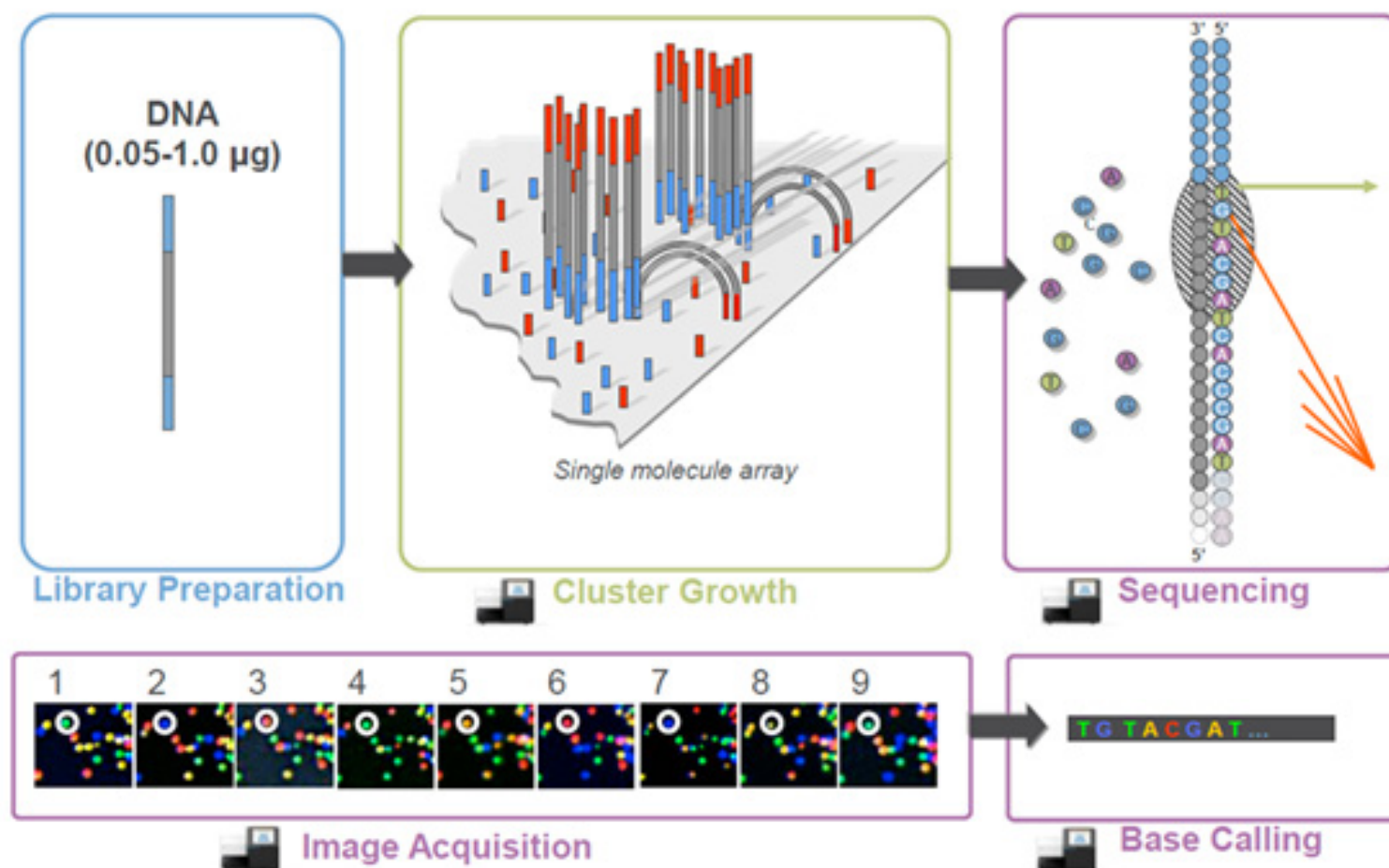
## 概况

---

- 高通量测序技术(High-throughput sequencing, HTS)又称下一代测序技术(Next-generation sequencing, NGS)，相对于传统的一代测序技术桑格测序(Sanger sequencing)而言，高通量测序技术是一次能对几十万到几百万条核酸分子进行序列测定的技术。同时高通量测序使得对一个物种的转录组和基因组进行细致全貌的分析成为可能，所以又被称为深度测序(Deep sequencing)。
- 高通量测序技术大约在2006-2007年左右进入中国市场，当时有454、Illumina、ABI Solid等品牌，后来由于技术优势，Illumina公司逐渐占据上风，独领市场。目前，Illumina公司的HiSeq X 10和Nova-Seq等机型是市场的主流。另外ABI(现属Thermo Fisher)后来收购Ion Torrent的测序技术由于测序周期较短，也占一定市场，但由于相对通量较低，成本比较高。近几年技术日趋成熟的还有第三代单分子测序技术，其特点是测序前不用进行PCR扩增，并且读长较长，目前可达9kb以上，其代表产商有Pacbio公司。

