

如何设计 PCR 引物

查找目的基因序列并不能直接用于我们研究所用，因为查找到的基因序列只能算是基因的电子序列，并不是我们研究要用的现实中的序列，如何获得现实的基因序列呢？这就需要我们根据基因的电子序列来设计引物了，通过引物借助 PCR 方法才能获得我们的目的基因序列。接下来，我们将介绍一下 PCR 方法并讲述如何设计引物。

PCR (polymerase chain reaction)，即聚合酶链反应，又称基因体外扩增技术，是由一对引物介导、能在体外对特定 DNA 片段进行快速酶促扩增的技术。PCR 能把很微量的遗传物质在数小时内扩增数百万倍达到检测水平。使原来无法进行分析和检测的许多项目得以完成。PCR 引物设计的目的是找到一对合适的核苷酸片段，使其能有效地扩增模板 DNA 序列。引物的优劣直接关系到 PCR 的特异性与成功与否。对引物的设计不可能有一种包罗万象的规则确保 PCR 的成功，但遵循某些原则，则有助于引物的设计。

1. 引物的特异性

引物与非特异扩增序列的同源性不要超过 70%或有连续 8 个互补碱基同源。

2. 避开产物的二级结构区

某些引物无效的主要原因是引物重复区 DNA 二级结构的影响，选择扩增片段时最好避开二级结构区域。用有关计算机软件可以预测估计 mRNA 的稳定二级结构，有助于选择模板。实验表明，待扩区域自由能 (ΔG°) 小于 58.6kJ/mol 时，扩增往往不能成功。若不能避

开这一区域时，用 7-deaza-2'-脱氧 GTP 取代 dGTP 对扩增的成功是有帮助的。

3. 长度

寡核苷酸引物长度为 15~30bp，一般为 20~27mer。引物的有效长度： $Ln=2(G+C) + (A+T)$ Ln 值不能大于 38，因为 >38 时，最适延伸温度会超过 Taq DNA 聚合酶的最适温度 (74℃)，不能保证产物的特异性。

4. G+C 含量

G+C 含量一般为 40%~60%。其 Tm 值是寡核苷酸的解链温度，即在一定盐浓度条件下，50%寡核苷酸双链解链的温度，有效启动温度，一般高于 Tm 值 5~10℃。若按公式 $Tm=4(G+C)$

+2 (A+T) 估计引物的 T_m 值，则有效引物的 T_m 为 55~80°C，其 T_m 值最好接近 72°C 以使复性条件最佳。

5. 碱基随机分布

引物中四种碱基的分布最好是随机的，不要有聚嘌呤或聚嘧啶的存在。尤其 3' 端不应超过 3 个连续的 G 或 C，因这样会使引物在 G+C 富集序列区错误引发。

6. 引物自身

引物自身不应存在互补序列，否则引物自身会折叠成发夹状结构而引物本身复性。这种二级结构会因空间位阻而影响引物与模板的复性结合。若用人工判断，引物自身连续互补碱基不能大于 3bp。

7. 引物之间

两引物之间不应有互补性，尤应避免 3' 端的互补重叠以防引物二聚体的形成。一对引物间不应多于 4 个连续碱基的同源性或互补性。

8. 引物的 3' 端

引物的延伸是从 3' 端开始的，不能进行任何修饰。3' 端也不能有形成任何二级结构可能，除在特殊的 PCR (AS-PCR) 反应中，引物 3' 端不能发生错配。

在标准 PCR 反应体系中，用 2U Taq DNA 聚合酶和 800 mol/L dNTP (四种 dNTP 各 200 mol/L) 以质粒 (103 拷贝) 为模板，按 95°C, 25s; 55°C, 25s; 72°C, 1min 的循环参数扩增 HIV-1 gag 基因区的条件下，引物 3' 端错配对扩增产物的影响是有一定规律的。A : A 错配使产量下降至 1/20, A : G 和 C : C 错七下下降至 1/100。引物 A : 模板 G 与引物 G : 模板 A 错配对 PCR 影响是等同的。

9. 引物的 5' 端

引物的 5' 端限定着 PCR 产物的长度，它对扩增特异性影响不大。因此，可以被修饰而不影响扩增的特异性。引物 5' 端修饰包括：加酶切位点；标记生物素、荧光、地高辛、Eu3+ 等；引入蛋白质结合 DNA 序列；引入突变位点、插入与缺失突变序列和引入一启动子序列等。

