

PCR 引物设计及软件使用技巧

张新宇, 高燕宁*

(中国协和医科大学中国医学科学院肿瘤研究所, 北京 100021)

摘要: 本文旨在介绍使用软件设计 PCR 引物的技巧。在 PCR 引物设计原则的基础上, 详细介绍了两种常用引物设计软件的基本使用方法, 并对其各自的优缺点进行了比较。一般性引物自动搜索可采用“Premier Primer 5”软件, 而引物的评价分析则可采用“Oligo 6”软件。

关键词: PCR; 引物设计

中图分类号: Q524

To design PCR primers with Oligo 6 and Primer Premier 5

Zhang Xinyu, Zhu Youkang, Gao Yanning

(Cancer Institute, Chinese Academy of Medical Sciences, Beijing, 100021, China)

Abstract: The skill of PCR primer design with software is introduced in this paper. Based on the principal of PCR-primer design, the usage of two kinds of popular primer-design software was reviewed detailedly, including their merit, demerit and usage skills. It is recommended that to use Premier Primer 5 does the usual automatic primers search, but Oligo 6 primers analysis.

Key Words: PCR primer; design; usage skill;

自从 1985 年 Kary Mullis 发明了聚合酶链式反应以来, PCR 技术已成为分子生物学研究中使用最多、最广泛的手段之一^[1], 而引物设计是 PCR 技术中至关重要的一环。使用不合适的 PCR 引物容易导致实验失败: 表现为扩增出目的带之外的多条带 (如形成引物二聚体带), 不出带或出带很弱, 等等。

现在 PCR 引物设计大都通过计算机软件进行。可以直接提交模板序列到特定网页, 得到设计好的引物, 也可以在本地计算机上运行引物设计专业软件。一般来说, 专门进行 PCR 引物设计的专业软件功能更为强大, 但使用起来却不太容易。本文将就引物设计原则及软件使用问题进行探讨。

1. 引物设计的原则

引物设计有 3 条基本原则: 首先引物与模板的序列要紧密互补, 其次引物与引物之间避免形成稳定的二聚体或发夹结构, 再次引物不能在模板的非目的位点引发 DNA 聚合反应 (即错配)。

具体实现这 3 条基本原则需要考虑到诸多因素, 如引物长度 (primer length), 产物长度

(product length), 序列 Tm 值 (melting temperature), 引物与模板形成双链的内部稳定性 (internal stability, 用 ΔG 值反映), 形成引物二聚体 (primer dimer) 及发夹结构 (duplex formation and hairpin) 的能值, 在错配位点 (false priming site) 的引发效率, 引物及产物的 GC 含量 (composition), 等等。必要时还需对引物进行修饰, 如增加限制性内切酶位点, 引进突变等。根据有关参考资料和笔者在实践中的总结, 引物设计应注意如下要点:

1. 引物的长度一般为 15-30 bp, 常用的是 18-27 bp, 但不应大于 38, 因为过长会导致其延伸温度大于 74°C, 不适于 Taq DNA 聚合酶进行反应^[2]。
2. 引物序列在模板内应当没有相似性较高, 尤其是 3'端相似性较高的序列, 否则容易导致错配。引物 3'端出现 3 个以上的连续碱基, 如 GGG 或 CCC, 也会使错误引发机率增加^[2]。
3. 引物 3'端的末位碱基对 Taq 酶的 DNA 合成效率有较大的影响。不同的末位碱基在错配位置导致不同的扩增效率, 末位碱基为 A 的错配效率明显高于其他 3 个碱基, 因此应当避免在引物的 3'端使用碱基 A^{[3][4]}。另外, 引物二聚体或发夹结构也可能导致 PCR 反应失败。5'端序列对 PCR 影响不太大, 因此常用来引进修饰位点或标记物^[2]。
4. 引物序列的 GC 含量一般为 40-60%, 过高或过低都不利于引发反应。上下游引物的 GC 含量不能相差太大^{[2][5]}。
5. 引物所对应模板位置序列的 Tm 值在 72°C 左右可使复性条件最佳。Tm 值的计算有多种方法, 如按公式 $Tm = 4(G+C) + 2(A+T)$, 在 Oligo 软件中使用的是最邻近法 (the nearest neighbor method)^{[6][7]}。
6. ΔG 值是指 DNA 双链形成所需的自由能, 该值反映了双链结构内部碱基对的相对稳定性。应当选用 3'端 ΔG 值较低 (绝对值不超过 9), 而 5'端和中间 ΔG 值相对较高的引物。引物的 3'端的 ΔG 值过高, 容易在错配位点形成双链结构并引发 DNA 聚合反应^[6]。
7. 引物二聚体及发夹结构的能值过高 (超过 4.5kcal/mol) 易导致产生引物二聚体带, 并且降低引物有效浓度而使 PCR 反应不能正常进行^[8]。
8. 对引物的修饰一般是在 5'端增加酶切位点, 应根据下一步实验中要插入 PCR 产物的载体的相应序列而确定。