# 高通量测序技术概况

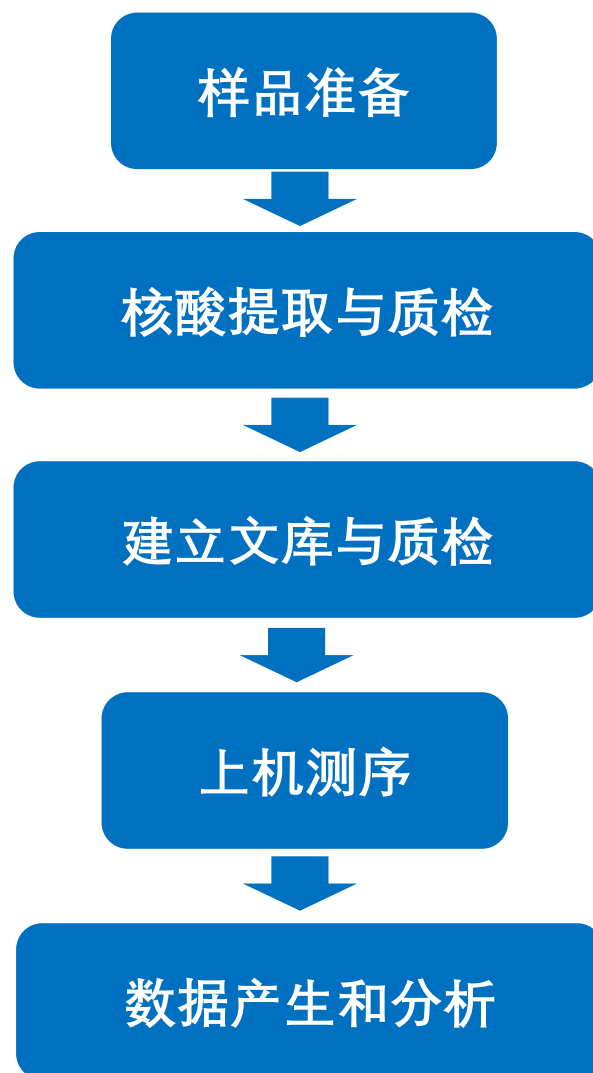
## 高通量测序主要步骤



# 高通量测序技术概况

## 技术流程

---



# 全基因组重测序

*Whole Genome Sequencing, WGS*

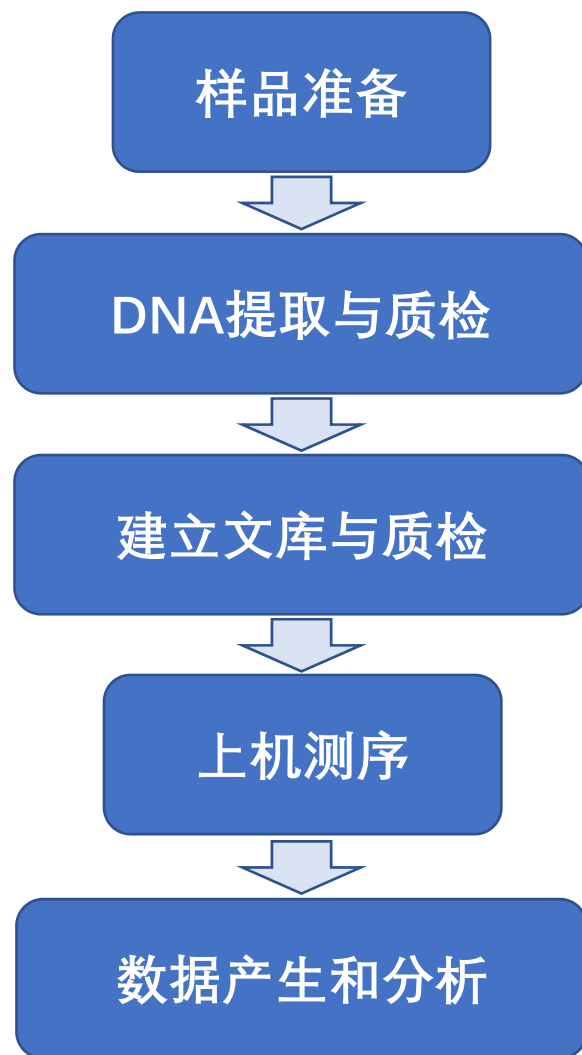
---

- 全基因组重测序是对已知基因组序列的物种进行不同个体的基因组测序，并在此基础上对个体或群体进行差异性分析。
- 它可以获取最全的基因组信息，找到大量的单核苷酸多态性位点 (Single Nucleotide Polymorphism, SNP)、插入缺失位点 (Insertion/Deletion, InDel)、杂合性缺失 (Loss of Heterozygosity, LOH)、拷贝数变异 (Copy Number Variation, CNV) 以及基因组重排导致的结构变异位点 (Structure Variation, SV)。
- 配合比较基因组学分析、群体遗传学分析、进化分析和计算生物学分析方法既可以用于深入探索疾病基因组的奥秘，也可以高效识别动植物的性状连锁区域，为遗传育种提供科学依据和技术支撑。

# 全基因组重测序

## 技术流程

---



# 全基因组重测序

## 技术参数

---

- 测序平台与方式：HiSeq, PE150
- 测序深度：30-50×或更高
- 周期：45个工作日

# 全基因组重测序

## 样本要求

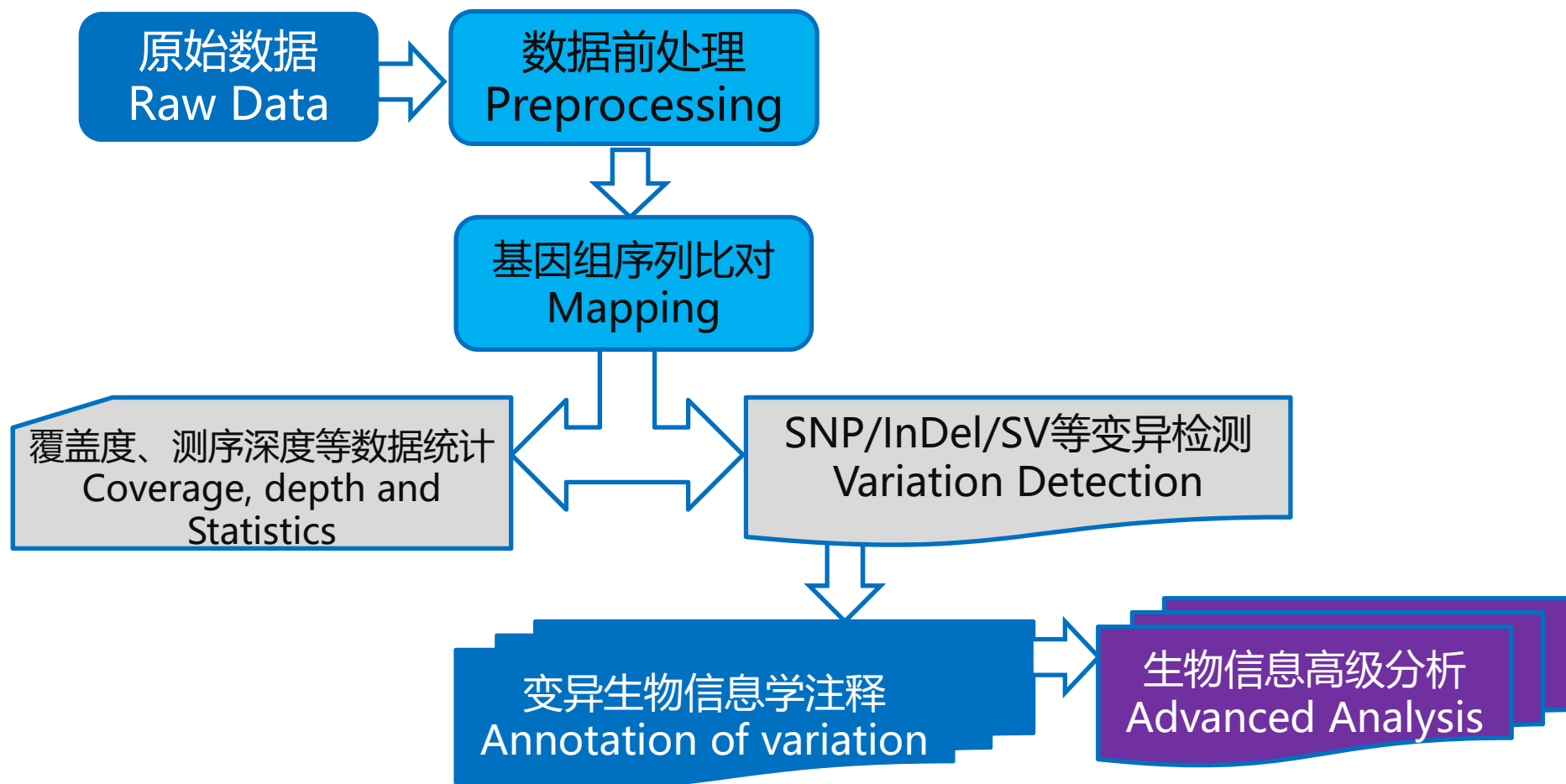
---

- 1. 样品纯度：OD 260/280值应在1.8 ~ 2.0 之间；RNA 应该去除干净。
- 2. 样品浓度： $\geq 50\text{ng}/\mu\text{l}$ 。
- 3. 样品总量：每个样品总量 $\geq 0.5 \mu\text{g}$ 。
- 4. 样品溶剂：要溶解在超纯水或Low-EDTA TE (10 mM Tris-HCl, 0.1mM EDTA, pH 8.0)中。
- 5. DNA完整纯净，无降解，无细胞凋亡，无RNA、蛋白质、多糖、试剂成分等残留或污染
- 6. 样品运输：DNA低温运输( $-20^{\circ}\text{C}$ )；且在运输过程中请用parafilm将管口密封好，以防出现污染；收到样品后，甲方需要对样品进行检测，最终样品的量和纯度，以甲方的检测结果为准。



