

从头设计抗体湿实验指南

Clickmab抗体从头设计生成模型的训练数据来源于百奥赛图全人源共轻链RenLite小鼠，经计算生成、筛选后的设计抗体是人源，共轻链的，为方便进行湿实验验证，最终的计算报告将设计抗体的氨基酸序列分为：重链HCDR3、重链FR1-FR3和共轻链。后续通过基因合成、展示筛选及重组表达验证获得结合指定表位的设计抗体。

1. 基因合成线性化scFv片段

完成计算设计后，在最终计算报告中下载并获得生成抗体重链的HCDR3和匹配的抗体重链骨架FR1-FR3，报告也包含抗体共轻链序列，这三部拼接成scFv片段，如图1。

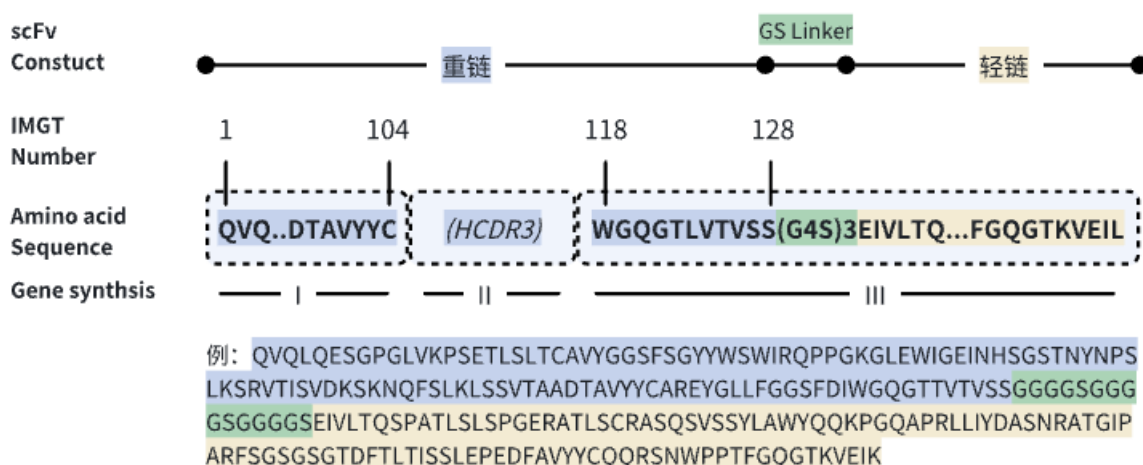


图1. 线性化scFv拼接图

基因合成可将scFv片段分为三部分合成，与图1中的编号对应，分别是：

- I：抗体重链FR1-FR3
- II：重链HCDR3
- III：重链FR4-GS Linker-VK

一个拼接好的线性化scFv的氨基酸序列，示例如下：

QVQLQESGPGLVKPSSETLSLTCAVYGGSFSGYYWSWIRQPPGKGLEWIGEINHSGSTNYPNPSLRSRTISVDKSKNQFSLKLSSVTAADTAVYYCAREYGLLFGGSFDIWGQGTITVTVSSGGGGSGGGGSEIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQQKPGQAPRLLIYDASNRATGIPARFSGSGSGTDFTLTSSLEPEDFAVYYCQQRSNWPPTFGQGTKVEIK

合成时两端需要添加酶切位点和保守碱基用于后续质粒构建，同时不同片段需要添加20bp左右的同源臂用于后续重叠延伸PCR拼接，此过程具体如下：

1. 对三部进行密码子优化，宿主为*E. coli*（注：以构建噬菌体展示筛选为例，理论上其它展示系统均可用于筛选，如酵母展示筛选，优化宿主为*Yeast*；另外，优化时，分片段优化，若拼接后优化，对于高量级