值得一提的是, 各种模板的引物设计难度不一。有的模板本身条件比较困难, 例如 GC 含量偏高或偏低, 导致找不到各种指标都十分合适的引物; 在用作克隆目的的 PCR 因为产物序列相对固定, 引物设计的选择自由度较低。在这种情况下只能退而求其次, 尽量去满足条件。

2. 引物的自动搜索和评价分析

软件的引物设计功能主要体现在两个方面: 首先是引物分析评价功能, 该功能只有少数商业版软件能够做到, 其中以“Oligo 6”最优秀; 其次是引物的自动搜索功能, 各种软件在这方面的侧重点不同, 因此自动搜索的结果也不尽相同。据笔者的经验, 自动搜索功能以“Premier Primer”为最强且方便使用, “Oligo 6”其次, 其他软件如“Vector NTI Suit”、“Dnasis”、“Omiga”和“Dnastar”都带有引物自动搜索功能, 但搜索结果不是十分理想。要想得到效果很好的引物, 在自动搜索的基础上还要辅以人工分析。笔者认为引物设计软件的最佳搭配是“Oligo”和“Premier”软件合并使用, 以“Premier”进行自动搜索, “Oligo”进行分析评价, 如此可快速设计出成功率很高的引物。

2.1 Primer Premier 5.0 的使用技巧简介

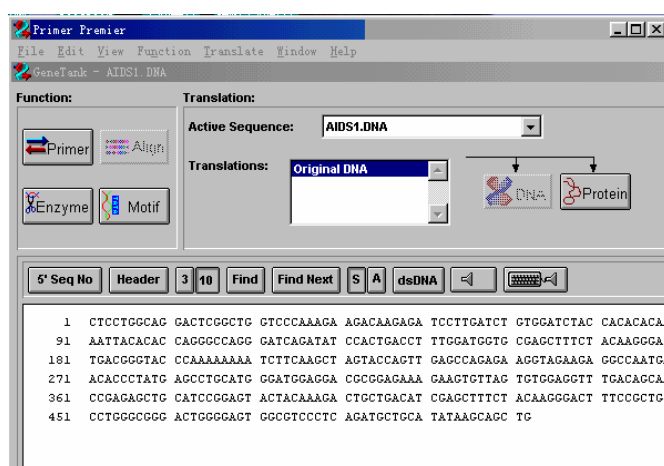
2.1.1 功能

“Premier”的主要功能分四大块，其中有三种功能比较常用，即引物设计 ()、限制性内切酶位点分析 () 和 DNA 基元 (motif) 查找 ()。“Premier”还具有同源性分析功能 ()，但并非其特长，在此略过。此外，该软件还有一些特殊功能，其中最重要的是设计简并引物，另外还有序列“朗读”、DNA 与蛋白序列的互换 ()、语音提示键盘输入 () 等等。