# 全基因组重测序

## 数据分析流程概况



# 全基因组重测序

## 数据分析内容

### 基本信息分析

1. 数据质控：去除接头污染和低质量数据
2. 与参考序列进行比对、统计测序深度及覆盖度
3. SNP/InDel/CNV/SV检测、注释及统计

### 高级信息分析

肿瘤	单基因病	复杂疾病
<ol style="list-style-type: none"><li>1. 已知的癌症基因的突变筛选</li><li>2. 高频突变基因总结</li><li>3. 突变位点分布情况分析</li><li>4. 基因组变异Circos 图展示</li><li>5. MRT 高频突变基因相关性分析</li><li>6. 高频CNV 分析</li><li>7. 肿瘤纯度及倍性分析</li><li>8. LOH 在基因组中的分布</li><li>9. 瘤内异质性及克隆结构分析</li><li>10. 融合基因及高频融合基因</li><li>11. 非编码区高频突变分析</li><li>12. 病毒整合分析</li><li>13. 其他个性化内容分析</li></ol>	<ol style="list-style-type: none"><li>1. 突变位点筛选</li><li>2. 显/隐性遗传模式筛选(纯合子或杂合子的筛选)</li><li>3. 候选基因功能注释</li><li>4. 候选基因功能富集分析 (GO , KEGG等)</li><li>5. 纯合子区域分析(连续纯合子变异区段)</li></ol>	<ol style="list-style-type: none"><li>1. 综合各种资源列举已知候选基因</li><li>2. 候选基因中突变的筛选和功能预测</li><li>3. 候选基因功能注释</li><li>4. 新生突变筛选和分析</li><li>5. 候选基因功能富集分析 (GO , KEGG等)</li><li>6. 蛋白互作网络分析(PPI)</li><li>7. 结合实验设计(家系或自然群体)进行遗传学分析(连锁关联分析)</li><li>8. 变异在全基因组范围内的共发生/互斥现象分析</li><li>9. 其他个性化分析</li></ol>

# 全基因组重测序

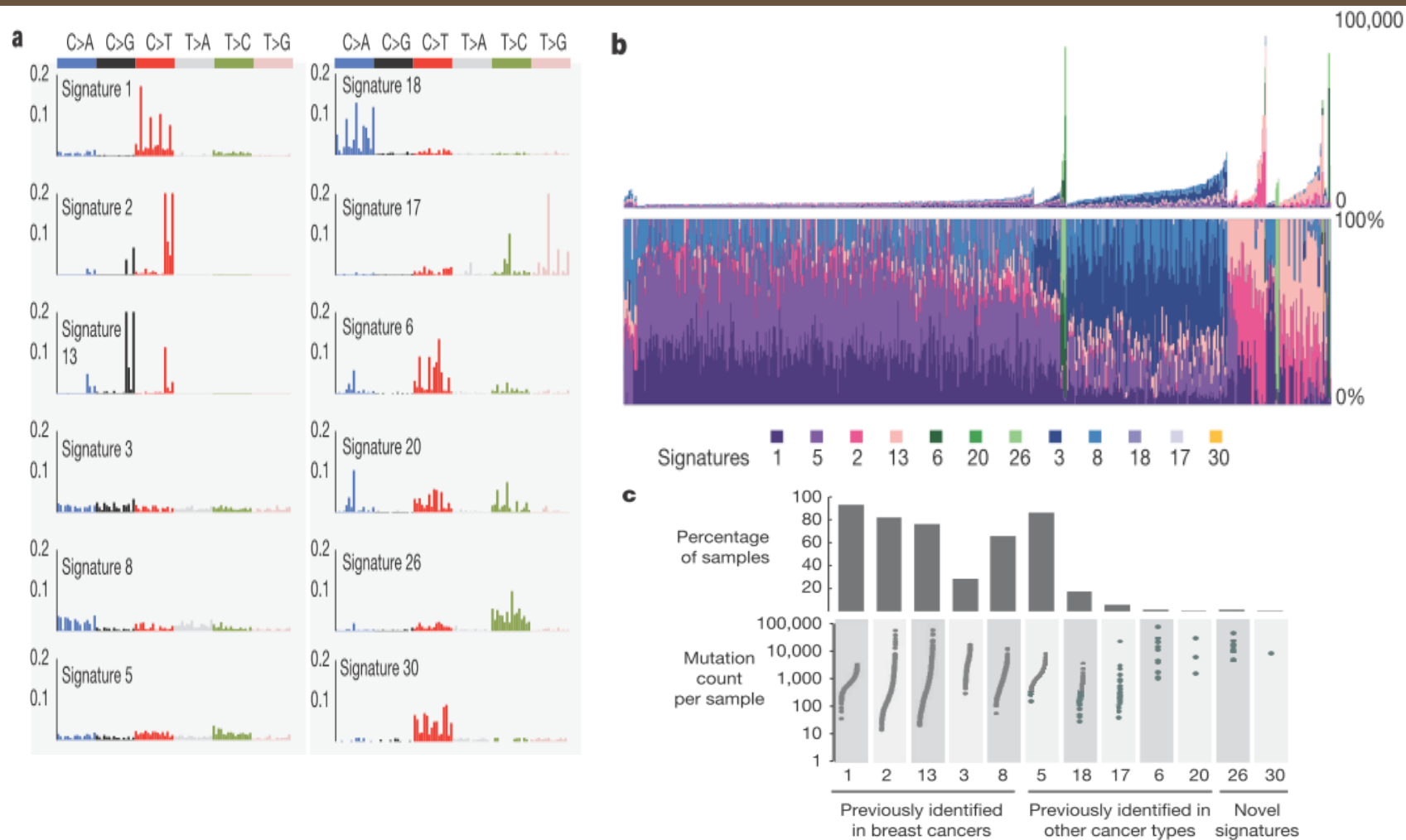
## 应用案例

---

- 全基因组测序描绘乳腺癌突变谱
- *Nik-Zainal, S., et al., Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature, 2016. 534(7605): p. 47-54.*