10. 密码子的简并

如扩增编码区域，引物 3' 端不要终止于密码子的第 3 位，因密码子的第 3 位易发生简并，会影响扩增特异性与效率。

随着人们对引物的认识，一些引物的计算机设计程序也应运而生，下面将讨论有关引物计算机设计方法。现有的引物设计软件有 primer5.0 比较好，现介绍一下 primer5.0 及其使用方法：

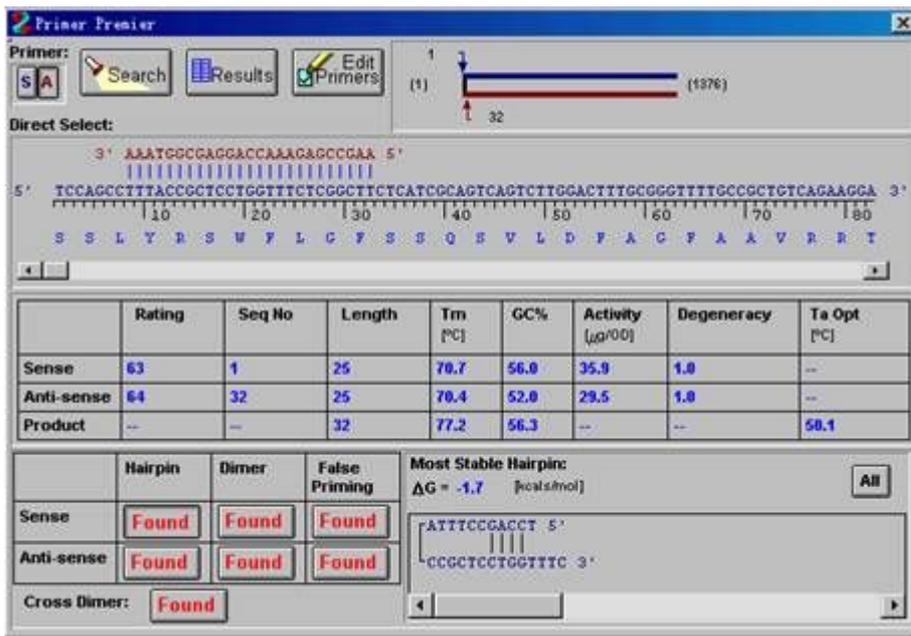
Primer5.0 中文使用说明

Primer Premier5.0 是由加拿大的 Premier 公司开发的专业用于 PCR 或测序引物以及杂交探针的设计，评估的**软件**，和 Plasmid Premier2.02 一起是该公司推出的最新的**软件**产品。其主要界面同样也是分为序列编辑窗口（Genetank），引物设计窗口（Primer Design），酶切分析窗口（Restriction Sites）和纹基分析窗口（Motif）。这里我们主要介绍其引物设计功能，其他功能的介绍请参看 Plasmid Premier2.02。



打开程序首先进入的是序列编辑参看，与 Plasmid Premier 相比，其多了一个语音校正的功能，即在输入序列的时候，程序自动将碱基读出，以使用户进行校正，保证输入的正确和快速。

点击该界面的按钮即可进入到程序的引物设计窗口。



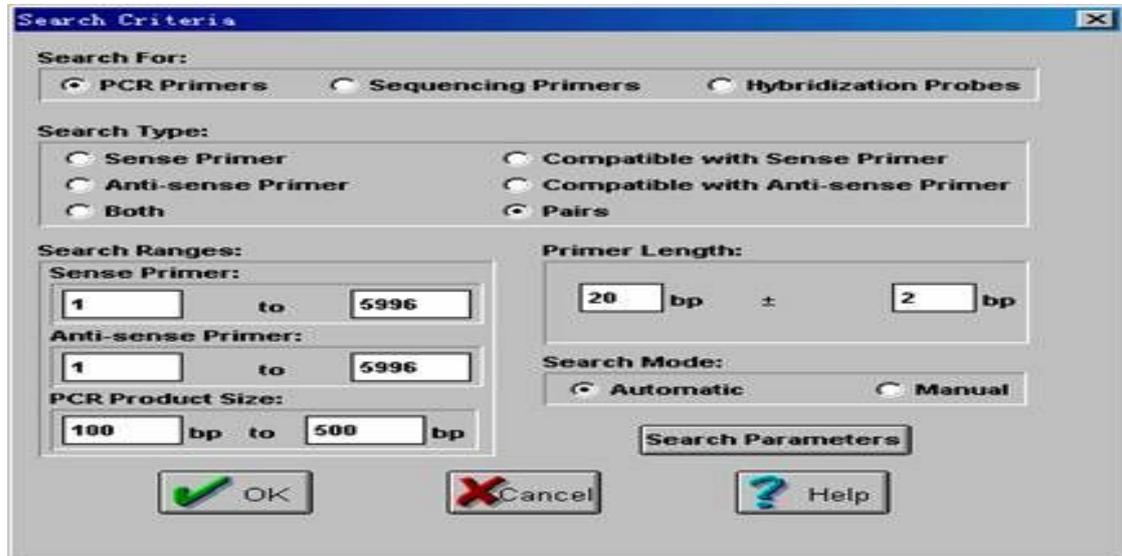
The screenshot shows the Primer Premier software interface. At the top, there are buttons for 'Search', 'Results', and 'Edit Primers'. Below this is a 'Direct Select' section with a sequence editor showing a DNA sequence: 3' AAATCGCGAGG&CCAAAC&GCGAA 5' and its complement 5' TCCAGCGTTTACCGCTCCTGGTTTCTCGGTTTCTCATCCGAGTCACTTTCGACTTTGGCGTTTTCGGCTCTCAGAAGGA 3'. A scale from 10 to 80 is visible below the sequence. Below the sequence is a table with the following data:

	Rating	Seq No	Length	Tm [°C]	GC%	Activity [μg/00]	Degeneracy	Ta Opt [°C]
Sense	63	1	25	70.7	56.0	35.9	1.0	--
Anti-sense	64	32	25	70.4	52.0	29.5	1.0	--
Product	--	--	32	77.2	56.3	--	--	50.1

Below the table are sections for 'Hairpin', 'Dimer', 'False Priming', and 'Cross Dimer', all marked as 'Found'. A 'Most Stable Hairpin' section shows a hairpin structure with $\Delta G = -1.7$ [kcal/mol].

该界面共分为四层，最上面一层左面是 5 个控制按钮，用于实现引物设计中的各种功能，包括引物自动寻找，寻找结果查看和引物编辑；右边是观察两个引物在模板上结合位置的直观图以及对正链还是负链引物进行选择；第二层是显示模板和引物序列及二者间的配对情况的显示；第三层是显示两个引物的各种参数，包括给引物的打分，引物以及产物的起始位置、

长度、T_m 值、GC%消光系数、简并性；最后一层是给出的有关于引物的二聚体结构、发卡结构、错配情况和引物间二聚体结构的预测，左边是显示是否存在以上各种对 PCR 扩增有影响的 结构，右边显示的是这些结构的位置，结构细节和稳定能，利用这些参数可以对引物作出可靠的评价。

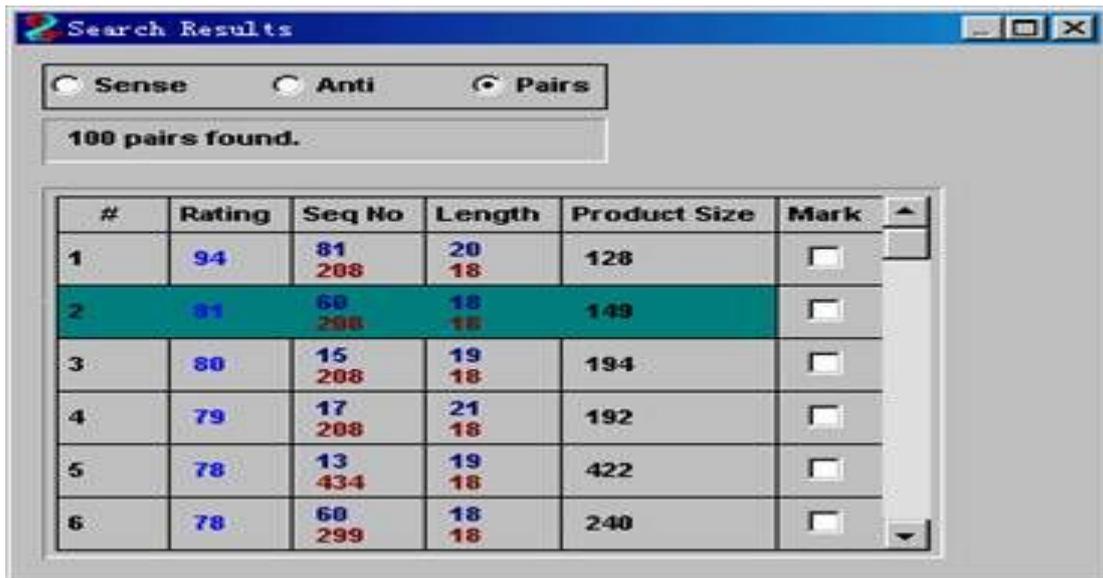


下面是根据模板序列寻找引物的界面，在该界面中可以设定所要搜索的引物的类型，包括 PCR 引物，测序引物和杂交探针以及引物所在的链；另外也能设定搜索引物的范围，以及最终 PCR 产物的长度和引物的长度等。