有时需要根据一段氨基酸序列反推到 DNA 来设计引物，由于大多数氨基酸 (20 种常见结构氨基酸中的 18 种) 的遗传密码不只一种，因此，由氨基酸序列反推 DNA 序列时，会遇到部分碱基的不确定性。这样设计并合成的引物实际上是多个序列的混和物，它们的序列组成大部分相同，但在某些位点有所变化，称之为简并引物。遗传密码规则因物种或细胞亚结构的不同而异，比如在线粒体内的遗传密码与细胞核是不一样的。“Premier”可以针对模板 DNA 的来源以相应的遗传密码规则转换 DNA 和氨基酸序列。软件共给出八种生物亚结构的不同遗传密码规则供用户选择，有纤毛虫大核 (Ciliate Macronuclear)、无脊椎动物线粒体 (Invertebrate Mitochondrion)、支原体 (Mycoplasma)、植物线粒体 (Plant Mitochondrion)、原生动物线粒体 (Protozoan Mitochondrion)、一般标准 (Standard)、脊椎动物线粒体 (Vertebrate Mitochondrion) 和酵母线粒体 (Yeast Mitochondrion)。

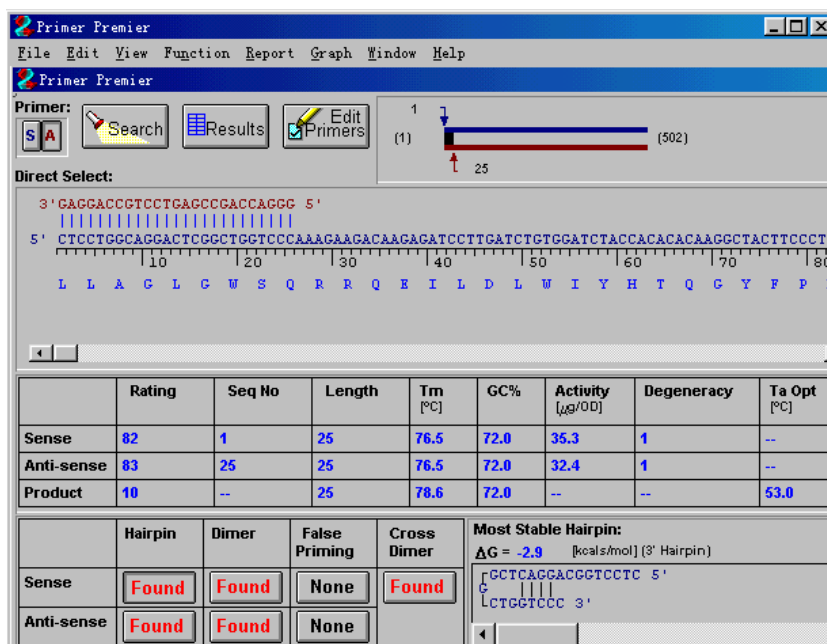
2.1.2 使用步骤及技巧




“Premier”软件启动界面如下：



其主要功能在主界面上一目了然 (按钮功能如上述)。限制性酶切点分析及基元查找功能比较简单，点击该功能按钮后，选择相应的限制性内切酶或基元 (如-10 序列，-35 序列等)，按确定即可。常见的限制性内切酶和基元一般都可以找到。你还可以编辑或者添加新限制性内切酶或基元。

进行引物设计时，点击  按钮，界面如下：



进一步点击  按钮，出现“search criteria”窗口，有多种参数可以调整。搜索目的 (Search For) 有三种选项，PCR 引物 (PCR Primers)，测序引物 (Sequencing Primers)，杂交探针 (Hybridization Probes)。搜索类型 (Search Type) 可选择分别或同时查找上、下游引物 (Sense/Anti-sense Primer，或 Both)，或者成对查找 (Pairs)，或者分别以适合上、下游引物为主 (Compatible with Sense/Anti-sense Primer)。另外还可改变选择区域 (Search Ranges)，引物长度 (Primer Length)，选择方式 (Search Mode)，参数选择 (Search Parameters) 等等。使用者可根据自己的需要设定各项参数。如果没有特殊要求，建议使用默认设置。然后按 ，随之出现的 Search Progress 窗口中显示 Search Completed 时，再按 ，这时搜索结果以表格的形式出现，有三种显示方式，上游引物 (Sense)，下游引物 (Anti-sense)，成对显示 (Pairs)。默认显示为成对方式，并按优劣次序 (Rating) 排列，满分为 100，即各指标基本都能达标 (如下图)。

Search Results

Sense
 Anti-sense
 Pairs

100 pairs found.