# 全基因组重测序

应用案例：560个乳腺癌样品中突变模式



# 全外显子测序

*Whole Exome Sequencing, WES*

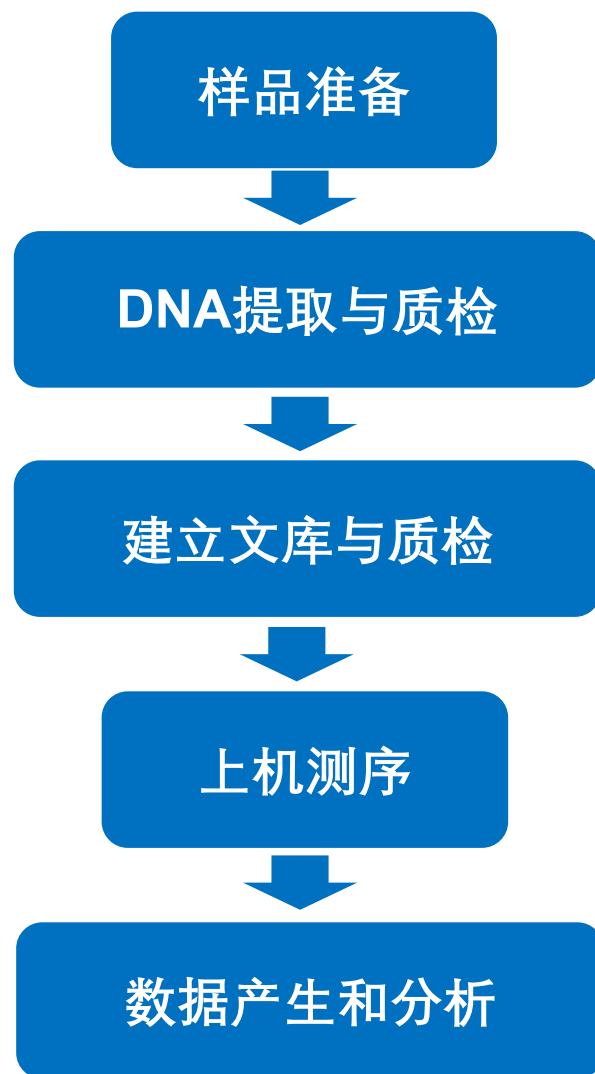
---

- 全外显子组测序，是利用目标序列捕获技术，将全基因组编码基因外显子区域的DNA捕获并富集后，进行高通量测序的基因组分析方法。
- 在具体应用方面和全基因组重测序类似，通过测序可以找到大量的单核苷酸多态性位点(Single Nucleotide Polymorphism, SNP)，插入缺失位点(Insertion/Deletion, InDel)、杂合性缺失(Loss of Heterozygosity, LOH)、拷贝数变异(Copy Number Variation, CNV)以及基因组重排导致的结构变异位点(Structure Variation, SV)。
- 配合比较基因组学分析、群体遗传学分析、进化分析和计算生物学分析方法既可以用于深入探索疾病基因组的奥秘。但在同样的测序通量下，外显子组测序针对目标区段的测序深度以及覆盖度都高于全基因组重测序，因此该技术对于发生在外显子区的常见和罕见的染色体变异具有很高的检测灵敏度。其简便、经济、高效等特点，使其成为近年的热门技术。

# 全外显子测序

## 技术流程

---



# 全外显子测序

## 技术参数

---

- 测序平台与方式：HiSeq, PE150
- 捕获平台：Agilent SureSelect
- 测序深度：100-200×或更高
- 周期：35个工作日

# 全外显子测序

## 样品要求

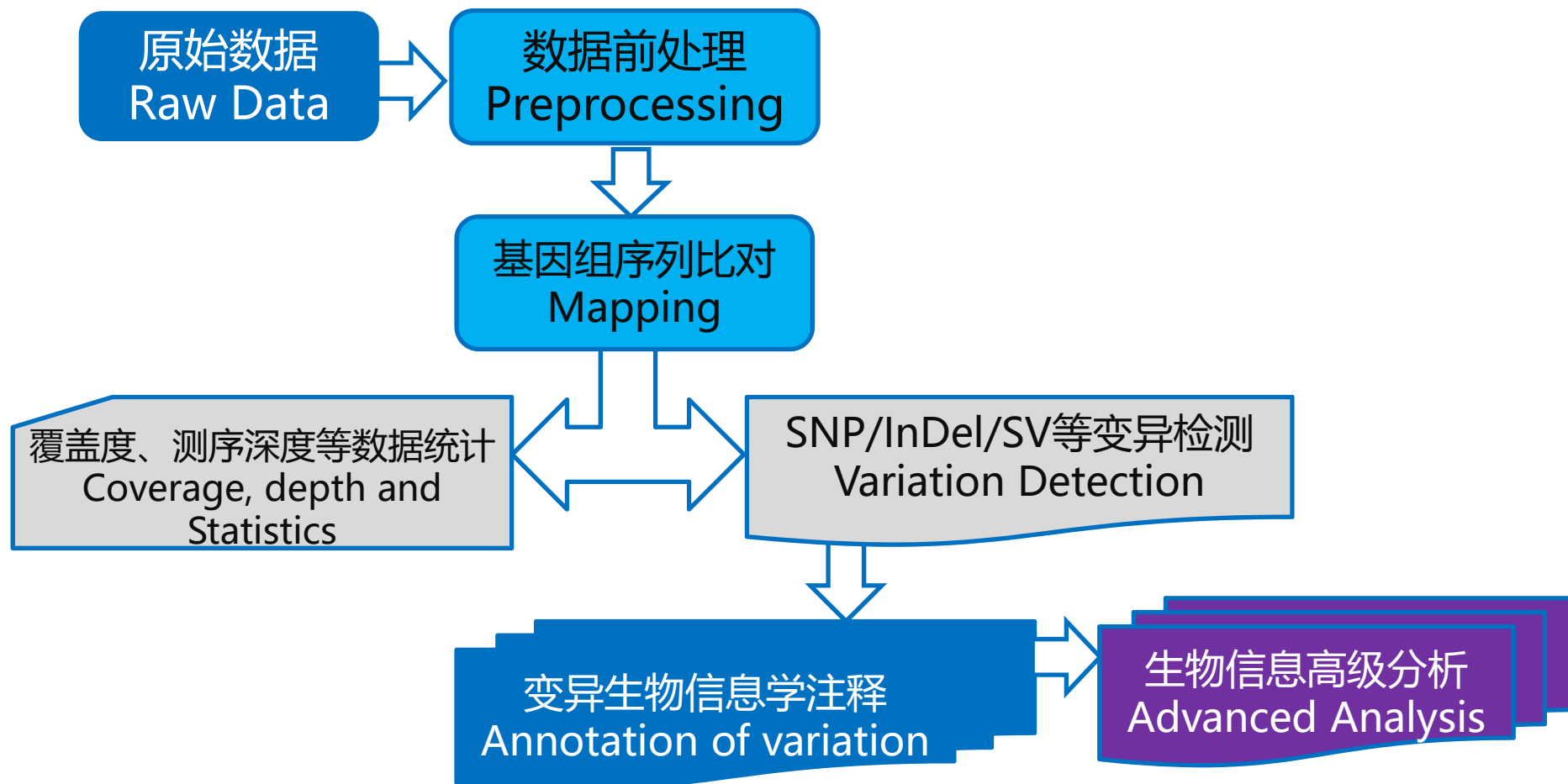
---

- 样本类型：DNA样品， $OD_{260/280}=1.8-2.0$ ，基因组DNA用TE溶解
- 样品浓度： $\geq 50\text{ng}/\mu\text{l}$
- 样品总量： $\geq 2\mu\text{g}$
- 样品运输：DNA低温运输( $-20^{\circ}\text{C}$ )；且在运输过程中请用parafilm将管口密封好，以防出现污染；收到样品后，甲方需要对样品进行检测，最终样品的量和纯度，以甲方的检测结果为准



