

# 转录组测序常见问题



上海欧易生物 技术支持部

# 转录组测序常见问题

## 目 录

1. 什么是转录组测序? .....	2
2. 为什么进行转录组研究而不是基因组研究? .....	2
3. 相比与其他公司, 欧易生物转录组测序服务的优势是什么? .....	2
4. 转录组测序可以同时检测 mRNA、miRNA 及其他非编码 RNA? .....	2
5. 测序深度和覆盖度是什么意思? .....	3
6. Q20、Q30 所代表的碱基质量含义? .....	3
7. Paired-end 与 Single-end 表示什么意思? .....	3
8. 原核生物与真核生物在进行转录组测序文库构建时有什么区别? .....	3
9. 利用 HiSeq 进行测序时, 一个 Lane 能输出多少数据? .....	3
10. 一般动植物进行转录组测序, 要测多少数据量, 如何估测? .....	3
11. 转录组测序一定要设置生物学重复吗, 如果样品间差异不大还需要设置吗? .....	4
12. 转录组测序对样品有什么要求吗? 样品用量是多少? .....	4
13. 样品该如何收集与处理? .....	4
14. 转录组测序的实验流程? .....	4
15. 关于转录组 <i>De novo</i> 分析: 你们采用什么软件进行拼接, 使用的参数是什么? .....	4
16. 转录组测序什么时候采用 <i>De novo</i> 分析测序, 什么时候采用 Reference 分析策略? .....	4
17. 欧易生物提供的转录组测序报告内容有哪些? .....	4
18. 转录组测序都能做哪些分析, 有哪些个性化的分析内容? .....	6
19. 测序原始数据提供的是 Raw data 还是 Clean data? .....	6
20. Raw data 如何读取? 为什么不提供原始图像数据和中间过程文件? .....	6
21. 转录组测序与表达谱芯片和 qPCR array 的区别是什么? .....	6
22. qPCR 定量结果与转录组测序结果不符原因? .....	6

# 转录组测序常见问题

## 1. 什么是转录组测序？

转录组测序的研究对象为特定细胞在某一功能状态下所能转录出来的所有 mRNA 的总和。转录组研究是基因功能及结构研究的基础和出发点，通过新一代高通量测序，能够全面快速地获得某一物种特定组织或器官在某一状态下的几乎所有转录本序列信息，已广泛应用于基础研究、临床诊断和药物研发等领域。

## 2. 为什么进行转录组研究而不是基因组研究？

转录组即特定材料在某一功能状态下所能转录出来的所有 RNA 的总和，包括 mRNA 和非编码 RNA。转录组研究是基因功能及结构研究的基础和出发点。应用新一代高通量测序，能够全面快速地获得某一物种特定器官或组织在某一状态下的几乎所有转录本。更为重要的是，转录组研究所需要的费用，与基因组研究相比，前者只有后者的几分之一，因此，进行全转录组方面的研究比起全基因组方面的研究具有更强的可实施性。尤其针对课题经费不足，又以没有 Reference sequence 物种为研究对象的科研工作者，更是如此。

## 3. 相比与其他公司，欧易生物转录组测序服务的优势是什么？

- ✧ **专业的技术操作人员：**具有十二年建库技术沉淀，具有丰富样本处理和测序实验操作经验；
- ✧ **强硬的硬件支撑：**CLIA 认证实验室，5S 实验规范管理，多套计算集群，保证数据产出质量和分析速度；
- ✧ **全面的分析内容：**具有专业、成熟的生物信息分析团队，熟悉各种分析算法和软件，保证标准、个性化和售后调整分析的全面支撑，并在分析过程中与客户充分互动；
- ✧ **更高的服务标准：**采取严格的质量标准、多重实验质控和完备的实验记录，确保更高的成功率、更低的样品损失及实验失败风险，并且客户可以参与整个芯片实验过程，接受客户监督，更加透明和可信；
- ✧ **完善的后续验证支持：**持续专注于 mRNA 的定量检测，为上千位客户提供过相关服务；
- ✧ **优秀的服务理念：**核心团队平均 12 年以上技术服务领域工作经验，均具备良好的职业操守及服务客户理念。