#	Rating	Seq No	Length	Tm [°C]	GC%	Product Size	Ta Opt [°C]	Degeneracy	Mark
1	94	302 493	19 18	49.8 50.6	47.4 50.0	192	52.8	1.0 1.0	<input type="checkbox"/>
2	88	17 294	18 18	51.9 53.7	55.6 50.0	278	52.4	1.0 1.0	<input type="checkbox"/>
3	88	17 287	18 19	51.9 53.3	55.6 52.6	271	52.4	1.0 1.0	<input type="checkbox"/>
4	85	231 407	18 18	49.3 50.1	50.0 50.0	177	51.0	1.0 1.0	<input type="checkbox"/>
5	84	13 145	18 18	58.8 59.3	61.1 61.1	133	53.8	1.0 1.0	<input type="checkbox"/>

点击其中一对引物，如第 1#引物，并把上述窗口挪开或退出，显示“Primer Premier”主窗口，如图所示：

Primer Premier

Direct Select:

(1)

3' CCGACTCTACGACCTATA 5'

5' TTCGCTGGGGACTTTCCAGGGAGCCCTGGCCCTGGGCGGACTGGGGAGTGGCGTCCCTCAGATGCTGCATATAAGCAGCTG 3'

430 440 450 460 470 480 490 500



F R W C L S R E A W P G R D W G V A S L R C C I - A A

	Rating	Seq No	Length	Tm [°C]	GC%	Activity [μg/00]	Degeneracy	Ta Opt [°C]
Sense	100	302	19	49.8	47.4	29.7	1	--
Anti-sense	88	493	18	50.6	50.0	31.2	1	--
Product	94	--	192	89.7	55.2	--	--	52.8

	Hairpin	Dimer	False Priming	Cross Dimer	No Hairpins Found
Sense	<input type="button" value="None"/>	<input type="button" value="None"/>	<input type="button" value="None"/>	<input type="button" value="None"/>	<input type="button" value="All"/>
Anti-sense	<input type="button" value="None"/>	<input type="button" value="Found"/>	<input type="button" value="None"/>	<input type="button" value="None"/>	

该图分三部分，最上面是图示 PCR 模板及产物位置，中间是所选的上下游引物的一些性质，最下面是四种重要指标的分析，包括发夹结构（Hairpin），二聚体（Dimer），错误引发情况（False Priming），及上下游引物之间二聚体形成情况（Cross Dimer）。当所分析的引物有这四种结构的形成可能时，按钮由 **None** 变成 **Found**，点击该按钮，在左下角的窗口中就会出现该结构的形成情况。一对理想的引物应当不存在任何一种上述结构，因此最好的情况是最下面的分析栏没有 **Found**，只有 **None**。值得注意的是中间一栏的末尾