# 全外显子测序

## 数据分析流程概况



# 全外显子测序

## 数据分析内容

### 基本信息分析

1. 数据质控：去除接头污染和低质量数据
2. 与参考序列进行比对、统计测序深度及覆盖度
3. SNP/InDel/CNV/SV检测、注释及统计

### 高级信息分析

肿瘤	单基因病	复杂疾病
<ol style="list-style-type: none"><li>1. 已知的癌症基因的突变筛选</li><li>2. 高频突变基因总结</li><li>3. 突变位点分布情况分析</li><li>4. 基因组变异Circos 图展示</li><li>5. MRT 高频突变基因相关性分析</li><li>6. 高频CNV 分析</li><li>7. 肿瘤纯度及倍性分析</li><li>8. LOH 在基因组中的分布</li><li>9. 瘤内异质性及克隆结构分析</li><li>10. 融合基因及高频融合基因</li><li>11. 非编码区高频突变分析</li><li>12. 病毒整合分析</li><li>13. 其他个性化内容分析</li></ol>	<ol style="list-style-type: none"><li>1. 突变位点筛选</li><li>2. 显/隐性遗传模式筛选(纯合子或杂合子的筛选)</li><li>3. 候选基因功能注释</li><li>4. 候选基因功能富集分析 (GO , KEGG等)</li><li>5. 纯合子区域分析(连续纯合子变异区段)</li></ol>	<ol style="list-style-type: none"><li>1. 综合各种资源列举已知候选基因</li><li>2. 候选基因中突变的筛选和功能预测</li><li>3. 候选基因功能注释</li><li>4. 新生突变筛选和分析</li><li>5. 候选基因功能富集分析 (GO , KEGG等)</li><li>6. 蛋白互作网络分析(PPI)</li><li>7. 结合实验设计(家系或自然群体)进行遗传学分析(连锁关联分析)</li><li>8. 变异在全基因组范围内的共发生/互斥现象分析</li><li>9. 其他个性化分析</li></ol>

# 全外显子测序

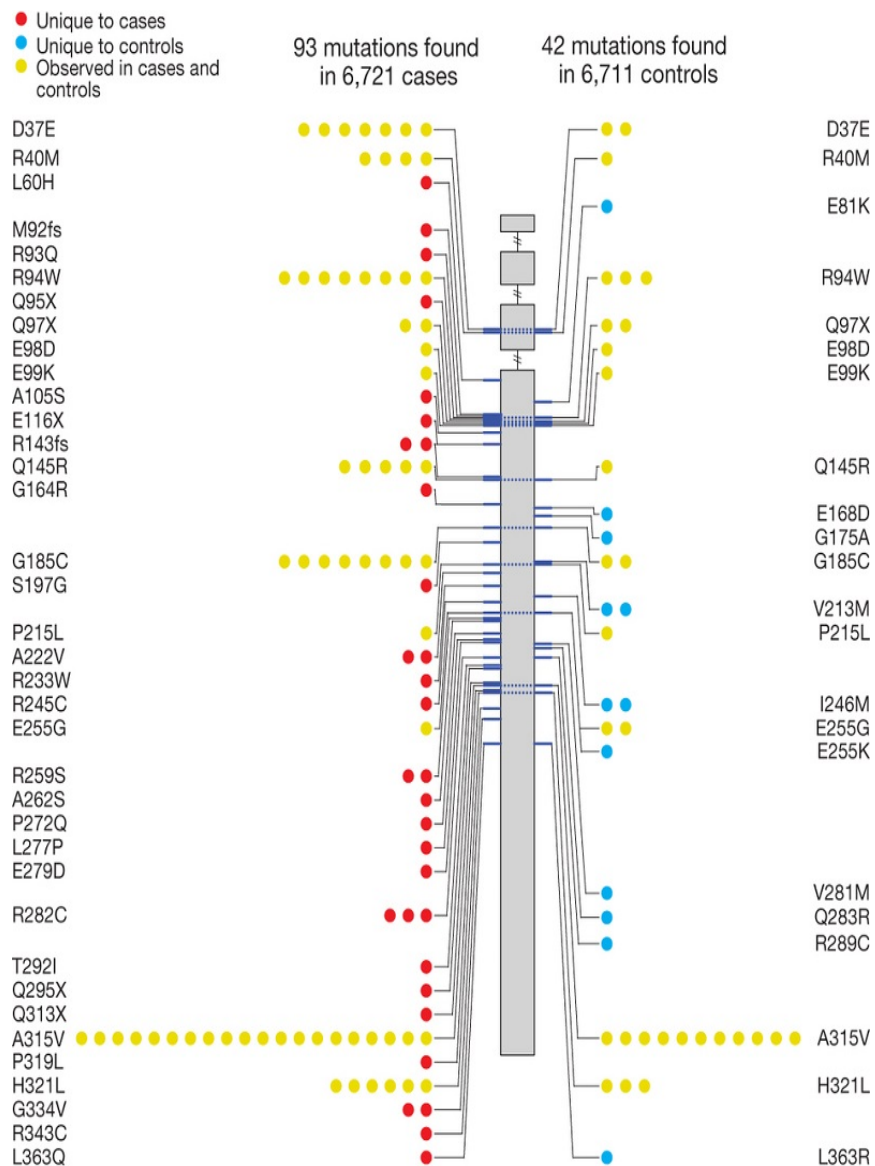
## 应用案例

---

- 全外显子测序发现心梗相关的罕见突变
- *Do, R., et al., Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. Nature, 2015. 518(7537): p. 102-106.*

# 全外显子测序

应用案例：在13432个个体中发现的APOA5基因突变



# 转录组测序

## RNA-Seq

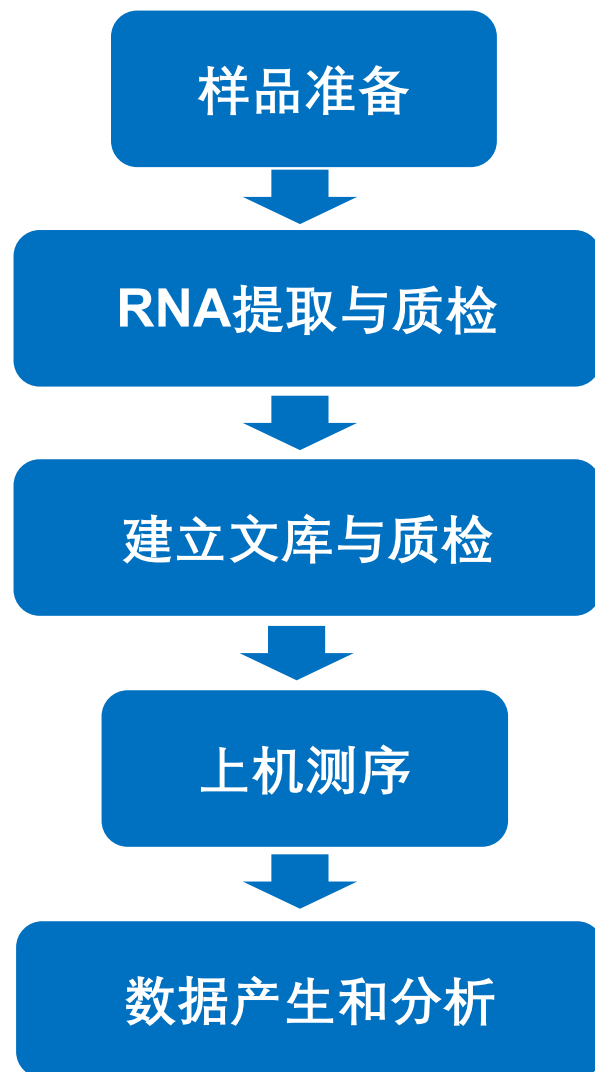
---

- 转录组测序RNA-Seq是基于二代测序技术的转录组学研究方法：首先提取生物样品的全部转录的RNA并进行mRNA富集，然后反转录为 cDNA后进行的高通量测序，在此基础上进行片段的拼接组装，从而可得到一个的转录本，进而可以形成对该生物样品当前发育状态的基因表达状况的全局了解。
- 不同阶段或部位的生物样品的RNA-Seq转录组进行比较分析，则可以在转录层面得到基因表达水平的变化，针对关键基因则可以进行代谢通路(Pathway)的构建。