## 4. 转录组测序可以同时检测 mRNA、miRNA 及其他非编码 RNA？

理论上技术是可行的，但是通常会根据测序对象长度的不同，在测序建库的时候会选择不同的片段大小，测序读长也会有不同。一般来讲，如果要进行 microRNA 测序的话，通常将 microRNA 分离出来，单独进行测序可。mRNA 测序，通常建库时选择 200-300bp 大小片段，采用 125PE/150PE 测序。而长链非编码 RNA (lncRNA) 存在正向转录和反向转录，所以常采用链特异性建库测序。

## 5. 测序深度和覆盖度是什么意思？

**测序深度：**测序得到的总碱基数与待测转录组大小的比值。

**覆盖率：**指测序获得的序列占整个转录组的比例。

## 6. Q20、Q30 所代表的碱基质量含义？

为了保证数据质量，要在信息分析前对原始数据进行质量评估。每个碱基测序错误率是通过测序碱基质量值（Phred score, Qphred）通过公式转化得到，而测序质量值是在碱基识别过程通过一种预测碱基判别发生错误概率模型计算得到的，对应关系如下表所显示：

Phred 分值	不正确的碱基识别	碱基正确识别率	Q-score
10	1/10	90%	Q10
20	1/100	99%	Q20
30	1/1000	99.9%	Q30
40	1/10000	99.99%	Q40

Q20: 原始数据中 Phred 数值大于 20 的碱基数量占总碱基数量的百分比。

Q30: 原始数据中 Phred 数值大于 30 的碱基数量占总碱基数量的百分比。

## 7. Paired-end 与 Single-end 表示什么意思？

**单端测序 (Single-read)：**是指测序引物结合位点只连接到待测片段的一端，然后末端加上接头，将片段固定在 flow cell 上形成簇，上机测序单端读取序列。

**两端测序 (Paired-end)：**是指在构建待测文库时在两端都加上测序引物结合位点，在第一轮测序完成后，去除第一轮测序的模板链，用对读测序模块引导互补链在原位置再生和扩增，以达到第二轮测序所用的模板量，进行第二轮互补链的合成测序。

通常情况下转录组测序都采用的是 Paired-end 测序法。

## 8. 原核生物与真核生物在进行转录组测序文库构建时有什么区别？

在原核生物中，mRNA 只占全部 RNA 的 1-5%，其余绝大部分是核糖体 RNA (rRNA)，因此若要测序 mRNA，首先必须先将 mRNA 纯化出来，然而，原核生物并不像真核生物 mRNA 具有 polyA 的结构，因此，无法直接利用 oligoT 将 mRNA 纯化出来，如果拿 total RNA 进行测序，那么测序的效率一定会非常差，因为大部分的序列都来自 rRNA。目前，提高原核生物中 mRNA 的量，最主要的方式是去除 total RNA 中 rRNA。

## 9. 利用 HiSeq 进行测序时，一个 Lane 能输出多少数据？

目前公司采用 HiSeq2500 或 4000 测序，HiSeq2500 一般一个 lane 至少可产生 60G clean data，HiSeq4000 一般一个 Lane 至少可产生 75G clean data。

## 10. 一般动植物进行转录组测序，要测多少数据量，如何估测？

由于转录组测序需要进行表达量的分析，因此不推荐使用覆盖度，在确定测序量时，我们以产生的 reads 数作为依据。转录组测序所需的测序量随物种转录组大小的不同而有所差

异。而转录组的大小受基因数目和丰度双重影响，不同物种间变化很大。因此在测序之前，需要对转录组的大小进行评估。针对有参考基因组的物种，可通过分析基因组信息，统计编码基因个数及其碱基数来评估转录组的大小，同时也可参考相近或相关物种转录组研究的文章，针对无参考基因组的物种，只能参考相近物种的转录组大小。**具体的数据量可以咨询当地销售。**

### 11. 转录组测序一定要设置生物学重复吗，如果样品间差异不大还需要设置吗？

是的。一般应该重复 3 次以上，具体情况需要根据具体实验来决定。样本数量越多对于统计筛选越有利。

### 12. 转录组测序对样品有什么要求吗？样品用量是多少？

总 RNA 量至少需要 5 $\mu$ g；组织量至少为 300mg，细胞量至少为 10<sup>6</sup>，全血样本建议 2ml 并分离白细胞。所有样品都采用干冰运输，偏远地区应保证干冰量足够。