给出该引物的最佳退火温度，可参考应用。

在需要对引物进行修饰编辑时，如在 5'端加入酶切位点，可点击 ，然后修改引物序列。若要回到搜索结果中，则点击  按钮。

如果要设计简并引物，只需根据源氨基酸序列的物种来源选择前述的八种遗传密码规则，反推至 DNA 序列即可。对简并引物的分析不需像一般引物那样严格。

总之，“Premier”有优秀的引物自动搜索功能，同时可进行部分指标的分析，而且容易使用，是一个相当不错的软件。

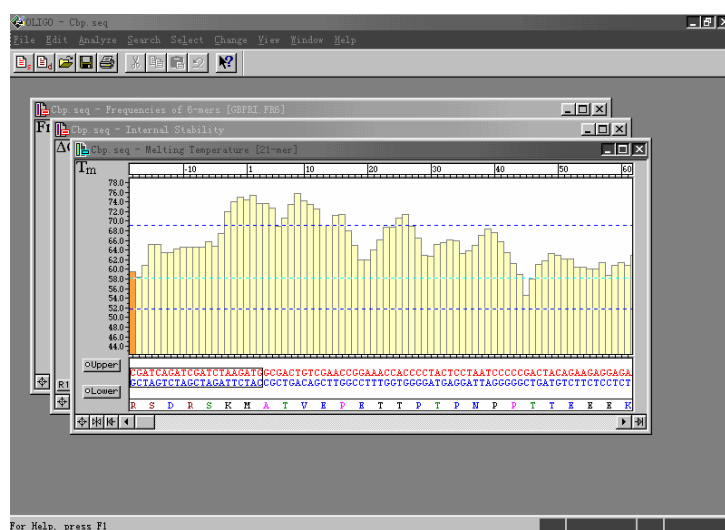
2.2 Oligo 6.22 使用技巧简介

2.2.1 功能

在专门的引物设计软件中，“Oligo”是最著名的。它的使用并不十分复杂，但初学者容易被其复杂的图表吓倒。Oligo 5.0 的初始界面是两个图：Tm 图和 ΔG 图；Oligo 6.22 的界面更复杂，出现三个图，加了个 Frq 图。“Oligo”的功能比“Premier”还要单一，就是引物设计。但它的引物分析功能如此强大以至于能风靡全世界。

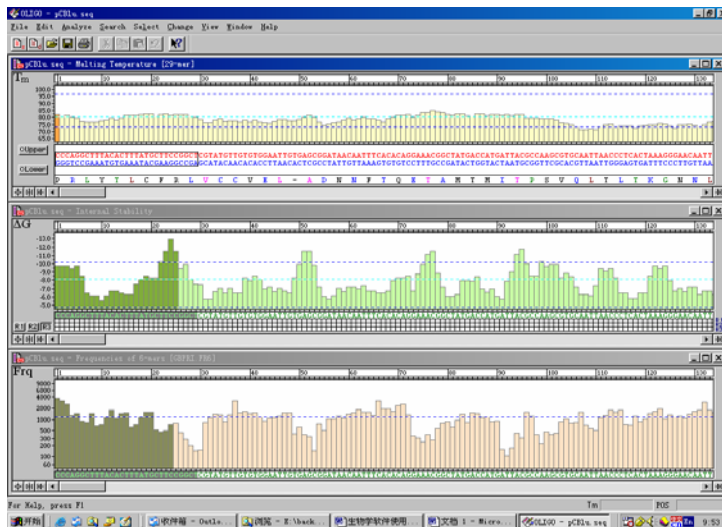
2.2.2 使用（以 Oligo 6.22 为例）

Oligo 6.22 的启动界面如下：

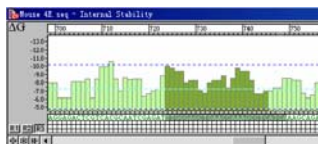


图中显示的三个指标分别为 Tm、 ΔG 和 Frq，其中 Frq 是 6.22 版本的新功能，为邻近 6 至 7 个碱基组成的亚单位在一个指定数据库文件中的出现频率。该频率高则可增加错误引发的可能性。因为分析要涉及多个指标，启动窗口的 cascade 排列方式不太方便，可从 windows 菜单改为 tili 方式。如果觉得太拥挤，可去掉一个指标，如 Frq，这样界面的结构同于 Oligo 5.0，只是显示更清楚了。

经过 Windows/Tili 项后的显示如图：



在设计时，可依据图上三种指标的信息选取序列，如果觉得合适，可点击 Tm 图块左上下角的 Upper 按钮 ，选好上游引物，此时该按钮变成 ，表示上游引物已选取好。下游引物的选取步骤基本同上，只是按钮变成 Lower。 ΔG 值反映了序列与模板的结合强度，最好引物的 ΔG 值在 5'端和中间值比较高，而在 3'端相对低（如图：）



