

KBSeq 质粒测序 Q&A

技术相关:

1. KBSeq 全质粒测序跟常规的一代测序测质粒有何区别? 能否一定测到全长质粒序列?

KBSeq 采用的是二代测序技术, 对全质粒进行建库测序, 与一代测序相比, KBSeq 无需设计合成测序引物, 单次测序可最高获得总长至 30K 的质粒序列。因为无需测序引物, 对于长质粒来说 (如 10K 以上) KBSeq 的周期会更快, 并且 KBSeq 还可以对未知质粒测序。不过 KBSeq 也继承了二代测序的缺点, 对高 GC、homopolymer、重复区测序效果不佳, 存在此区域的质粒可能会测不到全长。

样品相关:

2. 对质粒样品的送样要求是怎样的?

质粒浓度需 $\geq 2\text{ng}/\mu\text{l}$, 体积需 $\geq 10\ \mu\text{l}$, 且主带明显、无污染及降解, 样品请务必使用双蒸水进行溶解。

3. 对菌液样品的送样要求是怎样的?

需提供鲜菌液 $\geq 200\ \mu\text{l}$ 或培养的穿刺菌, 需冰袋运输, 不能干冰运输, 生工收到菌样后会摇床培养再提质粒, 生工提供氨苄青霉素和卡那霉素两种抗生素; 其余类型抗生素需自备并附使用说明;

4. KBSeq 怎么收费的?

KBSeq 按照质粒长度区间来收费, 收费区间分别是 0-6K, 6K-15K, 15K-30K, 30K 以上, 具体价格询问生工当地销售员。下单时候需要填一下预估长度, 最终以实际测出来的长度来收费。

5. 送样的时候需要提供质粒的一些信息吗?

提供质粒的话只要提供质粒长度, 若关注插入序列信息, 可以提供下插入序列上下游的引物序列, 不提供引物序列的话后续结果中只会给全质粒序列, 提供的话会再提供引物区域的插入序列。

6. 测序不用引物的话, 那是不是说质粒片段中高 GC 区域, 重复序列都能测通呢?

KBSeq 仅适用于非复杂序列, 即要求 GC 含量为 25-75%, 简单的二级结构, 无 100bp 以上直接或间接的重复序列, 无 100bp 以上长嘌呤 / 聚嘧啶区段。若满足以上要求的话基本上都能测通, 但高 GC 和重复序列会影响测序, 从而影响拼接效果, 导致不能测通。

7. 质粒未知序列, 但 GC 含量比较高, 一代测序重叠峰测不通, 大概 10 kb, 这种样品适用于 KBSeq 吗? 含 poly 结构的序列也能实现测通吗? 质粒有重复序列或者复杂结构可以测出来吗?

这类质粒测序效果也不太好, 基本上一代测序效果不好的样本用 KBSeq 做效果也不会好。

8. 提供引物和不提供引物的区别在哪?

提供引物可以快速找到关注目的片段, 不需要提供实体引物, 只需要提供引物序列; 若提供了引物序列结果中除了会给出全质粒序列, 还会给出引物之间的目标序列。

9. 参考序列是必须要提供吗?

KBSeq 测序后采用的是 denovo 拼接方式, 因此是不需要提供参考序列的。

10. 关于送样浓度, 一代测序要求质粒 100ng/ μl , KBSeq 则要求 2ng/ μl , 这个送样量真的可以测出来吗?

可以测出来, KBSeq 的建库方式采用的是微量建库, 因此对质粒的浓度要求不高。

11. 混合质粒样品可以吗? 还是必须是单一的样品?

理论上来说也是可以的, 但是混合样本可能会导致拼接效果不好, 因此不建议混合样品。

12. 送菌液会额外收取提取费用吗?

不额外收取提取费。

13. 一代测序菌液只能测大肠杆菌，KBSeq 测序菌液也是只能测大肠杆菌么？
不限于大肠杆菌。
14. 提供菌液的话抗性只能是氨苄和卡那吗？
不是的，目前生工免费提供氨苄青霉素和卡那霉素两种抗生素进行摇菌培养；若是其余类型抗生素的话需自备并附使用说明。
15. 酶切后的线状质粒也可以用 KBSeq 吗？
可以的。
16. 对质粒的浓度要求低，是否提供的菌浓度也可以很低？
菌液的浓度还是需要高一些，因为若菌浓度太低可能导致摇菌失败，没法提取到合格的质粒。
17. 普通测序的质粒样品，都可以做 KBSeq 测序嘛？
理论上都可以，不过二代也有一代一样的问题，高 GC，重复区，homopolymer 不好测。
18. 最终结果中除了质粒序列外，还会提供其他什么信息？
还会提供质粒圈图，突变列表。

