

序列模板填写指南

● 模板文件版本

通用模板

至多填写 65,536 条序列

下载: [TruPool_Universal.xlsx](#)

● 文件格式

文件格式为含有单张表单的 Excel 文件，您可以修改文件名，但请勿对表单名进行修改。

Oligo_ID_and_Seq 表单 含有四列信息：

Oligo_No	Oligo_ID	Oligo_Seq	Redundancy
----------	----------	-----------	------------

● 每种寡核苷酸的索引编号，0 代表示例，

1 至 4096/65536 代表用户需要合成的序列，
请勿对该列进行任何修改、增删。

● 每种寡核苷酸的名称，
相当于 FASTA 格式的 “>” 后的内容。

● 每种寡核苷酸的序列，
为 5' -> 3' 方向。

● 每种寡核苷酸的冗余度 / 拷贝数，
默认为 1。

● 填写和校验规则

Oligo_No	需为连续、不重复的正整数。
	不应该被修改。
Oligo_ID	如不想设置寡核苷酸的名称，可将所有行全部留空（序列将自动赋予 ID），但 不允许 部分填写、部分留空。
	不允许出现重复 的 Oligo_ID ，若该序列需要被合成多次，请设置 Redundancy ，而非将该序列的信息复制多行。
	允许出现的字符 包括大小写字母 (A-Z、a-z)、阿拉伯数字 (0-9)、部分字符 (+-=_.:)，禁止出现空格、全角字符（含中文）、换行符、制表符和其他特殊字符。
	基于多数生物信息软件的限制， Oligo_ID 的 第一个字符 必须为大小写字母，不可以是数字或特殊字符。
Oligo_Seq	允许出现的字符 包括大小写字母 (A-Z、a-z)、阿拉伯数字 (0-9)。
	如无特殊要求，尽量使用大写 A、C、G、T 表示 天然碱基 ，默认情况下，小写字母将被转换为大写。
	大写 R、Y、S、W、K、M、B、D、H、V、N 代表 简并碱基 ，默认情况下，小写字母将被转换为大写。
	您可以使用 N 代表 随机简并碱基 ，但如果需要合成非 N 的简并碱基类型（如 R、Y、S、W、K、M、B、D、H、V），请联系我们的客户经理进行说明和确认。
	非 A、C、G、T、R、Y、S、W、K、M、B、D、H、V、N 字符（含数字）可代表 特殊单体、特殊合成工艺或修饰基团 ，如需使用，请联系我们的客户经理进行说明和确认。
	如您的 大小写字母代表不同 的单体或合成工艺，请联系我们的客户经理进行说明和确认。
Redundancy	需为正整数。
	如不填写，则 Redundancy 默认设置为 1。
	Redundancy 列的和不应超通量上限（如 4096 或 65536）。

● 其它建议和规则

- [1] **Oligo_ID_and_Seq** 表单中，**每行代表一种唯一的寡核苷酸序列**；
比如用户需要使用 4K 通量芯片合成 1800 种不同序列，每种序列在芯片的不同坐标处合成 2 个拷贝，该表单中应填写 **Oligo_No** 为 1 至 1800 的行，
并将 **Redandancy** 设置为 2，
而非填写 **Oligo_No** 为 1 至 3600 的行。
- [2] **Oligo_No** 为 “0”、**Oligo_ID** 为 “Demo_Oligo” 的行为示例行，请勿进行修改；
该行内容不参与合成，不会出现在合成产物中。
- [3] 原则上，**Oligo_Seq** 应尽可能与 **Oligo_ID** 一一对应、保持无重复；
同一 **Oligo_ID** 只能对应一种 **Oligo_Seq**；
不同 **Oligo_ID** 对应相同 **Oligo_Seq** 的情况是被允许的，但我们依然建议您进行核实。
- [4] 应避免出现 **Oligo_ID** 为空但 **Oligo_Seq** 不为空的情况。

● 产物质量相关提示

- [1] 如果您对产物的产量有一定要求，或是下游实验涉及扩增，我们强烈建议您设置 **通用引物序列**。
- [2] 使用同一对引物进行扩增的一个寡核苷酸池或者一个子池 (subpool) 内，如果不同序列 **长度差异** 较大，可能会造成合成载量的差异，以及较严重的 **扩增偏好性** 差异，使扩增产物的 **均一性** 受到较大影响。
- [3] 如果您需要使用 **polyX** 序列补齐序列长度，**dT** 是较为推荐的碱基类型。
- [4] 过多的 **dA** 或 **dG** 碱基含量，可能对合成质量造成影响。
- [5] 过高、过低、不同序列间差异较大的 **CG% 含量**，可能使扩增产物的 **均一性** 受到较大影响。