Tm 值曲线以选取 72°C 附近为佳，5'到 3'的下降形状也有利于引物引发聚合反应。Frq 曲线为“Oligo 6”新引进的一个指标，揭示了序列片段存在的重复机率大小。选取引物时，宜选用 3'端 Frq 值相对较低的片段。

当上下游引物全选好以后，需要对引物进行评价并根据评价对引物进行修改。首先检查引物二聚体尤其是 3'端二聚体形成的可能性。需要注意的是，引物二聚体有可能是上游或下游引物自身形成，也有可能是在上下游引物之间形成（cross dimer）。二聚体形成的能值越高，越不符合要求。一般的检测（非克隆）性 PCR，对引物位置、产物大小要求较低，因而应尽可能选取不形成二聚体或其能值较低的引物。第二项检查是发夹结构（hairpin）；与二聚体相同，发夹结构的能值越低越好。一般来说，这两项结构的能值以不超过 4.5 为好。当然，在设计克隆目的的 PCR 引物时，引物两端一般都添加酶切位点，必然存在发夹结构，而且能值不会太低。这种 PCR 需要通过灵活调控退火温度以达到最好效果，对引物的发夹结构的检测就不应要求太高。第三项检查为 GC 含量，以 45-55%为宜。有一些模板本身的 GC 含量偏低或偏高，导致引物的 GC 含量不能被控制在上述范围内，这时应尽量使上下游引物的 GC 含量以及 Tm 值保持接近，以有利于退火温度的选择。如果 PCR 的模板不是基因组 DNA，而是一个特定模板序列，那么最好还进行 False priming site 的检测。这项检查可以看出引物在非目的位点引发 PCR 反应的可能性。如果引物在错配位点的引发效率比较高，就可能出假阳性的 PCR 结果。一般在错配引发效率以不超过 100 为好，但对于特定的模板序列，还应结合比较其在正确位点的引发效率。如果两者相差很大，比如在正确位点的引发效率为 450 以上，而在错误位点的引发效率为 130，那么这对引物也是可以接受的。

当我们结束以上四项检测，按 Alt+P 键弹出 PCR 窗口，其中总结性地显示该引物的位置、产物大小、Tm 值等参数，最有用的是还给出了推荐的最佳退火温度和简单的评价。

由于“Oligo”软件的引物自动搜索功能与“Primer Premier 5”的相类似，并且似乎并

不比后者更好用，在此不再赘述。其实，使用软件自动索引引物就是让计算机按照人的要求去寻找最佳引物，如果参数设置得当将大大提高工作效率。

除了本地引物设计软件之外，现在还有一些网上引物设计软件，如由 Whitehead Institute 开发的“Primer 3”等（本网站 <http://210.72.11.60> 已引进并调试好该软件，欢迎使用）。该软件的独特之处在于，对全基因组 PCR 的引物设计；可以将设计好的引物对后台核酸数据库进行比对，发现并排除可引发错配的引物。因此建议经常做全基因组 PCR 的用户试用。

参考文献 (References):

- (1) H.A.艾得希主编,田丁等译. PCR 技术——DNA 扩增的原理与应用. 北京:北京医科大学中国协和医科大学联合出版社,1991.
- (2) 林万明著. PCR 技术操作与应用指南. 北京:人民军医出版社,1993.
- (3) 朱平主编. PCR 基因扩增实验操作手册. 北京: 中国科学技术出版社, 1992
- (4) 郑仲承. 寡核苷酸的优化设计. 生命的化学, 2001,21(3):254~256.
- (5) George H.Keller Mark M.Manak. DNA probes. 1992.
- (6) Oligo 软件帮助文件(Oligo.HLP).
- (7) Martin, F.H., Castro, M.M., and Tinoco, Jr. I. Base pairing involving deoxyinosine, implications for probe design. Nucleic Acids Res, 1985, 13: 8927~8938.
- (8) Rychlik, W. and Rhoads, R.E. A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA. Nucleic Acids Res, 1989,17: 8543~8551.