### 13. 样品该如何收集与处理？

针对不同类型的样品收集与处理，我们专门整理了《样品采集操作指南》，客户可以自行在我们的官网下载（<http://www.oebiotech.com/index.php?s=/Index/down/id/0.html>）或者向当地的销售索取。

### 14. 转录组测序的实验流程？

转录组测序实验流程详情请见：

<http://www.oebiotech.com/index.php?s=/Index/detail/id/36.html>

### 15. 关于转录组 *De novo* 分析：你们采用什么软件进行拼接，使用的参数是什么？

*De novo* 拼接是指在不依赖参考基因组的情况下，将有 overlap 的 reads 连接成一个更长的序列，经过不断的延伸，拼接成 transcript。我们使用 Trinity(version:trinityrnaseq\_r20131110) 软件 paired-end 的拼接方法，对样本的有效 reads 合并进行 *de novo* 拼接，取每个 Loci (comp\*\_c\*\_ ) 下最长的转录本作为 Unigene，以此作为后续分析的参考序列。

使用参数为：Trinity.pl --seqType fq --min\_contig\_length 200 --JM 400G --left \$R1 --right \$R2 --SS\_lib\_type RF --output trinity\_out\_dir --CPU 80。

### 16. 转录组测序什么时候采用 *De novo* 分析测序，什么时候采用 Reference 分析策略？

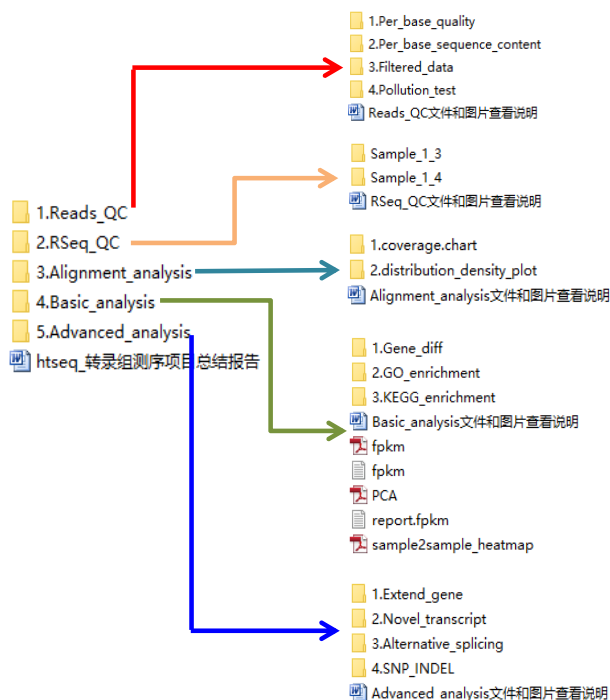
对于有参考基因组信息，而且拼接质量较好，注释信息较完整的物种可采用 Reference 分析策略；对于无参考基因组信息，或者参考基因组拼接质量不好，注释信息不完整的物种，采用 *De novo* 分析策略。

### 17. 欧易生物提供的转录组测序报告内容有哪些？

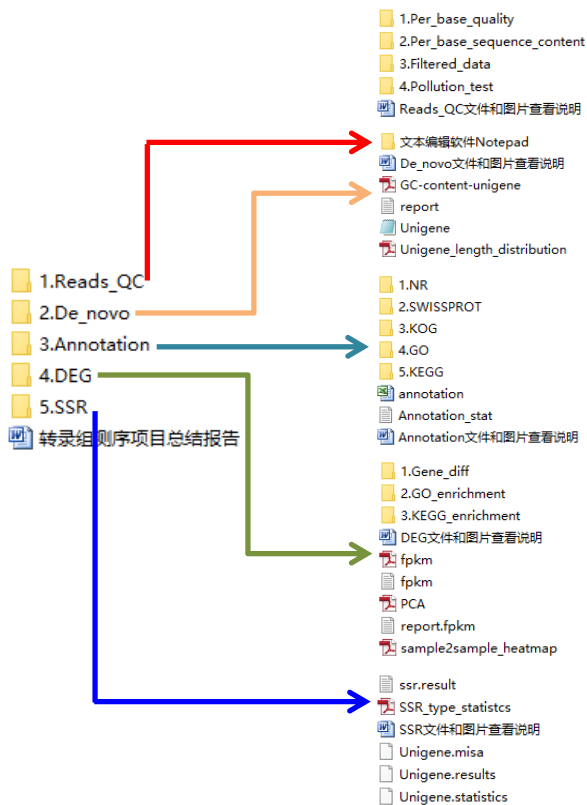
欧易生物提供的转录组测序报告分为 Reference 和 *De novo* 两种。Reference 转录组测序报告由 Reads QC、RSeq QC、Alignment analysis、Basic analysis、Advanced analysis 和项目