HCDR3库的密码子优化需要花费很长的时间)。

2. **基因合成第I部分**，即抗体重链FR1-FR3，计算通常会推荐出十几种、或二十几种匹配度高的Germline基因，采用常规基因合成方法进行合成，需注意基因合成时，5'端需要添加用于后续构建的酶切位点（如SfiI，以噬菌体载体pCANTAB5e为例，用SfiI和NotI酶切位点构建载体）。
3. **基因合成第II部分**，即抗体重链HCDR3，计算通常会生成并筛选出1万、20万和100万三种量级，可以根据需要选择不同量级的HCDR3进行合成，需注意更大的库容量筛选成功率更高，相应的合成成本会更高一些。考虑到重链HCDR3多样性高，且长度不是很长（<100nt），可以采用oligopool合成，合成时两端添加用于拼接的同源臂，5'端为抗体重链骨架FR3结尾的7个氨基酸，即DTAVYYC，（第I部不同FR1-FR3的结尾7个氨基酸都是一样的），3'端的同源臂选择Human抗体重链使用频率最高的J基因组成的FR4的前7个氨基酸，即WGQGTLV。合成完成后，用5'端和3'端的同源臂引物，即图2中的引物F1和引物R1，扩增单链oligopool产物得到双链DNA。需注意，oligopool的基因合成往往伴随一些随机错误，正常情况下错误率1/2000，如果扩增的循环次数过多，会带来偏好性和错误率增多的问题，因此，oligopool扩增双链后的产物满足后续重叠延伸的需求即可，不宜循环数过多。
4. **基因合成第III部分**，即抗体重链FR4-GS Linker-VK，这部分序列由使用频率最高的J基因组成的重链FR4、GS linker和共轻链组成，氨基酸序列是固定的，采用常规基因合成。需注意，5'端需要添加拼接使用的同源臂，即WGQGTLV，3'端需要添加用于后续构建的酶切位点（如NotI，以噬菌体载体pCANTAB5e为例，用SfiI和NotI酶切位点构建载体）。
5. **重叠延伸扩增拼接线性化的scFv片段**，将I、II和III这三部分通过不同片段之间的同源臂经重叠延伸PCR拼接起来，5'端扩增引物使用第I部分含SfiI酶切位点的序列，3'端引物使用第III部分含NotI酶切位点的序列，分别为图2中引物F和引物R。同oligopool双链扩增，过度扩增往往会产生偏好性与突变的问题，此步骤不宜扩增过度，满足后续酶切建库即可，一般达到15-30μg即可。

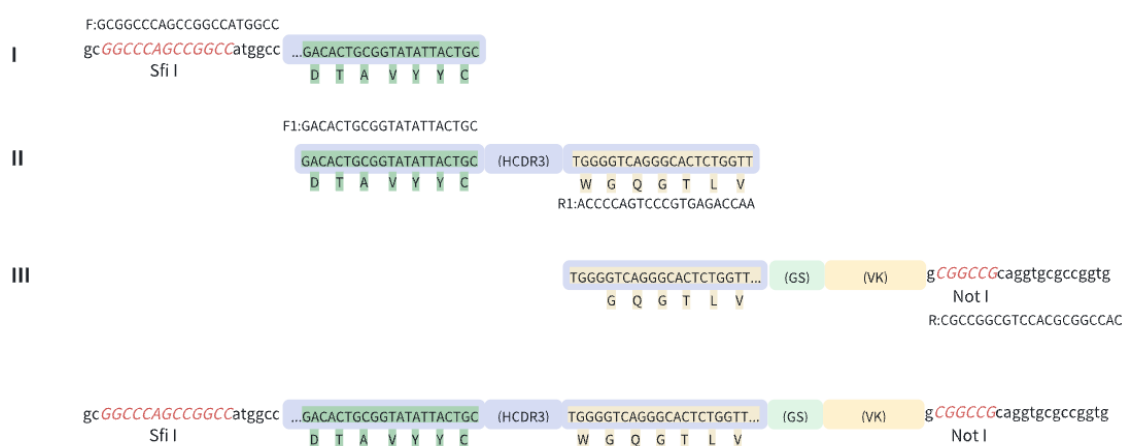


图2. 重叠延伸扩增线性化scFv片段示意图

6. **质控线性化scFv片段**，为保证筛选质量，线性化scFv片段需进行质控，估算线性化库中正确构建scFv的比例。oligo合成和拼接过程会出现阅读框移动、终止密码子和突变等问题，正常情况下正确构建的scFv比例达到60%以上才能保证后续筛选的质量，少量的错误构建对展示筛选的影响比较小。若进行Sanger测序，将线性化的片段酶切连接或平末端连接载体，铺平板后挑96个单克隆进行Sanger测序，再分析其中正确构建的比例，该方法比较简单快速，费用更低。也可以采用NGS测序评估，将线性化片段进行扩增子建库、测序及分析，该方法可以全面评估线性化库的正确构建比例，但耗时更长且费用更高。