订购相关：

19. KBSeq 可以线上预约取样吗？
KBSeq 可以线上预约取样，并且跟生工一代测序预约取样流程完全一样，目前还开通了微信端预约取样，预约更方便。
20. 结果是直接发邮箱还是上传系统，线上预约享受积分活动吗？
结果给出的方式和一代测序一样上传至系统并邮箱通知客户，线上预约的订单也享受 10%积分，非线上预约订单无积分。
21. 如果不知道质粒大小，不填写可以吗？
可以先填写一个预估的大小，因为 KBSeq 建库是需要按照质粒片段大小来混样的，所以需要提供一个预估的大小，费用方面后面系统会根据实际测出来的长度重新计算费用。
22. 我有 KBSeq 相关的问题需要咨询，请问如何联系你们呢？
KBSeq 全质粒测序服务咨询电话：021-57072096/021-57072057；KBSeq 全质粒测序服务咨询邮箱：kbseq@sangon.com

周期相关：

23. 正常接单后，多久时间能出结果？
KBSeq 的周期为收到样本后 2-7 个工作日，因二代测序高通量的要求，需累计大量样本一次性上机。故在当日样本数足够，无需累计等待上机的情况下，可实现最快 2 个工作日完成从样本接收到发送报告全部流程。
24. 同一份订单中，样品长度差别大的，比如 3kb 与 10kb 的，出结果时间是否一致，还是说长度更长的会慢一点？
出结果的时间是一致的。

结果相关：

25. 最终交付的结果形式是怎样的？结果中包括哪些数据？
最终结果通过邮箱交付，可从邮箱中下载每个样本结果，每个样本包含三个文件，分别如下：
.fasta 文件：测序结果文件，fasta 格式，可能一个样本会拼接出来多条序列，该多条序列可能来自同

一质粒，当质粒中含有高 GC、重复区等复杂结构导致不能拼接成一条成环的质粒；也可能是杂质序列，比如宿主污染、其它质粒污染等。若有多条序列，按照 Contig_1、Contig_2 依次编号。

.pdf 文件：质粒圈图文件，当拼接出多条序列的时候会给每一条序列作图，图中包含酶切位点、基因注释等信息。若客户需要自行作图也可下载 SnapGene，用该软件绘制圈图。

.xlsx 文件：突变信息文件，当序列中没有突变，则该文件只有表头，没有具体信息。表头解释如下：

INDEX	序列编号，1 表示 Contig_1,2 表示 Contig_2
POS	突变发生的位置
REF	该位置测出来最多的碱基
ALT	该位置存在的突变碱基
TYPE	突变类型，SUB 表示碱基替换，INS 表示插入、DEL 表示缺失
DP(read_depth)	该位置的总测序深度，就是该位置测了多少次
AD(allele_depth)	突变碱基测到的次数
AF(allele_frequency)	突变比例

26. 我在贵公司做的全质粒测序结果中出现了两个圈图，我想知道这是为什么？哪一个才是我的质粒呢？

圈图文件中出现了多个圈图说明最终拼接结果中不止一条序列，可能是样本中本身含有多个质粒，然后每个质粒绘制了一个圈图。也可能是样本中确是单个质粒，但由于高 GC、重复区等复杂结构导致不能拼接成一条序列，最终产生多条序列，每条序列分别为质粒的一部分，这些序列都是有效序列，我们对每条序列分别绘制了圈图，不过此时的每个圈图不是完整的质粒图谱。

27. 我刚刚收到了贵公司全质粒测序结果的邮件，打开后提示实际测得的大小跟预期不符，但是必须要确认价格后才可以下载结果，是要必须点击才可以吗？

KBSeq 是按照最终实际测到的序列大小来收费的，当您收到此邮件的话说明实际测到的质粒大小跟下订单时的预期大小不一致，此时会跟您确认价格，只有您点击确认才能下载结果。因此需要您先点击确认价格，下载结果，如果后期结果有什么问题的话，可以电话我们沟通协商。电话:021-57072096/2057

28. fasta 文件中，序列里面有除了 ATCG 四个碱基以外的字母，是否是测序错误？

不是测序错误，这些碱基是简并碱基，比如说 K 碱基，说明这个位置的碱基既测到 G 也测到了 T，出现简并碱基说明该位置可能存在突变，最好需要一代测序验证下。下图是简并碱基对照表。