总结报告组成。*De novo* 转录组测序报告由 Reads QC、*De novo*、Annotation、DEG、SSR 和项目总结报告组成。项目总结报告是这个项目的完整说明。具体报告模板请向当地销售索取。

Reference 转录组测序报告:



*De novo* 转录组测序报告:



## 18. 转录组测序都能做哪些分析，有哪些个性化的分析内容？

欧易生物提供的转录组测序的数据分析内容包括：测序数据质控、基本信息分析、基因表达分析和高级信息分析。个性化分析内容包括基因结构优化、新转录本预测及编码能力分析、可变剪切分析、SNP 和 Indel 分析以及 SSR 分析等。我们专门整理了两套转录组测序报告模板（有参和无参），里面详细介绍了我们所提供的数据分析内容，客户可向当地销售索取。

## 19. 测序原始数据提供的是 Raw data 还是 Clean data？

我们提供的测序原始数据是 raw data，这样老师可以对测序质量进行评估，而且一般文献期刊要求上传的数据也是 raw data。

## 20. Raw data 如何读取？为什么不提供原始图像数据和中间过程文件？

测序数据文件以 txt 文本格式为主。对于 Windows 用户，推荐使用 Editplus 或 UltraEdit 作为浏览程序，否则会因文件过大造成死机。Unix 或 Linux 系统比较适于浏览较大的文本文件。由于 Illumina 更新了 pipeline，测序过程中生成的图像文件将实时转化为中间过程文件，这一步完成后图像文件将自动删除，得到中间过程文件（此文件为二进制文件，只有在 Illumina 机器上才能读取，通常不保留），在 Illumina 分析软件下将其转化为序列文件，即所说的 raw data。目前各公共数据库接受 fastq 文件，所以我们提供的 raw data 都是 fastq 文件。

## 21. 转录组测序与表达谱芯片和 qPCR array 的区别是什么？

	表达谱芯片	qPCR array	转录组测序
基因数	上万个基因	某个通路的 84/336 个关键基因	上万个基因
物种	常见的研究物种	主要是人、大小鼠、恒河猴，其他物种需要定制	理论上所有物种
是否需要 qPCR 验证	需要	不需要	需要
能否发现新的基因	不能	不能	理论上可以
研究可变剪切转录本	部分产品可以	不能	可以
实验周期	20-25 个工作日	25 个工作日	55-60 个工作日
最低样品量 (Total RNA)	Agilent: 200ng Affymetrix: 250ng	96 孔板: 1μg 384 孔板: 4*400ng	5μg
所用仪器	芯片扫描系统	qPCR 仪	测序仪

## 22. qPCR 定量结果与转录组测序结果不符原因？

答：

- 1) 用于 qPCR 定量检测的基因在样本中表达量偏低，Ct 值偏高；
- 2) 用于 qPCR 定量检测的样本与转录组测序样本不一致；

3) 转录组测序毕竟是高通量差异筛选手段，所以跟低通量的 qPCR 定量结果可能会存在不一致的情况。



联系我们:



公司总部：上海市浦东新区祖冲之路 1505 弄 138 号 6 楼

服务热线：4006-4008-26

欧易官网：[www.oebiotech.com](http://www.oebiotech.com)

市场联络邮箱：[market@oebiotech.com](mailto:market@oebiotech.com)

技术服务咨询邮箱：[service@oebiotech.com](mailto:service@oebiotech.com)

邮编：201210

微信：