Search Parameters 并过来点击按钮来设置一些搜寻参数：

这些参数包括引物的 T_m 值，GC 比，有简并性碱基，3'端稳定性，引物的稳定性，重复序列，二聚体/发卡结构和与模板及可能的杂质 DNA（需要从另外的序列文件中读入）之间的错配情况，这些参数的设定可以根据要求变化，程序本身根据一定的标准分成从极高严谨性到极低严谨性 5 个档次。该程序对引物的自动搜索过程是采用的排除法，根据用户设定的引物范围和产物长度所规定区域内，所有的符合设定的引物长度的寡核苷酸都被视为可能的引物，然后根据以上各个参数逐步去除不符合条件的寡核苷酸，经过这种层层剔除的筛选，最终找到符合用户所设定标准的引物，如果找不到合适的引物可以逐步降低要求来进行进一步搜寻。

搜寻到引物结果最终显示在下面的窗口中：



#	Rating	Seq No	Length	Product Size	Mark
1	94	81 208	20 18	128	<input type="checkbox"/>
2	81	60 288	18 18	148	<input type="checkbox"/>
3	80	15 208	19 18	194	<input type="checkbox"/>
4	79	17 208	21 18	192	<input type="checkbox"/>
5	78	13 434	19 18	422	<input type="checkbox"/>
6	78	60 299	18 18	240	<input type="checkbox"/>

在结果窗口中给出了程序给该对引物的打分（rating）和上下游引物的起始位置和长度以及产物的长度。通过直接点击各对引物在相应引物搜寻界面中相应的显示引物的各种信息。包括其各种参数和各种可能存在的不利结构。

Single Primer Rating Formula:

$$\text{Rating} = 100 - ((\Delta G (\text{Dimer}) * 1.7) + (\Delta G (\text{Hairpin}) * 1.4) + (\Delta G (\text{False Priming}) * 1.5))$$

Primer Pair Rating Formula:

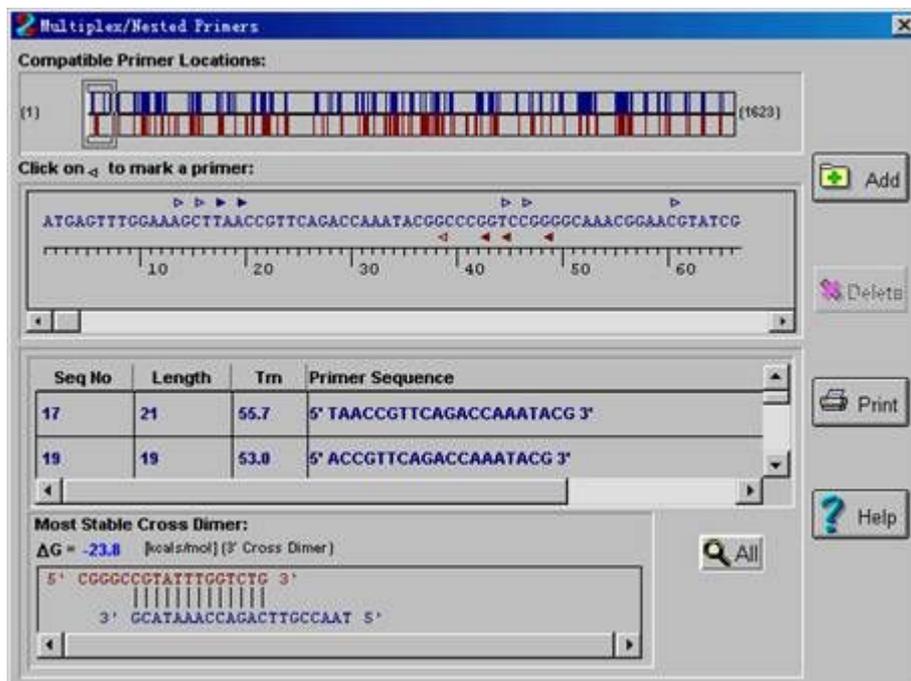
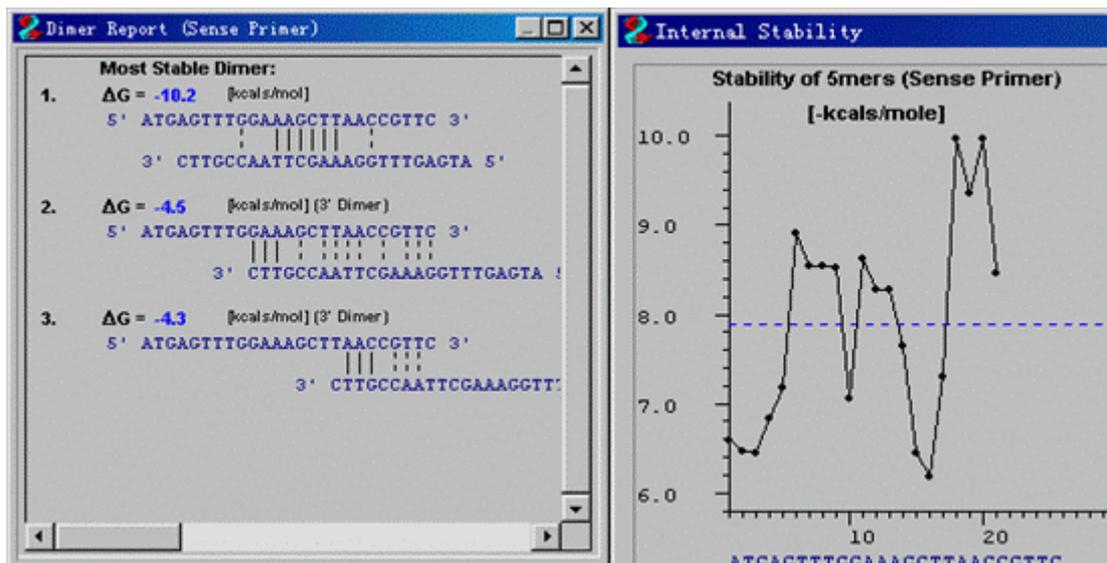
$$\text{Rating} = (\text{Average of sense \& anti-sense primer rating} - (\text{Tm} * 1.7)) - (\Delta G (\text{CrossDimer}) * 1.3)$$

其中用 File 菜单中的 Parameter 命令可以对系统参数，PCR 反应条件和程序打分系统进行设置，以帮助用户根据自己的特殊要求来进行引物设计，其中反应条件按照引物浓度，单价离子浓度，Mg 离子浓度，Na 盐浓度等。利用该打分系统可以为对引物进行有效评估和在引物之间进行对比选择提供有效的有效的依据。



最后还可以在利用引物编辑窗口对程序自动找到的或手工找到的引物序列进行编辑，以便用户能在引物引入突变及加入特定的酶切位点，并对修改后的序列的二聚体结构、发卡结构和错配进一步评估。其中是对修改后的引物再次搜寻找出其最合适的引发位置。

除了以上的直观地对引物进行基本的设计和评估功能外，该程序还提供了多种数据报告，主要通过 Report 菜单和 Graph 菜单中命令来实现：



另外，Primer Premier 在引物设计上一个比较独到的功能，就是能辅助进行巢式 PCR 引物设计，在首先进行了引物自动搜索后，即可通过 Function 菜单中 Multiplex/Nested Primer 命令进入巢式 PCR 引物设计，通过点中代表各个引物的三角号来选择引物，然后根据最下面的窗口中各个引物之间的二聚体形成情况来判断引物是否适合于作为巢式 PCR 的引物。另外该菜单中的 Database 命令可以让用户建立自己的引物数据库，方便于查找和对照。

整体而言，Primer Premier 这个软件在引物设计方面还是比较优秀的，能够对全面的给出一

个引物的各种参数，并在多个必需的方面对引物作出有效的评价，但是在程序设计中还有一些不足，例如在错配情况中，只对是否存在错配作出了预测，也给处理错配的稳定性，但是错配对 PCR 的影响还需要根据错配是否发生在引物的 3' 来判断，3' 发生的错配要比其他位置的错配对引物的引发效率的影响大的多，所以在使用中应当结合着错配的具体配对情况来综合分析，这方面程序未能进行量化，只有通过用户根据经验判断；另外，程序对于产物的 Tm 值和引物 Tm 值之间的差异也为能给出评价，如果二者差异较大（如 >22℃），则容易造成退火时引物—模板和模板—模板退火之间的竞争。而在结果输出方面，该程序也只能以图文格式打印，无法转换成文本格式。