简并碱基	正常碱基
R	A/G
Y	C/T
M	A/C
K	G/T
S	G/C
W	A/T
H	A/T/C
B	G/T/C
V	G/A/C
D	G/A/T
N	A/T/C/G

29. 我的测序结果有样本显示测序失败，无信号或者信号弱，是什么原因，这部分样本收费吗？

显示无信号跟信号弱均表明测序失败，未测到质粒序列，可能的原因是质粒拷贝数太低未提取到合格质粒，或者其他原因导致的建库失败，这部分样本不收费，若要重做的话只能重新送样再重做一次。

30. 为什么拼接出来的序列跟我的参考序列方向不一致，是反向互补的？

这是正常现象，因为 KBBSeq 采用的是 NGS 建库测序方式，序列中会测到正向的序列，也会测到反向的序列，导致最终拼接出来的序列即可能是正向序列，也可能为反向互补序列，在做下游分析的的时候需要注意区分。

31. 我拿到的序列文件中有 3 条序列，这是为什么？

两种可能，1 可能是样本中本身含有多个质粒或者杂质序列，导致拼接出来多条序列。2 可能是样本中确是单个质粒，但由于高 GC、重复区等复杂结构导致不能拼接成一条序列，最终产生多条序列，每条序列分别为质粒的一部分，这些序列都是有效序列。

32. 我的结果中有多条序列，这几条序列的前后关系是什么样的？是否是按照 Contig_1, Contig_2 的顺序来的？

我们给到您的 contig 编号并不代表顺序，是随机产生的，需要拿拼接到的序列跟参考序列比对才能确定前后顺序以及序列正反向。

33. 结果中的多条序列能否再帮我拼接一下看能否拼接成一条序列吗？

不可以的，结果中之所以会有多条序列，原因可能就是质粒中存在高 GC、重复区等复杂结构导致拼接不全，序列间存在 gap，所以基于现有数据是没办法再做进一步的拼接了。后续需要再拼接成完整序列的话需要对 gap 区域设计引物，之后进行一代测序。这部分工作我们公司也可以做，不过需要额外收费。

34. 拼接出来的质粒序列跟我的预期完全不一致，这是为什么？

这种情况可能是您样品给错了或者我们内部建库的时候出现了交叉污染，我们可以用原始样本再重新做一次，若新的结果跟第一次保持一致说明不是我们内部污染的问题，此时需要收两份样本的费用，若第二次结果跟第一次不一致且与您的预期一致，则就按第二次的结果来收费。

35. 拼接出来的序列部分区域缺失或者序列不对，这是为什么？

KBBSeq 不能保证重复序列、Homopolymer 区域的准确性，当在结果中发现存在疑似不准确的区域时，可通过常规一代测序来进行验证。生工也可做此工作，但是需要收费，并告知需要验证的区段位置及参考序列。

36. 拼接出来的序列大小跟我预期相差较大，为什么？

我们在实验之前会对每个质粒跑胶并把胶图留档，当出现此问题时看实际测出来的质粒大小跟胶图中的大小是否一致，若一致的话说明实验没有问题，此时需要正常收费。若实测大小跟胶图中大小不一致，此时我们可免费再重做一次，若还是有异常说明样本存在问题，实测质粒偏小的话可能是质粒浓度太低或者质粒存在复杂结构导致只测到部分序列；偏大的话可能是样本质粒不纯，含有宿主核酸或者其它别的核酸，杂质序列均会被测到，此时需要在结果中筛选一下目的序列。这两种情况均需要按照实测大小进行收费。

37. 出现拼接不全的问题，如何收费？

这种情况还是会收费，按照最终拼接出来的序列长度收费。

38. 突变能知道具体的碱基突变位置和突变序列吗？

突变结果见*.xlsx 文件，若存在突变，文件中会给出突变位置，突变碱基，还有突变比率等信息。若出现突变位点，最好用 Sanger 测序验证一下。

39. 给到的都是结果文件，能否把原始数据给我们呢？

可以的，我们默认不提供原始数据，若需要原始数据请联系我们技术支持，KBBSeq 的原始数据是二代测序的 Fastq 格式文件。

40. 能否提供给我们峰图文件？

不能的，因为 KBBSeq 采用的是二代测序，没有 Sanger 测序的峰图文件。