# 转录组测序

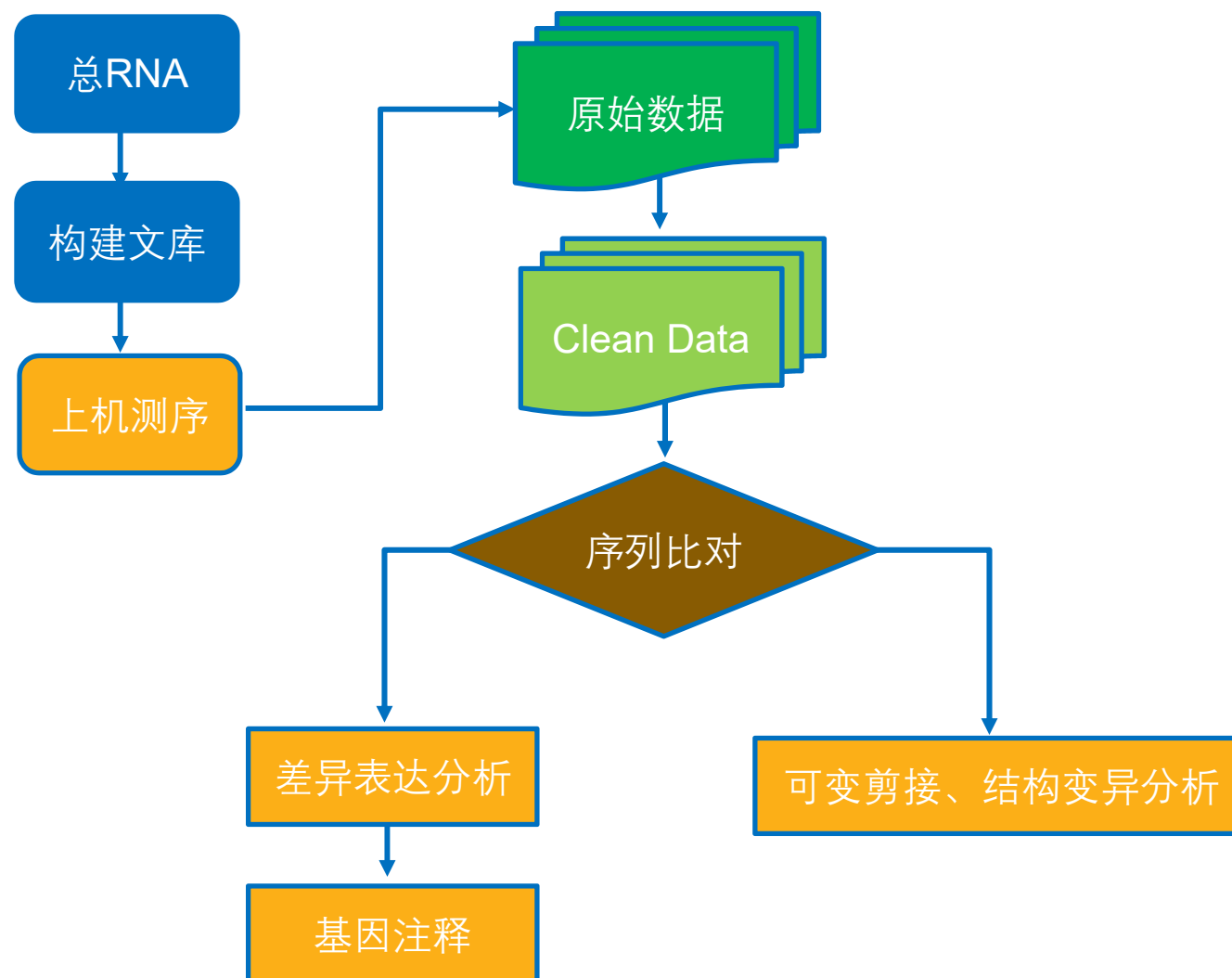
## 技术流程

---



# 转录组测序

## 技术流程



# 转录组测序

## 技术参数

---

- 测序平台与方式：HiSeq, PE150
- 测序数据量： $\geq 6$  Gb/样本
- 项目周期：35个工作日



# 转录组测序

## 样品要求

---

- 样本类型：total RNA，溶解于DEPC-水或者RNase-free的缓冲液中
- 样本总量： $\geq 2 \mu\text{g}$
- 样本浓度： $\geq 100 \text{ ng}/\mu\text{L}$
- $\text{OD}_{260/280} \geq 2.0$ ；完整性 $\text{RIN} \geq 8.0$
- 样品运输：样品置于1.5 ml管中，封口膜封好，干冰运输

# 转录组测序

## 数据分析内容

---

- 1. Reads QC/QA
- 2. 比对到参考基因组
- 3. mRNA 表达水平分析
- 4. 基于mRNA 表达水平的QC
- 5. 基因差异表达分析；差异表达基因筛选
- 6. 差异表达基因的功能富集分析(GO,KEGG)
- 7. 差异表达基因的蛋白功能富集分析(Pfam , KOG等)
- 8. 差异表达基因PPI分析
- 9. 特定功能的基因的差异表达情况
- 10. 基因共表达分析
- 11. 可变剪切分析
- 12. 新转录本预测
- 13. SNP/InDel 分析
- 14. 融合基因分析