2. 噬菌体建库及筛选

基因合成线性化scFv库后构建噬菌体展示库并进行噬菌体展示筛选，采用常规展示筛选的方法，简述如下：

1. **SfiI/NotI双酶切噬菌体载体和合成的线性化scFv片段**，切胶回收相应条带，然后连接载体与线性化片段。
2. **电转构建噬菌体库**，需注意电转库容尽量覆盖多样性的 10–100 倍，比如文库理论多样性为 10^6 ，构建的噬菌体库容应达到 10^7 – 10^8 。
3. **展示筛选**，可根据已有的材料进行海选方案的设计，建议多种筛选方法并行进行筛选，比如固、液或者交叉筛选，如果有细胞也可以同时进行细胞筛选。
4. **设置阳性菌库**，建议在筛选时设置阳性菌库，将已知能结合的抗体或Benchmark抗体构建展示噬菌体，与构建的筛选库以一定比例混合。若阳性抗体的亲和力很强，建议混合比例更高一些，比如1: 10000，反之，混合比例可以适当降低一些。
5. **筛选轮次及检测**，进行3轮以上筛选后，检测Titer，并通过phageELISA判断是否有富集；若富集不明显可以适当增加几轮筛选，若存在明显富集可进行单克隆筛选或（和）NGS测序获得阳性binder序列。
 - **5.1 单克隆筛选**：对有富集的库进行单克隆phage展示及筛选，对于阳性的phage进行Sanger测序，以获得筛选阳性的binder序列。需注意筛选阳性的序列在真核表达系统中可能存在不表达或表达后不结合的情况，如果实验流程正常，假阳性的比例应该小于10%。
 - **5.2 NGS测序**：对富集的库也可以采用NGS测序，由于筛选库是共轻链抗体，只需要扩增测序重链即可。测序完成后分析富集程度比较高的重链抗体，这些抗体大概率是有结合活性的抗体，但需进行进一步真核表达重组验证。

3. 设计抗体重组表达

通过噬菌体展示筛选获得的阳性binder序列需要进一步重组表达及验证，才能确认是阳性抗体。

1. **基因合成**：将阳性binder抗体的可变区序列进行基因合成并构建表达载体，通常情况下构建全长IgG的形式，比如hulgG1，kappa骨架，也可以根据实际情况构建需要的抗体形式（如scFv或Fab等）和适合的Fc。质粒载体通常使用pcDNA3.4、pTT5等常用表达载体。
2. **重组表达**：获得表达载体后进行CHO或HEK293真核表达系统表达。对于抗体初筛及鉴定，4mL CHO表达可以获得满足前期检测的抗体量（正常情况下，4mL CHO表达能获得200ng以上的纯化抗体，SDS/SEC纯度能达到90%以上，部分抗体表达量可能比较低）。对于表达量很低的抗体，如果能获得50ng以上的纯化抗体，也足够进行BLI或ELISA检测确认结合活性；需注意，对于重组抗体，可以通过表达量，纯度等理化性质粗略判断抗体的可开发性。
3. **结合活性检测**：获得纯化抗体后，进一步结合活性的检测可以根据靶点材料的情况进行实验。若靶点是可溶性抗原，可以进行ELISA/BLI/SPR检测，若靶点是多次跨膜抗原，也可以用过表达细胞系或肿瘤细胞进行FACS检测判断结合活性。

4. **表位验证**：若存在结合设计表位的Benchmark抗体或Ligand，可以采用竞争ELISA或Epitope Binning快速分析获得的阳性抗体是否结合在指定表位上；更进一步，可以通过Cryo-EM或X-ray等结构生物学方法判断设计抗体是否结合在指定表位上。