# 转录组测序

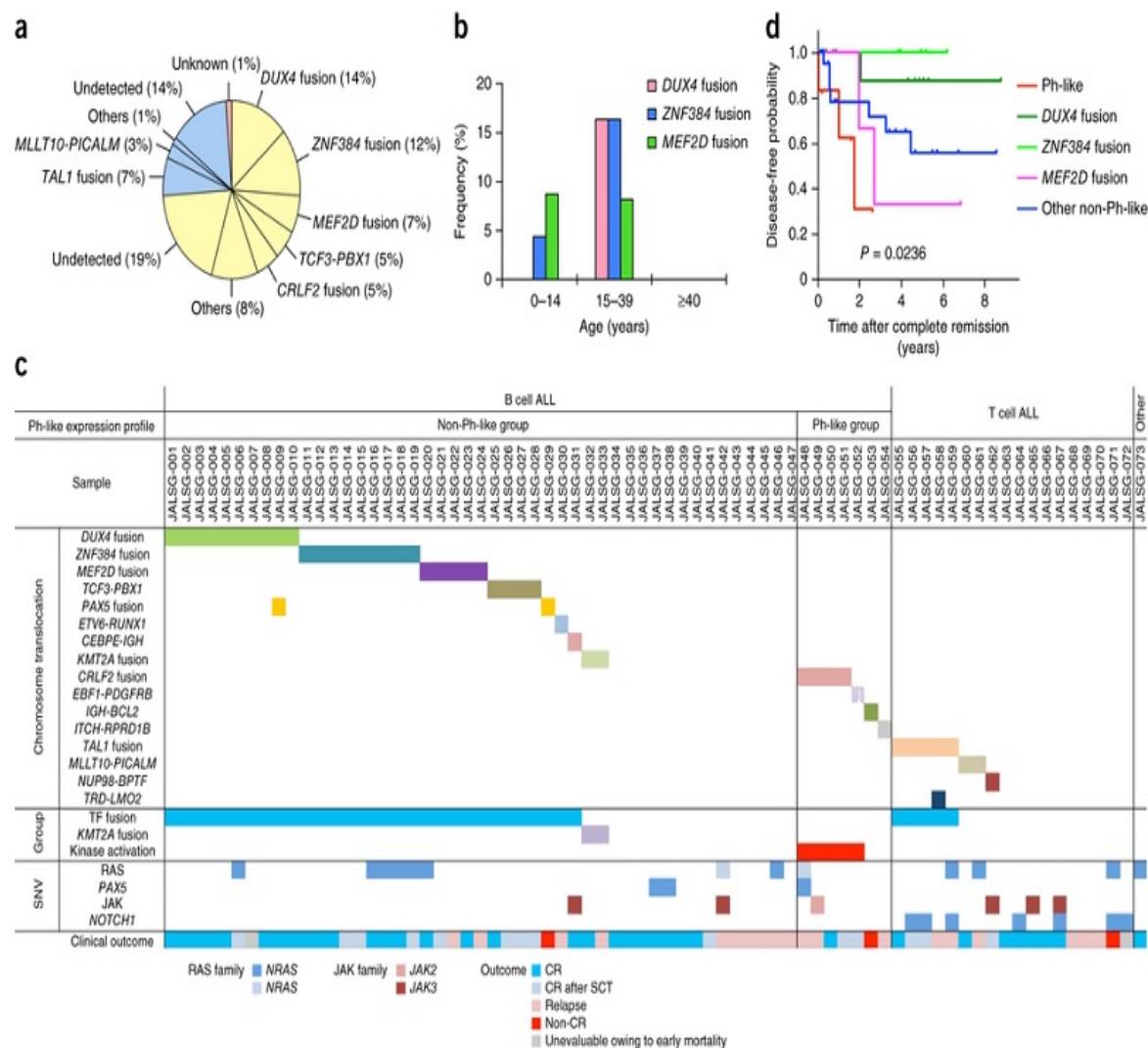
## 应用案例

---

- 低龄急性成淋巴细胞性白血病病例中发现新的DUX4融合基因
- *Yasuda, T., et al., Recurrent DUX4 fusions in B cell acute lymphoblastic leukemia of adolescents and young adults. Nature genetics, 2016.*

# 转录组测序

## 应用案例：AYA-ALL 中的融合基因



# 小RNA测序

## Small RNA Sequencing

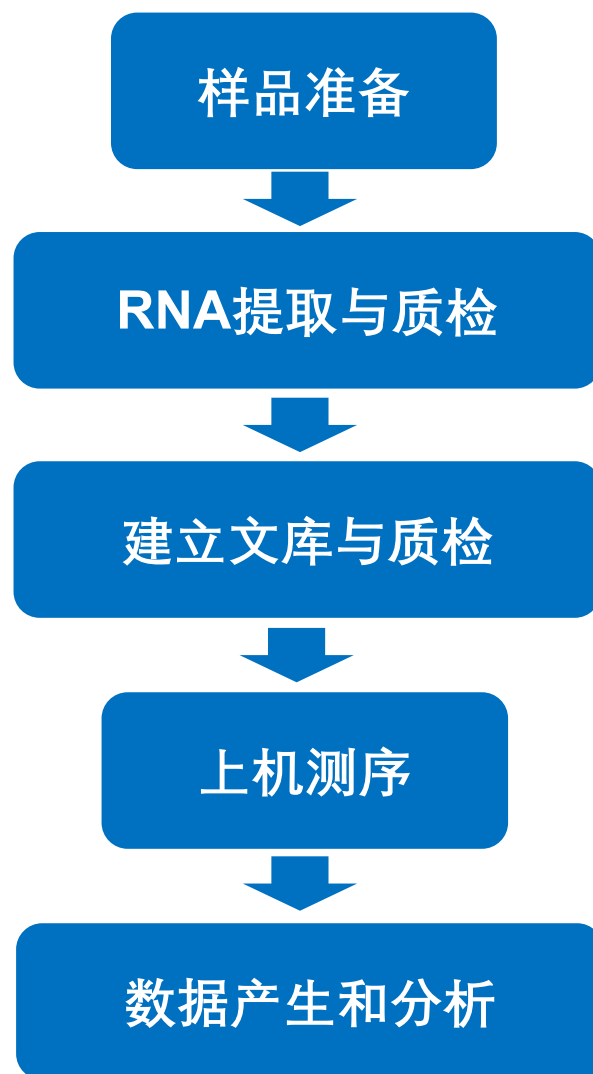
---

- Small RNA是一大类调控分子，几乎存在于所有的生物体中。Small RNA包括：miRNA、ncRNA、siRNA、snoRNA、piRNA、rasiRNA等。Small RNA有多种多样的作用途径，包括mRNA降解、翻译抑制、异染色质形成以及DNA去除，来调控生物体的生长发育和疾病发生。
- MicroRNA (miRNA)是一类内生的、长度约为20-24个核苷酸的小RNA，其在细胞内具有多种重要的调节作用。每个miRNA可以有多个靶基因，几个miRNA也可以靶向同一个基因。这种复杂的调节网络通过一个miRNA来调控多个基因的表达，也可以通过几个miRNA来精细调控某个基因的表达。

# 小RNA测序

## 技术路线

---



# 小RNA测序

## 技术参数

---

- 测序平台与方式：HiSeq, SE50
- 测序数据量：10-20M Reads/样本
- 项目周期：45个工作日

# 小RNA测序

## 样品要求

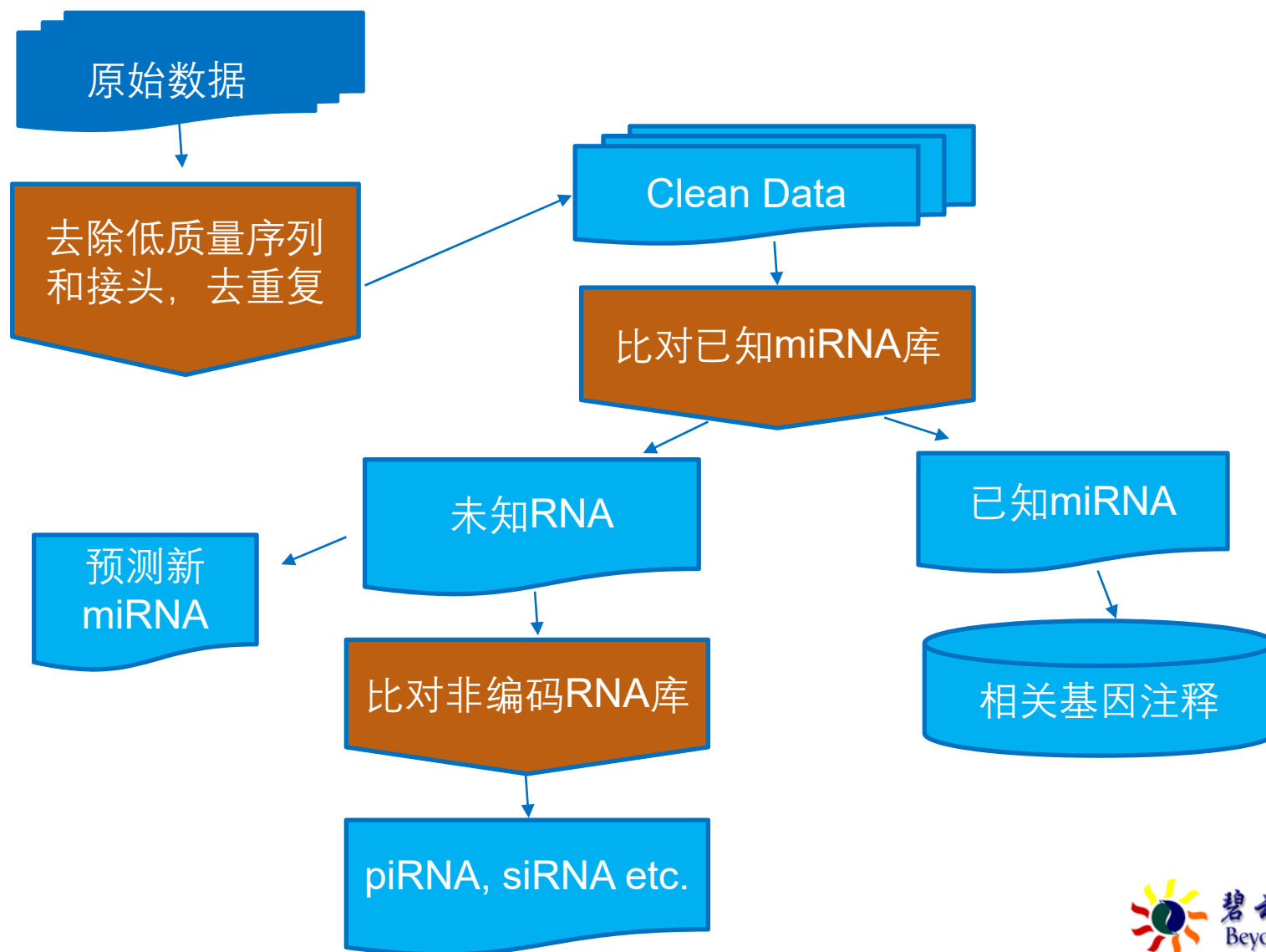
---

- 1. 样品类型：细胞、新鲜组织或RNA样品。
- 2. 样品量：细胞样品请提供至少 $1 \times 10^7$ 个细胞，组织样品请提供至少300mg的组织块或切片，RNA样品请提供2  $\mu$ g以上的总RNA。
- 3. 样品质量：RNA无明显降解，提取的总RNA OD260/280值在1.8~2.2之间，浓度 $\geq 100 \text{ ng}/\mu\text{l}$ ，28S:18S  $\geq 1.5$ ，RIN  $\geq 7$ 。
- 4. 样品保存：细胞样品或新鲜组织块(切成~50mg的小块)可用TRIZOL或RNA保护剂处理或液氮冻存后，-80°C保存。RNA样品可溶于乙醇或RNA-free的超纯水中，-80°C保存。样品保存期间避免反复冻融。
- 5. 样品运输：样品置于1.5 ml管中，封口膜封好，干冰运输。



# 小RNA测序

## 数据分析流程



# 小RNA测序

## 数据分析内容

---

- 1. Reads QC/QA
- 2. 公共序列和特异序列的分析
- 3. 小RNA在选定的参考基因组上的分布
- 4. 已知miRNA筛选
- 5. 新miRNA预测
- 6. 分类注释
- 7. 靶基因预测
- 8. 靶向转录本分析

# 小RNA测序

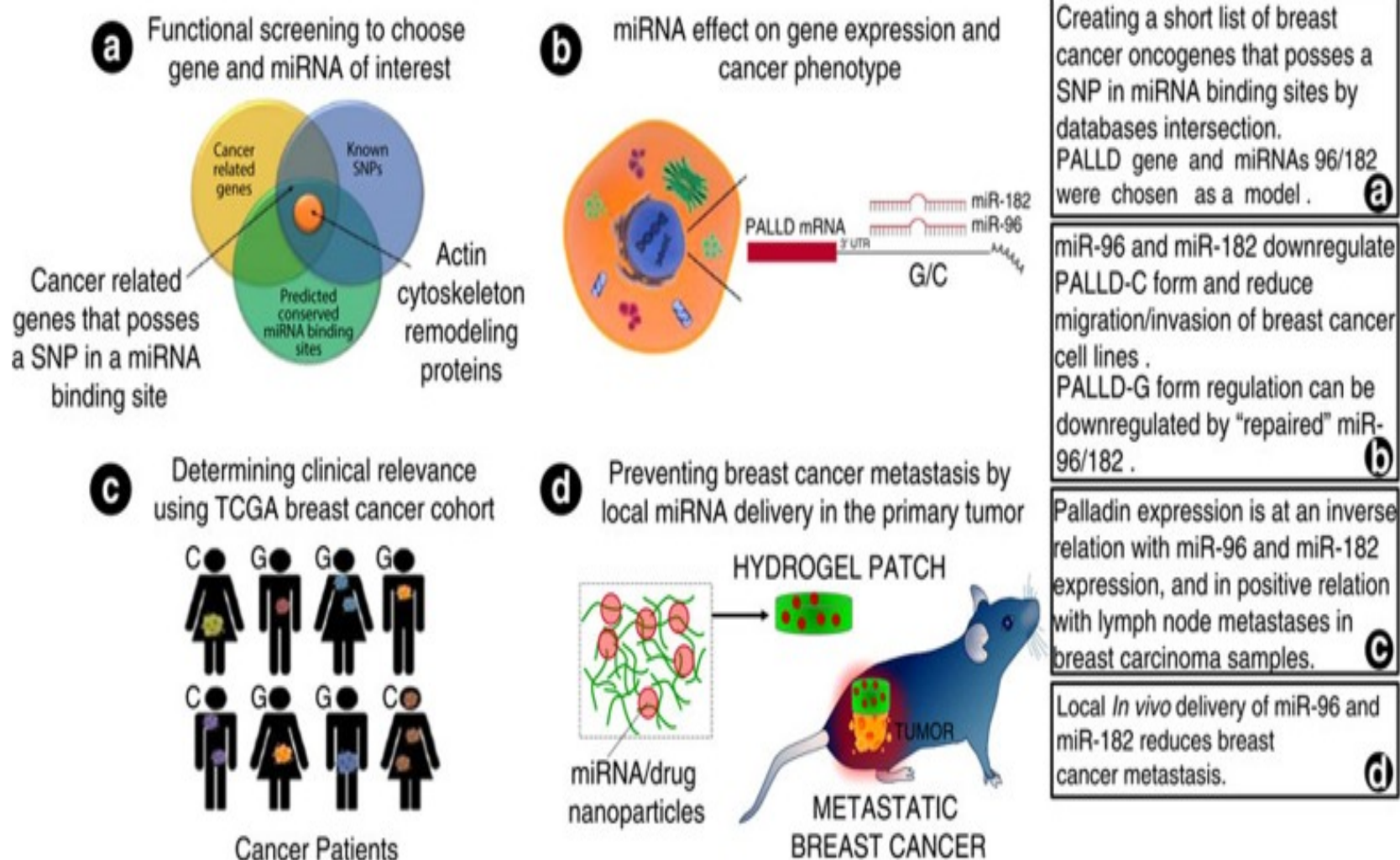
## 应用案例

---

- microRNA调节Palladin的表达从而控制乳腺癌转移
- *Gilam, A., et al., Local microRNA delivery targets Palladin and prevents metastatic breast cancer. Nature Communications, 2016. 7: p. 12868.*

# 小RNA测序

应用案例：多组学数据整合揭示体内治疗新方法



# 谢谢!



碧云天网站



微信公众号

碧云天生物技术/Beyotime Biotechnology

订购热线：400-168-3301或800-8283301

技术咨询：info@beyotime.com

高通量测序服务：ngs@beyotime.com

网址：<http://www.beyotime.com>